

遥感影像分类方法在水体面积估算中的比较研究

张成才¹, 窦小楠¹, 张楠², 张颖¹

(1. 郑州大学水利与环境学院, 郑州 450001; 2. 民航河南空管分局, 郑州 450000)

摘要: 随着遥感技术的广泛应用, 利用遥感影像提取水体信息为水文研究提供了基础数据。目前进行水体提取所使用的遥感数据分辨率较低, 影响了水体提取的精度。Landsat TM 遥感影像主要依据水体在 7 个波段上光谱的不同特征以及其他地物与水体的区别, 通过分析水体及背景地物的光谱值, 利用单个波段或多个波段组合来提取水体信息。以昭平台水库的 TM 数据为例, 对其进行了几种提取水体信息方法的研究。通过总体精度及 *Kappa* 系数的对比, 选择最优分类方法, 并将该方法用于水体面积的估算。

关键词: 遥感; 分类方法; TM 影像; 面积估算

中图分类号: TP79

文献标识码: A

文章编号: 1673-7148(2008)03-0024-05

引言

随着航空、航天技术的飞速发展, 遥感技术已成为研究全球变化、人口资源环境不可替代的先进手段。遥感技术是能够快速获得大范围地面数据的有效手段。其覆盖面大, 同步效果好, 信息丰富, 有利于观察和研究各种地面现象的连续空间分布。遥感数据具有感测范围大、信息量大、更新时间快以及实时性和动态性且对被调查对象不产生破坏的特点, 适用于大范围的资源调查与研究。卫星遥感影像中快速、准确地提取水体信息已经成为水资源调查、水资源宏观监测、城市绿化度调查及湿地保护的重要手段。水体提取是进行水体悬浮固体浓度及水深计算的基础, 水体提取的好坏直接决定各种计算的精度, 因此水体信息提取具有不可忽视的作用。目前, 利用卫星遥感数据进行水体自动提取得到了广泛的研究, 各种方法相继提出并得到应用。

Shih(1985)利用 Landsat MSS 提取水体, 认为单波段(近红外波段 7)密度分割法与波段 5 和波段 7 组合的非监督分类法所获得的水体表面积仅差 3%^[1]。Steven(2002)在利用 MT 影像、航空影像等数据对美国 10 个主要城市的湖泊透明度进行评估时, 利用非监督分类的方法将 TM 影像分为 10 类, 然后将分类后影像聚类成水体和陆地两类, 并利

用聚类后的影像作为二值掩膜进一步生成只有水体的影像^[2]。

邓劲松等利用波段运算得到特征波段(PRWI), 经过图像增强处理后提取水体和居民地混合信息, 在此基础上通过分析水体的光谱特征, 发现水体和居民地在近红外和短波红外上有显著差异, 采用决策树模型将水体专题信息提取出来^[3]。

周成虎、杜云艳等提出了基于水体光谱知识的 AVHRR 影像水体自动提取识别的水体描述模型, 并将该模型应用于太湖、淮河、渤海等地区^[4]。

1 理论基础

1.1 遥感分类原理

影像分类的目的是将影像中每个像元根据其所在不同波段的光谱亮度、空间结构特征或者其他信息, 按照某种规则或算法划分为不同的类别, 从而进行更进一步的研究, 比如土地利用变化、影像融合效果比较等。

每种地物有其固有的光谱特征, 它们位于光谱空间中的某一点。由于干扰的存在, 环境条件的不同, 例如阴影、地形上的变化、扫描仪视角、干湿条件、不同时间拍摄及测量误差等, 同一类物质的各个样本不是集中到光谱空间的一点, 而是围绕某一点呈概率分布^[5], 即使如此也可以根据边界、阈值进

行分类。

1.2 水体光谱特性

水体属于光滑表面,以镜面反射为主,后向散射弱,在图像上呈暗色或黑色,所以在影像上很容易被识别出来。同时,水体内所含的物质如泥沙、矿物质、氧化物等,可以使遥感影像发生变化,也能被识别出来;与水有关的土壤、岩石、植物等因受水体影响,也容易在遥感图像上识别出来^[6]。因而,水文学在遥感技术应用中有得天独厚的优势。

2 遥感影像分类过程

2.1 分类方法

遥感影像的分类可以分为非监督分类(unsupervised)与监督分类(supervised)两种。非监督分类,也称为聚类分析或点群分析,即在多光谱图像中搜寻、定义其自然相似光谱集群的过程^[7]。非监督分类不需要人工选择训练样本,仅需要极少的人工初始输入,计算机按照一定规则自动根据像元光谱或空间等特征组成集群组,然后分析者将每个组和参考数据比较,将其划分到某一类别中去,如 ISO-DATA 方法、K-均值方法等。

监督分类又称训练分类法,即用被确认类别的样本像元去识别其他未知类别像元的过程。已被确认类别的样本像元是指那些位于训练区的像元。在这种分类中,分析者在图像上对每一类别选取一定数量的训练区,计算机计算每种训练样区的统计或其他信息,把每个像元和训练样本作比较,按照不同规则将其划分到和其最相似的样本类,如平行六面体法、最小距离法、最大似然法和二进制编码法等。

2.2 训练样本的选取

训练样本的选择需要分析者对要分类的图像所在的区域有所了解,或进行过初步的野外调查。其最终选择的训练样本应能准确代表整个区域内每个类别的光谱特性差异。同一类别训练样本必须是均质的,不能包含其他类别。训练样本的大小、形状和位置必须能同时在图像和实地容易识别和定位。必须要考虑每一类别训练样本的总数量,如果图像有 N 波段,则每一类别应该至少有 $10N$ 个训练样本,才能满足一些分类算法中计算方差及协方差矩阵的要求。

2.3 分类精度

对于分类过的影像,要对其精度进行评价,以检验选择的分类方法是否适合。分类误差主要分为两

类:一类是位置误差,即各类别边界的不准确;另一类是属性误差,即类别识别错误。

精度评价包括很多指标,其中最主要的是总体精度和 Kappa 系数。总体精度由误差矩阵计算得出。误差矩阵主要是将各个像元正确分类、错误分类和漏分的个数进行统计,矩阵的对角线上列出的是正确分类的像元。也就是说,误差矩阵对角线的数值越大,总体精度越高。Kappa 系数则考虑到了对角线上被正确分类的像元,同时也考虑了不在对角线上的各种漏分和错分的像元。所以这两个指标往往并不一致,在精度评价中,应同时计算以上两个指标,以便能得到更多的精度信息。

2.4 分类过程

影像分类有 3 个核心处理部分,分别为样本的提取、分类方法的选择和精度评价(见图 1)。首先对要分类的影像进行噪声去除、影像配准等预处理,然后选择非监督分类中的某种方法进行分类,与此同时,可以按照一定的准则进行训练样本的提取,进而选择监督分类的某种合适算法进行分类。最后把分类后的影像进行精度评价,如果精度满足要求,可进行下一步的实际应用;如精度不满足要求,则重复分类步骤,直至满足要求为止。

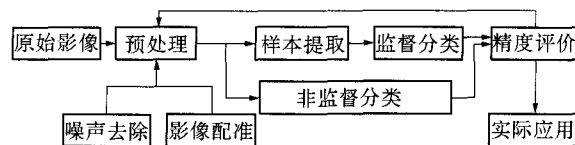


图1 影像分类的基本流程

3 影像对比实验与分类精度分析

昭平台水库位于淮河流域沙颍河水系的沙河干流上,距河南省鲁山县县城西约 13 km。1959 年建成水库,控制流域 1 430 km²,蓄水量 7.27 × 10⁸ m³,是集防洪、灌溉、发电、养鱼、旅游为一体的一座大(Ⅱ)型水库枢纽。

分类实验采用的数据是美国 NASA 的陆地卫星(Landsat)提供的 1999 年的 TM 影像。该影像的大小为 400 像元 × 398 像元。TM 影像有 7 个波段,其中第 2、3、4 波段为最佳观测波段,因此本实验选择该 3 波段进行。

实验的目的:用遥感影像的多种分类方法进行水体面积的提取,通过精度分析选择最优的分类结果,用其结果进行水体面积的估算。

为了目视效果明显,选择 TM2、3、4 波段(B、G、R)进行假彩色合成。影像如图 2。



图 2 原始假彩色合成影像

3.1 非监督分类

由于非监督分类不用选择训练样本,所以先进行非监督分类中的 K-均值法,并且在分类数量上选择两类即可满足水面提取的要求,结果如图 3。



图 3 K-均值分类结果

3.2 监督分类

进行监督分类时,要选择训练样本,而且要对训练样本进行分析,从而检验样本选择是否合适。从样本反射率(图 4)可以看出,选定的水体与陆地的训练样本反射率差异很大,因此满足分类要求。如果选择的训练样本的反射率相差不大,甚至有交叉情况,则说明训练样本选择不当,应重新选择。

选择训练样本后,就可以进行监督分类,从而提取水面面积。选择方法与分类结果如图 5 所示。

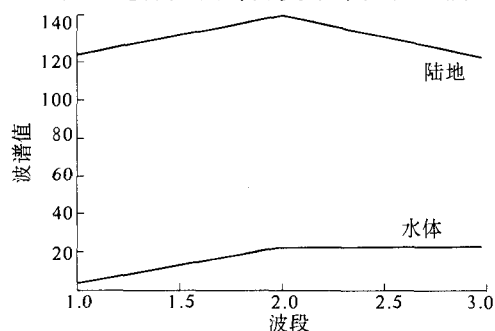


图 4 两类训练样本的波谱值

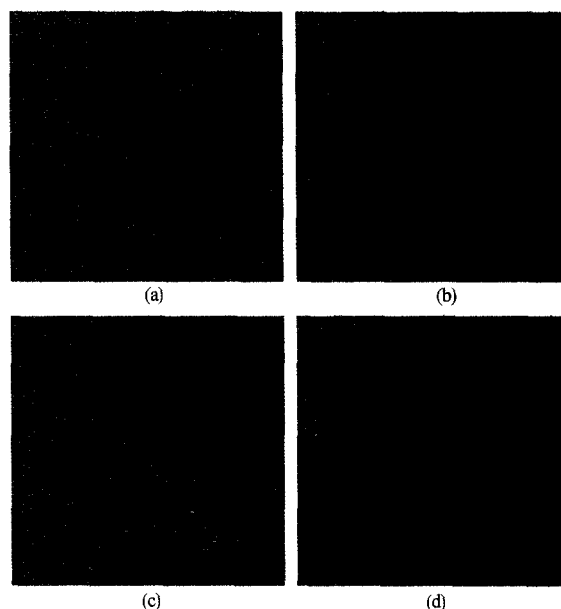


图 5 4 种监督分类结果

(a)二进制编码法 (b)马氏距离法
(c)最大似然法 (d)最小距离法

3.3 精度评价

如果把各种分类结果与原始图像的单个波段(如波段 4)进行叠加,则分类结果更有利于目视比较,结果如图 6 所示。

由图 6 目视可以得出,5 种方法都能够比较理想地把水体提取出来,但是二进制编码法与 K-均值法对于陆地像元错分为水体像元的部分比较严重,而马氏距离法、最大似然法、最小距离法的分类效果较好,基本能够达到实验要求,其中又以最大似然法的效果最理想。对以上各分类方法精度评价结果如表 1 所示。

由表 1 可以得出,马氏距离法、最大似然法、最小距离法用于该影像的分类结果的总体精度都达到了 90% 以上,是满足实验要求的,这与图 5 的目视结果相一致。其中,最大似然法最优,总体精度为 95.361%, Kappa 系数为 0.9353。因而,选择采用最大似然法的分类结果做该影像的水体面积提取与估算。

3.4 面积提取计算

TM 影像的空间分辨率为 $30 \text{ m} \times 30 \text{ m}$,代表的含义为影像上一个像元点所对应的实际面积为 900 m^2 。由最大似然分类法统计得出该影像水面的像元数为 36 397 个,所以计算得出昭平台水库面积为 32.8 km^2 。资料显示,1999 年昭平台水库的实际水面面积为 36 km^2 ,因此,采用以上方法能满足实际要求。

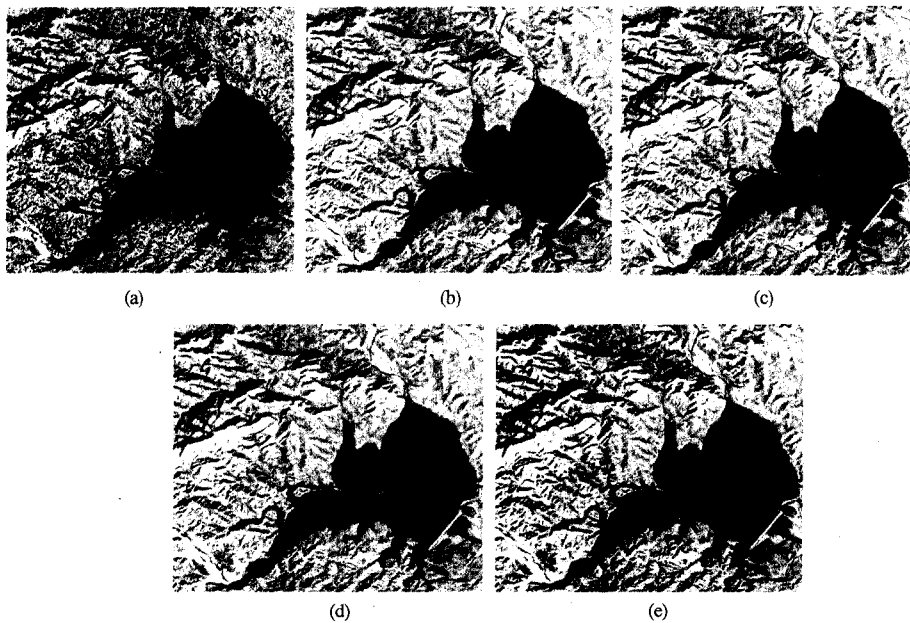


图6 分类结果与原始影像的叠加

(a)二进制编码 (b)马氏距离 (c)最大似然 (d)最小距离 (e)K-均值

表1 不同分类方法分类结果精度比较

分类方法	K-均值	二进制编码法	马氏距离法	最大似然法	最小距离法
总体精度/%	75.687	81.232	92.106	95.361	93.128
Kappa 系数	0.4282	0.5136	0.7456	0.9353	0.7733

4 结 语

随着卫星技术的迅速发展,遥感被越来越多地应用到了各个交叉领域。本文利用遥感中的分类方法,实现了对水体面积的提取与估算。通过研究发现,在实现的过程中需要注意以下几点:

①假彩色合成的影像有利于目视判别,并有利于训练样本选点的正确性。

②一幅影像在未知分类个数的情况下,可以先进行非监督分类,然后根据非监督分类的结果,确定训练样本的个数。

③训练样本的选点很重要,应该遵循以下几点:同一类别训练样本必须是均质的,不能包含其他类别;大小、形状和位置必须能同时在图像和实地容易识别和定位;如果图像有 N 波段,则每一类别应该至少有 $10N$ 个训练样本,才能满足一些分类算法中计算方差及协方差矩阵的要求;每间隔 N 个像元采集一个像元样本能得到较高的分类精度。

④分类后的影像可以与原始影像的单个波段进行叠加,从而能更有利于目视的比较。

⑤本文得出的最大似然法仅为该影像的最优分类方法,如果在其他区域,则不一定为最优。

参考文献

- [1] Shih S F. Comparison of ELAS classifications and density slicing Landsat data for water surface area assessment[C]//Johnson A I. In Hydrologic Applications of Space Technology (Publication No.160). Intl. Assoc. Hydrologicalsci, 1985.
- [2] Steven M K. Application of Landsat imagery to regional - scale assessments of lake clarity[J]. Water Research, 2002, 36: 4330 - 4340.
- [3] 邓劲松,王珂,邓艳华,等. PSOT-5 卫星影像中水体信息自动提取的一种有效方法[J]. 上海交通大学学报:农业科学版, 2005, 23(2): 198 - 201.
- [4] 周成虎,杜云艳,骆剑承. 基于知识的 AHVRR 影像的水体自动识别方法与模型研究[J]. 自然灾害学报, 1996, 5(3): 100 - 108.
- [5] 王一达,沈熙玲,谢炯. 遥感图像分类方法综述[J]. 遥感信息, 2006(5): 67 - 70.
- [6] 魏文秋,张利平. 水文信息技术[M]. 武汉:武汉大学出版社, 2003: 106 - 107.
- [7] 赵英时. 遥感应用分析原理与方法[M]. 北京:科学出版社, 2003: 194 - 196.

Comparative Study of the Remote Sensing Image Classification Method Based on Water Area Estimation

Zhang Chengcai¹, Dou Xiaonan¹, Zhang Nan², Zhang Ying¹

(1. School of Water Conservancy and Environment Engineering of Zhengzhou University, Zhengzhou 450001, China;

2. Henan Air Traffic Management Sub-bureau of CAAC, Zhengzhou 450000, China)

Abstract: With the extensive application of remote sensing technology, it provides basic information for hydrology by using remote sensing images from the water body. At present, low-resolution of remote sensing data that is used for water extraction influences the precision of water extraction. Based on the different characteristics in water 7-band spectrum and the difference between the water body and other landmark, Landsat TM image extracts water information by utilizing single-band or band combinations after analyzing the spectra of background feature. Taking the TM data of Zhao reservoir for example, several kinds of methods for water extraction are studied. Through the overall accuracy and contrast of the Kappa coefficients, the optimal choice of classification method is obtained, and then it is used to estimate the area of water.

Key words: remote sensing; classification method; TM image; area estimate