

基于随机森林分类方法和多源遥感数据的湿地植被精细分类

武思媛

中交和美环境生态建设有限公司 湖北 武汉 430000

[摘要]以随机森林分类方法以及多源遥感数据为基础对湿地植被进行精细分类的相关研究相对较少。在此次研究过程中,主要是以随机森林分类方法为主,探讨多源遥感数据在湿地植被进行分类过程中的应用要点。在研究时,面向对象分类方法有利于减少椒盐效应,而随机森林分类算法的准确度相对较高,具有较强的应用价值。为了保证湿地植被分类效果,可以对面向对象的分割参数进行调整。研究发现,该精细分类方法的应用准确性较高,可以构建最优的面向对象随机森林分类模型。

[关键词]随机森林分类方法;多源遥感数据;湿地植被分类

【DOI】10.12252/j.issn.2096-6261.2021.08.1478

0 引言

目前,国内外的很多学者都对湿地植被精细分类方法进行了深入研究。现阶段,比较常用的湿地植被分类方法包括决策树分类方法以及随机森林分类方法、面向对象分类方法等。在具体的研究过程中,有研究人员以崇明动态湿地作为主要对象,通过TM影像识别信息的方法,完成湿地植被神经网络监督分类。这一分类方法需要的专业知识比较少,能够掌握湿地植被的分布状况和动态变化趋势。在实际研究中结合面向对象分类方法完成湿地植物群落分类。可以确定面向对象方法利用尺度合适的对象作为分类单元,可以对具体的类型进行科学分类,同时能够解决在传统分类方法应用过程中存在的分类板块破碎问题。但是在应用过程中对人工依赖度相对较强,因此,需要对基于随机森林分类方法和多源遥感数据的植被精细分类进行进一步研究。

1 森林湿地分类方法研究意义

湿地本身是重要的生态系统,可以防洪蓄水、调节气温、美化环境,被称为地球之肾和物种基因库,与海洋、森林组成三大生态系统。森林湿地是湿地的主要组成部分之一,主要是在有积水或者地面比较湿的地段集中喜湿和水生植物作为主要物种。在此基础上,构建构成的森林植被。森林湿地的面积比较大,具有较强的代表性。森林湿地主要分布在寒带或者寒温带湿润气候区,例如西伯利亚是地球上最大的泥炭沼泽区,森林湿地的分布面积也比较广,主要以针叶林湿地为主。除此之外,波兰、俄罗斯、德国、挪威等也是森林湿地分布面积比较广阔的国家。森林湿地可以为很多物种提供生存繁殖场所,能够维持生态系统平衡。如果对森林湿地进行过度开发利用,会影响珍稀物种面积以及森林湿地环境的生态平衡,并且会对生物的繁衍产生影响。对森林湿地进行科学分类,可以在掌握森林湿地具体情况的基础上进行开发和利用,有利于保护森林湿地系统。因此,对森林湿地分离方法进行研究具有至关重要的作用。

森林湿地受地形条件限制,对其进行实地研究时存在很大局限性。但是全球分布的森林湿地空间比较多,对先进的遥感技术进行应用,可以在短时间内完成大范围区域信息获

取。因此,在湿地系统检测以及森林湿地信息提取过程中,遥感技术具有至关重要的作用。20世纪60年代,美国和加拿大等国家已经开始对遥感技术进行应用完成森林湿地信息识别和监测;在20世纪50年代开始我国对森林湿地的相关研究进行普查,初步掌握了森林湿地的具体类型、形成原因和空间分布规律等。而在遥感技术不断发展的背景下,利用遥感手段完成森林湿地识别和信息提取,可以对森林湿地植被进行精确分类。在研究过程中,要加强森林湿地遥感信息提取方法分析,有利于明确森林湿地空间分布情况,对保护森林湿地环境有重要意义。

2 源数据分析与处理

在此次研究过程中主要是对某地区森林湿地植被进行精细分类。这一地区是典型森林湿地区域,气候条件比较优,也是亚洲东部最大的湿地植被资源库,物种比较丰富,以耐湿乔灌木和湿地草本植物为主。在该地区分布大量的湿地植被,包括互花米草、芦苇、盐蒿等,这些植被具有较强的耐盐碱度。在此次研究过程中需要完成源数据提取以及预处理工作。此次研究的数据源主要以机载LIDAR数据为主,搭载的机载激光雷达系统采集相关数据;高光谱遥感影像数据是同一时间搭载的机载高光谱扫描仪获取的高光谱影像数据信息。在对数据进行预处理时,需要将数据上传在平台上完成条带拼接,奇异点删除以及点云滤波处理。其中条带拼接主要是对不同航带之间的数据进行集合匹配;奇异点删除主要是因为数据采集过程中,空中的飞鸟、塑料袋等会对目标影像产生影响,在点云数据中生成异常值点,影响数据的整体质量;而点云滤波处理主要是滤波非地面点。

高光谱遥感影像在处理过程中,需要将激光雷达CHM作为主要参考影像,而高光谱影像是校正影像,在影像中可以采集控制点,确保其在影像区均匀分布,同时要对误差进行合理控制^[2]。为了保证湿地植被分类的精准度,降低在分类结果中存在的椒盐效应,可以利用面向对象分割算法完成影像分割。在具体的分割处理过程中,所有的波段权重为1,形状因子为0.1,紧凑度因子为0.5,分割尺度为100。

3 研究方法

3.1 数据信息提取

在对数据信息进行提取的过程中,需要完成以下两种信息处理工作:(1)样本信息处理。需要选择具有较强典型性、代表性的纯净对象作为样本。一般情况下,可以利用简单随机抽样、等距抽样、分层随机抽样等不同方法完成样本信息提取。因为RF算法分类精度对样本数量和空间分布的依赖度相对较大。因此,在此次研究中主要是以分层随机抽样为主完成样本信息提取。(2)完成对象特征信息提取工作。在对象特征提取过程中主要是对机载LIDAR数据以及高光谱影像特征信息进行提取。其中机载LIDAR数据提取主要是获取地物三维信息,在辅助影像分类与地物信息获取方面具有至关重要的作用。在高光谱影像特征信息提取过程中,主要是以光谱空间的降维处理为基础,在波段之间建立重新组合,提取的结果可以直接表示地物的特性。特征信息提取对提高高光谱数据处理效率有重要作用,可以保存高光谱数据中地物的主要信息,同时能够解决数据量比较大形成的计算机处理荷载大等问题^[1]。在实际分类过程中高光谱影像采取最小噪声分离变换可以集中信息,将其分布在有限的特征之内,并利用使用类特征向量汇集方式提高分类信息的精准度。在选择器选择过程中需要选择典型地类。基于面向对象分类方法要确保样本数量在合理范围内,随着样本数量不断增加,分类精度会逐渐趋于饱和状态。在选取训练样本后,要以训练样本光谱特征、形状特征以及纹理特征等确定分类阈值,最终获取高光谱影像的分类特征信息。

3.2 基于面向对象的湿地植被精细分类

在实际研究过程中需要掌握随机森林基本原理:随机森林主要是若干棵没有关联的决策树随机组成形成的整个森林。在随机森林方法应用中,变量以及数据可以被随机抽取,完成训练,生成更多的决策树,最终汇总决策树的结果。每一个决策树都是以根节点为基础从上到下递归分裂获取二叉树,二叉树的根节点数据利用训练方式将所有的数据按照节点程度最小的原子进行分裂,形成两个节点,训练数据也可以被分为两个子集。使用相同的方法使两个节点可以继续分裂,一直到满足分裂停止规则为止。

4 研究结果分析

4.1 分类精度检验

为了对随机森林算法的具体分类效果进行准确掌握,在此次实验过程中需要利用SVM算法进行对比,在实验中传统算法选择的特征和随机森林相同。经过分析,获取以下结论(1)与传统的支持向量机分类算法79.5%相比,随机森林分类算法的精度达到90.3%;随机森林的Kappa系数为0.87,提高大约19%左右。(2)在对互花米草以及芦苇的识别度进

行分析时,发现个别长势相对较好的芦苇和互花米草的特征与林地比较相似,存在误分情况。(3)随机森林分类方法与传统分类方法在对盐蒿、互花米草、芦苇、林地等进行识别时,分类精度明显更优,分别提升3.9%、10.6%、15.8%、23.6%左右。因此,在对湿地植被精细分类过程中,对随机森林分类方法进行应用可以大大提升分类精度。

4.2 决策树深度对分类精度的影响

利用随机森林模型进行湿地植被精细分类,需要掌握决策树深度对分类结果产生的具体影响。在验证过程中,决策树最大个数保持70、特征个数为17、深度数量按照1~150完成分类。分类模型的总体精度随着树深度不断增高,先出现上升,之后为下降的趋势。在树深度增加到30时,总体精度为最高值0.9,总体精度与Kappa系数的趋势相同。这意味着树深度设置会对总体精度产生极大影响,在树深度在为30时,可以大大提高随机森林模型的分类精度^[3-4]。

4.3 森林规模与分类精度之间的关系

在对森林规模和分类精度产之间的关系进行分析时,参与分类的特征个数、决策树的深度为30。经过实验确定决策树个数不断变化,总体精度和时间出现缓慢上升的趋势,决策树个数超过9时,分类时间不断增加,决策树个数达到30时,总体精度为最大值,分类时间也比较短,进一步证明了决策树个数为30时,能够提高森林随机模型的分类质量。

5 结束语

综上,在此次研究过程中可以确定利用随机森林模型完成湿地植被精细分类具有较强的分类精度。将其与传统分类方法进行对比验证,可以获取以下结论:随机森林方法在高光谱影像分类和湿地植被精细分类过程中具有明显的应用优势,分类速度相对较快,可以保证分类精度,是一种对湿地植被精细分类的有效方法。

参考文献

- [1]李方方,刘正军,徐强强,等.面向对象随机森林方法在湿地植被分类的应用[J].遥感信息,2018,33(1):6.
- [2]宗影,李玉凤,刘红玉.基于面向对象随机森林方法的滨海湿地植被分类研究[J].南京师范大学学报(工程技术版),2021,21(04):47-55.
- [3]侯蒙京,殷建鹏,葛静,等.基于随机森林的高寒湿地地区土地覆盖遥感分类方法[J].农业机械学报,2020,51(7):8.
- [4]张瑜伟,苗小莉,张泳.基于多源遥感数据的植被覆盖度反演方法比较研究[J].测绘与空间地理信息,2020,43(03):131-134+137.