

基于决策树的面向对象变化信息自动提取研究

张雨霁^{①②}, 李海涛^②, 顾海燕^②

(^① 辽宁工程技术大学, 阜新 123000; ^② 中国测绘科学研究院, 北京 100039)

摘要: 为了从不同时相的遥感影像数据中自动提取变化信息且保证其效率, 本文结合面向对象分析技术, 提出了一种基于决策树变化信息自动提取的新方法。该方法利用影像的特征指数及形状特征、光谱特征、纹理特征等作为特征集, 将其作为知识库应用到决策树控制模型中, 进而利用该模型实现自动分类。对所得到的分类后影像对象, 组织分析其综合属性并作为决策规则再次分类, 通过“双重分类”的方式实现面向对象的遥感影像变化信息自动提取。该方法为遥感影像变化信息自动提取提供了新的思路。

关键词: 变化检测; 面向对象; 决策树; 特征指数

doi: 10.3969/j.issn.1000-3177.2011.02.016

中图分类号: TP79 **文献标识码:** A **文章编号:** 1000-3177(2011)114-0091-04

A Research on Object-oriented Analysis of Automatic Change Information Extraction Based on Decision Tree

ZHANG Yurji^{①②}, LI Haitao^②, GU Haiyan^②

(^① Liaoning Technical University, Fuxin 123000; ^② Chinese Academy of Surveying and Mapping, Beijing 100830)

Abstract: In order to extract change information automatically from the different time remote sensing images and to ensure the efficiency, a new automatic extraction method which is combined with the object-oriented image analysis technology based on the decision tree is presented in this paper. This method uses the feature index and shape, spectral, texture of the image as a feature set to establish the decision tree model for automatic classification. Organize and analyze the synthesized attributes of image objects classified above, then use it as the decision rule to make classification the second time. We can bring about the automatic extraction of image change information by ‘double classification’ which is based on the object-oriented image analysis. This method offers a new thought for automatic extraction of change information from remote sensing images.

Key words: change detection; object-based; decision tree; feature index

1 引言

变化检测技术问世以来, 以应用为导向, 技术方法与手段日臻完善, 应用深度和广度也在不断拓展。空间遥感获取技术的迅猛发展, 使遥感数据的获取趋向三多(多传感器、多平台、多角度)和三高(高空间分辨率、高光谱分辨率、高时相分辨率)^[1], 遥感数据的不断更新及实用化极大促进了变化检测理论和

方法的变革。但与遥感数据获取技术取得的巨大进步相比, 其相配套的数据处理、分析和应用方法相对落后, 大部分还只是停留在基于像元级别的影像分析处理, 突出问题展现在卫星获取的遥感数据未能得到充分利用, 对遥感信息认识不足和对遥感专题信息提取水平的滞后, 自动化、智能化程度不高, 效率及精度低下, 这就造成了遥感信息资源的巨大浪费及应用价值的降低, 使得变化检测的实效性难以

收稿日期: 2010-01-29 修订日期: 2010-05-05

基金项目: 国家科技支撑项目(2008BA C34B07_04)。

作者简介: 张雨霁(1985~), 女, 在读硕士, 主要研究方向为遥感图像处理。

E-mail: YUJI-911@163.com

满足。针对这些不足, 本文结合 GIS 空间数据挖掘和遥感技术, 以如何得到变化属性类型, 分析并提取出地物变化信息为目标, 突破传统意义上像元级的检测方法, 提出了对同一地区不同时相的遥感影像利用特征指数及分割后的对象特征建立决策树模型, 实现面向对象的变化检测新方法, 实验证明了该方法的有效性与鲁棒性。

2 面向对象的变化信息自动提取

2.1 多尺度分割

对影像进行合理的分割得到影像对象, 是面向对象变化信息提取的前提。本文采用多尺度分割技术, 通过不同尺度的选择可将影像方便灵活地分割成由同质不规则多边形组成的不同分辨率影像对象层。形成的由多边形矢量边界包围的影像对象区域, 达到了矢量与栅格的融合处理, 为进一步的影像分析处理提供了大量可用信息。分割多采用区域增长算法^[2], 每一层都由它直接的子对象构成, 在下一个高层上, 子对象合并为较大对象, 不断分离合并的过程中自动建立起了一个与地表实体相似的层次等级结构, 层次中的每个影像对象都明确其自身与上级、下级及同级对象之间的拓扑关系, 实现了原始像元信息在不同空间尺度间的传递。

2.2 特征提取指数

(1) 归一化植被指数 (NDVI)

Rouse J W^[3] 等人提出的 NDVI 作为目前最常用的植被指数之一主要用于植被信息提取、植被监测和植被物候等研究。其计算公式为

$$NDVI = \frac{nir - red}{nir + red} \quad (1)$$

nir 代表近红外波段, *red* 代表红波段。

(2) 归一化水体差异指数 (NDWI)

Mcfeters^[4] 提出的 NDWI, 是一种根据波段间比值进行分类的方法。它主要用于提取水体信息, 同时削弱土壤、陆生植物在图像中的亮度值, 勾画水域边界等。其计算公式为

$$NDWI = \frac{green - nir}{green + nir} \quad (2)$$

green 代表绿波段, *nir* 代表近红波段。

(3) 归一化建筑物差异指数 (NDBI)

ZHA Y^[5] 提出的 NDBI 主要用于提取由各种不透水建筑材料所覆盖的表面, 如由瓦片、沥青、水泥混凝土等材料构成的建筑物、路面和停车场等。

其计算公式为

$$NDBI = \frac{mir - nir}{mir + nir} \quad (3)$$

mir 代表中红外波段, *nir* 代表近红外波段。本文选取 TM 影像中的第 5 波段(短红外) 替代 *mir*。

2.3 面向对象的决策树分类

面向对象决策树分类方法是一种建立在先验知识基础上的分层次处理结构。该方法利用总结的特征提取指数及分割后所得对象的形状、语义等特征作为建立决策树所描述的多项判断准则, 对影像中各对象进行逐层识别和归类, 逐步将待提取目标从地物中分离出来, 避免此目标对其他目标提取时造成的干扰和影响, 通过若干次中间判别最终将所有数据图层复合以实现图像的自动分类。即通过一组独立变量, 将一个复杂数据集逐步分解为更纯、更同质的子集的过程。其基本思想是通过一些判断条件对原始数据集逐步进行二分和细化, 其中, 每一个分叉点代表一个决策判断条件, 每个分叉点下有两个叶结点, 分别代表满足和不满足条件的类别^[6-8]。

2.4 变化信息自动提取过程

本试验利用对影像多尺度分割后影像对象的形状指数、特征提取指数等特征, 建立面向对象决策树模型分别对不同时相的 TM 影像进行面向对象分类, 对分类后的结果, 利用影像层次间相应类别及对应位置对象间的语义关系进行再分类, 双重分类后最终实现变化信息的自动提取。具体过程如下:

①对输入影像进行多尺度分割获取影像对象(试验中尺度选择 10, 形状异质性权重系数选择 0.1, 光谱异质性权重系数选择 0.9, 紧致度指数权重系数选择 0.5, 光滑度指数权重系数选择 0.5)。

②利用对象的特征提取指数及形状指数建立面向对象决策树分类层次模型(主要分析提取水体、裸地、稀疏植被及茂密植被 4 个类别变化信息如图 1)。

③利用面向对象决策树分类模型分别对前后时相的影像进行面向对象的分类处理, 分别得到分类后前后时相单波段影像层。试验分类后结果如图 5 (b)、(d) 所示。

④将前后时相单波段分类后影像分别按类别赋予层次属性值(前一时相赋予 2, 后一时相赋予 1), 使各对象都明确各自属于哪层哪类。

⑤用前一时相的影像对象层分割后的各均质区域去限定相应位置后一时相影像对象层分割后的各

均质区域,使其形状与前一时期保持一致,但区域内的属性值各自保留,如图2。

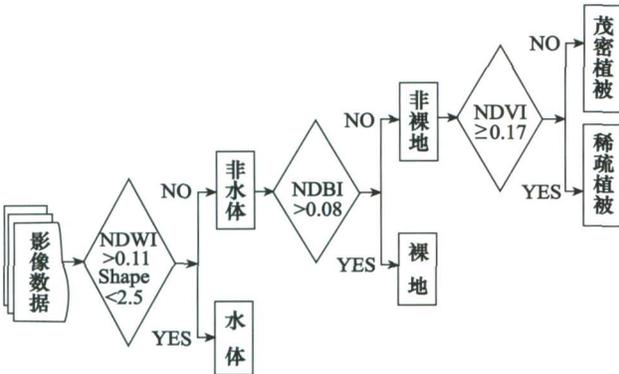


图1 面向对象决策树分类模型

分类的方式,实现变化信息的自动提取。图3描述了面向对象变化信息提取的过程。图4为变化信息提取结果,图5为面向对象决策树分类最终结果。

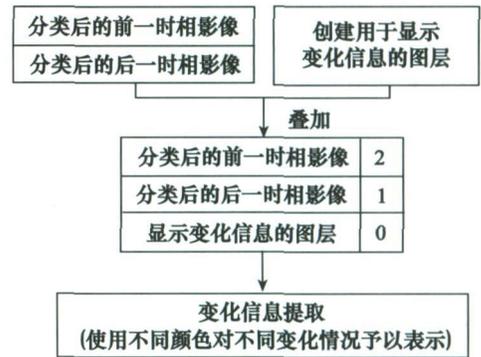


图3 面向对象变化信息提取流程

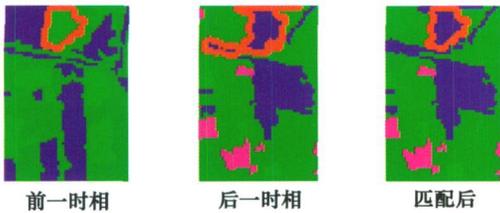


图2 影像对象匹配化

统计分析相同形状区域内的变化情况(与前一时期相的均值区域相比后一时相异质情况),所选类别影像对象存在赋值为1,不存在赋值为0,如表1:

表1 地物类别变化信息分析表

变化状况	所选类别对象 (层次属性) 状态	对象存在与否
增加	2000年(1) 存在	1
	1999年(2) 不存在	0
减少	2000年(1) 不存在	0
	1999年(2) 存在	1
不变	2000年(1) 存在	1
	1999年(2) 存在	1
不存在	2000年(1) 不存在	0
	1999年(2) 不存在	0

于是每个影像对象就具备了层次、类别、位置及存在与否这4种属性,通过四重属性的索引,就可以进行不同层次不同类别间对应位置影像对象的对比分析。

⑥创建用于显示变化信息的显示层,并赋予层次属性值0。将该层与面向对象决策树分类后的层次属性值分别为1、2的后、前时相影像对象层进行叠加处理。

⑦对于变化信息提取各个类别(水体、裸地、稀疏植被及茂密植被)采用表1的方式,即可表示出地物的变化情况。又可将不同的变化情况利用不同颜色属性值在0层中予以显示表达,通过分类后再次

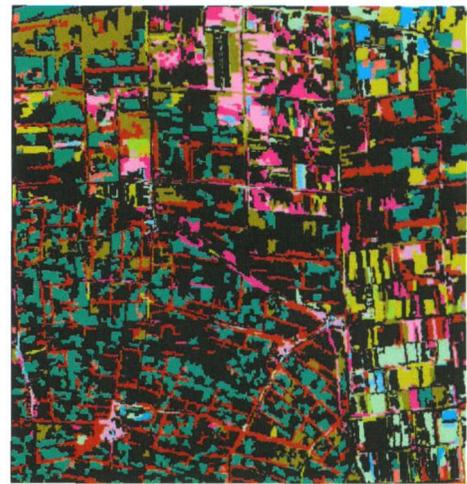


图4 变化信息提取结果

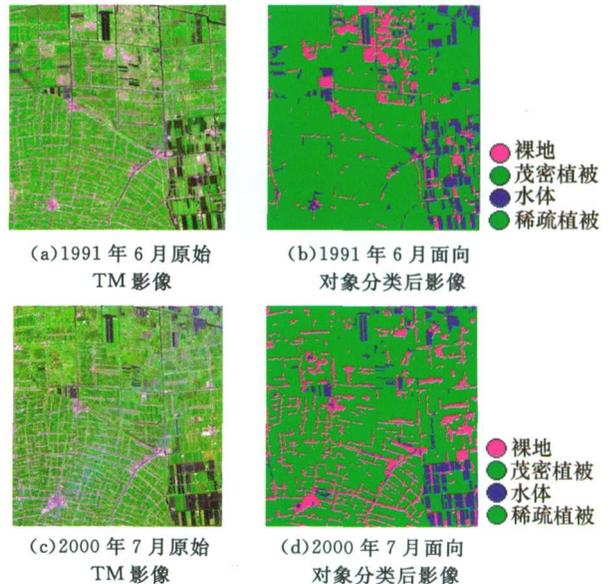


图5 面向对象决策树分类最终结果

3 精度评价

本文选用经过前期几何校正、直方图匹配等预处理后的 1991 年 6 月及 2000 年 7 月前后时相 TM 影像作为实验数据源,将 TM1、TM2、TM3、TM4、TM5、TM7 六个波段(由于 TM6 波段信息量较少)进行波段组合作为分析数据,采用 543 波段进行假彩色合成,合成影像上茂密植被表现为深绿色,稀疏植被表现为绿色,裸地表现为淡紫色,水体表现为深蓝色,方便目视判读。研究区域内,相对与 1991 年影像,2000 年影像中各类型地物均发生不同程度变化,各部分地物的变化情况明显。根据图 4 提取结果,对地物变化情况进行定量分析,结果见表 2:

表 2 地物类别变化信息统计表

类别	像元数	比率(%)
裸地→水体	1049	0.56
裸地→茂密植被	7334	3.89
裸地→稀疏植被	4114	2.19
茂密植被→裸地	21109	11.21
茂密植被→水体	5130	2.72
茂密植被→稀疏植被	11941	6.34
稀疏植被→裸地	2113	1.12
稀疏植被→水体	1096	0.58
稀疏植被→茂密植被	25285	13.43
水体→裸地	2322	1.23
水体→稀疏植被	1072	0.57
水体→茂密植被	7841	4.16
未变化	97901	52.00
总计	188307	100

对于变化区域的检测状况,本试验将信息提取后的结果与原始影像在 ENVI 中进行 Link 显示,目视判读检验其提取效果,认为该方法的总体提取效果较为理想。为了评价其精度,利用 ENVI 中的样本选择工具,随机选取 200 个样本区域,其中共有 174 个区域发生变化,结合遥感影像进行对比分析,发现检测出的 174 个图斑中实际发生变化的为 163 个,有 11 个未发生变化。另外有 3 个图斑发生了变化但是没

有被检测出来。以前后时相遥感影像数据为参照,比对分析变化信息提取后的图像,结果见表 3。

表 3 变化信息验证误差混淆矩阵

检测结果 \ 参考数据	变化	未变化
	变化	163
未变化	3	23

根据检测结果计算,漏检率为 11.5%,错检率为 6.3%,总体检测精度为 93%。

但需要说明的是:本文采用的基于决策树的面向对象自动分类技术,属于分类后的变化信息提取范畴,不同时相分类结果的分类精度通过误差传递直接会影响到最后的提取结果。决策树分类过程中没有人工参与,分类后错分、漏分现象在所难免,本文旨在介绍该理论方法自动化的流程思路,若想进一步达到精度要求,在分类阶段可适当加入人工参与步骤,进行分类后的检校,在半自动化下使精度满足要求。

4 结束语

传统的像元级变化信息提取方法,“椒盐”现象严重,所利用的区分特征往往集中在光谱信息上,精度及效率都很难满足要求。本文在研究了多种变化检测方法的基础上,提出了基于决策树的面向对象变化信息自动提取的方法过程,并做了相关试验。在多尺度分割的基础上,利用多种提取特征建立决策树模型,进行面向对象的分类,考虑其分割后的影像对象(均质的栅格区域),利用对象所具有的层次网络结构关系,语义关系去锁定待确定变化情况的影像对象区域,继而根据规则判定前后时相影像的对应区域发生的变化情况,实现变化信息的自动提取。试验证明了该方法的可行性,影像对象级的变化信息提取可以提高变化类和非变化类的可分性,在一定程度上消除影像中“椒盐”噪声,充分利用影像对象中的各种特征信息,较高精度地完成遥感影像中变化信息的自动提取。

参考文献

[1] 李德仁,沈欣.论智能化对地观测系统[J].测绘科学,2005,30(4):9-11.

[2] U. C. Benz, P. Hofmann, G. Willhauck, Iris Lingenfelder, M. Heynen. Multi resolution, object oriented fuzzy analysis of remote sensing data for GIS ready information[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2004(58): 239- 258.

[3] Rouse J W, Haas R H, Schell J A, et al. Monitoring vegetation systems in the great plains with ERTS[C]//In: The Third ERTS Symposium. 1973:309- 317.

[4] McFeeters S K. The use of the Normalized Difference Water Index(NDWI) in the delineation of open water features[J]. International Journal of Remote Sensing, 1996, 17(7): 1425- 1432.

(下转第 97 页)

4 对 InSAR 应用的影响

新方案实施后,除了 SAR 干涉测量(InSAR)外,原有的 Envisat 应用不受影响,而前者的应用要受到一些限制,其根源是由于停止对轨道倾角的控制。从 2010 年 10 月开始,轨道倾角变化不再得到补偿,轨道倾角将逐渐变小(图 2)。基线随季节的变化也十分显著,在北半球两个相邻的周期内,基线在夏季小于冬季,故在实际应用中要考虑季节对 InSAR 测量的影响。

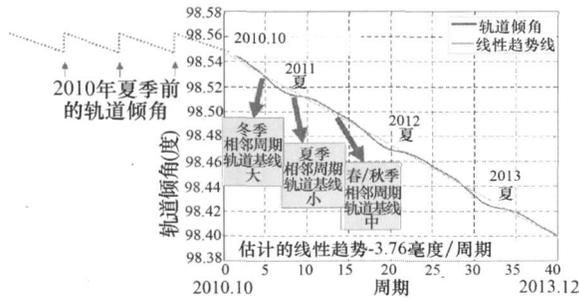


图 2 E3 阶段轨道倾角的变化示意图

轨道倾角变化对 InSAR 应用有着较大的影响,它直接导致了 InSAR 基线的增大。(垂直)基线是干涉测量的关键因素,它决定了空间去相关的程度,以及差分干涉(DInSAR)处理中应消除的地形影响。为了降低对 SAR 干涉测量带来的负面影响,使 InSAR 基线在预定的纬度变化最小,设定该纬度为北纬 38°(降轨)和南纬 38°(升轨)(图 3),而不再是在赤道上。这样的设置使得 DInSAR 应用仍能用于地质构造与火山覆盖最为丰富的北纬 38°纬度带内,这也是过去 InSAR 技术用于监测地质构造、火山、地表沉降等最多的地区,包括意大利南部、希腊、土耳其、伊朗北部、日本和美国加州北部地区。

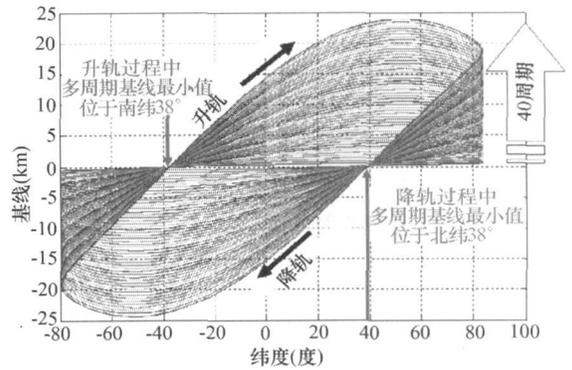


图 3 垂直基线随时间和纬度的变化示意图

由于近端波束(如 IS1、IS2)的去相干效应比远端波束(如 IS6、IS7)更明显,基线较小的区域范围很强烈地依赖于数据获取过程中的几何关系。由 InSAR 理论知,干涉(垂直)基线如果超出了临界值,信号不相干,不能进行干涉测量。因而,可以通过长度小到能够进行 DInSAR 处理的多个轨道周期的基线,来估算适于 DInSAR 的纬度带。该纬度带被定义为在整个 E3 阶段(40 个周期,即 40 个月)内长度小于临界基线一半的多个周期的基线所对应的纬度值。覆盖 40 个周期适于 DInSAR 的纬度带估计为: 38° ± 1.3° (对于 IS2); 38° ± 4.0° (对于 IS6)。在 DInSAR 纬度带以外,仍有可能找到合适的干涉像对,但基本上是在相邻的轨道周期内,而多个轨道周期的基线对于 InSAR 则显得太大。

图 3 反映了 E3 阶段第 1 周期轨道与后续 40 周期同轨道间的垂直基线随时间和纬度的变化情况,40 个周期时间即为 2010 年 11 月开始后的 40 个月,越接近 38°纬度,基线越小。

(编译自:《Impact of the Envisat Mission Extension on SAR data》,欧空局于 2010.10.12 发布

<http://earth.esa.int/category/index.cfm?categoryid=42>)

参考文献

[1] Macdonald, M., McKay, R. J., Vasile, M., Bosquillon de Frescheville, Francois. Extension of the sur synchronous Orbit [J]. Journal of Guidance, Control and Dynamics. 2010. ISSN 0731- 5090.
 [2] ASAR Product Handbook[A]. European Space Agency[Z]. 2007.

(上接第 94 页)

[5] Zha Y, Gao J, Ni S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery [J]. International Journal of Remote Sensing, 2003, 24(3) : 583- 594.
 [6] Volgamann J E, Sohl T, Howard S M. Regional characterization of land cover using multiple sources of data [J]. Photogrammetric Engineering and Remote Sensing, 1998, 64(1) : 45- 57.
 [7] Pielke R, Lee T J, Copeland J H, East man J L, etc. The use of USGS provided data to improve weather and climate simulations [J]. Ecological Applications, 1997, 7(1) : 3- 21.
 [8] Fennessy M J, Xue Y. Impact of USGS vegetation ma Pon GCM simulation over the United States [J]. Ecological Applications, 1997, 7(1) : 22- 33.