

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

协同随机森林方法和无人机LiDAR空谱数据的盐沼植被“精灵圈”识别

韩江涛, 谭凯, 张卫国, 周若彤, 刘帅

引用本文:

韩江涛, 谭凯, 张卫国, 等. 协同随机森林方法和无人机LiDAR空谱数据的盐沼植被“精灵圈”识别[J]. 光电工程, 2024, 51(3): 230188.

Han J T, Tan K, Zhang W G, et al. Identification of salt marsh vegetation "fairy circles" using random forest method and spatial-spectral data of unmanned aerial vehicle LiDAR[J]. *Opto-Electron Eng*, 2024, 51(3): 230188.

<https://doi.org/10.12086/oe.2024.230188>

收稿日期: 2023-07-27; 修改日期: 2023-11-20; 录用日期: 2023-11-23

相关论文

基于直接散射光谱的多环境要素激光遥感方法

郭圆新, 梁琨, 徐杨睿, 赵延鹏, 王元庆, 王蓓

光电工程 2024, 51(3): 240003 doi: 10.12086/oe.2024.240003

合成孔径激光雷达成像技术研究进展

徐晨, 晋凯, 魏凯

光电工程 2024, 51(3): 240007 doi: 10.12086/oe.2024.240007

结合遥感卫星及深度神经决策树的夜间海雾识别

李涛, 金炜, 符冉迪, 李纲, 尹曹谦

光电工程 2022, 49(9): 220007 doi: 10.12086/oe.2022.220007

基于改进YOLOv5s的无人机图像实时目标检测

陈旭, 彭冬亮, 谷雨

光电工程 2022, 49(3): 210372 doi: 10.12086/oe.2022.210372

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



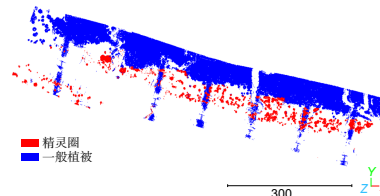
OE_Journal



Website

DOI: 10.12086/oe.2024.230188

协同随机森林方法和无人机 LiDAR 空谱数据的盐沼植被 “精灵圈”识别



韩江涛, 谭凯*, 张卫国, 周若彤, 刘帅

华东师范大学河口海岸学国家重点实验室, 上海 200241

摘要: “精灵圈”作为一种典型的空间自组织结构, 对盐沼植被生态系统及其功能有重要影响。获取“精灵圈”的空间格局及时空变化, 可为厘清其生态演化机理提供重要科学支撑。本文基于随机森林机器学习方法, 结合无人机激光雷达 (LiDAR) 点云的空间信息与光谱信息, 对盐沼植被“精灵圈”进行智能识别与提取。首先, 利用激光雷达方程和 Phong 模型, 消除距离、入射角以及镜面反射效应对强度数据的影响, 并且通过校正后强度数据滤波分离植被点云与地面点云。然后, 构造系列空间特征及几何变量, 利用随机森林算法, 对植被点云中的正常植被和“精灵圈”进行分类。结果表明: 该方法无需人工经验设置参数, 能够精确地从无人机 LiDAR 三维点云数据中快速自动识别“精灵圈”, 总体精度为 83.9%。本文为“精灵圈”时空分布反演提供了一种高精度的方法, 也为基于机器学习的三维点云数据处理提供了技术借鉴。

关键词: 随机森林; 无人机; 激光雷达; 空间自组织; 点云分类**中图分类号:** P237**文献标志码:** A

韩江涛, 谭凯, 张卫国, 等. 协同随机森林方法和无人机 LiDAR 空谱数据的盐沼植被“精灵圈”识别 [J]. 光电工程, 2024, 51(3): 230188

Han J T, Tan K, Zhang W G, et al. Identification of salt marsh vegetation "fairy circles" using random forest method and spatial-spectral data of unmanned aerial vehicle LiDAR[J]. *Opto-Electron Eng*, 2024, 51(3): 230188

Identification of salt marsh vegetation "fairy circles" using random forest method and spatial-spectral data of unmanned aerial vehicle LiDAR

Han Jiangtao, Tan Kai*, Zhang Weiguo, Zhou Ruotong, Liu Shuai

State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200241, China

Abstract: Spatial self-organization is a common phenomenon in many natural ecosystems. The "fairy circle" is a typical spatial self-organization structure that has significant impacts on the ecological functions of the salt marsh vegetation ecosystems. Obtaining the spatial pattern and spatiotemporal changes of the "fairy circle" can provide important scientific support for clarifying its ecological evolution mechanism. In this study, a machine learning method based on random forest is used to intelligently identify and extract the "fairy circle" in salt marsh vegetation

收稿日期: 2023-07-27; 修回日期: 2023-11-20; 录用日期: 2023-11-23

基金项目: 国家自然科学基金资助项目 (4217010220, 41901399); 上海市科学技术委员会资助项目 (22ZR1420900, 20DZ1204700); 重庆市自然科学基金项目 (CSTB2022NSCQ-MSX1254); 测绘遥感信息工程湖南省重点实验室开放基金 (E22335)

*通信作者: 谭凯, ktan@sklec.ecnu.edu.cn。

版权所有©2024 中国科学院光电技术研究所

using the spatial-spectral information from unmanned aerial vehicle (UAV) LiDAR. First, the effects of the distance, incident angle, and specular reflection on intensity data are eliminated using the laser radar equation and the Phong model. Second, the corrected intensity data are filtered to separate the vegetation from the ground. Third, a series of spatial features and geometric variables are used to classify the normal vegetation and "fairy circles" using the random forest algorithm. The results demonstrate that the proposed method can accurately extract "fairy circles" from UAV LiDAR 3D point cloud data without requiring manual experience-based parameter settings. The overall accuracy of the proposed method is 83.9%, providing a high-precision method for the spatiotemporal distribution inversion of "fairy circles" and technical references for 3D point cloud data processing based on machine learning.

Keywords: random forest; unmanned aerial vehicle; LiDAR; spatial self-organization; point cloud classification

1 引言

为了应对滩涂淤涨减缓、外来物种入侵、滩涂生态系统破碎和功能退化等多重挑战,以及充分发挥消浪滞沙、固堤保滩和防洪减灾等生态防护价值,盐沼植被会自发地触发“空间自组织”现象,即其个体单元会通过短距离内的相互作用自发地形成比个体尺度大很多倍的空间有序圆状/环状规则结构,即“精灵圈”^[1-2]。“精灵圈”广泛存在于许多生态系统,并被认为可以调节生态系统的功能,如生产力和生态修复能力,对于研究盐沼植被的生长、演替、扰动、入侵、退化、恢复等过程具有重要的科学价值^[3-4]。因此,获取“精灵圈”的总体数量、尺寸大小以及空间分布,可为厘清其分布规律及演化机理提供重要科学支撑。

传统的植被调查方法(样方调查、实地测量、预测估算、模型模拟等)存在工作量大、自动化水平差、观测范围有限、研究尺度小、数据不连续等缺陷,无法满足大范围“精灵圈”时空监测的需求。此外,盐沼植被位于近海前缘地带,出露和淹没交错发生,滩面泥泞,潮沟纵横,水文气象多变,生境复杂,对台站设立、现场监测、设备仪器等构成多重挑战。因此,传统的植被调查方法难以实现“精灵圈”的时空监测及演化机理研究。目前盐沼植被监测常用的遥感数据源有光学卫星/航空影像和微波雷达影像等^[5-6]。常规的光学影像空间分辨率最高只能达到米级,且只能提供盐沼植被冠层的二维信息,无法满足对“精灵圈”三维空间结构的精细观测与分析。此外,光学遥感影像受太阳光照、天气状况影响较为严重,高质量影像数据的获取较为困难。合成孔径雷达虽然具有全天候、受云雾影响小、穿透性强等优势,但其后向散射系数在植被各类参数估算中存在诸多问题(饱和效应、下垫面干扰等)。

激光雷达遥感(Light detection and ranging, LiDAR)作为一种新型、快速、实时的三维空间信息获取手段,具有主动性、植被穿透性、三维探测、空谱同步获取、不受环境光照条件限制等优势^[7-8],可以高密度、高分辨率、高效率地获取植被和地面的几何(空间)信息和强度(光谱)信息^[9-11]。利用LiDAR高精度点云数据能够真实地反映盐沼植被的空间形态和物理特性,从而实现“精灵圈”的精准提取和参数反演。根据搭载平台不同,LiDAR可以分为星载、机载以及地面。星载LiDAR观测尺度大,适用于全球或区域尺度的观测,但是其光斑和采样间距非常大,无法实现精细尺度“精灵圈”的观测。地面LiDAR成本低、精度高、作业灵活,但其观测视野较小,受遮挡较为严重,且难以在淤泥质滩涂上进行仪器架设。利用无人机搭载LiDAR传感器,能同时兼顾数据采集效率、数据精度和空间尺度等各类需求,是“精灵圈”观测的理想手段。

但是,使用无人机LiDAR技术进行“精灵圈”提取仍面临诸多技术难点。首先,强度信息校正存在较大困难,由于潮滩独特的光谱反射特性及无人机数据获取的动态性,常规的强度校正方法往往难以取得良好的校正结果;其次,三维点云数据量巨大,且数据是非结构化的,常规点云数据处理方法通常需要手动调整参数,普适性较差,在处理海量三维点云数据时极为耗时;再次,正常植被和“精灵圈”植被的物种组成和光谱特性几乎没有差别,仅依据强度数据无法实现二者的精确识别,二者的区别主要来源于几何形态与空间分布(例如:点云密度、垂直度等)方面的细微差异,但常规点云数据处理方法难以实现这些细微差异的精准探测。机器学习方法能够从数据中选取合适的特征,自动地归纳逻辑和规则,并根据归纳的结果与新数据来进行预测,实现不同类别数据之间细微

特征的识别, 在提高三维点云数据处理的智能性和效率方面具有巨大优势。本文基于无人机航高和 GPS 时间恢复了点的瞬时观测几何构造, 利用激光雷达方程和 Phong 模型消除了距离、入射角以及镜面反射效应对强度数据的影响, 解决了强度数据的校正难题, 实现了植被点云的精准滤波。在此基础上, 构造系列空间特征及几何变量, 利用随机森林 (random forest, RF) 机器学习算法, 避免了三维点云数据常规处理方法的复杂计算和参数设置, 实现了一种正常植被和“精灵圈”的智能识别方法, 快速、高精度地对“精灵圈”进行提取, 极大地提高了算法运行效率和普适性, 为“精灵圈”时空分布研究及其生态机制解译提供了技术支撑。

2 研究区域与数据获取

本文选取上海市崇明岛东滩鸟类自然保护区北部太阳村西北侧一淤泥质潮间带为研究区域 (图 1)。该区域盐沼植被主要包括芦苇、互花米草等。本文选用的无人机平台为 ZR-M66, 搭载的传感器为“蜂鸟”Genius 多线激光雷达。“蜂鸟”激光雷达采用双回波技术, 视场角为 $360^{\circ} \times 30^{\circ}$, 最大扫描距离为 250 m, 激光波长为 905 nm。在数据采集过程中, 无人机飞行高度为 80 m, 飞行速度为 7 m/s, 条带宽度为 120 m,

相邻条带的旁向重叠率为 40%。最终共获取约 2.9 亿个点, 总体点云密度约为 110 pt/m^2 , 垂直和水平精度分别约为 0.15 m 和 0.10 m。 x 、 y 坐标采用 WGS84 坐标系, 高程 z 采用大地坐标。原始 LAS 格式点云数据中主要包括三维坐标、原始强度以及 GPS 时间等信息。由于“精灵圈”主要分布于光滩与植被边缘区域, 因此本文选择实验区域中间的光滩及周边植被区域作为研究区域 (图 1 橙黄色矩形框)。为了对“精灵圈”提取结果进行验证分析, 本文实验中还利用 ZR-M66 搭载 Sony RX1R II 相机进行了实验区域正射影像的获取。

3 研究方法

如图 2 所示, 本文所采用的方法主要分为模型计算和模型应用两部分。模型计算包括强度校正、几何特征计算、强度滤波、手动标注训练区域样本等步骤, 最终实现对随机森林模型的构建与优化。模型应用是运用前一步骤得出的模型, 对其他区域进行分类, 并计算其分类精度。本文方法最重要的步骤包含强度校正与随机森林模型的构建。其中, 校正强度数据既对点云滤波的精度起着关键性的作用, 同时也是随机森林模型的重要输入特征之一, 对随机森林模型的分精度亦有影响。随机森林模型的构建需要前续步骤对

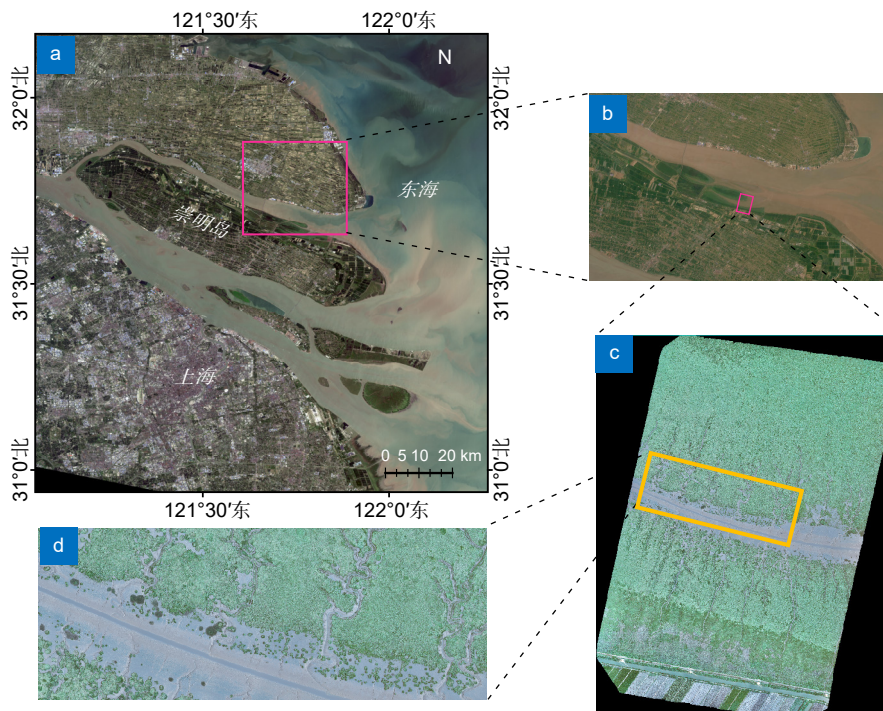


图 1 (a, b) 研究区域概况; (c, d) 研究区域正射影像

Fig. 1 (a, b) Overview of the study area; (c, d) Orthophoto of the study area

数据的准确处理, 还需要对模型参数不断调整, 以达到最佳的分类效果。下面将对这两个步骤详细描述。

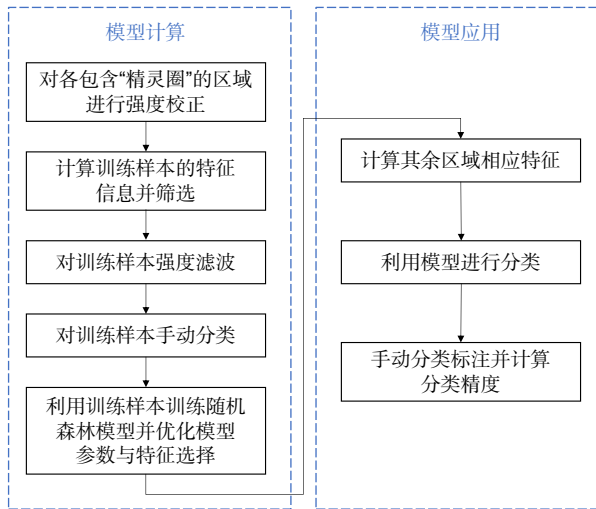


图 2 基于随机森林的“精灵圈”和正常植被识别技术路线
Fig. 2 "Fairy circle" and normal vegetation identification technology route based on random forest algorithm

3.1 强度滤波

为了提取“精灵圈”, 首先需要分离无人机 LiDAR 点云数据中的地面点与植被点。潮间带地形较为平坦, 植被覆盖密集, 导致现有基于点云空间信息的滤波方法难以准确提取植被点。由于植被与滩涂淤泥底质物质组成不同, 其在光谱反射特性上会表现出较大差异; 此外, 潮滩淤泥底质含水量极高, 水对近红外激光吸收作用强, 进一步加剧了植被与潮滩的反射特性差异^[12]。因此, 可使用无人机 LiDAR 的强度数据提取植被点。但是, 原始强度数据受到距离和入射角的强烈影响; 此外, 由于无人机 LiDAR 扫描角度较小, 淤泥含水量较高, 镜面反射较强。本文利用文献 [13] 中提出的模型对无人机 LiDAR 强度数据进行校正 (式 (1))。该模型结合激光雷达方程和 Phong 模型, 综合考虑了距离、入射角及镜面反射效应对强度数据的影响。利用该模型得到较为精确的校正后强度, 不仅能较好地分离地面点和植被点, 而且能为本文方法后续随机森林模型提供较为准确的强度输入特征, 提高随机森林模型预测精度。

$$I_s = \begin{cases} \left(I \cdot \frac{R^2}{R_s^2} - K_0 K_s \cos^n(2\theta) \right) \cdot \frac{\cos \theta_s}{\cos \theta}, & \theta \leq 45^\circ \\ I \cdot \frac{R^2}{R_s^2} \cdot \frac{\cos \theta_s}{\cos \theta}, & \theta > 45^\circ \end{cases}, \quad (1)$$

式中: I 为原始强度, I_s 为校正强度, R 和 R_s 分别代表距离和自定义的标准距离, θ 与 θ_s 分别代表入射角和

自定义的标准入射角, K_s 为镜面反射百分比, K_0 为常数 (与目标反射率和系统特性相关), n 为物体表面的光泽度系数。在本文中, 取 $R_s = 100$, $\theta_s = 0$ 。模型的具体推导及参数估计方法详见文献 [13]。

由于提供的原始点云文件中不包含距离与入射角, 所以强度校正之前还需要对观测几何 (扫描距离与入射角) 进行计算。相关研究表明, 在飞行高度一定且地形相对平坦的情况下, 扫描角可以近似代替入射角^[14]。本文采用了一种基于点云 GPS 时间和无人机飞行高度的距离与扫描角计算方法。具体原理为: 按照一定时间间隔将原始点云数据分段, 用该段时间内的中心点近似表示无人机在地面投影的位置, 再根据飞行高度则可以计算扫描瞬间距离和扫描角^[13]:

$$\begin{cases} R_i = \sqrt{D_i^2 + H^2} \\ \theta_i = \arctan \frac{D_i}{H} \end{cases}, \quad (2)$$

式中: R_i 为估算的 i 点的瞬时扫描距离, θ_i 为估算的 i 点的瞬时扫描角, D_i 为无人机地面投影点与 i 点的距离, H 为飞行高度。

3.2 “精灵圈”提取

通过强度数据 K 均值滤波去除地面点后, 正常植被与“精灵圈”植被在空间上相互分离。常规的“精灵圈”提取方法通过先验知识设置阈值和参数, 直接对点云数据进行三维空间聚类等处理^[13], 运行非常耗时, 且参数和阈值普适性较差。机器学习方法能够概括出超出训练样本的数据, 这为利用该方法提取“精灵圈”提供了较好的普适性。随机森林是一种集成学习方法, 用于解决分类和回归问题, 最早由 Breiman 提出^[15]。集成学习是一种通过集成多个模型来解决某一问题, 以此提高准确性的机器学习方法, 目前广泛使用的集成学习方法是 Boosting^[16] 和 Bagging^[17], 这两者的区别在于集成的方式是并行还是串行。随机森林就是通过集成学习的 Bagging 思想将多个决策树集成为“森林”的一种算法, 其基本思想是^[18]: 首先, 利用 bootstrap 抽样从原始训练集抽取 n 个样本; 然后, 对 n 个样本分别建立决策树, 得到 n 种分类结果; 最后, 投票表决出最优分类 (图 3)。其中分类决策可表示为

$$H(x) = \arg \max_Y \sum_{i=1}^n F(h_i(x) = Y), \quad (3)$$

式中: $H(x)$ 代表组合分类模型, $h_i(x)$ 为单个决策树的分类模型, Y 表示目标变量, $F()$ 为示性函数。

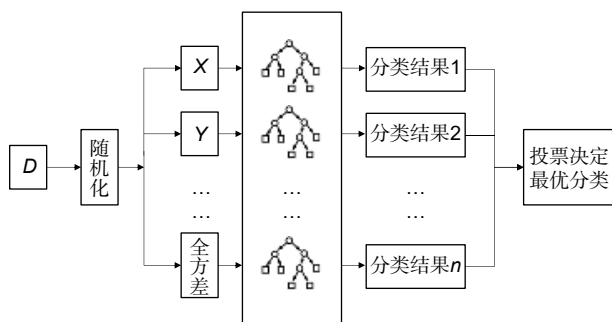


图 3 随机森林示意图

Fig. 3 Schematic diagram of the random forest algorithm

自提出以来, 随机森林被广泛应用于分类、预测、变量选择和异常值监测等方面, 其优点可以概括为^[19]: 1) 可处理多维数据, 且无需降维, 也无需进行特征选择; 2) 可在不增加预测偏差的前提下减小误差; 3) 可判断不同特征之间的相互影响; 4) 对异常值和噪声具有鲁棒性; 5) 与其他集成学习方法相比计算效率更高。因此, 本文利用随机森林, 实现正常植被和“精灵圈”植被的分类和识别, 具体方法介绍如下。

如表 1 所示, 除点云的三维空间坐标(X,Y,Z)和校正后强度 I_c 之外, 本文还选取了 7 个不同的点云几何特征用于构造随机森林模型, 包括: 粗糙度 (roughness)、密度 (density)、全方差 (ominvariance)、特征熵 (eigenentropy)、各向异性 (anisotropy)、垂直度 (verticality)、第三特征值 (3rd eigenvalue), 其中: L_1 、 L_2 和 L_3 分别为点的第一、二、三特征值。已有研究表明, 这些几何特征能够较好地反映“精灵圈”和普通植被之间的细微差异^[20-21]。如图 2 所示, 利用随机森林实现正常植被和“精灵圈”分类的具体步骤为: 1) 选取“精灵圈”植被的训练区域, 计算该区域的 11 个特征 (表 1); 2) 从该区域有放回地随机抽取 70% 的点作为训练集用来训练模型, 随机森林算法会通过上述的过程选择出当前最优分类模型, 30% 的点作为验证集, 评估模型的一般错误率, 并基于此调整参数, 最终获得最优模型; 3) 将研究区域中其他区域作为测试集, 进行“精灵圈”提取, 并计算其精度。

4 结果与讨论

4.1 实验结果与分析

研究区域原始强度和式 (1) 校正强度如图 4 所示。由镜面反射引起的条带效应得到了较好的校正, 植被点与地面点校正强度存在显著差异, 因此利用 I_c 可以

较好地提取植被点。在对校正后强度数据进行 K-means 聚类后, 仍有部分地面点被误分为植被点, 这是 K-means 聚类本身的弊端导致的, 但这部分地面点并不会影响“精灵圈”提取的效果, 因为通过随机森林算法中选取的特征可以将这些点进一步剔除。

表 1 随机森林点云特征

Table 1 Point cloud features in random forests

特征	含义
X	点云数据x坐标
Y	点云数据y坐标
Z	点云数据z坐标
校正强度	式(1)校正后的强度值
粗糙度	该点与最近邻最佳拟合平面之间的距离
密度	指定半径内包含的邻域点数量
全方差	$\sqrt[3]{L_1 \cdot L_2 \cdot L_3}$
特征熵	$-(L_1 \cdot \ln(L_1) + L_2 \cdot \ln(L_2) + L_3 \cdot \ln(L_3))$
各向异性	$L_1 - L_3 / L_1$
垂直度	$1 - Z, e_3 $
第三特征值	L_3

在“精灵圈”提取时, 首先, 选取图 4(a) 所示的训练区域 (红色矩形框) 进行模型训练。该区域共包含 333041 个点, “精灵圈”几何特征与空间分布典型, 利用 70% 的点进行模型训练后, 将模型分类的结果与手动分类的结果进行对比, 得到训练集的分类精度为 99.90%, 分类效果极佳。接下来, 将模型应用于训练区域的验证集, 得到验证集的分类精度为 99.83%, 误差曲线收敛, 混淆矩阵中各类别的分类精度较高, 无需对参数进行过多调整。最后, 将模型应用于其他研究区域, 进行“精灵圈”的提取。最终提取结果如图 5(b) 和表 2 所示。本文提取的总体精度为 83.90%。在 441 个“精灵圈”中, 漏分 63 个, 漏分率为 14.29%, 误分 8 个, 误分率为 1.81%。本文方法在研究区域内大部分局部区域的分类精度都高于 90%, 但最左侧区域由于“精灵圈”分布格局特殊, 有较多“精灵圈”存在于密集植被区域, 而在训练集中并不存在这种分布格局的“精灵圈”, 因此模型的识别精度较差, 仅为 61.76%。在此区域的 102 个“精灵圈”中, 漏分 37 个, 漏分率为 36.27%, 误分 2 个, 误分率为 1.96%。误分存在两种情况, 其一是将密集植被区域边缘误识别为“精灵圈”, 另一种情况是误将与“精灵圈”几何特征与辐射特征相似的区域识别为“精灵圈”。

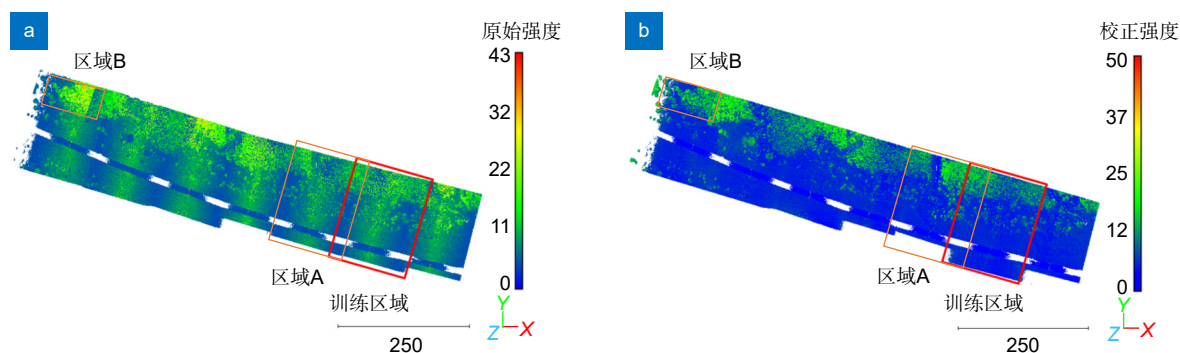


图 4 (a) 原始强度; (b) 校正强度。其中: 红色矩形框代表本文模型训练区域, 橙色矩形框代表替换的其他训练区域
Fig. 4 (a) Original intensity, and (b) corrected intensity, where the red box represents the training set and the orange box represents the replacement training set

4.2 消融实验

为了进一步说明本文所选特征的重要性和必要性, 分别在模型训练过程中去除空间坐标、校正强度和几何参量这三大类特征, 通过消融实验测试其分类精度。如表 3 所示, 去除任意类型的特征后提取精度均有所下降。

1) 去除几何参量后分类精度下降最多, 训练集分类精度约为 93.57%, 验证集分类精度约为 90.65%, 其中 70.1% 的“精灵圈”点云被错分, 这主要是由于去除几何特征后, 参与模型训练的特征过少, 无法正确识别“精灵圈”导致模型不够精确。

2) 去除空间坐标后训练集分类精度约为 96.71%, 验证集分类精度约为 93.67%, 通过混淆矩阵可知约 44.5% 的“精灵圈”点云被误分为普通植被, 在模型的推广过程中其错分比例将进一步扩大, 因此空间坐标特征必不可缺。

3) 去除校正强度后分类精度下降最少, 训练集和验证集分类精度分别为 98.64% 和 96.51%, 其中约 25.7% 的“精灵圈”点云被误分为普通植被。虽然校正强度数据已经用于植被点和地面点滤波, 但在随机森林模型中对“精灵圈”和普通植被分类仍有一定贡献。

消融实验去除三类特征后得到的模型应用于整个区域的提取结果分别如图 5(c-e) 所示。

4.3 不同方法对比实验

在利用校正后强度得到植被点云之后, 文献 [13] 中利用三维点云连通性聚类以及线性、尺寸和圆度特征进行了“精灵圈”的提取, 其结果如图 5(a) 所示和表 3 所示。该方法的提取精度为 77.55%, 在 441 个“精灵圈”中, 漏分 83 个, 漏分率为 18.82%, 误分 16 个, 误分率为 3.63%。由于三维点云连通性聚类涉及到点

的逐点判断, 该算法在包含正常植被和“精灵圈”植被的点云中直接运行, 非常耗时。虽然该方法也可以较好地提取精灵圈, 但是在植被离群分布的区域, 存在部分漏分的情况, 这是由于聚类算法并不能完全保证将一个完整的“精灵圈”分为一类, 从而导致部分“精灵圈”点云无法被正确识别, 且该误差会影响后续的提取精度, 造成误差的累积。该方法还会将部分与“精灵圈”几何形态相似的普通植被或斑块误分为“精灵圈”, 这是因为该方法仅通过圆度、几何大小以及线性程度识别“精灵圈”, 而没有使用强度信息进一步区分。此外, 该方法直接在三维空间对点云数据进行运算处理, 计算量巨大, 需要依据先验知识手动设置参数阈值, 具有较强的局限性与地域性。与之相比, 本文方法只需对数据计算相应特征后输入已训练的模型, 即可实现“精灵圈”的提取, 具有较好的普适性, 且大大缩短了运行时间。但需要注意的是, 在本方法中, 手动标注的准确性、特征的选取以及训练样本的选择对该方法的提取精度有较大的影响, 因此, 为了达到良好的提取效果, 需要选取具有区分度的特征和特征明显的区域进行模型训练。

本文选取图 4(a) 所示红色矩形框区域作为训练区, 该区域“精灵圈”与普通植被的几何特征与空间分布较为典型。当选择其他区域为训练区域后, 精度有可能会发生变化。1) 当所选训练区域包含足够典型的“精灵圈”和普通植被时(如图 4 区域 A), 训练得到的随机森林模型精度与本文所选区域相当, 所训练出的模型可以推广应用到整个研究区域。2) 当所选区域不包含典型的“精灵圈”和普通植被时(如图 4 区域 B), 训练得到的随机森林模型精度有可能降低, 也难以推广到整个研究区域“精灵圈”的提取。

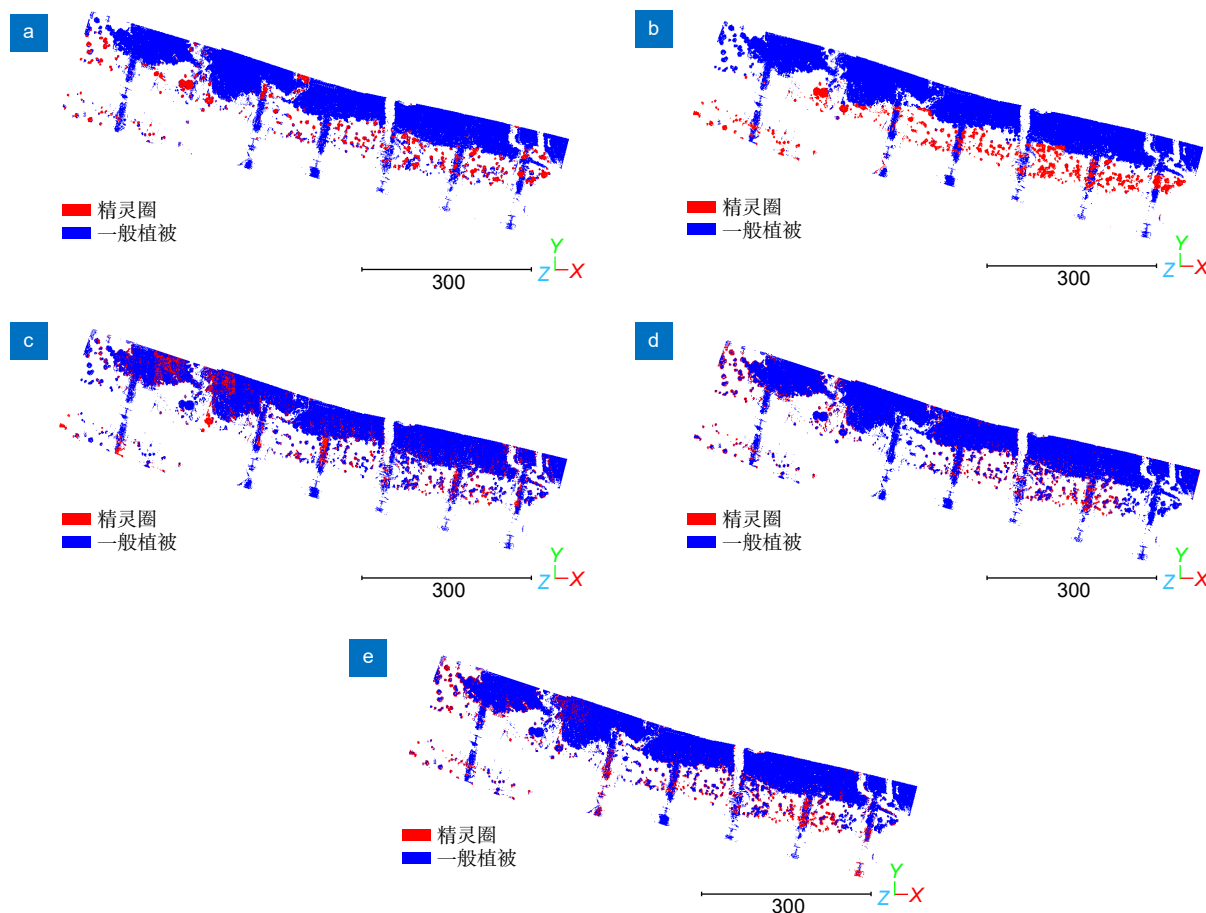


图 5 (a) 文献 [13] 方法提取结果; (b) 本文方法提取结果; (c) 去除几何参量提取结果; (d) 去除空间坐标提取结果; (e) 去除校正强度提取结果

Fig. 5 (a) Extracted result obtained by reference [13]; (b) Extracted result obtained by the proposed method; (c) Extracted result obtained by the proposed method without geometric features; (d) Extracted result obtained by the proposed method without spatial coordinates; (e) Extracted result obtained by the proposed method without corrected intensity

表 2 本文方法与文献 [13] 方法对比

Table 2 Results comparison between the proposed method and that in reference [13]

方法	总体精度/%	漏分率/%	误分率/%
随机森林方法	83.90	14.29	1.81
文献[13]方法	77.55	18.82	3.63

表 3 消融实验结果

Table 3 Results of the ablation study

消融实验	训练集精度/%	测试集精度/%
去除几何参量	93.57	90.65
去除空间坐标	96.71	93.67
去除校正强度	98.64	96.51

5 结论与展望

本文基于无人机 LiDAR 数据和随机森林模型, 提出了一种盐沼植被“精灵圈”的快速高精度识别方法。该方法的总体精度为 83.90%, 漏分率为 14.29%, 误分率为 1.81%。研究表明, 无人机 LiDAR 技术在“精灵圈”时空多维监测中具有巨大潜力。在未来的研究中可以考虑进一步优化强度校正算法, 以解决本研究中强度滤波的分离效果欠佳的问题。此外, 在利用随机森林算法时, 可以选取更加具有代表性的

特征用以训练模型, 以此来提高提取精度。多源遥感数据的融合是遥应用的趋势, 融合 LiDAR 点云数据与遥感影像数据, 实现对“精灵圈”的联合提取、属性分析、动态监测, 同样是未来研究的重大课题。

参考文献

[1] Tarnita C E, Bonachela J A, Sheffer E, et al. A theoretical foundation for multi-scale regular vegetation patterns[J]. *Nature*, 2017, 541(7637): 398-401.
 [2] Zhao L X, Zhang K, Siteur K, et al. Fairy circles reveal the resilience of self-organized salt marshes[J]. *Sci Adv*, 2021, 7(6):

- eabe1100.
- [3] Rietkerk M, Bastiaansen R, Banerjee S, et al. Evasion of tipping in complex systems through spatial pattern formation[J]. *Science*, 2021, **374**(6564): eabj0359.
- [4] Ruiz-Reynés D, Gomila D, Sintés T, et al. Fairy circle landscapes under the sea[J]. *Sci Adv*, 2017, **3**(8): e1603262.
- [5] Shi L, Liu Q S, Huang C, et al. Mapping quasi-circular vegetation patch dynamics in the Yellow River Delta, China, between 1994 and 2016[J]. *Ecol Indic*, 2021, **126**: 107656.
- [6] Liu Q S, Huang C, Liu G H, et al. Comparison of CBERS-04, GF-1, and GF-2 satellite panchromatic images for mapping quasi-circular vegetation patches in the Yellow River Delta, China[J]. *Sensors*, 2018, **18**(8): 2733.
- [7] Yang B S, Liang F X, Huang R G. Progress, challenges and perspectives of 3D LiDAR point cloud processing[J]. *Acta Geod Cartogr Sin*, 2017, **46**(10): 1509–1516.
杨必胜, 梁福逊, 黄荣刚. 三维激光扫描点云数据处理研究进展、挑战与趋势[J]. *测绘学报*, 2017, **46**(10): 1509–1516.
- [8] Li Y M, Guo Q H, Wan B, et al. Current status and prospect of three-dimensional dynamic monitoring of natural resources based on LiDAR[J]. *Natl Remote Sens Bull*, 2021, **25**(1): 381–402.
李玉美, 郭庆华, 万波, 等. 基于激光雷达的自然资源三维动态监测现状与展望[J]. *遥感学报*, 2021, **25**(1): 381–402.
- [9] Liang X L, Kankare V, Hyypää J, et al. Terrestrial laser scanning in forest inventories[J]. *ISPRS J Photogramm Remote Sens*, 2016, **115**: 63–77.
- [10] Xie W M, Guo L C, Wang X Y, et al. Detection of seasonal changes in vegetation and morphology on coastal salt marshes using terrestrial laser scanning[J]. *Geomorphology*, 2021, **380**: 107621.
- [11] Yang B S, Dong Z. Progress and perspective of point cloud intelligence[J]. *Acta Geod Cartogr Sin*, 2019, **48**(12): 1575–1585.
杨必胜, 董震. 点云智能研究进展与趋势[J]. *测绘学报*, 2019, **48**(12): 1575–1585.
- [12] Jin J L, Verbeurgt J, De Sloover L, et al. Support vector regression for high-resolution beach surface moisture estimation from terrestrial LiDAR intensity data[J]. *Int J Appl Earth Obs Geoinf*, 2021, **102**: 102458.
- [13] Tao P J, Tan K, Ke T, et al. Recognition of ecological vegetation fairy circles in intertidal salt marshes from UAV LiDAR point clouds[J]. *Int J Appl Earth Obs Geoinf*, 2022, **114**: 103029.
- [14] Yan W Y, Shaker A. Radiometric correction and normalization of airborne LiDAR intensity data for improving land-cover classification[J]. *IEEE Trans Geosci Remote Sens*, 2014, **52**(12): 7658–7673.
- [15] Breiman L. Random forests[J]. *Mach Learn*, 2001, **45**(1): 5–32.
- [16] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. *J Comput Syst Sci*, 1997, **55**(1): 119–139.
- [17] Breiman L. Bagging predictors[J]. *Mach Learn*, 1996, **24**(2): 123–140.
- [18] Fang K N, Wu J B, Zhu J P, et al. A review of technologies on random forests[J]. *Stat Inf Forum*, 2011, **26**(3): 32–38.
方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. *统计与信息论坛*, 2011, **26**(3): 32–38.
- [19] Sheykhmousa M, Mahdianpari M, Ghanbari H, et al. Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review[J]. *IEEE J Sel Top Appl Earth Obs Remote Sens*, 2020, **13**: 6308–6325.
- [20] Hackel T, Wegner J D, Schindler K. Contour detection in unstructured 3D point clouds[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 2016: 1610–1618. <https://doi.org/10.1109/CVPR.2016.178>.
- [21] Liu G D, Liu J, Liu L. A mountain road extraction method based on airborne LiDAR data[J]. *Laser Technol*, 2022, **46**(4): 466–473.
刘国栋, 刘佳, 刘浪. 一种基于机载LiDAR数据的山区道路提取方法[J]. *激光技术*, 2022, **46**(4): 466–473.

作者简介



韩江涛(2000-), 男, 硕士研究生, 研究方向为海岸带遥感应用。

E-mail: 51253904012@stu.ecnu.edu.cn



【通信作者】谭侃(1987-), 男, 副研究员, 研究方向为海岸带多源遥感数据处理及应用。

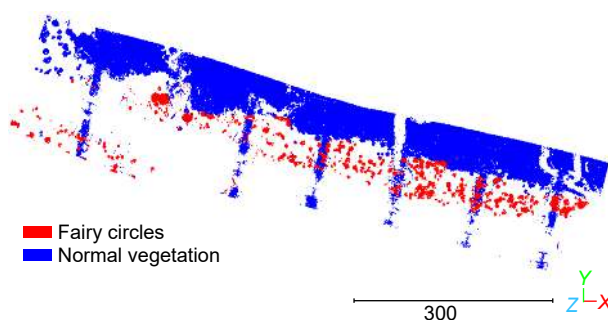
E-mail: ktan@sklec.ecnu.edu.cn



扫描二维码, 获取PDF全文

Identification of salt marsh vegetation "fairy circles" using random forest method and spatial-spectral data of unmanned aerial vehicle LiDAR

Han Jiangtao, Tan Kai*, Zhang Weiguo, Zhou Ruotong, Liu Shuai



The extracted result obtained by the proposed method

Overview: Spatial self-organization is a fascinating and widespread phenomenon observed in various natural ecosystems. One such intriguing structure is the "fairy circle", known for its significant influence on the functioning of salt marsh vegetation ecosystems. "Fairy circles" are known to play a crucial role in shaping salt marsh vegetation ecosystems, and their identification and understanding can offer valuable insights into ecological processes and dynamics. Understanding and identifying these "fairy circles" is of utmost importance for ecological research and conservation efforts. To address this, the present study employs a sophisticated machine learning technique called random forest to intelligently identify and extract "fairy circles" within salt marsh vegetation using data from unmanned aerial vehicle (UAV) LiDAR point clouds. The initial step in this research involves addressing potential complications arising from distance, incident angle, and specular reflection effects on the intensity data obtained from the UAV LiDAR. By applying the laser radar equation and the Phong model, these confounding factors are successfully eliminated to obtain the corrected intensity data. A filtering process is employed on the corrected intensity data to separate the vegetation from the ground points. To effectively distinguish between the normal vegetation and the "fairy circles," a set of spatial features and geometric variables are employed, and a random forest model is constructed using these features and variables. The results demonstrate the extraordinary capability of the proposed method to accurately identify and extract "fairy circles" from UAV 3D point cloud data, achieving an overall accuracy rate of 83.9%. The study represents a groundbreaking advancement in the study of "fairy circles" and paves the way for spatiotemporal distribution inversion of these intriguing structures. Additionally, the application of machine learning, particularly the random forest algorithm, in combination with UAV LiDAR technology, demonstrates the potential of artificial intelligence and remote sensing in ecological research. The implications of this research extend beyond salt marsh ecosystems. The methodological approach presented here can be adapted and applied to other natural ecosystems with spatial self-organization phenomena. By integrating machine learning and advanced remote sensing techniques, researchers can explore and comprehend various spatial structures, contributing to a deeper understanding of ecological patterns and processes on a broader scale.

Han J T, Tan K, Zhang W G, et al. Identification of salt marsh vegetation "fairy circles" using random forest method and spatial-spectral data of unmanned aerial vehicle LiDAR[J]. *Opto-Electron Eng*, 2024, 51(3): 230188; DOI: [10.12086/oe.2024.230188](https://doi.org/10.12086/oe.2024.230188)

Foundation item: National Natural Science Foundation of China (42171425, 41901399), Science and Technology Commission of Shanghai Municipality (22ZR1420900, 20DZ1204700), Chongqing Municipal Bureau of Science and Technology (CSTB2022NSCQ-MSX1254), and Hunan Provincial Key Laboratory of Geo-Information Engineering in Surveying, Mapping and Remote Sensing, Hunan University of Science and Technology (E22335)

State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200241, China

* E-mail: ktan@sklec.ecnu.edu.cn