

Modified Water Information Extraction Method Based on Remote Sensing Images*

Hongxu Ma, Shenglian Guo, Yanlai Zhou

State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan
Email: mahongxu@whu.edu.cn

Received: Feb. 16th, 2013; revised: Feb. 27th, 2013; accepted: Mar. 9th, 2013

Copyright © 2013 Hongxu Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: By analyzing the error extracted areas which is caused by using NDWI and NDVI methods individually, a modified water information extraction method was proposed and tested. The method interconnects the NDWI, NDVI and mathematical morphology methods together. The HJ-1A sequence satellite images in the Poyang Lake during 2010 to 2011 were selected as case study. The water information were recognized and extracted by the modified method and compared with three conventional methods based on confusion matrix. The results show that modified method is more accurate than three conventional methods by improving 9.43% on average.

Keywords: Water Recognition; NDVI; NDWI; Satellite Images

基于遥感影像水体信息提取的改进方法*

马鸿旭, 郭生练, 周研来

武汉大学水资源与水电工程科学国家重点实验室, 武汉
Email: mahongxu@whu.edu.cn

收稿日期: 2013年2月16日; 修回日期: 2013年2月27日; 录用日期: 2013年3月9日

摘要: 本文分析了单独使用归一化差分植被指数法和归一化差分水指数法所引起的误提取和漏提取, 总结了导致错误的原因及规律。耦合了归一化差分植被指数法、归一化差分水指数法和数学形态学理论, 并提出了一种水域信息提取的改进方法。以鄱阳湖水域为研究对象, 选取2010年3月至2011年12月的HJ-1A时序卫星影像图进行水域信息提取, 并采用混淆矩阵对提取结果进行了比较和评价。结果表明: 改进的方法比三种常规方法(NDVI, NDWI, 监督分类)的水体提取精度平均提高了9.43%。

关键词: 水体提取; 归一化差分植被指数; 归一化差分水指数; 卫星影像

1. 引言

对于水域信息的研究是水文测验及分析计算的基础内容之一。通过时序遥感数据进行水域信息提取及变化检测, 具有及时、快速、准确的特点。马国明^[1]

*基金项目: 国家自然科学基金(51079100, 51190094)。

作者简介: 马鸿旭(1989-), 男, 吉林省吉林市人, 硕士研究生, 主要从事遥感技术在水文中的应用等研究。

等利用多时相的高分辨率遥感数据检测了若羌湖泊群的动态变化特征。Li 和 Narayanan^[2]选用多时相的 Landsat MSS 数据对 Nebraska 地区建立了基于空间模型边界形状提取的湖泊变迁分析方法。在水域信息提取方法领域, 由 McFeeters^[3]提出归一化水指数 NDWI 法。Xu^[4]对 NDWI 进行了优化, 提高了指数对水体信

息的敏感度。曹荣龙等^[5]对 NDWI 法进行了进一步的改进, 使水体提取精度进一步提高。Ma^[6]基于 SPOT 序列影像, 综合应用 NDVI 和 NDWI 提取艾比湖水体, 证明了指数联合提取水体的可行性。翟辉琴^[7]将数学形态学理论应用到遥感影像的水体识别中, 取得了很好的效果。现行的水域信息提取主要存在以下几点不足: 1) 水体提取所使用的方法单一, NDVI 或 NDWI 都各自存在导致误提取的地物类别, 这种状况在汛期尤为严重。2) 采用 NDVI 或 NDWI 提取的水体信息具有噪声, 即偶然符合指数的非水体像点也被归为水体。

本文在分析现有方法优点和不足的基础上, 耦合归一化差分植被指数法 NDVI(Normalized Difference Vegetation Index)、归一化差分水指数法 NDWI(Normalized Difference Water Index)和数学形态学理论, 提出了一种水域信息提取的改进方法。该法通过对 NDVI 及 NDWI 水体提取的联合处理, 剔除现有方法各自的误提取区域, 并通过数学形态学理论对提取结果进行进一步噪声剔除处理。通过实例分析和比较, 以论证所提方法的计算精度和适用性。

2. 水体提取方法原理

水域提取主要是对水域信息的识别和提取, 即通过一定的规则和方法将水域信息与非水域信息相互区别开来。而两类信息的区别的分隔程度及准确性即是评价水域信息提取的主要判别标准^[8]。

由于水体和陆地在接受了太阳辐射的相互作用之后, 对于太阳辐射的反射、吸收、散射、透射的特征差异非常大, 而使其在遥感图像上的反映截然不同, 对于不同波段的不同反映形成了对于水域信息唯一的波谱响应特征, 成为区分水域信息和其他非水域信息的重要基本原理^[9]。

2.1. 归一化差分植被指数法

归一化差分植被指数, NDVI(Normalized Difference Vegetation Index)。

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

式中: NIR 为近红外波段, RED 为红光波段。

因为红光波段中的水体的反射率要高于植被, 而

对于近红外波段, 植被的反射率要明显高于水体的反射率。因此, 采用 NDVI 指数来处理, 以增强水陆的反差。NDVI 是目前应用最广的植被指数, 常用于指示植被的数量特征并用于检测植被的季节变化和对土壤覆盖的研究, 也可用于对植被区域的检测与非植被区域的识别。在植被覆盖密集的地区, NDVI 值较高, 而在植被覆盖稀疏的地区, NDVI 指数受到土壤背景影响较大, 如沙漠、水体的 NDVI 指数值很低。利用这个特点, 可以通过选用阈值建立识别水体和植被、土壤的模型。

2.2. 归一化差分水指数法

归一化差分水指数 NDWI(Normalized Difference Water Index)。

$$NDWI = \frac{NIR - GREEN}{NIR + GREEN} \quad (2)$$

式中: NIR 为近红外波段, GREEN 为绿光波段。

水体在此两波段光谱特征有着明显的差异, 而且由于水体的反射从可见光到红外波段逐渐减弱, 在近红外和中红外波长范围内的吸收性最强, 而植被在近红外波段的反射率最强。因此, 用绿光波段和近红外波段的反差构成 NDWI 指数, 可以最大程度上地抑制植被的信息, 从而突出水体特征, 进而提取水体信息。

2.3. 监督分类法

监督分类(supervised classification)又称训练场法, 是以建立统计识别函数为理论基础, 依据典型样本训练方法进行分类的技术。即根据已知训练区提供的样本, 通过选择特征参数, 求出特征参数作为决策规则, 建立判别函数以对各个待分类影像进行的图像分类, 是模式识别的一种方法。

3. 水体提取改进方法

3.1. 常规方法存在的主要问题分析

NDVI 指数法和 NDWI 指数法都已被证明可以很好地对一些流域的水体信息进行较高精度的信息提取^[3]。但由于 NDVI 指数法和 NDWI 指数法都只分别采用了两个波段的遥感图像进行波段运算以进行水体信息提取, 导致无法准确地区分一些波段运算后指数相近的地物。

在应用 NDVI 方法提取水体信息时, 发现由于建筑物同水体信息在红色波段及近红外波段的响应幅度和程度相似, 导致明显的建筑物信息也被错误地识别为水体信息。如图 1 所示。左图为包含水体和城区建筑物信息的遥感图像。右图为 NDVI 水体信息提取结果图, 黑色为被识别的水体, 由图可见部分城区建筑及裸地信息也被错误识别为水体信息。

通过分析研究经过二值化处理的影像和对应的水体信息提取混淆矩阵发现, 影响水体提取信息准确度的主要误差分为线误差和点误差。点误差主要存在于使用 NDVI 法对水体信息图中, 误差的形式为错误提取误差, 即将本来不属于水体信息的某些像元错误分类为水体信息。如图 2(b)所示。线误差主要存在于使用 NDWI 法提取的水体信息图中, 误差的形式为漏提取误差, 即 NDWI 法忽略了很多细小的支流信息。如图 2(c)所示。

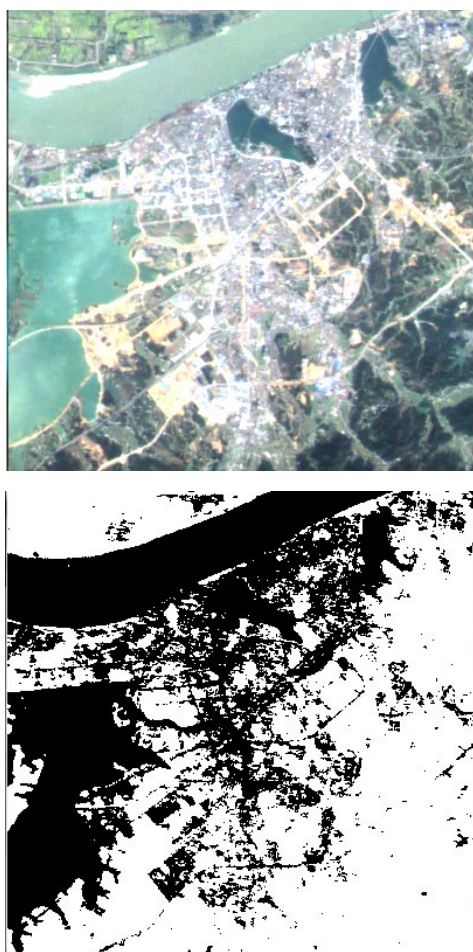


Figure 1. NDVI extracts the land and thebuilding by mistake
图 1. NDVI 对裸地及建筑物的误提取效果

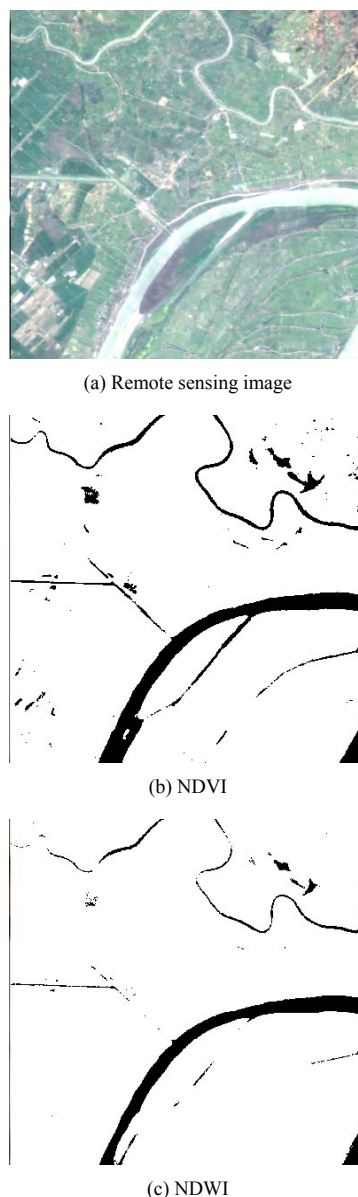


Figure 2. Water extraction using NDVI and NDWI method
图 2. NDVI 及 NDWI 方法提取的水体信息图

NDWI 指数法和 NDVI 指数法在提取水体信息的计算时都具有不同的优势及明显的劣势。在详尽分析了各自方法在水体信息提取结果后, 提出以下几点方法的改进思路。

1) NDVI 对建筑信息的过分敏感导致误提取了很多建筑物信息, 与此同时 NDWI 对细小支流的不敏感性导致了大量小面积水体的遗漏。因此, 通过 NDWI 指数与 NDVI 指数的联合运算, 用以拉大水体信息与建筑信息的反差的同时, 增大对小面积水体的敏感性。

2) 在水体提取影像图中, 总会出现某一些非水体的像元满足了水体信息的识别条件, 形成了误差。而此类误差往往具有小范围、无规律的信息特征, 可采用基于数学形态学的方式进行消除, 计算其形状指数, 通过进行腐蚀运算和膨胀运算进行剔除。

3.2. 水体提取改进方法原理

方法先对同一遥感图像分别进行 NDVI 指数及 NDWI 指数的运算, 并通过自动解算最佳阈值或人工判读的方式确定最佳阈值。然后再采用最佳阈值提取及超阈值提取两种方式进行水体信息的提取计算。将最佳阈值提取得到的水体信息命名为 A(NDWI), 将超阈值提取得到的水体信息分别命名为 O(NDWI)及 O(NDVI)。通过以下公式进行计算:

$$(O(NDVI) \cup A(NDWI)) \cap O(NDWI) \quad (3)$$

式中: \cup 表示像元并集运算, \cap 表示像元交集运算。通过上式可以将 NDVI 法错误提取的建筑物像元剔除, 并通过过度提取, 将 NDWI 并不擅长的细小支流的提取强化, 同时不增加其他错误的提取像元。

经过交并集运算后仍有一部分错误提取的水体信息混杂在水体计算值的值域中, 但由于形态上的不同, 具有小面积性及无规律性的特点, 选择基于数学形态学的方式, 计算其形状指数, 通过进行腐蚀运算后膨胀运算的算法将其进一步剔除, 以获得最佳的提取效果。改进算法流程图如图 3 所示。

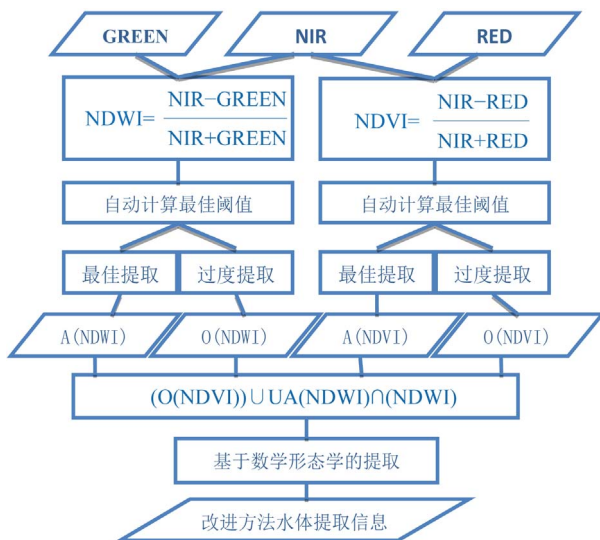


Figure 3. Flowchart of modified water recognition method
图 3. 水体信息提取改进方法流程图

4. 实例研究

实验研究区域为鄱阳湖, 位于江西省北部、长江南岸, 是我国第一大淡水湖, 地理坐标东经 115°49' 到 116°46'、北纬 28°24'到 29°46'。鄱阳湖承纳赣江、抚河、信江、饶河、修水五河及博阳河、清丰山溪、漳河、西河的来水, 经调蓄后由湖口北注入长江。鄱阳湖流域面积 162,225 km², 南北长 173 km, 东西平均宽度 16.9 km, 最宽处约 74 km, 长江水道最窄处的屏峰卡口, 宽约 2.8 km。当湖口水文站水位 21.0 m(吴淞基面, 下同)时, 湖水面积 3840 km², 容积 263 亿 m³, 但到了冬季, 湖水下降, 湖面缩小, 湖泊面积仅由 140 余 km²。鄱阳湖水位变化受五河及长江来水的双重影响, 高水位维持时间长, 水位年变幅大, 最高达 9.59~14.85 m, 最小为 3.54~9.59 m, 水位变幅自北向南递减。

此次研究主要使用的是 2010 年 3 月至 2011 年 12 月的 HJ-1-A 卫星数据及鄱阳湖地区的范围矢量图。

4.1. 水体信息的提取

对实验区域的各个时段遥感图像分别进行了 NDWI 指数法、NDVI 指数法、监督分类法及改进方法的水体信息的提取计算。并针对提取得到的图像进行分析比对。

使用 NDWI 指数法提取的水域信息图像如图 4 所示, 图像中的陆地部分在提取水体信息时没有受到植物信息的影响, 所以陆地与水体区域划分明显, 没有过多的小面积噪点, 但支流信息被抑制得很严重, 有些区域甚至无法对细小的支流进行响应。



Figure 4. Water extraction using NDWI method
图 4. NDWI 方法提取的水体图像图

使用 NDVI 指数法提取的水域信息图像如图 5 所示, 可以看到对于细小的支流部分提取效果很好, 但由于陆地上的裸地信息的影像导致存在很多小块的噪点, 严重影响水体提取效果及精度。

使用文中所述的改进方法提取的水域信息图像, 如图 6 所示, 细小的支流部分信息得到了充分的提取, 而错误的植被信息也被大部分抑制。这就直接结合了 NDWI 指数和 NDVI 指数的优点, 又通过相互协作, 抑制了彼此的劣势。

通过上述改进方法的提取后, 由于图像上还残存有部分的小面积的噪点。此类噪点的产生有两个主要原因: 第一是由于特殊值恰好满足提取条件, 第二是由于 NDWI 指数对于建筑物信息的错误提取及 NDVI 指数对于植物信息的错误提取。但无论由哪种原因引起的噪点, 都具有小面积的特点。即仅由小于 20 个像元组成, 这就可以通过数学形态学的概念进行提



Figure 5. Water extraction using NDVI method
图 5. NDVI 方法提取的水体影像图



Figure 6. Water extraction using modified method
图 6. 改进方法提取的水体信息图像

取, 即计算灰度值距离或波段间距离, 将距离相近的像元归为一个组, 当组元素提取完毕后, 即其周围不再有连续的能被收入的像元时, 计算组内像元数量, 当像元数量小于阈值时剔除计算。如图 7, 图 8 所示。

4.2. 水体提取方法的精度比较分析

本文采用混淆矩阵进行水体提取结果的验证。首先对遥感红、绿、蓝三个波段图像组合而成的假彩色图像进行显示, 通过人工目视判别在整幅图像上进行人工地物信息提取, 分为水体信息和非水体信息两类, 并确定人工提取的区域分布均匀且具有代表性实验中共提取了 10,000 个兴趣点进行检验计算, 实验中将水体信息值置为 1, 非水体信息值置为-1, 其他非检验点设置为 0。

检测时将通过基本方法获得的水体信息提取图像预处理, 将提取得到的水体信息值置为 255, 非水体信息置为-1。检验方法即将检验图像与待检验图像按照像素位置对应相乘获得新的值组成的精度图像。



Figure 7. Band classify as a pixels based on the distance between each wave
图 7. 根据波段间距离分类将相近程度小于阈值的像元化为一类

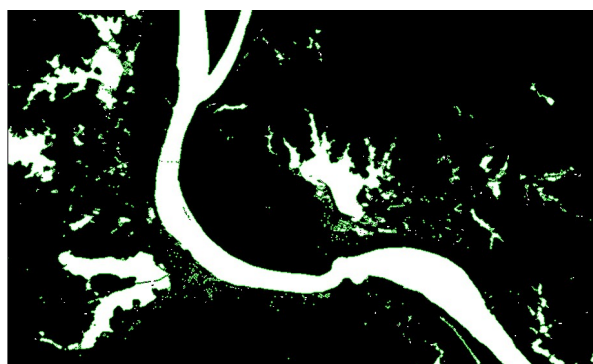


Figure 8. The small area pixels group using the modified extraction method
图 8. 经过改进方法计算后的小面积像元提取

统计精度图像中各个值的数量即可获得精度结果混淆矩阵。

由表 1 可知, 正确提取的水体信息计算后值为 255, 正确提取的非水体信息计算后值为 1, 即正确提取的信息计算后值为正值。相反, 误提取的水体信息(将非水体信息提取为水体信息)的值为-255, 漏提取的水体信息(将水体信息提取为非水体信息)的值为 -1, 即错提取的信息计算值为负值。未参与计算的值为 0。

通过统计正值与负值的像元数和占兴趣检测点总数的百分比即可获得水域提取的精度混淆矩阵, 结果见表 2。统计正确分类像元数量占整体统计计算像元的百分比, 得出水体提取的精度。各方法提取精度, 结果见表 3。

使用改进方法进行水体新体的提取相比于 NDVI 指数法精度提升了 15%, 相比于 NDWI 指数法精度提

升了 11%, 相比于监督分类法精度提升了 8%。

5. 结语

本文耦合归一化差分植被指数法、归一化差分水指数法和数学形态学理论, 提出了新的水域信息提取方法。并以鄱阳湖流域为研究对象, 通过使用 HJ-1A 卫星时序影像, 进行水体信息的提取计算。主要结论如下:

1) NDWI 法及 NDVI 法单一使用时都由于波段选取的局限性易产生严重的误提取和漏提取地物类别, NDWI 对于细小支流的敏感度不强, 为了充分提取而降低阈值会导致误提取及噪点增多。NDVI 受裸地影响较为严重, 对城市及居民用地易产生误提取。综合两种指数优点联合提取, 相互比照, 可剔除误提取和错提取区域。

2) 综合数学形态学的主要理论, 引入腐蚀和膨胀算法对提取后的图像进行进一步处理, 消除噪点。经过与 NDWI 法、NDVI 法及监督分类法的对比分析, 得出该改进方法更适用于混合了城市建筑及多细小支流的水体面积的信息提取。比照常规方法, 改进方法提取精度平均提高了 9.43%。

Table 1. The results of confusion matrix used to assess accuracy of each method

表 1. 精度检测生成的混淆矩阵值对应表

项目	水体信息 (1)	非水体信息 (-1)	非检测信息 (0)
水体信息(255)	255	-255	0
非水体信息(-1)	-1	1	0

Table 2. The confusion matrix of each method

表 2. 各方法水体提取精度混淆矩阵

混淆矩阵统计数		真实水体像元	真实非水体像元
NDWI 法	检测为水体的像元	5617	833
NDWI 法	检测为非水体的像元	1126	2724
NDVI 法	检测为水体的像元	5406	612
NDVI 法	检测为非水体的像元	1337	2645
监督分类法	检测为水体的像元	5778	466
监督分类法	检测为非水体的像元	965	2791
改进方法	检测为水体的像元	6270	262
改进方法	检测为非水体的像元	473	2995

Table 3. Comparison of errors between modified and conventional methods

表 3. 改进方法与常规方法的误差对比

水体提取方法	改进方法	NDVI 指数法	NDWI 指数法	监督分类法
精度	92.65%	80.51%	83.41%	85.69%

参考文献 (References)

- [1] 马国明, 宋怡, 王雪梅. 新疆若羌湖泊遥感动态监测研究[J]. 冰川冻土, 2008, 30(2): 189-195. MA Mingguo, SONG Yi and WANG Xuemei. Dynamically monitoring the lake group in Ruoqiang county, Xinjiang region. Journal of Glaciology and Geocryology, 2008, 30(2): 189-195. (in Chinese)
- [2] JIANG L., NARAYANAN, R. M. A shape-based approach to change detection of lakes using time series remote sensing images. Transactions on Geosciences and Remote Sensing November, 2003, 41(11): 2466-2477.
- [3] MCFEETERS, S. K. The use of normalized difference water index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 1996, 17(7): 1425-1432.
- [4] XU, H. Q. Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, 2006, 27(14): 3025-3033.
- [5] 曹荣龙, 李存军, 刘良云. 基于水体指数的密云水库面积提取及变化监测[J]. 测绘科学, 2008, 33(2): 158-160. CAO Ronglong, LI Cunjun and LIU Liangyun. Extracting miyun reservoir's water area and monitoring its change based on a revised normalized different water index. Science of Surveying and Mapping, 2008, 33(2): 158-160. (in Chinese)
- [6] MA, M. Change in area of Ebinur lake during the 1998-2005 period. International Journal of Remote Sensing, 2007, 28(24): 5523-5533.
- [7] 翟辉琴. 基于数学形态学的遥感影像水域提取方法[J]. 测绘科学, 2006, 31(1): 22-24.

- ZHAI Huiqin. Anew algorithm for water area extracting from the remote sensing image based on morphologic method. *Science of Surveying and Mapping*, 2006, 31(1): 22-24. (in Chinese)
- [8] 柯长清. 湖泊遥感研究进展[J]. *海洋湖泊通报*, 2004, 23(4): 81-86.
- KE Changqing. A review of monitoring Lake Environment change by mass of remote sensing. *Transactions of Oceanology and Limnology*, 2004, 23(4): 81-86. (in Chinese)
- [9] 吴赛, 张秋文. 基于MODIS遥感数据的水体提取方法及模型研究[J]. *计算机与数字工程*, 2005, 22(7): 1-4.
- WU Sai, ZHANG Qiuwen. Method and model of water body extraction based on remote sensing data of MODIS. *Computer and Data Engineering*, 2005, 22(7): 1-4. (in Chinese)