

基于地面激光雷达点云数据的树种识别方法

王 佳^{1,2} 张隆裕^{1,2} 吕春东^{1,2} 牛利伟^{1,2}

(1. 北京林业大学精准林业北京市重点实验室, 北京 100083; 2. 北京林业大学测绘与 3S 技术中心, 北京 100083)

摘要: 为了能够更有效地利用地面激光雷达的点云数据识别树种,以北京林业大学为研究区域,利用 FARO Photon 120 型地面激光雷达在研究区内获取 4 个树种、共 92 棵树木的点云数据。依据点云的三维坐标值提取研究区内立木的胸径、枝下高、树高、冠高、最长冠幅、垂直最长方向冠幅 6 个测树因子,同时提取由测树因子组合而成具有鲁棒性的 6 个树形特征参数,包括冠长树高比、胸径树高比、冠高树高比、分枝角、冠长最大冠幅之比、最长冠幅与垂直方向冠幅之比。分别使用测树因子和组合特征参数,采用支持向量机、分类回归决策树和随机森林的方法,对树种进行冠幅自动识别。研究结果表明:使用测树因子树木识别方法,识别平均准确率为 0.765,平均召回率为 0.778,3 种识别方法中,分类效果较好的依次为分类回归决策树、随机森林、支持向量机;使用组合特征参数树木识别方法,识别平均准确率为 0.891,平均召回率为 0.896,分类效果较好的方法是随机森林和支持向量机,其次是分类回归决策树;总体上来看,不论是对于单个树种还是总体的准确率和召回率,组合特征参数法均高于测树因子法,而对于 3 种不同的分类方法,随机森林相对最好。研究结果表明,结合地面激光雷达获取的点云和不同机器学习分类方法进行树种识别分类可以达到满意的效果,且能节省大量时间和人力。

关键词: 树种识别; 地面激光雷达; 点云; 测树因子; 组合特征参数

中图分类号: S758 **文献标识码:** A **文章编号:** 1000-1298(2018)11-0180-09

Tree Species Identification Methods Based on Point Cloud Data
Using Ground-based LiDAR

WANG Jia^{1,2} ZHANG Longyu^{1,2} LÜ Chundong^{1,2} NIU Liwei^{1,2}

(1. Beijing Key Laboratory of Precision Forestry, Beijing Forestry University, Beijing 100083, China

2. Surveying and 3S Technology Center, Beijing Forestry University, Beijing 100083, China)

Abstract: The traditional tree species identification depends on time-consuming and labor-intensive efficiency of artificial field survey. In order to more effectively utilize the point cloud data identification tree of ground-based LiDAR, taking Beijing Forestry University as the research area, and FARO Photon 120 ground-based LiDAR was used to obtain point cloud data of a sample set of 92 trees, four tree species in the study area. According to the three-dimensional coordinate values of point cloud, the six tree-measuring factors of breast diameter, height of branches, height of tree, height of crown, width of crown, and the longest direction of vertical trees in the study area were extracted, and the extracted tree-measuring factors were combined. The robust tree features six parameters, namely crown length tree height ratio, DBH height ratio, crown height tree height ratio, branch angle, crown length ratio, maximum crown width and vertical direction. For the ratio of crown width, the tree species were automatically identified by using the tree-measuring factor and the combined feature point parameters to support the tree sample by using the support vector machine, the classification regression decision tree and the random forest. The results showed that for the tree identification method using tree-measuring factor, the average accuracy of recognition was 0.765, and the average recall rate was 0.778. Among the three identification methods, the best effect was classification regression decision, followed by random forest, and finally support vector. Using the combined feature parameter tree identification method, the average accuracy of recognition was 0.891, and the average recall rate was 0.896. The best method was

收稿日期: 2018-07-13 修回日期: 2018-09-18

基金项目: 国家自然科学基金项目(41401650)、北京市自然科学基金项目(8182038)和中央高校基本科研业务费专项资金项目(2015ZCQ-LX-01)

作者简介: 王佳(1980—),男,副教授,博士,主要从事激光雷达技术在林业应用研究,E-mail: Wangjia2002_0@163.com

random forest and support vector machine, followed by classification regression decision. In general, the combined feature parameter method had higher accuracy and recall rate of single tree species or overall than those of the tree-measuring factor method, random forests were relatively the best for three different classification methods. The research result showed that the tree species identification classification combining the point cloud obtained by ground-based LiDAR and different machine learning classification methods could achieve satisfactory results and save a lot of time and manpower.

Key words: tree species classification; ground-based LiDAR; point cloud; tree measuring factor; combined feature parameters

0 引言

森林在维持生态平衡中发挥着重大作用,为了保护森林资源,树种识别是其中重要一环,而传统意义上的人工实地调查方法费时费力,激光雷达作为一种主动遥感技术,近年来在国际上发展迅速,使得遥感技术用于树种分类和识别的效率大大提高^[1-3]。通过激光雷达进行数据扫描,运用点云数据软件处理,可以提取一系列基于激光雷达点云数据的森林参数和变量,进而实现树种识别。在提取森林参数方面,LIANG 等^[4]通过将分离出来的树干点云数据分割成切片,达到自动识别树干位置的目的。STRAHLER 等^[5]利用红外激光雷达获取林分影像,从中识别单木树高及胸径提取叶面积指数,效果良好。KOUKOULAS 等^[6]使用 LiDAR 和多光谱遥感数据对阔叶落叶林的单木位置、高度和树种进行制图,采样间隔约为 2 m。CLARKM 等^[7]用小光斑 LiDAR 系统估测热带雨林的地形高程和树高,采样密度为每平方米 9 个回波点。刘清旺等^[8]利用机载激光点云数据提出了一种双正切角树冠识别算法,获取树冠和树高数据。树种识别方面,余超^[9]使用分类回归树、神经网络和支持向量机算法进行研究区内主要风景林树种的分类,分类回归树的总体分类精度可以达到 87.1%,支持向量机算法为 76.91%,神经网络法只有 73.85%。魏田^[10]从三维点云数据提取单木信息,选用树干收缩度为树种分类的因素之一,基于核函数三次多项式的支持向量机的算法分类,并用单因素分类、双因素分类与多因素分类的正确率进行对比,得出单因素和双因素的分类精度低的结论。PLAZA 等^[11]从光谱特征角度出发,利用波谱角的不同对城区中的树种进行识别,分类效果较好;GEORGE 等^[12]基于 Hyperion 影像,采取支持向量机法和面向对象分类法对树种进行识别,结果显示,面向对象分类法的精度高于支持向量机法。文献[13-15]均是利用高光谱和机载雷达相融合的方法对复杂森林树种识别,发现总体分类精度和树种分类精度比单一高光谱数据有所提高。ALONZO 等^[16]利用激光雷达数据在基于冠层尺度

上对常见树种进行分类,分类精度提高了 4.2%。李永亮等^[17]建立了 BP 神经网络模型,输入高光谱特征参数,输出森林树种类别,分类综合精度可达 93.3%。

从上述研究可以看出,目前大多数树种识别研究主要集中于利用多光谱、高光谱遥感结合机载激光雷达,对大区域森林垂直结构的获取能力较强,但对林冠下层结构描述不详细,树种识别精度也有待提高,而地面激光雷达的研究仍主要集中在森林参数的提取方面,主要包括树高、冠幅、郁闭度和生物量等。本文通过地面激光雷达获取树木点云数据,利用机器学习较为主流的 3 种方法,通过对样本学习生成分类器,依照校园立木的形态参数对其进行树种的分类识别。

1 立木特征因子提取方法

1.1 胸径提取算法

胸径为树干距地面 1.3 m 处的直径。设单木点云中 Z 坐标最小的为 $Z_{\min}$,在点云数据中检索出 Z 坐标为 $Z_{\min} + 1.3$ m 的所有点,设为集合 P 。求出 P 中所有点的重心 O ,对于 $\forall P_i \in P$,求其到 O 的距离,最大的距离则为胸径 D_{BH} (Diameter at breast height),单位为 cm。

$$D_{BH} = \max \{ \sqrt{(X_{P_i} - X_O)^2 + (Y_{P_i} - Y_O)^2}, P_i \in P \} \quad (1)$$

1.2 枝下高提取算法

枝下高位置是树木主干与树冠的分割点,对枝下高的测量一般是测量离地面最近的明显大枝的高度,枝下高即为树木的主干高度。如果树木从 1.3 m 以下分叉,则认为是两棵树,1.3 m 以上分叉就认为是枝干。所以枝下高大于 1.3 m。对于每棵树,监测不同高度处树干的干径和重心位置。在每棵树的点云中取出该树从 1.3 m 至最高处的点云。从 1.3 m 处开始对点云数据水平分层,单层厚度 10 cm。采用 Hough 变换和圆拟合方法得到该层树干圆心和半径。以获取的圆心为圆心,提取半径为 $2r$ 范围内点云数据(r 为每层树干处拟合圆的半径),

将该薄层数据体元化,体元尺寸为 $5\text{ cm} \times 5\text{ cm} \times 10\text{ cm}$,若体元内回波点个数 $n > 2$,则认为该体元被覆盖。遍历所有层,得到各层的覆盖体元个数,即为树干处垂直剖面。在树干处点云分布比较集中,而且各层的覆盖度变化不大,到枝下高处,覆盖度明显变大。根据这个特点,可以得到枝下高位置。该位置所在高度与树干最低端所在高度之差即是要提取出的枝下高 B_H ,单位为 m 。

$$B_H = Z_{\max b} - Z_{\min b}$$

(2)

式中 $Z_{\max b}$ ——第 1 分枝树干位置坐标
 $Z_{\min b}$ ——立木底部坐标

1.3 冠高提取算法

冠高的提取过程比冠幅的提取更方便。冠高定义为树冠顶端到树冠最低端间的竖直距离。在传统测量当中,冠高的测量难点在于分辨树冠最低端。在冠层底部较复杂的情况下,找到最低端的过程需要进行仔细观测。利用地面激光雷达可以对单木的点云进行三维检视,可以有效解决枝叶间互相遮挡的问题,方便寻找树冠最低端,准确地提取冠层部分点云。

在树冠点云数据中,树冠顶部即是其 Z 坐标最大值 $Z_{\max c}$,树干第 1 分枝的位置是 Z 坐标最小值 $Z_{\min c}$,其二者的差值即是树冠的冠高 C_H ,单位为 m 。

$$C_H = Z_{\max c} - Z_{\min c}$$

(3)

1.4 树高提取算法

在完整的立树点云数据中,立木顶部即是其 Z 坐标最大值 $Z_{\max t}$,立木底部的的位置是 Z 坐标最小值 $Z_{\min t}$,其二者的差值即是树高 T_H ,单位为 m 。

$$T_H = Z_{\max t} - Z_{\min t}$$

(4)

枝下高为从地面算起的第 1 个分枝与地面之间的竖直距离,冠高为冠层顶端到冠层最低端的竖直距离,树高为树的顶端到地面的距离。三者的关系如图 1 所示。

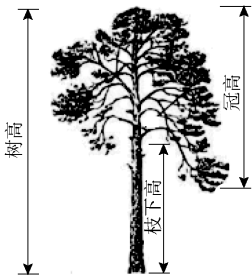


图 1 枝下高、冠高、树高的关系

Fig.1 Relationship between heights of branch, crown and tree

1.5 最长冠幅和垂直最长方向冠幅提取算法

点集的凸包是指一个最小凸多边形,满足点集 Q 中的点或者在多边形内或者在多边形上。本文中

提取冠幅的先导步骤是采用二维凸包的快速算法获得树冠的外轮廓,如图 2 所示。

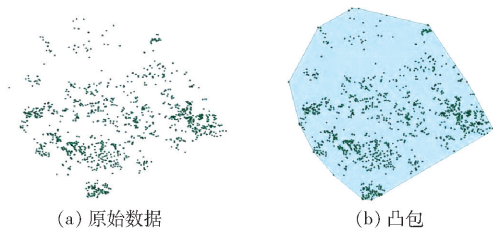


图 2 冠层点云的凸包

Fig.2 Convex hull of canopy point cloud

设凸包顶点的集合为 V 。对于凸包的每一个顶点 V_i ,计算其在二维坐标内与其他顶点 V_j 的欧氏距离,距离最长的两个顶点之间的距离就是该树的最长冠幅 L_S ,单位为 m 。

$$L_S = \max \{ \sqrt{(X_{V_i} - X_{V_j})^2 + (Y_{V_i} - Y_{V_j})^2}, V_i, V_j \in V \wedge i \neq j \}$$

(5)

式中 $X_{V_i}, Y_{V_i}, X_{V_j}, Y_{V_j}$ ——顶点 V_i, V_j 的坐标
设直线 $l: y = kx + b$,令 l 穿过最长冠幅的 2 个顶点,将顶点 V_i, V_j 代入 l 中,有

$$y = \frac{Y_{V_i} - Y_{V_j}}{X_{V_i} - X_{V_j}}x + Y_{V_i} - \frac{X_{V_i}(Y_{V_i} - Y_{V_j})}{X_{V_i} - X_{V_j}}$$

(6)

直线 l 过 V_i, V_j 中点的垂线 l' 可表示为

$$y = \frac{X_{V_i} - X_{V_j}}{Y_{V_i} - Y_{V_j}}x + \frac{Y_{V_i}^2 - Y_{V_j}^2 - X_{V_i}^2 + X_{V_j}^2}{2(Y_{V_i} - Y_{V_j})}$$

(7)

设凸包内的点的集合为 P ,对于 $\forall P_i \in P$,判断其是否在 l' 上。设在 l' 上的 P_i 的集合为 P' , P' 中欧氏距离最长的两点的距离即为需要的垂直最长方向冠幅 L_{CS} ,单位为 m 。

$$L_{CS} = \max \{ \sqrt{(X_{P'_i} - X_{P'_j})^2 + (Y_{P'_i} - Y_{P'_j})^2}, P'_i, P'_j \in P' \wedge i \neq j \}$$

(8)

式中 $X_{P'_i}, Y_{P'_i}, X_{P'_j}, Y_{P'_j}$ —— P_i, P_j 点的坐标
最长冠幅(L_S)与垂直最长方向冠幅(L_{CS})的关系如图 3 所示。

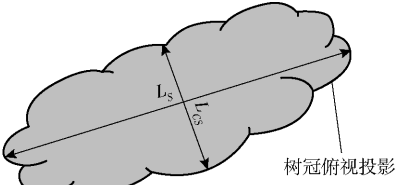


图 3 最长冠幅(L_S)与垂直最长方向冠幅(L_{CS})

Fig.3 The longest spread of crown cover (L_S) and the longest cross-spread of crown cover (L_{CS})

2 树种识别方法

2.1 支持向量机方法

支持向量机产生自在线性可分前提下最优分类面的发展,分类线的方程是 $x\omega + b = 0$,对分类线方

程进行归一化处理,得到样本的集合 $(x_i,y_i),i=1,2,\cdots,n,x\in\mathbf{R}^d,y\in(-1,1)$,该样本集线性可分,且满足

$$y_i[l(x_i,\omega)+b]-1\geq 0 \quad (i=1,2,\cdots,n) \quad (9)$$

支持向量机的核函数有 3 类,分别是神经网络核函数、径向基核函数和多项式核函数。每一种核函数对应不同类型的非线性映射问题。

2.2 分类回归决策树方法

分类回归决策树方法 (CART) 是 BERIMAN 等^[18]在 1984 年提出的一种非参数方法。分类回归决策树依用途可以分为分类决策树和分类回归树两种。本文中根据多种测树因子对样地中选取的 4 种树种进行分类,主要用到的是分类决策树。分类决策树将分类变量的情况作为因变量 (又称目标变量、输出变量)。假设自变量 (又称输入变量、属性) 是随机向量 $X_{n\times m}(X_1,X_2,\cdots,X_m)$, X_i 既可以是连续变量,也可以是离散变量,设它的定义域为 $\text{Dom}(X_i)$ 。 Y 是随机变量,作为因变量,如果 Y 为分类变量,设其定义域是 $\text{Dom}(Y)=\{1,2,\cdots,J\}$,那么由自变量 X_i 和因变量 Y 构建的决策树就是分类决策树。分类决策树将分类变量的情况作为因变量。CART 是非参数方法,它不需要假设总体服从先验分布。CART 在计算过程中充分利用二叉树结构,是一种二分递归的分割方法。样本集在一定的分割规则下被分割为 2 个样本集,生成的决策树每个非叶子节点都有 2 个分支。此过程在样本集上始终重复进行,直到不可再分成叶子节点。

2.3 随机森林方法

BREIMAN^[19]在随机决策森林方法的基础上提出了把分类树组合成随机森林再汇总分类树的随机森林算法。不同的 CART 决策树由投票决策组合成随机森林^[20],随机森林算法对训练集的随机性策略分为训练样本选择和特征属性选择。随机森林的抽象表示如下:假设有 K 棵 CART 树组成随机森林,产生第 i 棵决策树的函数表示为: $f_i(x,\theta_i):X\rightarrow Y,i=1,2,\cdots,K$,这里 x 为输入向量, θ_i 为随机向量 (独立同分布),这个向量是作用在训练样本的机制,因此随机森林可以表示为

$$F=\{f_1,f_2,\cdots,f_K\}$$

其中 K 为森林的规模。在构造完学习器之后,使用随机森林对样本数据进行分类,并且随机森林和多数集成学习算法一样,均在决定类别的过程采用投票机制,表示为

$$y^* = y \in Y \sum_{t \in F} I(f(x|t) = y) \quad (10)$$

式中 $I(\cdot)$ ——示性函数,取值为 0 和 1,括号中条件成立时,示性函数值为 1,条件不成立时,示性函数值为 0

投票最多的类别被随机森林选为样本的最终类别。

2.4 树种识别精度评价方法

为了评价几种模式分类模型的精度,采用准确率、召回率和 F 值 3 个指标。准确率和召回率是广泛用于信息检索和统计学分类领域的两个度量值,用于评价结果的质量。准确率为提取出的正确信息条数与提取出的信息条数的比值,召回率为提取出的正确信息条数与样本中的信息条数的比值,两者取值在 0 和 1 之间,数值越接近 1,查准率或查全率就越高。 F 值为准确率、召回率的乘积与准确率、召回率之和比值的 2 倍,也是正确率和召回率两者之间的调和平均值。

为了避免验证样本影响模型精度的偶然性,本研究中采用交叉验证的策略选取验证样本。交叉验证的过程为:将原始数据分成 K 组 (一般是均分),将每个子集数据分别做一次验证集,其余的 $K-1$ 组子集数据作为训练集,这样会得到 K 个模型,用 K 个模型最终验证集分类准确率的平均数作交叉验证下分类器的性能指标。 K 一般大于等于 2,实际操作时一般从 3 开始取,只有在原始数据集合数据量小的时候才会尝试取 2。

3 实验结果及分析

3.1 实验数据获取与处理

在北京林业大学校园内选取树木,利用 FARO Photon 120 型地面激光雷达 (如图 4 所示,主要参数如表 1 所示) 对研究区内树木进行了激光扫描,获取树木的点云数据。对同一树木进行了多次设站扫描,首先需要对多次扫描的点云进行拼接。拼接后的点云图像中除目标树外还存在许多噪点,在立体视图中删除多余噪点。研究共采集到 26 棵银杏、20 棵鹅掌楸、26 棵臭椿、20 棵西府海棠的激光点云数据。



图 4 FARO Photon 120 型地面激光雷达
Fig. 4 Picture of FARO Photon 120 model ground-based LiDAR

表 1 FARO Photon 120 型地面激光雷达的主要参数
Tab. 1 Parameters of FARO Photon 120

参数	数值
最大测量距离/m	120
最小测量距离/m	0.6
分辨率/mm	0.1
数据获取速率/(像素·s ⁻¹)	≤50 800
25 m 内误差/mm	≤2
水平视野范围/(°)	0 ~ 360
垂直视野范围/(°)	0 ~ 320

数据采集方法:在树干胸高处朝南或朝北方向贴标靶纸,在样木周围无遮挡的地方均匀放置 3 个参考球,选择能同时见到 3 个参考球的均匀分布的 3 个位置设站,理想的测站位置间隔角度为 120°(以待扫描树为参考),安装上地面激光雷达并连接好相关设备,对扫描范围、扫描分辨率等参数进行设置之后,就可进行树木扫描。本研究设置扫描区域为水平方向 360°,垂直方向 155°,扫描分辨率每一圆周 10 000 点,每株样木需从不同角度扫描至少 3 次,完整扫描一株样木需 10 min 左右,扫描场景如图 5 所示。

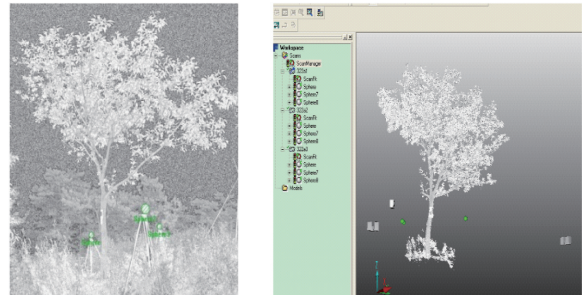


图 5 点云数据处理
Fig. 5 Process of point cloud data

将地面激光雷达扫描数据导入计算机中,通过其配套软件 Faro Scene,对点云数据进行加载、套准、过滤、剔除、导出等内业处理。打开 Faro Scene 软件,直接把测站数据加载到软件中,第一次打开软件时,扫描数据的 Mio 扫描点数值默认为 62,这个数值是根据计算机的内存自动显示的,本文所涉及的点云提取使用的计算机内存为 8 GB,经过在不同内存的计算机中进行点云数据提取实验,结果显示默认的 Mio 扫描点的数值均能满足精度要求。图 5 为加载的测站点云数据显示。

当对图 5 示例的立木扫描的 3 站数据都加载到统一的工作空间的窗口下,当窗口的显示标记为绿色的 3 个公共参考球,表示 3 站数据已经很好地拟合。在 3D 模式下,手动选取单木点云数据,剔除冗余数据,最终提取结果如图 6 所示,然后导出所需要后处理的数据格式. xyz。

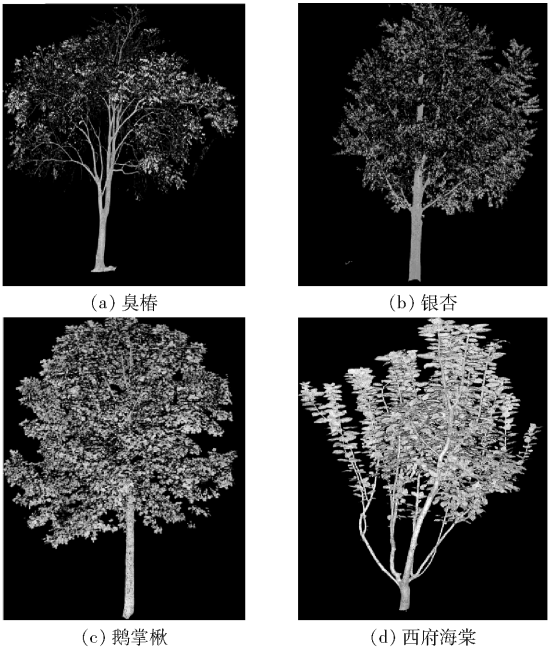


图 6 单木数据提取结果

Fig. 6 Extraction results of single tree for four species

3.2 测树因子提取结果及精度分析

为了验证点云数据提取的测树因子精度,将点云数据提取的测树因子和传统方法进行了对比,对比结果见表 2 ~ 5。其中表 2 中给出了银杏的各测树因子的绝对精度与相对精度。从总体来看,各测树因子中枝下高的相对误差均值是最小的,这主要是因为枝下高与胸径的量纲有所不同。从绝对精度来看的话,胸径的绝对精度误差分析最小,为 1.72 cm,枝下高的绝对精度均值为 0.17 m,比胸径绝对精度均值大一些,可以满足实际测量需求的程度。对于涉及高程的测量,常用的其他测量手段也会产生比较大的偏差,高密度的三维点云在高程提取上可以比其他手段有更好的效果。枝下高的误差最低,冠高的误差其次,树高的误差最大,产生该现象的原因主要有:①立木的最高点的确定有难度。②在单木最高处的点云密度较小,在前期点云预处理和后期树高提取的过程中都有被忽略掉的可能。另外,两个冠幅相关的测树因子的误差都要大一些,主要是因为冠幅会受风速、风向等环境因素影响产生变化,而且冠幅传统测量方法也是目视的方法,方法本身也有很大的不确定性,同时目视确定最长冠幅的方向有难度。垂直最长方向冠幅的误差相对最长冠幅误差大,主要原因是提取过程涉及垂直方向的确定,并且有可能遇到多棵树树冠相互遮挡的情况。

表 3 为臭椿的测树因子误差统计。各个测树因子的误差之间的相对规律与银杏测树因子的误差大致相同。皆为胸径的绝对误差最小,除了胸径外,枝

表 2 银杏测树因子误差统计

Tab. 2 Measuring factor error statistics of *Ginkgo biloba*

测树因子	样本量	绝对误差/m			相对误差/%		
		最小值	最大值	均值	最小值	最大值	均值
胸径	26	0.025 0	0.031 6	0.017 2	0.82	10.07	5.43
枝下高	26	0	0.33	0.17	0	9.23	5.31
树高	26	0	2.74	0.98	0	18.94	7.32
冠高	26	0	1.02	0.55	0	10.11	5.61
最长冠幅	26	0.09	1.80	0.74	1.24	20.22	11.20
垂直最长方向冠幅	26	0.20	1.86	0.93	3.70	30.03	16.24

表 3 臭椿测树因子误差统计

Tab. 3 Measuring factor error statistics of *Ailanthus altissima*

测树因子	样本量	绝对误差/m			相对误差/%		
		最小值	最大值	均值	最小值	最大值	均值
胸径	26	0.010 0	0.026 5	0.012 4	0.32	11.16	4.74
枝下高	26	0	0.28	0.19	0	12.65	6.19
树高	26	0.14	2.55	1.09	1.21	20.69	8.53
冠高	26	0.18	1.61	0.74	1.39	12.87	7.19
最长冠幅	26	0.10	2.50	0.81	1.10	24.39	9.39
垂直最长方向冠幅	26	0.12	4.00	1.99	1.39	31.54	13.94

表 4 鹅掌楸测树因子误差统计

Tab. 4 Measuring factor error statistics of *Liriodendron chinense*

测树因子	样本量	绝对误差/m			相对误差/%		
		最小值	最大值	均值	最小值	最大值	均值
胸径	20	0.026 0	0.027 5	0.014 3	1.07	11.40	6.44
枝下高	20	0.01	0.22	0.12	0.45	9.36	5.55
树高	20	0.09	2.63	1.56	0.71	19.65	11.87
冠高	20	0.22	1.39	0.94	2.12	11.89	8.26
最长冠幅	20	0.02	1.10	0.49	0.42	19.82	9.67
垂直最长方向冠幅	20	0.07	0.89	0.45	1.53	21.30	11.86

表 5 西府海棠测树因子误差统计

Tab. 5 Measuring factor error statistics of *Malus micromalus*

测树因子	样本量	绝对误差/m			相对误差/%		
		最小值	最大值	均值	最小值	最大值	均值
胸径	20	0.012 0	0.017 5	0.010 3	0.57	9.32	5.85
枝下高	20	0.02	0.31	0.18	0.34	7.66	5.55
树高	20	0.04	1.53	0.76	0.61	14.65	9.74
冠高	20	0.18	1.05	0.88	1.19	10.53	8.38
最长冠幅	20	0.05	1.01	0.56	0.23	16.80	9.65
垂直最长方向冠幅	20	0.09	0.91	0.53	0.88	19.62	8.90

下高和冠高的相对误差最小,树高相对误差相比而言较大,两个冠幅因子的相对误差最大,而其中垂直最长方向冠幅的相对误差比最长冠幅的相对误差大。不过与银杏不同的是,臭椿胸径的相对误差最小。

表 4 是鹅掌楸测树因子的误差统计。鹅掌楸的胸径相对误差没有枝下高相对误差小。冠幅的相对误差要小于树高的相对误差。

表 5 是西府海棠测树因子的误差统计。西府海

棠测树因子中,相对误差最大的是树高和最长冠幅,达到了 9% 以上,而胸径和枝下高相对误差最小,在 5% 左右。

3.3 树种识别结果及精度分析

选取训练样本的胸径、树高、枝下高、冠高、最长冠幅、垂直最长方向冠幅这 6 个参数,使用训练样本的多个属性训练分类器,可以使分类器兼顾样本的各方面特征,而不是像只是用一个属性时分类器只考虑一个属性中的特征,忽略其他特征。样本的属

性越多,其特征被描述得越全面,越能更好地训练分类器,从而使分类器可以产生更好的分类效果。本研究采集了立木的 6 个测树因子,较为全面地描述了样本立木的形态特征。将所有属性全部参与分类器的训练,分别采用支持向量机、分类回归树和随机森林方法进行树种分类。检验分类准确性时,采用交叉验证。分树种取样本的 80% 为训练样本,剩余 20% 的样本留作验证样本,在检验分类效果时使用。使用精准率和召回率评价分类的效果。结果如表 6 所示。

表 6 多参数分类评价指标(测树因子)
Tab. 6 Multi parameter classification evaluation index (tree measurement factor)

分类方法	树种	准确率	召回率	F 值
支持向量机	臭椿	0. 923	0. 923	0. 923
	鹅掌楸	0. 125	0. 050	0. 071
	银杏	0. 474	0. 692	0. 563
	西府海棠	0. 683	0. 683	0. 683
	总体	0. 551	0. 587	0. 556
分类回归决策树	臭椿	0. 923	0. 923	0. 923
	鹅掌楸	0. 938	0. 750	0. 833
	银杏	0. 800	0. 923	0. 857
	西府海棠	0. 832	0. 906	0. 863
	总体	0. 873	0. 875	0. 869
随机森林	臭椿	0. 923	0. 923	0. 923
	鹅掌楸	0. 800	0. 800	0. 800
	银杏	0. 846	0. 846	0. 846
	西府海棠	0. 917	0. 917	0. 917
	总体	0. 871	0. 871	0. 871
平均值		0. 765	0. 778	0. 765

从表 6 中可以看出,从总体准确率、召回率和 F 值来看,分类效果最好的是分类回归决策树法,其次是随机森林法,但两者相差不大,最差是支持向量机法;具体分析每个树种,3 种方法对臭椿的识别都较好,准确率都在 0.9 以上,而鹅掌楸的识别差异较大,最好的分类回归法准确率可以达到 0.938,而最差的支持向量机只有 0.125。银杏、鹅掌楸和西府海棠的情况均类似。

本文除了采用常见的 6 个测树因子之外,还通过组合 2 个常见测树因子得到更多的树形参数。组合而成的特征参数可以归结为 5 个类别,分别反映出立木的树干、树枝、树冠和全树的形态结构特征。其中全树的特征参数 1 个,树干的特征参数 2 个,树枝的特征参数 2 个,树冠的特征参数 1 个。具体见表 7。表中 L_c 为冠长, D 为胸高直径, L_{HB} 为在一定高度处的树枝长度, L_{HS} 为在相同高度处的树干长度。

表 7 组合特征参数		
Tab. 7 Combined feature parameters		
类型	参数	公式
单木特征	冠长树高比 P_{LcHt}	$P_{LcHt} = \frac{L_c}{T_H}$
树干特征	胸径树高比 P_{DsHt}	$P_{DsHt} = \frac{D}{T_H}$
	冠高树高比 P_{HcHT}	$P_{HcHT} = \frac{C_H}{T_H}$
树枝特征	分枝角 P_{ab}	$P_{ab} = \arccos \frac{L_{HS}}{L_{HB}}$
	冠长最大冠幅之比 P_{LCLS}	$P_{LCLS} = \frac{L_c}{L_s}$
树冠特征	最长冠幅与垂直方向冠幅之比 P_{LSLCS}	$P_{LSLCS} = \frac{L_s}{L_{CS}}$

参照陈国定等^[21] 和刘镇波等^[22] 提出的鲁棒性分析的方法,将提取的测树因子和组合特征参数分别在 SPSS 17 中进行识别分析,结果输出如图 7 所示。从图 7a 可以看出,利用测树因子的方法,4 个树种识别质心(图中质心 1、2、3、4 分别表示西府海棠、臭椿、鹅掌楸、银杏)聚集度高,树种之间辨识难度增加,特别是鹅掌楸和银杏,2 个树种基本混在一起,大大增加了识别难度。而组合特征参数可以很好地解决这一问题,从图 7b 可以看出相对于测树因子,4 个树种质心相距较远,树种识别容易,减少识别错误发生的几率,同时也说明了组合特征参数鲁棒性要优于测树因子。

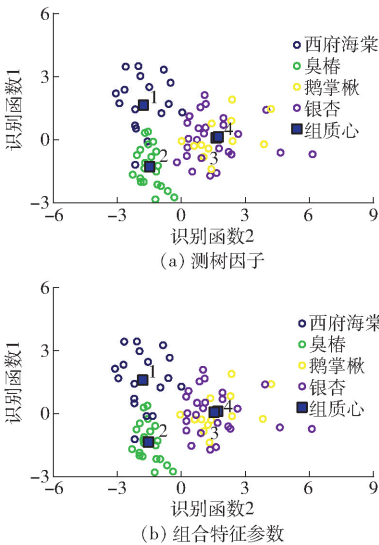


图 7 鲁棒性分析结果

Fig. 7 Results of robust analysis

4 个树种样本数据的冠长树高比、胸径树高比、冠高树高比、分枝角、冠长最大冠幅之比、最长冠幅与垂直方向冠幅之比 6 个指标,各参数对全部样本,分别采用支持向量机、分类回归决策树和随机森林方法进行树种分类。分树种取样本的 80% 为训练样本,剩余 20% 的样本留作验证样本,在检验分类

效果时使用。使用准确率和召回率评价分类的效果。

从表 8 可以看出,采用组合特征参数进行树种识别,3 种分类识别方法均取得了较为满意的结果,其中随机森林和支持向量机,总体准确率、召回率和 F 值都在 0.9 以上,而分类回归决策树略低于 0.9。其中臭椿识别效果最优,准确率达到 0.96,最差的鹅掌楸是 0.792,相对于单一特征参数分类而言,平均准确率、召回率和 F 值均大幅提高,识别效

果相对更优,这主要是因为多参数可以使分类器依据全面的立木结构情况对其进行分类,而不是只依据单一的结构情况。

4 结论

(1)利用地面激光雷达提取单木测树因子及识别树种是可行且有效的,与传统仪器和测量手段相比具有明显优势。利用一台地面激光雷达设备对单木进行一次扫描,获取的点云数据可用于提取树高、胸径、枝下高、冠高、冠幅等测树因子,减轻外业测量工作量,提升效率;在提取测树因子的同时,根据树木测树因子和组合特征参数可以准确识别不同树种,为今后建立树种识别信息库提供技术支撑。

(2)从提取的测树因子结果来看,胸径的绝对误差最小;一般而言枝下高的相对误差最小,其次是胸径、冠高、树高,冠幅的相对误差最大。从不同树种来看,臭椿提取胸径的相对误差最小,银杏提取枝下高、树高、冠高的相对误差最小,鹅掌楸提取冠幅的相对误差略小。

(3)对于树种识别问题,采用了树木测树因子和组合特征参数 2 种方式,分别利用支持向量机、分类回归决策树和随机森林方法,加入分类识别器,进行树种识别。通过分析发现,总体来说使用组合特征参数识别效果比单一使用测树因子优,不同分类方法相比,从平均准确率和召回率来看,随机森林方法优于支持向量机和分类回归决策树方法。

表 8 多参数分类评价指标 (组合特征参数)

Tab. 8 Multi parameter classification evaluation index (combined feature parameters)

分类方法	树种	准确率	召回率	F 值
支持向量机	臭椿	0.960	0.923	0.941
	鹅掌楸	0.905	0.950	0.927
	银杏	0.885	0.885	0.885
	西府海棠	0.892	0.892	0.892
	总体	0.910	0.912	0.911
分类回归决策树	臭椿	0.840	0.808	0.824
	鹅掌楸	0.792	0.950	0.864
	银杏	0.826	0.731	0.776
	西府海棠	0.905	0.950	0.927
	总体	0.841	0.859	0.847
随机森林	臭椿	0.960	0.923	0.941
	鹅掌楸	0.905	0.950	0.927
	银杏	0.885	0.885	0.885
	西府海棠	0.940	0.908	0.924
	总体	0.922	0.916	0.919
平均值		0.891	0.896	0.892

参 考 文 献

1 张煜星,王祝雄,武红敢,等. 遥感技术在森林资源清查中应用研究[M]. 北京:中国林业出版社,2007.

2 张丽云. 基于高光谱遥感数据的森林树种分类[D]. 北京:北京林业大学,2016.
ZHANG Liyun. The study on identification of forest tree species based on hyperspectral image [D]. Beijing: Beijing Forestry University,2016. (in Chinese)

3 王志辉. 高光谱遥感在森林树种识别中的应用[D]. 杭州:浙江农林大学,2011.
WANG Zhihui. Application of hyperspectral semote sensing in forest tree species identification [D]. Hangzhou:Zhejiang Agriculture and Forestry University,2011. (in Chinese)

4 LIANG X, LITKEY P, HYYPÄ J, et al. Automatic stem location mapping using TLS for plot-wise forest inventory [J]. Proceedings of the Silvi Laser, 2009;314 – 323.

5 STRAHLER A H, JUPP D L B, WOODCOCK C E, et al. Retrieval of forest structural parameters using a ground-based LiDAR instrument (Echidna) [J]. Canadian Journal of Remote Sensing, 2008, 34 (Supp. 2) : S426 – S440.

6 KOUKOULAS S, BLACKBURN G A. Mapping individual tree location, height and species in broadleaved deciduous forest using airborne LiDAR and multi-spectral remotely sensed data [J]. International Journal of Remote Sensing, 2005, 26 (3) :431 – 455.

7 CLARKM L, CLARK D B, ROBERTS D A. Small-footprint LiDAR estimation of sub-canopy elevation and tree height in a tropical rain forest landscape [J]. Remote Sensing of Environment, 2004, 91:68 – 89.

8 刘清旺,李增元,陈尔学,等. 利用机载激光雷达数据提取单株木树高和树冠 [J]. 北京林业大学学报,2008,30 (6) :83 – 89
LIU Qingwang, LI Zengyuan, CHEN Erxue, et al. Extracting individual tree heights and tree crowns using airborne LiDAR [J]. Journal of Beijing Forestry University,2008,30 (6) :83 – 89. (in Chinese)

9 余超. 基于 WorldView – 2 的森林公园主要风景树种信息提取及生物量估测 [D]. 南京:南京林业大学,2017.
YU Chao. Information extraction and biomass estimation of main landscape tree species in Forest Park based on the WorldView – 2 [D]. Nanjing:Nanjing Forestry University,2017. (in Chinese)

10 魏田. 基于静态地基激光扫描点云的单树结构表征 [D]. 唐山:华北理工大学,2015.

WEI Tian. Tree-level structure characterization based on static terrestrial laser scanning point clouds[D]. Tangshan:North China University of Science and Technology,2015. (in Chinese)

11 PLAZA A,BENEDIKTSSON J A,BOARDMAN J W,et al. Recent advance in techniques for hyperspectral image processing[J]. Remote Sensing of Environment,2009,113:110 – 122.

12 GEORGE P P,CHARITON K,KRISHNA P V. Support vector machines and object-based classification for obtaining Land-use/cover cartography from hyperion hyperspectral imagery[J]. Computers&Geosciences,2012,41:99 – 107.

13 刘丽娟,庞勇,范文义,等. 机载 LiDAR 和高光谱融合实现温带天然林树种识别[J]. 遥感学报,2013,17(3):679 – 695. LIU Lijuan,PANG Yong,FAN Wenyi, et al. Fused airborne LiDAR and hyperspectral data for tree species identification in a natural temperate forest[J]. Journal of Remote Sensing,2013,17(3):679 – 695. (in Chinese)

14 刘怡君,庞勇,廖声熙,等. 机载 LiDAR 和高光谱融合实现普洱山区树种分类[J]. 林业科学研究,2016,29(3):407 – 412. LIU Yijun,PANG Yong,LIAO Shengxi, et al. Merged airborne LiDAR and hyperspectral fusion for tree species classification in puer mountain[J]. Forestry Science Research,2016,29(3):407 – 412. (in Chinese)

15 刘清旺,谭炳香,胡凯龙,等. 机载激光雷达和高光谱组合系统的亚热带森林估测遥感试验[J]. 高技术通讯,2016,26(3):264 – 274. LIU Qingwang,TAN Bingxiang,HU Kailong, et al. The remote sensing experiment on airborne LiDAR and hyperspectral integrated system for subtropical forest estimation[J]. High-tech Communication,2016,26(3):264 – 274. (in Chinese)

16 ALONZO M,BOOKHAGEN B,ROBERTS D A. Urban tree species map-ping using hyperspectral and LiDAR datafusion[J]. Remote Sensing of Environment,2014,148(148):70 – 83.

17 李永亮,林辉,孙华,等. 基于 BP 神经网络的森林树种分类研究[J]. 中南林业科技大学学报,2010,30(11):43 – 46. LI Yongliang,LIN Hui,SUN Hua, et al. Study on forest species classification based on BP neural network[J]. Journal of Central South University of Forestry and Technology,2010,30(11):43 – 46. (in Chinese)

18 BREIMAN L, FRIEDMAN J, OLSHEN R, et al. Classification and regression trees[M]. New York:Wadsworth, Inc. ,1984.

19 BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5 – 32.

20 李欣海. 随机森林模型在分类与回归分析中的应用[J]. 应用昆虫学报,2013,50(4):1190 – 1197. LI Xinhai. Using “random forest” for classification and regression[J]. Chinese Journal of Applied Entomology,2013,50(4):1190 – 1197. (in Chinese)

21 陈国定,王涛,刘曼利,等. 轴承腔两相流动流型特征参数及鲁棒性分析[J]. 西北工业大学学报,2012,30(2):155 – 158. CHEN Guoding,WANG Tao,LIU Manli,et al. Exploring how to realize accurate and efficient lubrication design and heat transfer analysis for oil/air two-phase flow in bearing chamber [J]. Journal of Northwestern Polytechnical University, 2012,30(2):155 – 158. (in Chinese)

22 刘镇波,刘一星,于海鹏. 基于量化构造特征参数的树种计算机识别算法[J]. 福建林学院学报,2004,24(3):265 – 269. LIU Zhenbo,LIU Yixing,YU Haipeng. Study on the mold of wood computer identification system based on quantified anatomy properties [J]. Journal of Fujian College of Forestry, 2004,24(3):265 – 269. (in Chinese)