

基于机载激光雷达和高光谱数据的树种识别方法

陶江玥^{1,2}, 刘丽娟^{1,2}, 庞 勇³, 李登秋^{1,2}, 冯云云^{1,2}, 王 雪^{1,2},
丁友丽^{1,2}, 彭 琼^{1,2}, 肖文惠^{1,2}

(1. 浙江农林大学 浙江省森林生态系统碳循环与固碳减排重点实验室, 浙江 杭州 311300; 2. 浙江农林大学 省部共建亚热带森林培育国家重点实验室, 浙江 杭州 311300; 3. 中国林业科学研究院 资源信息研究所, 北京 100091)

摘要: 训练样本的选取是影响监督分类精度的直接原因之一, 数据空间分辨率越高, 训练样本要求越准确, 而人机交互训练样本选取推广力有限。利用机载高光谱(AISA)和激光雷达(LiDAR)主被动遥感数据, 探讨基于高分辨率影像的训练样本自动提取技术以及适合树种识别的遥感变量。根据树木的结构和高度差异, 开展树高分层掩膜试验, 并计算光谱间夹角, 在每个高度层中自动化优选树种的高纯度训练样本。计算植被指数、主成分分析等特征变量, 基于支持向量机分类器对研究区进行树种精细分类。实验表明: 通过对阔叶林、马尾松 *Pinus massoniana*, 毛竹 *Phyllostachys edulis*, 杉木 *Cunninghamia lanceolata*, 油茶 *Camellia oleifera* 的训练样本分层自动提取后再进行分类, 激光雷达和不敏感色素指数变量能有效提高树种分类精度。其中高光谱+激光雷达+结构不敏感色素指数变量组合的分类精度最高, 其总体精度和 Kappa 系数分别为 89.12%和 0.86, 阔叶林、马尾松、毛竹、杉木、油茶的用户精度分别为 75.00%, 100.00%, 86.36%, 90.91%和 96.55%。该方法对本研究区森林树种的识别是有效的。图 8 表 6 参 19

关键词: 森林测绘学; 高光谱; 激光雷达; 分层训练样本自动提取; 树种识别; 光谱角填图; 支持向量机
中图分类号: S758.5 **文献标志码:** A **文章编号:** 2095-0756(2018)02-0314-10

Automatic identification of tree species based on airborne LiDAR and hyperspectral data

TAO Jiangyue^{1,2}, LIU Lijuan^{1,2}, PANG Yong³, LI Dengqiu^{1,2}, FENG Yunyun^{1,2}, WANG Xue^{1,2},
DING Youli^{1,2}, PENG Qiong^{1,2}, XIAO Wenhui^{1,2}

(1. Zhejiang Provincial Key Laboratory of Carbon Cycling and Carbon Sequestration in Forest Ecosystems, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China; 2. State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China; 3. Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China)

Abstract: Selection of training samples, a direct factor affecting the accuracy of supervised classification, with a higher spatial resolution image, requires more accurate training samples, but the human-computer interaction capabilities in the selection of training samples is limited. Therefore, in this study, an algorithm was provided for automatic extraction of training samples. Airborne hyperspectral data and LiDAR data were used in Gutian Mountain National Nature Reserve. The hyperspectral data were used to extract training samples automatically and variables of tree species were calculated. According to differences in structure and height of individual trees provided by the canopy height model of LiDAR, a tree height mask was made to help circumvent the

收稿日期: 2017-02-23; 修回日期: 2017-04-12

基金项目: 国家级大学生创新创业训练计划项目(201610341013); 国家自然科学基金资助项目(41201365); 浙江农林大学科研发展基金资助项目(2014FR004)

作者简介: 陶江玥, 从事林业遥感研究。E-mail: 954972267@qq.com。通信作者: 刘丽娟, 讲师, 博士, 从事林业遥感研究。E-mail: llj7885@163.com

problem of different objects with the same spectra and identical objects with different spectra, as far as possible. Then, the spectral angle between each pixel and training sample pixel was calculated and highly pure pixels at different heights were selected automatically. In addition, a vegetation index and principal component analysis were calculated. The precise classification of tree species was carried out by a support vector machine classifier in the study area. The experiment used a method of stratified-auto sample selection to extract the training samples of broadleaf, Masson pine, Moso bamboo, Chinese fir, and tea-oil tree forests, and then classified these five tree species. Results showed that the combination of hyperspectral data, LiDAR data, and the structure of the insensitive pigment index revealed an overall accuracy of 89.12% and a Kappa coefficient of 0.86. Using a combination of the best variables, the user accuracy was as follows: broadleaf forest—75.00%, Masson pine—100.00%, Moso bamboo—86.36%, Chinese fir—90.91%, and tea-oil tree—96.55%. Therefore, integration of different remote sensing data, stratified-auto sample selection, and hyperspectral variable selection using LiDAR and the structure insensitive pigment index were effective ways for improving tree species classification. [Ch, 8 fig. 6 tab. 19 ref.]

Key words: forest measurement; hyperspectral; LiDAR; stratified-auto samples selection; tree species identification; spectral angle mapping; support vector machine

精确的树种分类对于提取树木特征属性, 保护物种多样性, 改善区域生态环境, 以及建立林木生长模型都非常重要^[1-3]。目前应用较广泛的宽谱段遥感数据由于空间和光谱分辨率均较低, 难以获得精确的树种分类结果, 而高光谱数据波谱范围更窄, 能准确探测到具有细微光谱差异的各种地物类型, 并借助理等特征分析, 识别光谱相似树种^[4]。光学数据只能探测到冠层表面信息, 限制了树种识别的精度。激光雷达(LiDAR)能获取详细的树木冠层结构三维信息^[5], 这对于森林类型识别、森林结构特性以及冠层理化特征均具有明显优势。因此, 将表征林分垂直结构信息的 LiDAR 数据与表征冠层水平方向信息的高光谱(AISA)数据融合, 形成优势互补, 理论上可以提高树种的识别精度^[6]。刘丽娟等^[7]利用机载 LiDAR 和高光谱数据融合对北方复杂森林树种进行识别时, 发现融合数据树种分类精度高于仅高光谱数据的精度, 总体精度达到 83.88%, Kappa 系数为 0.80。董彦芳等^[8]将高光谱遥感图像和 LiDAR 数据融合, 利用归一化植被指数(NDVI)和主成分分析(PCA)法进行去噪和降维, 再进行监督分类, 实现了城市内民用房屋和树木的提取。CAO 等^[9]研究利用全波形激光雷达数据对亚热带森林树种分类, 结果表明, 6 类树种的总体分类精度为 68.60%, 4 类树种为 75.80%, 而针叶林和阔叶林 2 类为 86.20%。ALONZO 等^[10]将高分辨率高光谱图像与 LiDAR 数据融合, 在基于冠层尺度上对美国某些地区常见的 29 类树种进行分类, 结果表明融合 LiDAR 数据后, 分类精度提高了 4.2 个百分点。DALPONTE 等^[11]提出了高光谱和 LiDAR 数据有效结合对复杂森林地区分类的方法, 并证明了支持向量机(SVM)分类器在对多源数据分类的准确性。国内外研究初步表明, 将 LiDAR 数据与高光谱数据融合, 可以有效提高树种分类的精度。目前多是针对地形平坦且树种结构单一的林型开展研究, 对于地形复杂且树种多样的亚热带林区分类研究较少。在监督分类中, 训练样本选取的数量和质量尤为关键, 其影响往往大于分类算法的选择。王春来等^[12]研究得出, 基于像元分类的训练样本选取数量在 24~30 倍于波段数时分类精度才达到较高的水平。训练样本的选取主要采取野外实地调查法, 但此法综合成本高, 且采集数量有限, 往往难以满足监督分类所需样本数, 另外转绘到图像上时也存在人为判断的误差。若能实现训练样本的自动选取, 将大大提高树种精细识别的效率和精度。因此, 本研究拟在 LiDAR 与高光谱数据融合的基础上, 试验训练样本自动化优选的方法, 并通过 SVM 分类器基于小样本进行分类, 比较不同特征变量的组合对树种分类效果的差异。

1 研究区概况

古田山国家级自然保护区(29°10′32.12″~29°17′44.33″N, 118°03′56.25″~118°10′56.51″E)位于浙江省开化县西北 30 km 处的苏庄境内, 与江西省婺源县、德兴市毗邻, 总面积为 81.07 km², 属于南岭山系怀玉山脉的一部分。地处中亚热带东部, 受夏季风影响较大, 气候有明显的季节变化, 年均降水天数

为 142.5 d, 年均降水量为 1 963.7 mm, 相对湿度为 92.4%。年均气温为 15.3 ℃, 无霜期约为 250 d, 冬暖夏凉, 温暖湿润。由于其特殊的地理位置, 分布着典型的中亚热带常绿阔叶林、常绿落叶阔叶混交林和针叶林。经济树种主要为常绿小乔木油茶 *Camellia oleifera*, 分布于海拔较低的居民区附近, 高度多在 2 m 以上, 由于在研究区内分布面积较大, 因此本研究将油茶作为一个树种参与分类。考虑到研究区内阔叶林的树种种类繁多而样本受限, 未细分到树种, 统称为阔叶树种。本研究选择研究区的马尾松 *Pinus massoniana*, 杉木 *Cunninghamia lanceolata*, 毛竹 *Phyllostachys edulis*, 油茶以及阔叶树种等 5 个树种类型开展研究。图 1 为研究区的地理位置及其部分图像。

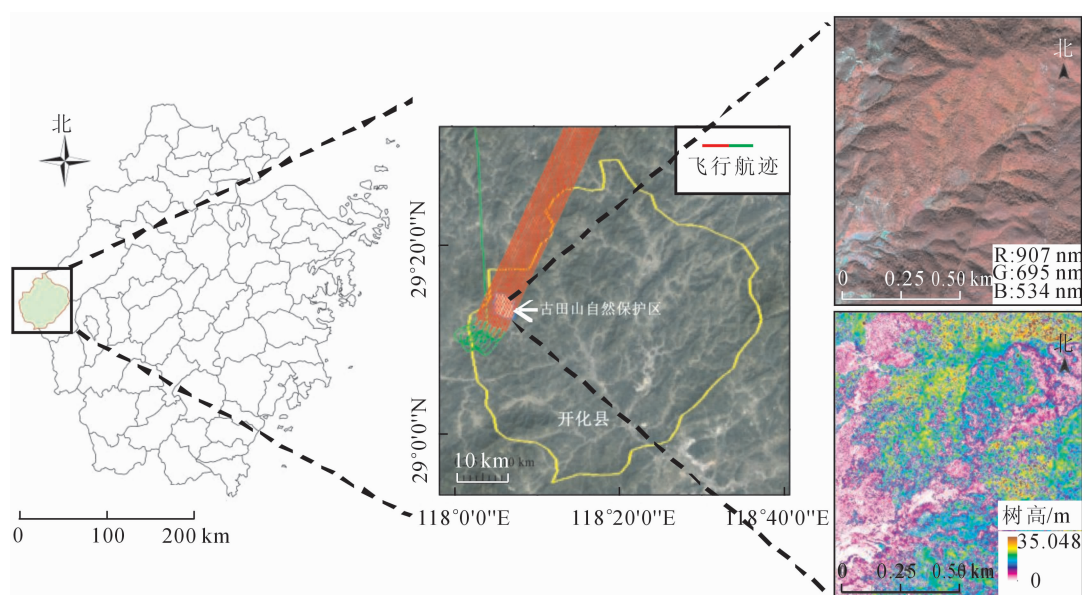


图 1 浙江省开化县地理位置示意图(左)、古田山国家级自然保护区地理位置和飞行轨迹(中)、高光谱图像(右上)和 LiDARCHM 图像(右下)

Figure 1 Location of Kaihua County in Zhejiang Province (left), location and flight trajectory of National Nature Reserve of Mount Gutian (middle), hyperspectral image (upper right) and LiDARCHM image (lower right)

2 数据收集及预处理

2.1 数据获取

机载数据航飞于 2014 年 10 月古田山国家级自然保护区, 获取了研究区高空间分辨率的高光谱和高密度 LiDAR 点云数据, 数据获取当天晴朗少云。同步开展了部分样地的踏勘工作, 同年 12 月又进行了样地补充调查。此次飞行的高光谱传感器为 AISA Eagle II, 由超光谱探头、小型 GPS/INS 探头及数据采集系统组成, 所得 64 个波段高光谱遥感图像可获得单木尺度精细的光谱信息。飞行同步搭载 1 台 Riegl LMS-Q680i LiDAR 传感器, 可获得单木尺度的三维及强度信息。高光谱传感器 AISA Eagle II 参数: 光谱范围为 400~970 nm, 帧频为 160 帧·s⁻¹, 焦距为 18.1 mm, 光谱分辨率为 3.3 nm, 视场角为 37.7°, 瞬时视场角为 0.037°。LiDAR 传感器 Riegl LMS-Q680i 参数: 波长为 1 550 nm, 激光发射角为 0.5 mrad, 激光脉冲长度为 3 ns, 视场角为 ±30°, 最大激光脉冲重复率为 400 kHz, 波形采样间隔为 1 ns, 1 000 m 高度的点密度为 3.6 pts·m⁻², 垂直分辨率为 1.15 m。

2.2 数据处理

2.2.1 机载高光谱数据的预处理 数据提供商已经对机载高光谱数据进行了系统辐射定标、几何校正和正射校正, 还需对数据进行大气校正以及与 LiDAR 数据的几何配准。由于机载数据受大气影响相对较小, 因此, 本研究采用 FLAASH(ENVI, 美国)对高光谱数据进行大气校正。该模块用 MODTRAN 4+ 辐射传输模型, 可以校正由于漫反射引起的连带效应, 还可以调整由于人为抑制而导致的波谱平滑^[14]。大气校正后, 消除了成像时光照和大气对地物反射率的影响, 植被光谱曲线突出了谷底和峰值, 并且近红外波段形成明显的高反射峰(图 2)。高光谱数据由于波段多, 易造成信息冗余, 因此在大气校正之后用

Wilks' Lambda 判别分析法选择具有代表性的特征波段。Wilks' Lambda 判别分析法是基于多变量的方差分析方法，以分类样本为基础，计算组间与组内差异，综合考虑其组内和组间差异，最后得出能够反映各个波段重要性的统计量 W 。Wilks' Lambda 计算公式具体如下^[15]：

$$\left\{ \begin{aligned} W &= \frac{S_{SE}}{S_{ST}} \\ S_{SE} &= \sum_{i=1}^r \sum_{j=1}^p (x_{ij} - \bar{x}_i)^2 \\ S_{ST} &= \sum_{i=1}^r \sum_{j=1}^p (x_{ij} - \bar{x})^2 \end{aligned} \right. \quad (1)$$

式(1)中： S_s 为样本离差矩阵， $x_{ij}(i=1 \dots r, j=1 \dots p)$ ， S_{SE} 为组内平方和， S_{ST} 为总平方和(即组间平方和与组内平方和的总和)。 W 值在 0 到 1 之间变化，值越接近 0 则说明组内的平均值差别越大，反之，值接近于 1 说明组内平均值越相近^[16]。由 Wilks' Lambda 法选择的波段组合作为高光谱特征变量，记作 AISA。

2.2.2 机载 LiDAR 数据处理 数字冠层高度模型(canopy height model, CHM)是一个重要的特征变量，通过对点云数据进行滤波分类(TerraScan, TerraSolid, 芬兰)，将它进行地面点和非地面点的分离，对已分类点云数据中的地面点进行 TIN 插值运算生成数字高程模型(digital elevation model, DEM)，首回波点插值生成数字表面模型(digital surface model, DSM)。DSM 与 DEM 进行差值运算即得到高程归一化后的 CHM。另外还开展了基于形态学滤波的单木冠层分割(SEAL, 激光雷达生态应用软件，中国林业科学研究院)，用于分类结果的验证。

2.2.3 高光谱与 LiDAR 数据配准 选择 20 个同名点，建立 3 次多项式纠正方程，对高光谱数据与 LiDAR 的 CHM 数据进行几何互配准。校正后坐标位置平均误差在 2 m 以内，即 1 个像元。

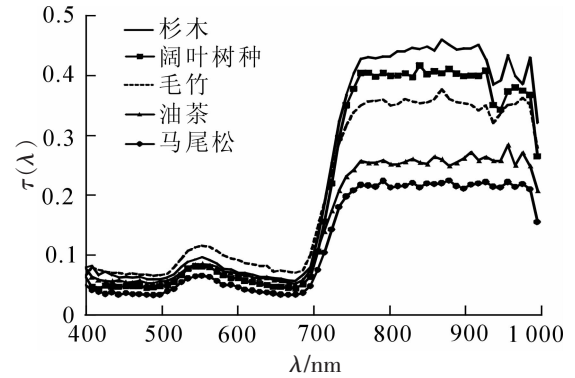


图 2 各树种典型光谱曲线

Figure 2 Spectral curves of five tree species

3 研究方法

“异物同谱”现象在地形复杂和树种多样的林区中普遍存在，增大了训练样本的选取难度，导致光谱相似树种的识别率偏低。LiDAR 数据可以描述复杂林型垂直分布特征。本研究基于分层分类的思想，在剔除非林地后，结合 LiDAR 的 CHM 与样地统计数据，获得各树种的高度分布，针对“异物同谱”树种像元，通过比较不同树种的高度聚集情况差异，提取各树种的高度掩膜层；在各掩膜层中计算各像元与参考样本的光谱夹角。进行训练样本的自动提取；最后，计算不同的特征变量[植被指数(vegetation index, VI)，PCA 降维波段、高光谱 AISA 与 LiDAR 的 CHM 数据]利用 SVM 分类器对不同变量的组合进行树种分类和精度比较。技术路线见图 3。

3.1 去除非林地

归一化植被指数(normalized differential vegetation index, NDVI)对土壤背景的变化较为敏感。计算本研究区 NDVI，通过比较林地与非林地的 NDVI 值，并结合 LiDAR 的 CHM 数据，设置 NDVI 大于 0.3，高度大于 2 m 的像元为林地，去除非林地光谱信息的干扰。

3.2 特征变量选择

3.2.1 高光谱 采用 Wilks' Lambda 波段选择法对 64 个波段的高光谱进行降维，选出数据冗余小、噪声少且能够表达植被光谱特征的 14 个波段^[17]作为特征变量 AISA，这 14 个优选波段为波段 7，波段 13，波段 16，波段 18，波段 21，波段 22，波段 24，波段 31，波段 33，波段 35，波段 37，波段 42，波段 55 和波段 58，它们的中心波长分别为 452.11，506.76，534.44，553.02，581.07，590.46，609.29，675.78，694.91，714.09，733.32，781.54，907.45 和 936.49 nm。

3.2.2 植被指数 以往研究多基于 NDVI，EVI(enhanced vegetation index)等宽波段指数，而反映色素含

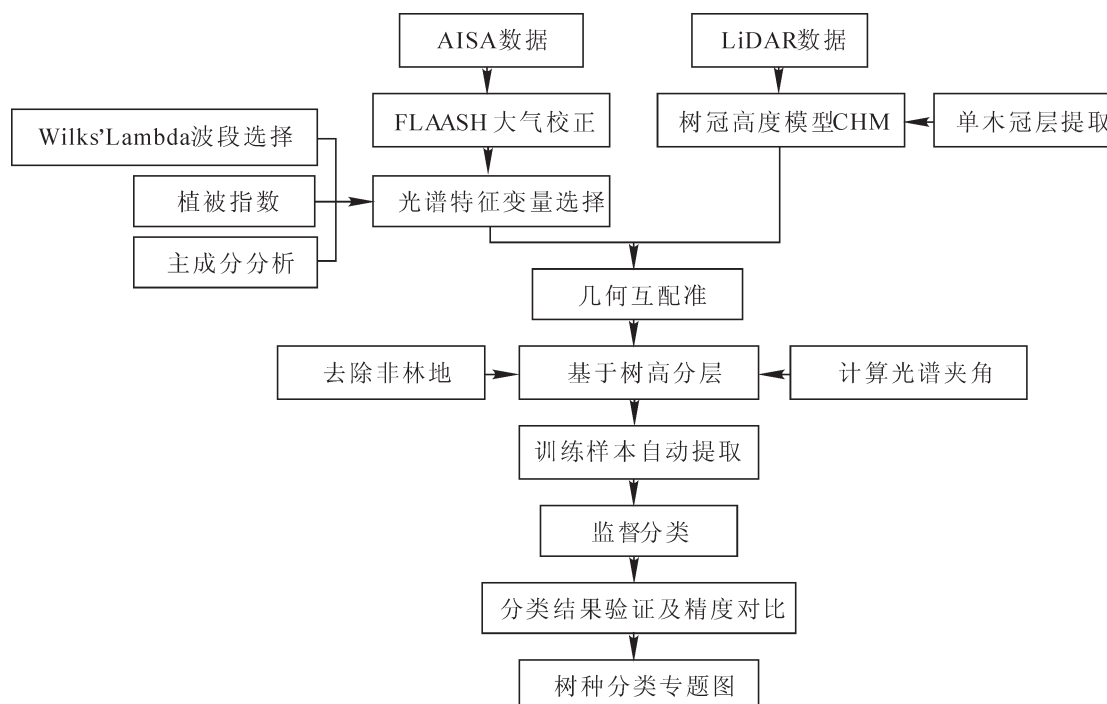


图 3 LiDAR 与高光谱融合的树种分类路线图

Figure 3 Flowchart of tree species classification based on LiDAR and AISA data

量的指数使用较少。因此,本研究利用随机森林法对 16 个高光谱植被指数变量按重要性排序^[18]。平均精度减少量(mean decrease accuracy, D_{MA})和平均基尼指数减少量(mean decrease gini, D_{MG})是重要性评价的指标,一般两者的值越大表示该变量的重要性越大。结合研究内容,本研究依据平均精度减少量选取排名前 8 的植被指数作为待选特征变量(表 1)。

表 1 植被指数重要性排序(前 8)

Table 1 Importance sorting of vegetation index (No.1~8)

变量	D_{MA}	D_{MG}	变量	D_{MA}	D_{MG}
结构不敏感色素指数 SIPI	16.781 221 7	1.059 741 828	花青素反射指数 2ARI2	8.631 932 8	0.931 436 794
改进红边比值植被指数 MRESRI	14.090 665 4	0.921 555 203	Vogelmann 红边指数 VREI	4.864 365 4	0.638 686 021
改进红边归一化植被指数 MRENDVI	13.784 894 3	0.890 267 344	归一化植被指数 NDVI	4.070 067 8	0.755 367 097
红绿比值指数 RGRI	10.034 705 4	0.887 654 227	花青素反射指数 1ARI1	3.909 291 8	0.358 767 829

3.2.3 主成分分析 主成分分析(PCA)也是一种常用的波段降维手段。PCA 变换后的前 4 个主成分包含了所有波段中 95%以上的信息量。前 4 个 PCA 的标准差分别是 8 121.92, 2 056.43, 448.07 和 336.14, PCA1 和 PCA2 的值远大于 PCA3 和 PCA4。因此,PCA1 和 PCA2 作为待选特征变量参与分类研究。

3.2.4 激光雷达 数字冠层高度模型作为 LiDAR 的特征变量,将该变量记作 CHM。

3.3 训练样本自动提取

3.3.1 基于树高分层 树种的结构和高度分布因树木生长习性不同而有差异。由于“异物同谱”现象使得光谱特征近似的不同树种的训练样本难以选择。为解决这一问题,本研究基于样地统计数据,根据各树种的高度分布,对数字冠层高度模型进行分层掩膜,提取出各树种高度集中分布层的高光谱像元作为训练样本。本研究各树种高度分布频率如图 4, 0.5 m 的区间归为一级。其中阔叶树种与杉木的高度分布曲线相近,毛竹和马尾松在高度层上与其他树种差距较大。根据各树高分布频率,最终选取掩膜高度为:阔叶树种 4.0~4.5 m 和 6.0~7.0 m;杉木 7.5~8.5 m;毛竹 10.0~10.5 m;马尾松 12.0~14.0 m。由于外业未采集油茶样地信息,因此没有列出,通常油茶高度为 2~3 m。

3.3.2 计算光谱角 光谱角填图法(spectral angle mapping, SAM)可以计算出 2 个光谱之间的夹角,夹角越小,光谱越相近,两者属于同一类的可能性越大。

$$\alpha=\cos^{-1}\left[\frac{\sum_{i=1}^N A_i B_i}{\left(\sum_{i=1}^N A_i^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^N B_i^2\right)^{\frac{1}{2}}}\right]。$$

(2)

式(2)中： α 为空间向量 A 和 B 的夹角，取值 $0\sim\frac{\pi}{2}$ ， N 表示数据的波段数，其中： A_i, B_i 表示第 i 个波段上的数值并分别代表测试光谱和参考光谱。计算每个高度层内各像元与参考光谱之间的光谱夹角，并调整夹角阈值，阈值越小，两光谱向量匹配越好，是同类的可能性越大。如设置阈值为 0.03，代表两像元光谱夹角 α 小于 0.03 弧度时，认为该像元与参考像元为同一类，实现训练样本的自动识别。

3.4 增补验证样本

研究区地形起伏较大，大部分区域数据难以获取，导致验证样本数量有限。本研究结合样地实测数据和高空间分辨率图像，增补阔叶树种、马尾松、毛竹、杉木、油茶的验证样本(表 2)，使其均匀分布在研究区内。后续各变量组合的分类精度验证使用同一套验证样本。

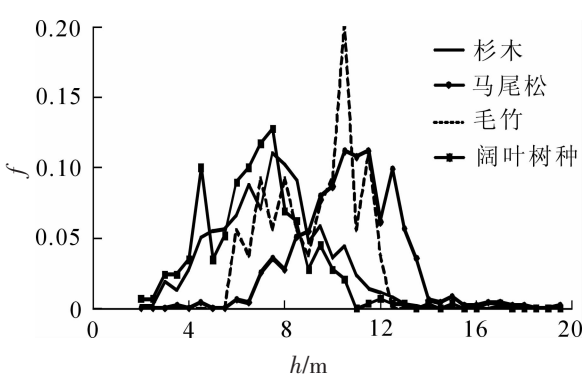


图 4 各树种高度分布频率

Figure 4 Distribution frequency chart of height for tree species

表 2 各树种样本点数量表

Table 2 Number of sample points of five tree species

项目	样本点数/个				
	阔叶树种	马尾松	毛竹	杉木	油茶
增补前	17	16	13	19	18
增补量	14	15	7	15	13

4 结果与分析

4.1 训练样本分层自动提取

以错分毛竹样本为例。研究区毛竹林为人工林，与阔叶树种混生，当人机交互直接在影像上选择训练样本时，边缘像元往往是混合像元，获得的是 2 个类型的均值光谱，造成部分毛竹与阔叶树种混分。图 5 是错分毛竹样本在高光谱图像上的地理位置，像元植被类型应为毛竹，但由于人机交互训练样本选择误差，错分为阔叶树种，下文称此类像元为错分毛竹样本。

图 6 为分层前后选取的阔叶树种、毛竹训练样本以及错分毛竹样本的光谱特征图。可以看出未分层前选取的阔叶树种与毛竹训练样本光谱特征曲线比较接近，而分层后两者光谱曲线之间差异较大。

计算错分毛竹样本与分层前后阔叶树种、毛竹样本之间的光谱角(表 3)，分层后训练样本间区分度更大，更精确。本研究利用数字冠层高度模型分层掩膜并计算光谱夹角自动提取训练样本的方法，降低了阔叶树种与毛竹的混分概率。

4.2 分类精度比较

经自动化分层选取的训练样本，充分利用 CHM 的高度信息，减少因光谱相似而高度不同的树种光谱信息的相互干扰，同时使训练样本的选择高效且高精度。本研究使用混淆矩阵对各特征变量组合的 SVM 分类结果进行精度评价。

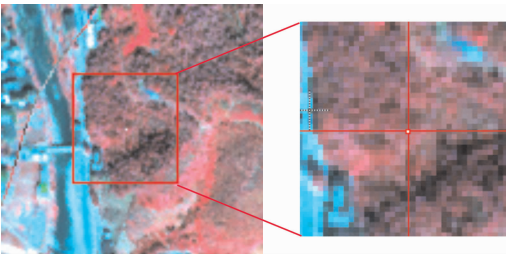


图 5 错分毛竹样本在高光谱图像上的地理位置

Figure 5 Location of mis-classification on the hyper spectral image

4.2.1 AISA+VI 变量 结果表明：AISA+SIPI 分类总体精度可达 77.55%，高于仅 AISA 的总体精度65.31%。其余 VI 变量与 AISA 融合后的分类精度均低于 SIPI。分析原因发现，当叶片相对含水量较低时，SIPI 与类胡萝卜素/叶绿素的比值具有较好的线性关系。水分缺失导致叶绿素因降解而减少，类胡萝卜素/叶绿素的比值升高，SIPI 也随之升高^[19]。本研究的数据获取于 10 月，此时植物体内水分开始流失。因此，在本研究中 SIPI 能做为树种识别的指示性变量，且另一方面证明重要性排序的可靠性。

4.2.2 AISA+PCA 变量 对 AISA+PCA 变量进行精度分析时，加入第 1 主成分 PCA1 的总体精度是70.07%，比加入 PCA2 高 5.44%。所以 PCA1 比 PCA2 更适合树种识别。

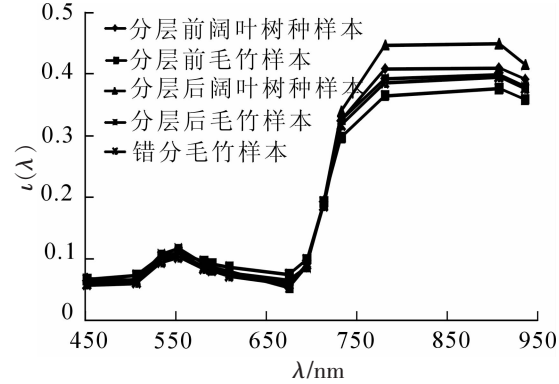


图 6 阔叶树种与毛竹光谱特征对照图
Figure 6 Comparison of spectral characteristics between broad leaved forest and moso bamboo

表 3 分层前后选取的阔叶树种、毛竹样本之间的光谱角

Table 3 Spectral angle between broad leaved forest and moso bamboo before and after stratification

项目	分层前夹角		分层后夹角	
	阔叶树种样本	毛竹样本	阔叶树种样本	毛竹样本
错分毛竹样本	0.029	0.038	0.060	0.014
分类结果	阔叶树种		毛竹	

4.2.3 4 类特征变量组合 特征变量 AISA，CHM，SIPI 和 PCA1 组合的分类结果总体精度和 Kappa 系数如表 4 所示。每小组中融合了 CHM 变量比未融合 CHM 变量的分类精度高；AISA+SIPI 比仅 AISA 的总体精度有明显提高，而 AISA+CHM+SIPI 的总体分类精度和 Kappa 系数达到最高，故 SIPI 有利于树种分类。然而添加 PCA1 这个变量后，最终分类效果均不优。这说明本研究选用 Wilks’ Lambda 法选择的波段，在树种识别应用中降维效果优于 PCA 法。因此，本研究的最佳分类变量的组合为 AISA+CHM+SIPI。由于阔叶林内树种种类复杂，常绿和落叶的树种间光谱相互干扰，加之空间上纯林区较少，阔叶树种仍存在混分现象(表 5)，与油茶、毛竹、针叶林混杂生长，影响了阔叶树种的分类精度。由于杉木和阔叶树种样本的高度分布频率比较近似，两者的训练样本提取困难，混分的可能性比较高；油茶为常绿小乔

表 4 分类结果的总体精度和 Kappa 系数

Table 4 Overall accuracy and Kappa coefficient of the classification results

特征变量	总体精度/%	Kappa 系数	特征变量	总体精度/%	Kappa 系数
AISA	65.31	0.56	AISA+PCA1	70.07	0.62
AISA+CHM	76.87	0.71	AISA+CHM+PCA1	74.15	0.68
AISA+SIPI	77.55	0.72	AISA+SIPI+PCA1	74.15	0.68
AISA+CHM+SIPI	89.12	0.86	AISA+CHM+SIPI+PCA1	78.23	0.72

表 5 AISA+CHM+SIPI 各树种分类的混淆矩阵及分类精度

Table 5 Accuracy and confusion matrix of five tree species classification under the combination of the AISA+CHM+SIPI

分类样本	参考样本					
	阔叶树种	马尾松	毛竹	杉木	油茶	总计
阔叶树种	27	4	0	2	3	36
马尾松	0	27	0	0	0	27
毛竹	1	0	19	2	0	22
杉木	3	0	0	30	0	33
油茶	0	0	1	0	28	29
非林地	0	0	0	0	0	0
制图精度/%	87.10	87.10	95.00	88.24	90.32	—
用户精度/%	75.00	100.00	86.36	90.91	96.55	—

木，其冠形和叶片与很多常绿阔叶树种相似，特别是高度相近的像元，仍会存在一小部分与阔叶树种混分；毛竹与阔叶树种的光谱比较近似，虽然经过分层分类的毛竹与阔叶树种混分较少，但由于少许像元距离相近，仍然有混合像元存在，产生少量混分。马尾松的光谱近红外平台峰值低于其他 4 类，较易区分。

4.2.4 单木尺度的精度 分析本文高光谱的空间分辨率为 2 m，一般单木树冠范围内的像元应为同一类，然而由于树冠间相互遮挡，单木树冠内不可避免出现多个树种的情况。叠加经单木冠层分割而得的树冠矢量图比较仅 AISA 和 AISA+CHM+SIPI 分类结果(图 7)，进一步得出 AISA+CHM+SIPI 结果中树冠内均质性更高。通过计算树冠内树种所占比例(正确分类像元数/树冠内总像元数)可说明分类效果，比例越大，分类效果越好(表 6)。AISA+CHM+SIPI 的分类结果在单木尺度上也优于仅 AISA 的结果，树种

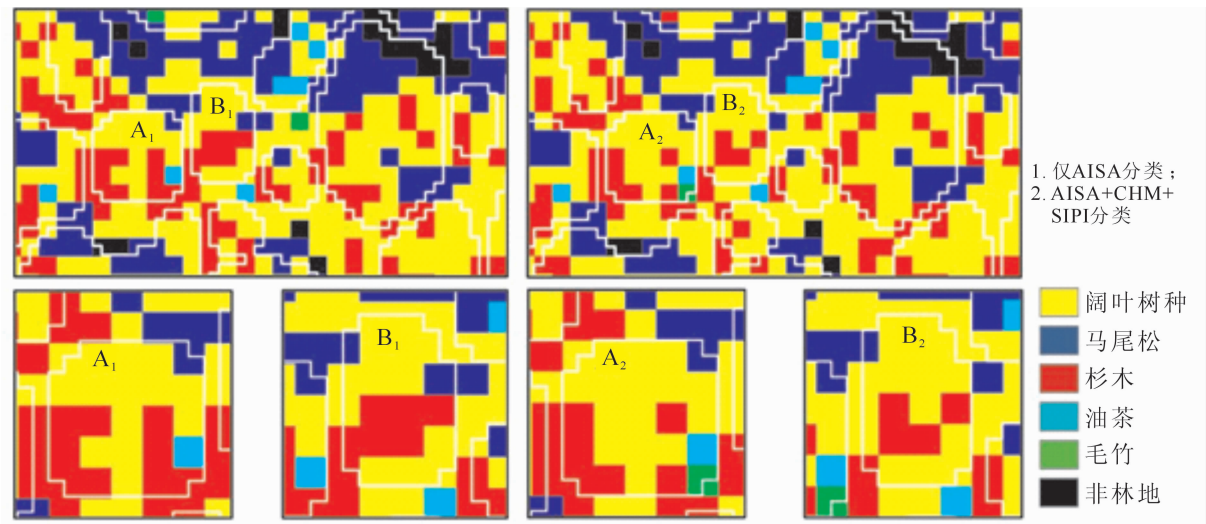


图 7 基于单木树冠的分类结果对比

Figure 7 Comparison of classification results based on individual canopy

纯度可达 70.00%以上。

4.3 分类结果专题图

基于 SVM 的高光谱 AISA+LiDAR 变量 CHM+植被指数 SIPI 的分类结果制专题图(图 8)。

5 结论与讨论

高光谱是光学被动遥感数据，其窄波段特性在较小的空间尺度上能区分地表细微变化，在树种识别方面有显著优势。但由于“同物异谱、异物同谱”现象的存在，导致分类精度受限。机载 LiDAR 是主动遥感数据，可获得树种垂直结构及强度信息，与高光谱优势互补，有效解决不同高度下不同地物因具有相似光谱特征而导致的混分问题。采用分层训练样本自动提取技术，不仅提高了训练样本选取的速度还有效提高训练样本选取精度，更从一定程度上尽可能地避免混分现象。

本研究结合高光谱与机载 LiDAR 的数据优势，为评估 LiDAR 垂直结构信息与特征变量参与分类的贡献，比较了基于 AISA，CHM，SIPI 和 PCA1 这 4 种不同变量组合的分类精度。其中 AISA+CHM+SIPI 变量组合的分类精度最高，其总体精度和 Kappa 系数分别为 89.12%

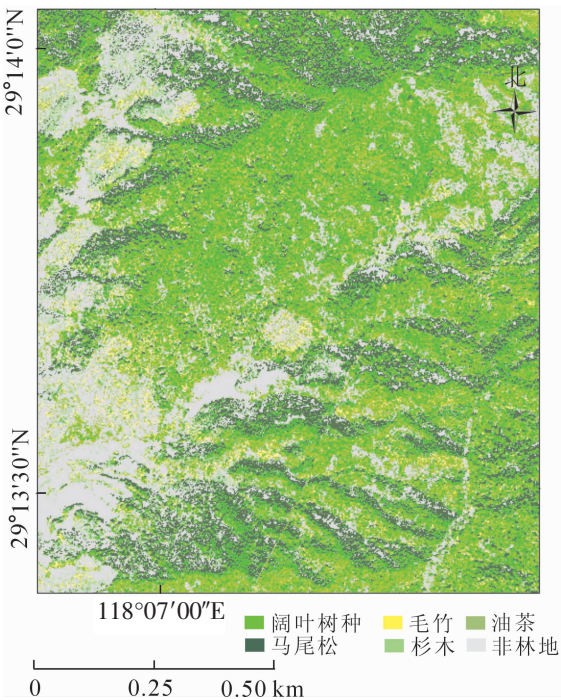


图 8 基于 SVM 的 AISA+CHM+SIPI 融合数据分类专题图

Figure 8 Thematic map of the classification for AISA + CHM+SIPI data with SVM

表 6 树冠内的成数对比

Table 6 Comparison of the percentage in the canopy

参考类别	序号	组别	总像元数	分类类别							比例/%
				阔叶树种	马尾松	毛竹	杉木	油茶	非林地	合计	
阔叶树种	1	A ₁	25	14	1	0	9	1	0	11	56.00
		A ₂		19	1	0	5	0	0	6	76.00
	2	B ₁	15	6	1	1	6	1	0	9	40.00
		B ₂		11	0	0	4	0	0	4	73.00

和 0.86，比仅 AISA 分类的总体精度高 23.81%，比 AISA+CHM 高 12.25%，比 AISA+SIPI 高 11.57%。但结果同时表明，PCA 降维变量的分类贡献要明显弱于 SIPI。本研究区为典型的亚热带森林，其中阔叶林内树种种类繁多，与其他类型的树种混杂生长，所以纯林区较少，易产生混合像元。在 AISA+CHM+SIPI 的分类结果中阔叶树种的制图精度和用户精度最高，分别为 87.10%和 75.00%，优于 AISA+CHM (70.97%，66.67%)以及仅 AISA 的分类结果(41.94%，61.90%)。这说明将机载 LiDAR 数据 CHM 与高光谱 AISA 融合，并添加植被指数 SIPI 能有效区分混合像元并提高分类精度，对古田山国家级自然保护区进行树种类型的精细分类具有可行性。但由于阴影区域的存在、树冠间相互遮挡、少部分边缘像元的光谱混合等，对树种分类的精度有一定影响。后续拟研究基于高空间分辨率数据的像元解混技术，期望能有效提高复杂林区的树种识别精度。

6 参考文献

- [1] JONES T G, COOPS N C, SHARMA T. Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada[J]. *Remote Sens Environ*, 2010, **114**(12): 2841 – 2852.
- [2] FALKOWSKI M J, HUDAK A T, CROOKSTON N L, *et al.* Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data [J]. *Can J For Res*, 2010, **40**(2): 184 – 199.
- [3] YU Xiaowei, HYYPPÄ J, KAARTINEN H, *et al.* Automatic detection of harvested trees and determination of forest growth using airborne laser scanning [J]. *Remote Sens Environ*, 2004, **90**(4): 451 – 462.
- [4] 王妮，彭世揆，李明诗. 基于树种分类的高分辨率遥感数据纹理特征分析[J]. *浙江农林大学学报*, 2012, **29**(2): 210 – 217.
WANG Ni, PENG Shikui, LI Mingshi. High-resolution remote sensing of textural images for tree species classification [J]. *J Zhejiang A&F Univ*, 2012, **29**(2): 210 – 217.
- [5] LEFSKY M A, COHEN W B, PARKER G G, *et al.* LiDAR remote sensing for ecosystem studies [J]. *BioScience*, 2002, **52**(1): 19 – 30.
- [6] ASNER G P, KNAPP D E, KENNEDY-BOWDOIN T, *et al.* Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR [J]. *Remote Sens Environ*, 2008, **112**(5): 1942 – 1955.
- [7] 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 trees pecies identification in a natural temperate forest [J]. *J Remote Sens*, 2013, **17**(3): 679 – 695.
- [8] 董彦芳，庞勇，许丽娜，等. 高光谱遥感影像与机载 LiDAR 数据融合的地物提取方法研究[J]. *遥感信息*, 2014, **29**(6): 73 – 76.
DONG Yanfang, PANG Yong, XU Lina, *et al.* Feature extraction based on fusion of hyperspectral image and airborne LiDAR data [J]. *Remote Sens Inf*, 2014, **29**(6): 73 – 76.
- [9] CAO Lin, COOPS N C, INNES J L, *et al.* Tree species classification in subtropical forests using small-footprint full-waveform LiDAR data [J]. *Int J Appl Earth Obs*, 2016, **49**: 39 – 51.
- [10] ALONZO M, BOOKHAGEN B, ROBERTS D A. Urban tree species mapping using hyperspectral and lidar data fusion [J]. *Remote Sens Environ*, 2014, **148**: 70 – 83.
- [11] DALPONTE M, BRUZZONE L, GIANELLE D. Fusion of hyperspectral and LIDAR remote sensing data for classifi-

- cation of complex forest areas [J]. *IEEE Trans Geosci Remote Sens*, 2008, **46**(5): 1416 – 1427.
- [12] 王春来, 张森原, 崔璐, 等. 训练样本数量选择和总体分类精度的关系研究[J]. 河南城建学院学报, 2015, **24**(3): 51 – 55.
- WANG Chunlai, ZHANG Senyuan, CUI Lu, *et al.* Relationship between number of training samples selection and accuracy of overall classification [J]. *J Henan Univ Urban Construct*, 2015, **24**(3): 51 – 55.
- [13] PANG Yong, LI Zengyuan, JU Hongbo, *et al.* LiCHy: The CAF's LiDAR, CCD and hyperspectral integrated airborne observation system [J]. *Remote Sens*, 2016, **8**(5): 398. doi: 10.3390/rs8050398.
- [14] 郝建亭, 杨武年, 李玉霞, 等. 基于 FLAASH 的多光谱影像大气校正应用研究[J]. 遥感信息, 2008(1): 78 – 81.
- HAO Jianting, YANG Wunian, LI Yuxia, *et al.* Atmospheric correction of multi-spectral imagery ASTER [J]. *Remote Sens Inf*, 2008(1): 78 – 81.
- [15] GEORGE R, PADALIA H, KUSHWAHA S P S. Forest tree species discrimination in western Himalaya using EO-1 Hyperion [J]. *Int J Appl Earth Obs*, 2014, **28**(1): 140 – 149.
- [16] 冯云云, 刘丽娟, 陆灯盛, 等. 基于曲线误差指数的高光谱降维方法评价指标研究[J]. 浙江农林大学学报, 2017, **34**(5): 765 – 774.
- FENG Yunyun, LIU Lijuan, LU Dengsheng, *et al.* A comparative study of reduction methods for airborne hyperspectral images [J]. *J Zhejiang A&F Univ*, 2017, **34**(5): 765 – 774.
- [17] GREEN P E, CAROLL J D. *Analyzing Multivariate Data* [M]. Illinois: The Dryden Press, 1978.
- [18] DYE M, MUTANGA O, ISMAIL R. Examining the utility of random forest and AISA Eagle hyperspectral image data to predict *Pinus patula* age in KwaZulu-Natal, South Africa [J]. *Geocarto Int*, 2011, **26**(4): 275 – 289.
- [19] 杜朋选. 川滇高山栎生理活动对不同程度水分胁迫的反射光谱响应[D]. 北京: 中国林业科学研究院, 2011.
- DU Pengxuan. *Quercus aquifolioides Physical Activity in Response to Different Degrees of Water Stress on the Reflection Spectral* [D]. Beijing: Chinese Academy of Forestry, 2011.