论文

GlobeLand30 遥感制图创新与大数据分析专题

顾及多元知识的 GlobeLand30 检核优化模型

张委伟^{①②*}, 陈军^①, 廖安平^①, 韩刚^①, 陈学泓^③, 陈利军^①, 彭舒^①, 武昊^①, 张俊^①

① 国家基础地理信息中心, 北京 100830;

② 苏州科技大学环境科学与工程学院, 苏州 215011;

③ 北京师范大学地表过程与资源生态国家重点实验室, 北京 100875

* E-mail: zww9820@126.com

收稿日期: 2015-10-19; 接受日期: 2016-03-28; 网络版发表日期: 2016-08-25

国家自然科学基金重点项目(批准号: 41231172)、测绘地理信息公益性行业科技专项项目(编号: 201512028)和国家高技术研究发展计划项目(编号: 2013AA122802)资助

摘要 综合运用自然、生态和人文等多元知识辅助计算机遥感自动分类, 确保大区域地表覆盖遥感数据产品质量是多年来国际上致力于解决的科技难题. 在全球 30m 地表覆盖遥感数据产品 GlobeLand30 的研制过程中, 作者提出了一种顾及多元知识的检核优化模型, 对基于“像元-对象”的遥感自动分类结果进行知识化检核, 发现和修改错分/漏分问题. 首先归纳总结全球地表覆盖分布相关的先验知识, 凝练出有关地表覆盖的地域地形分布规律、时间连续性和空间关系等质量检核知识规则; 继而发展了基于网络化服务平台的交互式协同检核技术, 辅助发现与定位分类异常区域, 并依据各类别最小图斑规定判定错分/漏分, 进而标注和发布错误信息, 供研制人员用于修改优化. 以林灌草和裸地数据为例, 抽取 10% 的图幅, 进行优化前后的分类精度对比分析, 得知质量明显改善, 精度提高最多达到 17.47%. 运用该模型, 对 GlobeLand30 数据进行了逐类检核与优化处理, 有效地减少了错漏分问题, 2010 期产品的总体精度达到 83.50%. 这一模型还可用于其他全球或区域的地表覆盖数据产品研制.

关键词 GlobeLand30, 地表覆盖, 数据优化, 多元知识, 服务平台, 交互式协同检核

1 引言

研制高质量的地表覆盖数据产品, 满足全球环境变化研究(气候变化、生物多样性的变化等)、资源管理和可持续发展规划等的急需, 是中国全球30m地表覆盖遥感制图研究的主要目的之一(陈军等, 2011, 2014; 徐冠华等, 2013; 施建成, 2014; 冉有华和李新, 2015;

Chen等, 2016). 这里地表覆盖数据产品的质量主要是指单期的空间一致性(分类图斑属性及其范围与真实世界的相吻合程度)和多期产品的时间连续性等.

近年来基于遥感的地表覆盖自动分类技术取得了长足的进步(Lu和Weng, 2007), 但大区域分类精度依然不高(宫鹏, 2009; Hansen和Loveland, 2012; 胡运等, 2014). 对于全球30m地表覆盖遥感制图来说,

中文引用格式: 张委伟, 陈军, 廖安平, 韩刚, 陈学泓, 陈利军, 彭舒, 武昊, 张俊. 2016. 顾及多元知识的 GlobeLand30 检核优化模型. 中国科学: 地球科学, 46: 1149–1161, doi: 10.1360/N072015-00286

英文引用格式: Zhang W W, Chen J, Liao A P, Han G, Chen X H, Chen L J, Peng S, Wu H, Zhang J. 2016. Geospatial knowledge-based verification and improvement of GlobeLand30. Science China Earth Sciences, 59: 1709–1719, doi: 10.1007/s11430-016-5318-4

由于地表景观在30m分辨率遥感影像上表现出极为复杂多样的光谱和纹理特征,同物异谱和异物同谱现象更为严重,再加上许多区域难以获取时相适宜的影像数据以及获取的影像数据质量难以保证等诸多问题(Giri等, 2013; 陈军等, 2014; 廖安平等, 2014),导致单纯遥感自动分类制图错分和漏分问题十分突出,难以实现高精度要求,无法满足在地学、生态学和环境学等领域对遥感产品的精度需求. 尚需要集成各种地学相关知识并建立适宜于规模化遥感制图的方法,来提高遥感分类精度. 地表覆盖数据产品质量检核优化就是综合利用各种知识与技术,查找和改正遥感自动分类的错分/漏分问题,提高数据质量. 值得指出的是,一方面,现有的商业遥感软件(如ENVI和ERDAS等)仅提供了对分类后的图斑进行简单的小图斑去除、类别合并等几何处理功能以及基于检验样本的分类成果精度验证,缺乏对图斑属性正确与否进行检查的方法与机制. 另一方面,以往全球地表覆盖遥感制图研究大多关注自动分类算法的改进与数据的验证,对分类数据的质量检核优化缺乏系统深入的研究,尚未形成可操作的规模化检核优化方法(Loveland等, 2000; Hansen等, 2000; Friedl等, 2002; Mayaux等, 2006; Bontemps等, 2011; Gong等, 2013; Hansen等, 2013). 如何综合利用地表覆盖分布相关的遥感、地理和生态等多元知识,形成一套可应用于大规模工程实践的行之有效的技术方法,

检核优化全球地表覆盖数据,弥补单纯遥感自动分类技术的不足,确保产品质量,是中国全球30m地表覆盖遥感制图面临的一大挑战.

国家高技术研究发展计划项目提出了基于“像元-对象-知识”(pixel-object-knowledge, POK)的分层分类方法(每一类型的提取是在利用已提取地类数据对分类影像掩膜基础上进行的)(Chen等, 2015, 2016),开展全球30m地表覆盖遥感制图. 本文重点是介绍K-顾及多元知识的GlobeLand30检核优化模型,其是集成各种地学相关知识,借助网络服务技术,可应用于规模化遥感制图以提高数据质量. 利用该模型,对全球30m地表覆盖遥感制图中基于“像元-对象”分类的初步结果,按照水体-湿地-冰雪-人造地表-耕地-森林、灌丛、草地和裸地(林灌草和裸地)的顺序,依次进行检核优化,取得了较好的结果.

2 总体思路

受自然、人文和生态等多种因素的制约,地表覆盖分布及变化具有一定的规律. 为了充分利用各种地表覆盖分布及变化相关的先验知识达到提高数据产品质量的目的,文章提出了适宜于规模化遥感制图的顾及多元知识的GlobeLand30数据检核优化模型(图1). 在征集的专家意见基础上,查找、收集地表覆

图1 顾及多元知识的 GlobeLand30 检核优化

盖分布相关的各种先验知识,并对其进行归纳总结,凝练出检核的知识规则;继而发展了基于网络化服务平台的交互式协同检核技术,以知识规则提炼-分类异常区域发现-错分/漏分问题判定-数据优化为主线,逐类实现对基于“像元-对像”的GlobeLand30自动分类结果的检核与优化。

通过系统地梳理收集到的全球地表覆盖分布相关的各种先验知识,总结出各地表覆盖类型宏观空间分布、地形分布规律(坡度、高程等)、地类分布相关性和变化规律等知识,提炼形成针对GlobeLand30数据的空间一致性、时间连续性等的知识规则。归纳地表覆盖类型的30m遥感影像特征和制图综合要求,制定30m尺度地表覆盖遥感制图各类型的最小图斑,作为错分/漏分问题是否需要修改的判断指标。

利用集成了多源异构参考资料,具备在线标报和发布功能的服务平台,发展了交互式协同检核技术。基于知识规则,通过统计数据对比、空间分析及交互式对比分析等方法,发现并定位分类异常区域(如水体和湿地坡度过大、两期数据变化较大、与已有地表覆盖数据大范围不一致等);然后,通过服务平台外接服务上的高分影像和最小图斑规定判断出需要修改的错分/漏分问题,由检查人员标注并发布在网络上,对应的修改人员订阅错误信息对其进行修改。从而,实现了数据检核优化工作的信息交互和多用户协同,提高了检核效率。

3 基于多元知识的检核规则提炼

长期以来,为了有效提高地表覆盖数据的质量,学者们在不断改进遥感自动分类方法的同时,开展了利用自然、人文和生态等各种先验知识对分类结果进行检核的研究(陈军, 1984; Harris和Ventura, 1995; Griguolo, 1999; Comber等, 2004; Judex等, 2006; Manandhar等, 2009; Kantakumar等, 2016)。自然地理规律特征(如气候和地形等)、一般性常识、已有地表覆盖数据和统计数据等辅助知识被用于对区域分类成果进行检核与优化。如Manandhar等(2009)在澳大利亚新南威尔士土地利用制图中,总结了葡萄园分布在250m高程以内的非军事区与非国家森林公园区域等知识,对MLC分类结果中的葡萄园、橄榄农场等地类进行了检核与修正。全球尺度上,对地表覆盖分

布相关知识的总结相对缺乏,利用地表覆盖分布相关知识进行数据检核的研究较少(Loveland等, 2000; Hansen等, 2000; Friedl等, 2002; Bontemps等, 2011; Gong等, 2013),这也是它们精度普遍不高,难以满足科研应用需求的原因之一(Herold等, 2008; 宫鹏, 2009; Verburg等, 2011)。

为了检核GlobeLand30数据的空间一致性、数据中各类型间的关系协调性和两期数据之间的时间连续性,本文对与全球地表覆盖相关的多元先验知识进行梳理总结,进而凝练出表征地表覆盖分布及变化规律的知识规则,作为GlobeLand30数据检核中可能分类问题区域的判定依据。

3.1 多元知识

这里的多元知识主要是指自然、人文和生态等领域与全球地表覆盖的分布及其变化相关的常识、规律和过程。通过征集遥感、地理和生态等领域专家的意见,收集已有参考数据、DEM和地类统计数据等资料,获取地表覆盖分布相关的信息,归纳总结以形成宏观分布、区域过渡和地形分布规律等的先验知识,可划分为如下几类:基于自然地理的地表覆盖分布知识、基于人文的地表覆盖分布知识和地表覆盖变化知识等。

3.1.1 基于自然地理的地表覆盖分布知识

地表覆盖各类型的空间分布与经纬度、高程、地形和气候等自然地理环境要素密切相关。各自然地理要素的组合与相互作用导致了地球表面各种各样的覆盖分布规律(Fang等, 2002)。通过各种参考资料对各种规律进行总结,形成基于自然地理的地表覆盖分布的先验知识。如受气候条件制约形成的分布规律。主要是受热量和水分条件的制约,植被在全球呈现出了纬度地带性、经度地带性和垂直地带性的分布(武吉华等, 2004),不同的生态分区形成了不同的优势地表覆盖类型;受地形条件影响,水体和湿地多分布在比较平缓的区域,绝大多数坡度小于5°(牛振国等, 2009);受高程影响,青藏高原地区耕地在4700m以上很少存在(张晶等, 2007),一般冰川和永久积雪只有山体高度超过某地的雪线高度才有可能存在等。

3.1.2 基于人文的地表覆盖分布知识

随着社会的发展,人类在地球表面的活动范围也越来越大,其文化传统、政治政策和社会经济事件等都影响着地表覆盖类型的空间分布(刘纪远等, 2014).同时,人类在各种活动中形成的地表覆盖图、地理名称标注和地表覆盖统计数据等信息也间接地反映了地表覆盖的分布特征.主要可以分为三类: (1) 人类活动形成的地表覆盖类型分布特征.如耕地分布区域一般应该伴随有人造地表,人造地表周围一般有耕地存在;耕地如葡萄园等一般不会出现军事管理区和自然保护区内(Manandhar等, 2009). (2) 各种地表覆盖统计数据.如中国统计年鉴中,都会有各个省的耕地统计数据;联合国FAO定期发布全球各国森林、耕地等统计数据、全球农业用地统计数据(Ramankutty等, 2008)和湿地手册中的湿地统计数据(Whigham, 2009)等. (3) 已有地表覆盖及相关数据.几十年来,已经有各种区域的、全球的土地覆盖/土地利用数据研制出来,如中国1:10万土地利用数据库(Zhang等, 2014),欧洲太空局的GlobCover2009 (Bontemps等, 2011)等;也有带有位置信息的地表覆盖名称数据,如世界冰川名录(Cogley和Graham, 2010)等.

3.1.3 地表覆盖变化知识

GlobeLand30数据产品共有两期: 2010期和2000期.一定时间间隔的地表覆盖数据之间的差异在空间上表征了两个时段地表覆盖的变化,但是有些变化可能不是现实世界真正的变化,而是来自于地表覆盖特征的随机的或系统性的错误,是“伪变化”.而在连续相继的地表覆盖数据之间存在着许多时相约束,总结出相关知识,提炼出知识规则,可以用来发现两期数据的异常区域,减少“伪变化”.如随着全球变暖,冰川和永久积雪表现出萎缩的趋势(Yao等, 2012);在绝大多数情况下,人造地表建成之后会存续相当长时间,不会突然消失;某一区域地表覆盖类型(尤其是受人类影响较小的)两期一般不会剧烈变化等.

3.2 知识规则的提炼

仅仅依靠总结的全球地表覆盖分布与变化相关的先验知识,难以支持GlobeLand30数据的检核.为了有效地利用多元知识辅助全球地表覆盖数据检核优化,将全球地表覆盖分布及变化相关的先验知识进行凝练,形成地表覆盖数据空间一致性、时间连续性和关系协调性的知识规则(表1).

表 1 先验知识与知识规则示例

知识类型	先验知识	分类异常区域判断的知识规则
基于自然地理的地表覆盖分布知识	耕地(cultivated land)一般分布在坡度(slope)较缓的地表	$Slope_{cultivated\ land} < 30^{\circ}$
	水体(water)存在于平坦或低缓地区的地表	$Slope_{water} < 5^{\circ}$
	不同生态地理分区(eco-geographical region)具有不同的优势地表覆盖类型(main land cover, MainLC)	$MainLC_{GlobeLand30} = MainLC_{eco-geographical\ region}$
	冰川和永久积雪(ice and snow)只有山体高度超过某地的雪线(snow-line)高度(Elevation),才有可能存在	$Elevation_{ice\ and\ snow} < Elevation_{snow-line}$
基于人文的地表覆盖分布知识	耕地与人造地表(artificial surfaces, AS)通常交错分布,相伴随而存在	Cultivated land && AS
	永久积雪和冰川可能与世界冰川名录(World Glacier Inventory, WGI)存在一致性	Ice and snow && WGI
	全球基础地理底图数据库(Global Basic Geographic Base Map Database, GBGBMD)的居民点图层(Settlement)是对人造地表的直观反映,二者应该是相互对应的	Settlement _{GBGBMD} && AS
	与地表覆盖相关的社会经济普查数据或是调查统计数据(statistic data, SD)一定程度上反映了地表覆盖类型的分布,其差异应在一定范围(threshold)内	$ SD_{GlobeLand30} - SD_{other} < Threshold$
地表覆盖变化知识	大部分人造覆盖修建后会持续存在	$Area_{AS-2010} - Area_{AS-2000} > 0$
	全球变暖,冰川和永久积雪分布有萎缩的趋势	$Area_{ice\ and\ snow-2010} - Area_{ice\ and\ snow-2000} < 0$

4 基于网络化服务平台的交互式协同检核

全球30m地表覆盖遥感制图项目研制的网络化服务平台(图2), 具备集成海量的多源异构、存在格式或坐标系的不统一的参考数据的功能, 并能够进行分屏浏览、在线标报和发布订阅等(Han等, 2015). 依托于此, 针对GlobeLand30遥感自动分类结果, 发展了基于网络化服务平台的交互式协同检核技术. 数据检核需要的各种参考资料和各种空间分析运算中间结果都可以集成在服务平台上, 辅助检查人员进行目视对比分析、异常区域定位、错分/漏分判别, 将发现的问题进行标报, 并及时反馈给修改人员. 利用服务平台, 较好地解决了规模化制图中参考资料多和众多参与人员相互配合的问题, 支持同时多人多任务在线操作, 提高了数据检核优化的效率和各种辅助参考资料的利用.

4.1 多源异构资料集成

在以往地表覆盖分类产品生产的过程中, 各生产机构和研究人员就已经开始充分收集并利用相关的各种地形地貌、文献资料以及网络在线资源, 形成地表覆盖生产辅助资料集和专家知识库, 辅助地表覆盖产品的分类生产、质量控制与精度提高都具有非常重要的作用(Muchoney等, 1999; Mayaux等,

2004; Bicheron, 2008). 但其使用上仍有诸多问题和不便: (1) 拷贝费时费力, 存在大量冗余; (2) 统一维护困难, 难以及时更新; (3) 数据复杂多样, 难以有效利用; (4) 各种在线资源, 缺乏集成手段.

服务平台能够集成检查需要的成果数据、需要使用的辅助数据(原始影像、生态分区、现有的地表覆盖数据和高分影像等)以及空间分析中间结果. 基于web service技术通过分析影像和各类参考数据的特点, 根据数据检查的应用需求确定数据用途, 进而设计出适合的集成处理与服务发布方法, 将异质的分类影像、已有地表覆盖数据、DEM等辅助数据和中间成果等集成发布在参考资料系统, 对数据使用者屏蔽了各类数据的格式、坐标和投影系统, 免去了大量数据传递和重复数据处理工作, 并接入谷歌高分、天地图等异构服务, 远程数据使用亦非常方便(Wu等, 2011).

4.2 在线标报与发布

服务平台具有数据分屏浏览和错误信息的空间化标报功能, 并能对标报内容进行发布/订阅等, 用于支持多用户协同的检查信息和修改结果的在线汇聚、有序传递. 利用数据浏览和分屏对比, 检核人员可以将GlobeLand30数据与天地图、Google等在线高

图2 GLOBELAND30 服务平台(www.globeland30.org)

分辨率影像数据或其他检核参考数据进行联动对比, 基于相关知识规则, 定位分类异常区域, 判别发现数据成果中存在的错提、漏提等质量问题, 并允许检核人员以点、线、面的方式标绘出现错误区域, 自动记录下错误类型、空间位置和错误描述等信息, 并发布在网络平台, 通过在线汇聚, 并有序的传递, 最终形成需要修改的错分/漏分信息. 在线标报具体流程见图3.

4.3 交互式协同检核

在服务平台的支撑下, 检查人员通过统计对比分析、空间分析和人机交互目视对比等手段发现并定位分类异常, 利用接入的高分影像进行错分/漏分判断, 并对错误信息的类型、位置、原因描述、修改意见以及图幅信息、检查时间和参考图片等信息进行标报. 管理人员负责分配任务并对检查人员发布的标注信息进行审核, 审核通过后的标注信息作为正式的检核结果, 发布到服务平台. 同时, 将检查人员的标注结果在线汇聚, 并有序的传递; 修改人员订阅矢量化格式的检核标注信息, 参考该信息的修改建议对数据进行改正, 并将修改结果在线汇聚. 基于服务平台的交互式协同检核, 实现了数据检核与数据优化工作之间的信息交互和多用户协同, 从而有效地实现了两期数据的检核与优化. 知识化检核的网络化流程如图4所示.

5 GlobeLand30 数据检核与优化

5.1 基于知识规则的分类异常区发现

基于地表覆盖分布及变化的知识规则, 通过数据统计分析、空间分析和人机交互目视对比等手段, 发现并定位分类异常区域.

5.1.1 统计数据对比

GlobeLand30的各类统计分析结果, 包括按照大洲和国家等统计总面积、覆盖比率、变化比率等与查找到的参考文献数据、官方发布的统计数据以及其他

图3 在线标报流程图

图4 交互式协同检核实现

土地利用/土地覆盖数据统计结果等应该具有一致性,其差值应该在一定的范围内.因而,将两者进行逻辑分析、变化对比,根据凝练的知识规则,研判GlobeLand30数据的正确合理与否,找出数据可能出现错误的区域,并通过服务平台进行定位.以人造地表为例,表2是按照国家对GlobeLand30的两期人造地表面积进行统计对比,发现安哥拉两期人造地表变化比率为70.06%,其变化不符合常规,在服务平台上,对安哥拉人造地表错分与漏分检核判断.对于面积较大的国家,则可以考虑按照其国内行政区划进行统计对比,如中国,则进行了分省对比.

5.1.2 空间分析筛查

对有些知识规则(如 $Slope_{water} < 5^\circ$ 等),则可以通过数据图层之间进行空间分析,如叠置分析和空间差异计算等来实现,直接将分类异常区域筛查出来,并将表征异常区域的数据集成到服务平台上进行错分漏分的判断.如:基于裸地定义为植被覆盖度小于10%的土地,利用MODIS NDVI计算植被覆盖度,与GlobeLand30裸地数据通过空间叠置功能,得到对应植被覆盖度大于10%的裸地确定为分类异常区域(Zhang等, 2015);人造地表一般建好后会持续存在,尤其是在发展中国家,人造地表整体趋势为增加,则可以对两期数据进行空间差值计算,生成人造地表变化数据(其中,人造地表成片减少区域则为分类异常区域),将其集成到服务平台上,对这些区域详细

表2 基于GlobeLand30各国两期人造地表面积对比(非洲)

国家	人造地表		
	2000年(km ²)	2010年(km ²)	变化率
阿尔及利亚	2770.50	3115.75	12.46%
安哥拉	1457.34	2478.35	70.06%
贝宁	557.84	631.61	13.22%
博茨瓦纳	1186.85	1343.76	13.22%
布隆迪	120.42	133.83	11.13%
喀麦隆	1169.80	1323.86	13.17%
加那利	295.54	314.12	6.29%
佛得角	35.32	35.83	1.44%
中非	460.07	514.42	11.81%
乍得	885.11	1049.49	18.57%
科摩罗	28.44	29.65	4.27%
刚果(金)	4213.10	4743.86	12.60%
科特迪瓦	1662.31	1949.55	17.28%
吉布提	49.59	56.78	14.51%

检查.图5为中国河北唐县附近两期人造地表变化图,可以发现该区域人造地表整体趋势为减少,确定为异常区域,进行分析判断.

5.1.3 交互式对比分析

对于基于已有地表覆盖数据的知识规则,则可以在服务平台上采用人机交互对比分析来实现.通过目视对比,依据知识规则,发现并定位异常区域.如图6,利用服务平台,进行了GlobeLand30-2010与GlobCover2009的目视对比,可以发现在巴西罗赖马州东北存在一片区域,前者地表覆盖以草地为主,而后者则主要是灌木地,确定为分类异常区域,再进一步进行错分/漏分判断.

5.2 基于最小图斑的错分/漏分判断

5.2.1 最小图斑设计

最小制图单元即最小图斑,可以对检核与修改进行控制.地表覆盖遥感数据在内容的表达和使用上,直接受到最小图斑的影响(张增祥等, 2012).在遥感制图中,选择合适的最小图斑,能够在保证地图数据准确描述真实地表覆盖的前提下,提高工作效率,减少人工编辑的工作量,并能在一定程度上消除混合像元引起的噪声误差.这对于大区域规模化的遥感制图尤为重要.如2001年美国国家土地覆盖数据库采用了1英亩(TM的5个像元,1英亩 $\approx 4046.8\text{m}^2$)作为最小制图单元,对于一些景观区域的农业类型则提高了最小制图单元(Homer等, 2007).但是,在全球地表覆盖制图中尚缺乏针对不同地表覆盖类型的最小图斑的设计.

基于全球尺度上的地表覆盖类型的复杂性、遥感影像特征的多样性与差异性,综合各地表覆盖类型在30m分辨率遥感影像上的特征、国家基本比例尺1:100000地形图图式上图要求等知识,并参考以往遥感制图的经验,研究制定了全球30m地表覆盖遥感制图的最小图斑(表3).

以最小图斑为下限,确定每个地表覆盖类型的容错指标,进行质量控制,以有效地减少检核中的识别、勾绘与编辑修改的工作量.在检核时,对于小于最小图斑的地表覆盖类型漏提是允许的,可不予考虑.如,在30m分辨率Landsat TM多光谱影像上,水体像元在4、3、2波段组合上影像特征呈现典型的深

图 5 中国河北唐县及周边地区人造地表变化情况

图 6 GlobeLand30-2010(a)和 GlobeCover2009(b)交互式对比分析发现分类异常区域

蓝色偏黑, 如图7中红色界线范围内的水体大于3×3个像元, 故需要提取并检核; 而黄框内光谱特征深蓝色的为水体, 易识别, 但其大小小于3×3个像元, 所以在提取与检核时可不予考虑。

5.2.2 错分/漏分判断

基于多元知识凝练出的知识规则, 发现成果的分类异常区域后, 利用服务平台, 分屏显示, 左右联动, 通过高分影像对该问题区域进行目视判读, 与GlobeLand30进行对比, 确定是否为错分/漏分。由于全球高分数据不能实现全面覆盖, 特定相就更加难以获得, 可以参考30m影像和其他数据, 进行解译。对于确定的错分/漏分区域, 则用点、线、面的方式标绘, 并发布在服务平台上。如果错分/漏分区域小于最小图斑或在容错指标范围内, 则可以不予标注。

5.3 数据优化

利用遥感处理软件(ERDAS等), 基于分割对象, 对标注的错分/漏分问题进行修改。通过服务接入功能, 可以将包括两期30m影像、高分影像、参考数据等信息与分幅成果, 分割数据进行整合。修改时, 对分割数据图层进行操作, 选中待修改的一块或者多块小区域, 然后用选中的区域生成一个与该区域范

围一致的AOI文件图层, 对待修改区域进行直接赋值, 保存修改结果即可完成修改(图8)。对于问题较严重的区域, 如影像判读出现较大错误, 则可能需要采取重分类的方法进行修改。

5.4 检核优化效果

全球30m地表覆盖遥感制图中, 基于顾及多元知识的GlobeLand30数据检核优化模型, 在服务平台的支持下, 按照水体、湿地、冰雪、人造地表、耕地、林灌草和裸地的顺序, 分层分类地进行了针对GlobeLand30的检核与修改, 提高了检核的效率, 达到了优化数据、生产高质量产品的目标。下文以2010年期GlobeLand30自动分类成果检核优化为例, 对基于该模型的检核优化工作进行了评价。

表 3 全球 30m 地表覆盖遥感制图各地类最小图斑

地表覆盖类型	水体		湿地	冰雪	人造地表
	湖泊、坑塘	河流(宽度)			
最小图斑 (像元个数)	3×3	3	9×9	3×3	4×4
地表覆盖类型	耕地	森林	灌木地	草地	裸地
最小图斑 (像元个数)	6×6	8×8	10×10	10×10	6×6

图 7 30m 影像上面积小于最小图斑的水体

图 8 基于对象的修改完善

(a) 分类结果+分割对象+影像(黄线范围内草地错分为森林); (b) 选取要修改的图斑对应的分割对象; (c) 错分区域 AOI 生成(白线范围内); (d) 分类结果修改(黄色区域修改为草地)

5.4.1 优化前后对比示例

通过图9中优化前后的对比,可以发现,检核优化后,数据的质量得到大幅改善.有图幅整体性趋势错误的改善,有地类判别错误的校正,有细节上的修改.

5.4.2 优化前后精度评价对比

对优化后的225幅数据(约占全球26%)GlobeLand30-2010的林灌草和裸地数据,随机抽取了10%(共23幅),每个图幅内采用优化抽样模型计算要素样本量,顾及地表覆盖类型间的空间相关性进行要素样本布设(Tong等, 2011).最终抽取并布设36556个样本点进行了优化前后的精度评价对比(表4),抽样图幅精度提高最多为17.47%.可见,通过顾及多元知识的GlobeLand30的检核与优化,发现并修改错分/漏分问题,有效地提高了数据质量.

6 讨论与结论

为解决“像元-对象”自动分类结果存在的错分/

表 4 抽样图幅修改前后精度对比

序号	图幅号	样本点数(个)	修改前精度	修改后精度	精度提高
1	N51_40	1701	74.96%	75.61%	0.65%
2	N35_40	1814	85.39%	85.98%	0.59%
3	N48_10	2968	61.24%	77.86%	16.62%
4	N30_40	1739	85.18%	91.66%	6.48%
5	N31_45	2182	86.39%	88.46%	2.07%
6	N34_45	2169	88.07%	88.80%	0.73%
7	N31_05	2765	77.70%	83.31%	5.61%
8	N31_35	1241	75.12%	75.42%	0.30%
9	N33_10	950	80.79%	83.03%	2.24%
10	N37_05	888	76.13%	78.94%	2.81%
11	N37_10	1207	84.02%	84.59%	0.57%
12	S33_10	1417	77.79%	79.83%	2.04%
13	S34_30	1004	81.66%	82.34%	0.68%
14	S35_10	2161	83.02%	84.98%	1.96%
15	S35_25	1098	79.15%	80.13%	0.98%
16	S36_00	1552	71.74%	76.93%	5.19%
17	S36_15	1359	52.78%	70.25%	17.47%
18	S38_15	1277	59.30%	74.43%	15.13%
19	N13_60	1261	73.33%	89.94%	16.61%
20	S20_30	653	81.12%	83.26%	2.14%
21	S23_20	1686	77.04%	78.10%	1.06%
22	S53_30	1241	83.44%	85.09%	1.65%
23	S60_35	2223	82.87%	84.48%	1.61%

图9 修改前后的对比

(a)和(b)、(c)和(d)、(e)和(f)、(g)和(h)、(i)和(j)、(k)和(l)分别为图幅 N28_25、N39_10、N51_40、S17_00、S36_05、S54_15 内修改前和修改后结果

漏分问题, GlobeLand30研制过程中充分征集了遥感、地理和生态等方面专家意见, 收集已有地表覆盖数据、DEM和地表覆盖统计数据等参考资料; 按照宏观分布、区域过渡和地表覆盖变化三个方面, 归纳了基于自然地理的地表覆盖分布知识、基于人文的地表覆盖分布知识和地表覆盖变化知识, 提炼出描述全球地表覆盖数据的空间一致性、时间连续性和关系协调性等的知识规则, 作为知识化检核的依据, 进行网络化交互式协同检核, 借助服务平台发现并定位分类异常区域, 对判定的错分/漏分问题进行优化。

利用该方法对GlobeLand30的各类数据进行了检核优化, 显著提高了数据质量。实验表明, 优化后的GlobeLand30数据的质量明显改善, 林灌草和裸地数据抽样图幅优化前后精度提高最多达到17.45%。经第三方验证, GlobeLand30-2010数据全球总体精度达到83.50%。另外, 根据国外学者开展了GlobeLand30的精度验证, 在意大利的总体精度超过80%(Brovelli等, 2015), 其水体数据在希腊Thessaly的精度则达到了91.9%(Manakos等, 2014), 这表明该方法是可行和可靠的(Ban等, 2015)。多元知识和网络化交互式协同

检核在大区域规模化遥感制图中的应用对各种土地利用/土地覆盖遥感制图具有重要的借鉴作用。其他全球或区域土地利用/土地覆盖遥感制图也可以依托该模型, 提炼相应尺度的知识规则, 把相关数据集成到服务平台上, 进行检核优化。

本研究是首次全球30m分辨率地表覆盖遥感制图中为解决如何保障数据产品质量这一国际科技难题进行的有益尝试。但是, 还仅是初步尝试, 多元知识在数据优化中的应用仍然面临完备的知识规则集的建立、知识化检核的自动实现和错分/漏分的自动完善等诸多许多问题。因而, 今后需要继续完善地表覆盖检核的知识规则, 建设规范化的知识规则集, 强化知识规则的形式化表达; 进一步发展知识化检核的网络化和自动化算法, 降低人工干预程度; 研究自动化修改漏分/错分问题方法, 提高数据优化效率。

参考文献

陈军. 1984. DTM在改善遥感影像分类精度中的应用. 武汉大学学报(信息科学版), 9: 69–81

陈军, 陈晋, 宫鹏, 廖安平, 何超英. 2011. 全球地表覆盖高分辨率遥感制图. 地理信息世界, 4: 12–14

陈军, 陈晋, 廖安平, 曹鑫, 陈利军, 陈学泓, 彭舒, 韩刚, 张宏伟, 何超英, 武昊, 陆苗. 2014. 全球30m地表覆盖遥感制图的总体技术. 测绘学报, 43: 551–557

宫鹏. 2009. 基于全球通量观测站的全球土地覆盖图精度检验. 自然科学进展, 19: 754–759

胡雯运, Chen Y L, 徐玥, 赵圆圆, 俞乐, 王杰, 宫鹏. 2014. 基于快速聚类方法的30m分辨率中国土地覆盖遥感制图. 中国科学: 地球科学, 44: 1621–1633

廖安平, 陈利军, 陈军, 何超英, 曹鑫, 陈晋, 彭舒, 孙芳蒂, 宫鹏. 2014. 全球陆表水体高分辨率遥感制图. 中国科学: 地球科学, 44: 1634–1645

刘纪远, 匡文慧, 张增祥, 徐新良, 秦元伟, 宁佳, 周万村, 张树文, 李仁东, 颜长珍, 吴世新, 史学正, 江南, 于东升, 潘贤章, 迟文峰. 2014. 20世纪80年代末以来中国土地利用变化的基本特征与空间格局. 地理学报, 69: 3–14

牛振国, 宫鹏, 程晓, 魏建宏, 王琳, 黄华兵, 沈少青, 吴昀昭, 王晓风, 王显威, 应清, 梁璐, 张丽娜, 王雷, 姚谦, 杨镇钟, 郭子祺, 戴永久. 2009. 中国湿地初步遥感制图及相关地理特征分析. 中国科学D辑: 地球科学, 39: 188–203

冉有华, 李新. 2015. 全球第一个综合高分辨率土地覆盖图——中国30m分辨率全球土地覆盖图评述. 中国科学: 地球科学, 45: 1243–1244

施建成. 2014. 编者按. 中国科学: 地球科学, 44: 1609–1609

武吉华, 张坤, 江源, 康慕谊, 邱扬. 2004. 植物地理学. 第4版.

北京: 高等教育出版社. 318–327

徐冠华, 葛全胜, 宫鹏, 方修琦, 程邦波, 何斌, 罗勇, 徐冰. 2013. 全球变化和人类可持续发展: 挑战与对策. 科学通报, 58: 2100–2106

张晶, 吴绍洪, 刘燕华, 杨勤业, 张月鸿. 2007. 土地利用和地形因子影响下的西藏农业产值空间化模拟. 农业工程学报, 23: 59–65

张增祥, 赵晓丽, 汪潇. 2012. 中国土地利用遥感监测. 北京: 星球地图出版社. 63–70

Ban Y F, Gong P, Giri C. 2015. Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. ISPRS-J Photogramm Remote Sens, 103: 1–6

Bicheron P, Defourny P, Brockmann C, Schouten L, Vancutsem C, Huc M, Bontemps S, Leroy M M, Frédéric A, Herold M, Ranera F, Arino O. 2008. GLOBCOVER: Product description and validation report. ESA GlobCover Project

Bontemps S, Defourny P, Van Bogaert E, Arino O, Kalogirou V, Perez J R. 2011. Globcover 2009 products description and validation report. ESA

Brovelli M A, Molinari M E, Hussein E, Chen J, Li R. 2015. The first comprehensive accuracy assessment of GlobeLand30 at a national level: Methodology and results. Remote Sens, 7: 4191–4212

Chen J, Chen J, Liao A P, Cao X, Chen L J, Chen X H, He C Y, Han G, Peng S, Lu M, Zhang W W, Tong X H, Mills J. 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS-J Photogramm Remote Sens, 103: 7–27

Chen, J, Dowman I, Li S N, Li Z L, Madden M, Mills J, Paparoditis N, Rottensteiner F, Sester M, Toth C, Trinder J, Heipk C. 2016. Information from imagery: ISPRS scientific vision and research agenda. ISPRS-J Photogram Remote Sens, 115: 3–21

Cogley J, Graham. 2010. A more complete version of the world glacier inventory. Ann Glaciol, 50: 32–33

Comber A, Fisher P, Wadsworth R. 2004. Integrating land-cover data with different ontologies: Identifying change from inconsistency. Int J Geogr Inf Sci, 18: 691–708

Fang J Y, Song Y C, Liu H Y, Piao S L. 2002. Vegetation-climate relationship and its application in the division of vegetation zone in china. Acta Bot Sin, 44: 1105–1122

Friedl M A, McIver D K, Hodges J C F, Zhang X Y, Muchoney D, Strahler A H, Woodcock C E, Gopal S, Schneider A, Cooper A, Baccini A, Gao F, Schaaf C. 2002. Global land cover mapping from MODIS: Algorithms and early results. Remote Sens Environ, 83: 287–302

Gong P, Wang J, Yu L, Zhao Y C, Zhao Y Y, Liang L, Niu Z G, Huang X M, Fu H H, Liu S, Li C C, Li, X Y, Fu W, Liu C X, Xu Y, Wang X Y, Cheng Q, Hu L Y, Yao W B, Zhang H, Zhu P, Zhao, Z Y, Zhang H Y, Zheng Y M, Ji L Y, Zhang Y W, Chen H, Yan A, Guo J H, Yu L, Wang L, Liu X J, Sh T T, Zhu M H, Chen Y L, Yang G W, Tang P, Xu B, Giri C, Clinton N, Zhu Z L, Chen J, Chen J. 2013. Finer resolution observation and monitoring of

- GLC: First mapping results with Landsat TM and ETM+ data. *Int J Remote Sens*, 34: 2607–264
- Giri C, Peng B, Long J, Loveland T R. 2013. Next generation of global land cover characterization, mapping, and monitoring. *Int J Appl Earth Obs*, 25: 30–37
- Griguolo S. 1999. Land cover recognition: An example of knowledge-based classification. In: *Process of International Symposium on Remote Sensing & Integrated Technologies*. Istanbul: Istanbul Technical University. 183–94
- Han G, Chen J, He C Y, Li S, Wu H, Liao A, Peng S. 2015. A web-based system for supporting GLC data production. *ISPRS-J Photogramm Remote Sens*, 103: 66–80
- Harris P M, Ventura S J. 1995. The integration of geographic data with remotely sensed imagery to improve classification in an urban area. *Photogramm Eng Remote Sens*, 61: 993–998
- Hansen M C, DeFries R S, Townshend J R G, Sohlberg R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int J Remote Sens*, 21: 1331–1364
- Hansen M C, Loveland T R. 2012. A review of large area monitoring of land cover change using landsat data. *Remote Sens Environ*, 122: 66–74
- Hansen M C, Potapov P V, Moore R, Hancher M, Turubanova S A, Tyukavina A, Thau D, Stehman S V, Goetz S J, Loveland T R, Kommareddy A, Egorov A, Chini L, Justice C O, Townshend J R G. 2013. High-resolution global maps of 21st-century forest cover change. *Science*, 342: 850–853
- Herold M, Mayaux P, Woodcock C E, Baccini A, Schmullius C. 2008. Some challenges in GLC mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens Environ*, 112: 2538–2556
- Homer C, Dewitz J, Fry J, Coan M, Hossain N, Larson C, Herold N, McKerrow A, VanDriel J N, Wickham J. 2007. Completion of the 2001 national land cover database for the conterminous United States. *Photogramm Eng Remote Sens*, 73: 337–341
- Judex M, Thamm H P, Menz G. 2006. Improving land-cover classification with a knowledge based approach and ancillary data. In: *Process of workshop of Earsel SIG on Land Use and Land Cover 2006*. Bonn, Germany, Sep 28–30. 184–191
- Kantakumar L N, Kumar S, Schneider K. 2016. Spatiotemporal urban expansion in Pune metropolis, India using remote sensing. *Habitat Int*, 5: 11–22
- Loveland T R, Reed B C, Brown J F, Ohlen D O, Zhu Z, Yang L, Merchant J W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int J Remote Sens*, 21: 1303–1330
- Lu D, Weng Q. 2007. A survey of image classification methods and techniques for improving classification performance. *Int J Remote Sens*, 28: 823–870
- Manandhar R, Odeh I O A, Ancev T. 2009. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sens*, 1: 330–344
- Manakos I, Chatzopoulos-Vouzoglanis K, Petrou Z I, Filchev L, Apostolakis A. 2014. Globalland30 mapping capacity of land surface water in Thessaly, Greece. *Land*, 4: 1–18
- Mayaux P, Bartholome E, Fritz S, Belward A. 2004. A new land-cover map of Africa for the year 2000. *J Biogeogr*, 31: 861–877
- Mayaux P, Eva H, Gallego J, Strahler A H, Herold M, Agrawal S, Naumov S, De Miranda E E, Di Bella C M, Ordoyne C, Kopin Y, Roy P S. 2006. Validation of the global land cover 2000 map. *IEEE Trans Geosci Remote Sens*, 44: 1728–1739
- Muchoney D M, Borak J, Strahler A, LoCastro J. 1999. The IGBP DISCover confidence sites and the system for terrestrial ecosystem parameterization: Tools for validating global land cover data. *Photogramm Eng Remote Sens*, 65: 1061–1067
- Ramankutty N, Evan A T, Monfreda C, Foley J A. 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Glob Biogeochem Cycle*, 22: 1–19
- Tong X H, Wang Z H, Xie H, Liang D, Jiang Z, Li J. 2011. Designing a two-rank acceptance sampling plan for quality inspection of geospatial data products. *Comput Geosci*, 37: 1570–1583
- Verburg P H, Neumann W, Linda N L. 2011. Challenges in using land use and land cover data for global change studies. *Glob Change Biol*, 17: 974–989
- Whigham D F. 2009. Global distribution, diversity and human alterations of wetland resources. In: Barker T, Maltby E, eds. *The Wetlands Handbook*. New Jersey: Blackwell Publishing Ltd. 43–64
- Wu H, Liao A P, He C Y, Peng S. 2011. A framework for integrating heterogeneous spatial information and applications adaptively based on multi-agent and web service. In: *Multimedia Information Networking and Security (MINES). Third International Conference, IEEE*. 253–257
- Yao T D, Thompson L, Yang W, Yu W S, Gao Y, Guo X J, Yang X X, Duan K Q, Zhao H B, Xu B Q, Pu J C, Lu A X, Xiang Y, Kattel D B, Joswiak, D. 2012. Different glacier status with atmospheric circulation in Tibetan Plateau and surroundings. *Nat Clim Change*, 2: 663–667
- Zhang W W, Liao A P, Peng S, Zheng X Y, Li M. 2015. Check and modification of GlobeLand30 with MODIS NDVI. In: *Process SPIE 9675. AOPC 2015: Image Processing and Analysis, 96750O*, doi: 10.1117/12.2197830
- Zhang Z X, Wang X, Zhao X L, Liu B, Yi L, Zuo L J, Wen Q K, Liu F, Xu J Y, Hu S G. 2014. A 2010 update of National Land Use/Cover Database of China at 1:100000 scale using medium spatial resolution satellite images. *Remote Sens Environ*, 149: 142–154