

## 基于遥感影像纹理信息的湖泊围网养殖区提取\*

李俊杰<sup>1,2</sup>, 何隆华<sup>1</sup>, 戴锦芳<sup>1</sup>, 李金莲<sup>1,2</sup>

(1: 中国科学院南京地理与湖泊研究所, 南京 210008)

(2: 中国科学院研究生院, 北京 100039)

**摘要:** 中国东部许多湖泊被人为围网养殖开发, 高密度的围网养殖区较容易导致湖泊富营养化和水质恶化, 因此人们利用遥感数据开展了湖泊围网养殖区的调查研究. 对于湖泊围网养殖区的提取大多采用多光谱分类或目视解译手工数字化, 多光谱分类围网养殖区和自然水体易于混淆, 而手工数字化对于大区域提取工作量比较大. 论文提出了基于遥感影像纹理信息的湖泊围网养殖区提取方法, 事实证明此方法易于实现且提取精度高. 该文选用中巴资源卫星 02 星多光谱数据, 以白马湖为试验区, 首先采用归一化差值植被指数提取水体及少部分光谱特征与之相似的人工建筑物和裸土, 再采用主成份分析对研究区影像进行数据压缩和几何信息增强, 利用灰度共生矩阵对影像纹理信息进行分析, 以均值为量化指标, 确定其最佳阈值, 最后以决策树分类方法提取湖泊围网养殖区.

**关键词:** 中巴资源卫星; 纹理; 湖泊围网养殖; 决策树

## Extract enclosure culture in lakes based on remote sensing image texture information

LI Junjie<sup>1,2</sup>, HE Longhua<sup>1</sup>, DAI Jingfang<sup>1</sup> & LI Jinlian<sup>1,2</sup>

(1: Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, P. R. China)

(2: Graduate School, Chinese Academy of Sciences, Beijing 100039, P. R. China)

**Abstract:** Most lakes in eastern China are exploited in the form of enclosure culture. The high density enclosure culture area is easy to cause lake eutrophication and deteriorate the water quality. So the lake enclosure culture area is focused by people and is usually extracted by using the multi-temporal remote sensing data and the multi-spectral classification or the manual digitization means. The multi-spectral classification is easy to confuse enclosure culture area with the natural water area and the manual digitization demands much time. This paper suggests a new method using the image texture information to extract lake enclosure area. The result is more accurate and the method demands less expense. Lake Baima is the study area and the image of the area is multi-spectral data of China Brazil Earth Resources Satellite 02. First the water area, the man-made area and the bare soil which has the similar spectral characteristics with the water are extracted by using the normalized difference water index. Then the principal component analysis was used to compress the multi-spectral data of the study area and enhance the texture information of the image. The third step Grey Level Co-occurrence Matrix (GLCM) was used to analyze the image, select a proper threshold of mean value which quantifies the GLCM. Finally, the lake enclosure culture area was extracted by using the decision tree.

**Keywords:** China Brazil Earth Resources Satellite; texture; enclosure culture in lake; decision tree

中国东部湖泊由于长期泥沙淤积面积日趋减少, 湖床渐被淤高, 洲滩广为发育, 普遍呈现浅水型湖泊的特点, 多数湖泊水深只有 2m 左右. 随着东部社会经济的发展, 人们对湖泊的开发利用强度也逐渐提高, 湖区围网养殖和围垦种植情况非常普遍. 很多湖泊被过度开发, 人工围网养殖面积要超过湖泊自身水面面积

\* 中国资源卫星应用中心项目、中国科学院知识创新工程项目 (INF105 - SDB - 1 - 22)、江苏省自然科学基金项目 (BK2002149) 联合资助. 2005 - 10 - 25 收稿; 2006 - 01 - 20 收修改稿.

李俊杰, 男, 1983 年生, 硕士研究生; E-mail: lijunjie299@126.com.

的一半甚至更多,而且围网阻滞水流,导致湖泊水质恶化,对原有生态系统造成严重破坏,甚至恶性循环.因此利用科学手段对湖泊资源要素进行监测,引导湖泊良性开发,对地区可持续发展有着十分重要的意义.

近年来,随着遥感科学的发展,遥感数据内容不断丰富,分辨率不断提高,为湖泊研究提供了有效、快捷的手段.不少学者利用遥感手段对东部湖泊进行了湖泊面积变化、围网养殖区变化、水质变化等多项监测研究,如沈芳 2003 年采用 1986 - 1998 年多时相 TM 数据对太湖流域中小湖群中的溇湖、洮湖围网养殖区域及其发展变化进行了研究,并利用遥感数据对养殖区的水质进行了分析<sup>[1]</sup>,杨英宝 2005 年对多时相的 TM 数据和航片进行数字化,研究了东太湖的围网养殖区域及其动态变化,并确定了超密度的围网养殖区域<sup>[2]</sup>.这些学者对湖泊围网养殖区域的研究一般采用多光谱数据分类或人工目视判读数字化.由于围网养殖区域和自然水体光谱特征比较接近,利用多光谱分类自然水体和湖泊围网养殖区域很容易混淆;而目视判读手工数字化虽然精度较高,工作量却较大.

针对以上方法的局限性,论文提出的基于遥感影像纹理信息的湖泊围网养殖区提取较好的解决了以上问题.论文选择白马湖为研究对象,采用中巴资源卫星多光谱数据,综合应用主成份降维处理、灰度共生矩阵、决策树分类等方法,并结合多光谱的 NDWI (Normalized Difference Water Index) 分析,系统地探讨了中巴资源卫星数据在信息提取应用中的技术方法,可以为应用中巴资源卫星进行大范围的湖泊围网养殖区提取,以及生态调查提供技术参考.

## 1 研究区概况及数据搜集

### 1.1 研究区概况

白马湖是运西湖群中位置最北的一个湖泊,跨金湖、宝应、洪泽、淮安四县市.湖盆浅碟形,人工湖岸,岸线规则,湖底平坦,淤泥深厚.湖中散布有大小土墩近百个,是 1855 年黄泛后湖区群众垦荒所遗弃的居民点.白马湖水位 6.50m,长 18.0km,最大宽 11.0km,平均宽 6.0km,面积 108.0km<sup>2</sup>,最大水深 2.0m,平均水深 0.97m,蓄水量 1.05 × 10<sup>8</sup> m<sup>3</sup>.白马湖原与宝应湖相连,1969 年于两湖间筑堤建闸,从此成为一个独立的湖泊<sup>[3]</sup>.

白马湖围网养殖和围垦比较严重,现在自然水面面积还不到湖泊面积的一半.白马湖区域自然水体在遥感影像上纹理表现比较光滑均一,而围网养殖区域纹理表现为小的比较规则的多边形.

### 1.2 数据搜集

遥感影像为中巴资源卫星 02 星 (China Brazil Earth Resources Satellite, CBERS - 2) 的 CCD 多光谱数据. CBERS - 2 多光谱数据包括蓝、绿、红三个可见光波段,一个近红外波段,一个全色波段,5 个波段星下点空间分辨率均为 19.5m.所用影像成像时间为 2004 年 4 月 20 日,行列号是 369 (PATH)、63 (ROW),影像的波段编号及对应波段光谱范围见表 1.用于遥感影像几何校正的参考图是扬州地区 1:100000 地形图.

表 1 CBERS - 2 波段编号及对应波段的光谱范围

Tab. 1 CBERS - 2 band number and the corresponding spectral range

波段编号	1	2	3	4	5
光谱范围(μm)	0.45 - 0.52	0.52 - 0.59	0.63 - 0.69	0.77 - 0.89	0.51 - 0.73

## 2 研究方法与技术流程

### 2.1 研究方法

2.1.1 阈值法 阈值法是数字图像处理中一种基本而且重要的方法,常用于图像分割.当使用阈值来进行遥感图像分割时,所有灰度值大于某阈值的像素都被判属于某类地物;所有灰度值小于或等于该阈值的像素被排除在此类地物之外.

2.1.2 NDWI 计算与主成分分析 归一化差异水体指数 (Normalized Difference Water Index, NDWI) 是 Mcfeeters 于 1996 年提出的<sup>[4]</sup>. NDWI 计算公式如下:

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (1)$$

式中 Green、NIR 分别表示多光谱数据中的绿波段和近红外波段。

NDWI 与 NDVI(Normalized Difference Vegetation Index) 计算公式相似,它是 Mcfeeters 根据 NDVI 公式改进得到,Mcfeeters 分析认为 NDVI 对低红光反射率和高近红外反射率的物体(如植被)有增强的作用,但对于低红光反射率和低近红外反射率的物体(如水)会产生抑制和去除的效果。为了更好的描述和提取地表水体,他提出了基于 NDVI 的改进 NDWI。绿光的水体反射率要大于红光,因此 NDWI 既保留了 NDVI 对植被和土壤的增强作用,又最大化了水体的特性。Mcfeeters 的研究表明在 NDWI 图像上水体将是正值,而植被和土壤将是负值和零值。

主成份分析(PCA, principal components analysis)是一种去除波段之间多余信息,将多波段的图像信息压缩到比原波段更有效的少数几个转换波段的方法。通过对多波段数据进行主成份分析并抽取第一主成份可以达到数据压缩和增强单波段数据纹理信息的效果。

2.1.3 纹理分析方法 从影像解译的观点来看,一个物体的空间特征往往是鉴别该物体的极为重要的特征;空间特征包括物体的大小、形状、纹理或线性特征,空间特征还不能达到光谱特征的有效利用水平<sup>[5]</sup>。

纹理是一种反映一个区域中像素灰度级的空间分布的属性,纹理表现为图像灰度在空间上的变化和重复或图像中反复出现的局部模式(纹理单元)及其排列规则。纹理分析的基本方法有三类:统计分析方法,结构分析方法和频谱分析方法。论文采用的纹理分析方法是统计分析方法中的灰度共生矩阵(Gray Level Co-occurrence Matrices, GLCM)方法,GLCM 是纹理分析中普遍常用的方法。参考有关论文表明 GLCM 是最适合于自然界影像的纹理分析<sup>[6]</sup>。

灰度共生矩阵提供了影像中像元与像元、像元与整体影像之间的空间关系,但还需进一步量化为单一数值的形式,才能提供给电脑直接分析使用。Haralick1973 年提出了 14 种用来量化 GLCM 矩阵的纹理统计指标<sup>[7]</sup>。根据权重大体可以把这些纹理统计指标分为 3 组:(1)对比(contrast)组,测量距离 GLCM 矩阵对角线的远近权重,度量影像中对比的强烈程度;(2)规则(orderliness)组,用于度量在一定的窗口范围内,像元灰度值的条理程度;(3)GLCM 纹理度量的描述统计量组。根据纹理指标的分组,在每一组中选取了一个指标进行后续纹理分析,选取的指标如下:

$$\text{Contrast: } Con = \sum_{i=0}^N \sum_{j=0}^N (i - j)^2 C_{ij} \quad (2)$$

$$\text{Entropy: } Ent = - \sum_{i=0}^N \sum_{j=0}^N C_{ij} \ln C_{ij} \quad (3)$$

$$\text{Mean: } Mea = \sum_{i=0}^N \sum_{j=0}^N i C_{ij} \text{ or } \sum_{i=0}^N \sum_{j=0}^N j C_{ij} \quad (4)$$

## 2.2 湖泊围网养殖区提取的技术流程

湖泊围网养殖区提取的技术流程如图 1 所示:首先是对数据的预处理:噪声去除;采用 1:100000 地形图进行几何校正,精度控制在一个像元以内;对研究区影像进行裁减。然后用数据的第 2 和第 4 波段计算 NDWI,同时对第 1 到第 4 波段进行主成分分析,选取第一主成份,第一主成份包含的原数据总方差百分比为 97.3%,它代表图像 4 个波段的高频部分和低频部分,体现了地物边缘纹理信息和基本色调特征,因此可选用第一主成分来代替其它波段来参与纹理特征提取。利用灰度共生矩阵的相应算法求出第一主成份的相应纹理统计指标,包括对比度、熵和均值,确定 mean 为量化指标。以 NDWI 和纹理统计指标均值为基础,利用决策树分类方法自动提取湖泊围网养殖区。最后分析结果并进行精度评价。

决策树(Decision Tree)是遥感图像分类中的一种分层处理结构,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子节点,分别代表和满足和不满足条件的类别。这种方法不仅不需要依赖任何先验的统计假设条件,而且可以方便地利用亮度值以外的其它知识,所以在遥感影像分类和专题信息提取中已有广泛的应用<sup>[8]</sup>。利用决策树方法提取湖泊围网养殖区的基本思路见图 1:(1)根据归一化差异水体指数  $NDWI > 0$  提取水体和少部分混入的人工建筑物和裸土,去除农作物、植被、大部分土壤和人工建筑物;(2)根据多波段数据主成分分析后

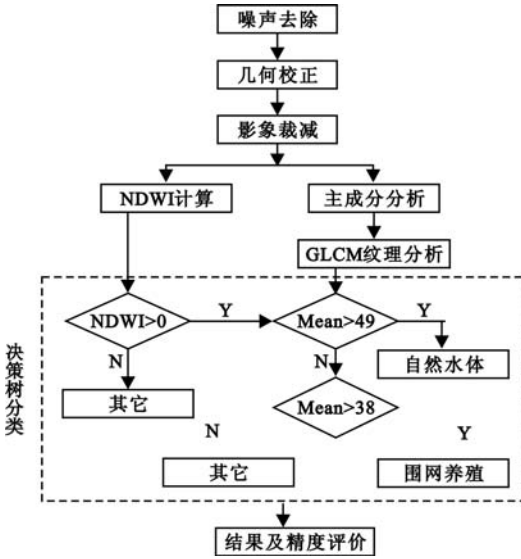


图1 湖泊围网养殖区提取的技术流程图

Fig. 1 The flow chart of extraction of the lake enclosure culture area

第一主成分计算的灰度共生矩阵的量化指标均值  $\text{mean} > 49$  提取(1)中肯定结果的自然水体;(3)根据  $\text{mean} > 38$  提取(2)中否定结果的围网养殖区域(判定肯定)和其它(判定否定)。

### 3 提取结果及精度分析

#### 3.1 纹理分析结果及最佳阈值的确定

用多光谱数据主成份分析后的第一主分量作为计算灰度共生矩阵的初始数据. 在 ENVI4.0 软件中计算灰度共生矩阵及其量化指标: 对比度 (contrast)、熵 (entropy) 和均值 (mean), 计算中选取的窗口大小为  $5 \times 5$ , 原输入的第一主成分数据灰度级重新量化到 64, 像元对  $x$  和  $y$  方向均偏置 1. 研究区纹理指标计算结果见图 2.

在纹理指标选择中论文只分析 NDWI 为正值的区域, NDWI 为正值区域地物类别主要为自然水体、围网养殖区域和其它 (包括人工建筑区和裸土). 对三种纹理指标图上三类要区分的类别进行样点采集分析, 判断三种指标对三类地物的可分性, 从图 3、图 4 可以看

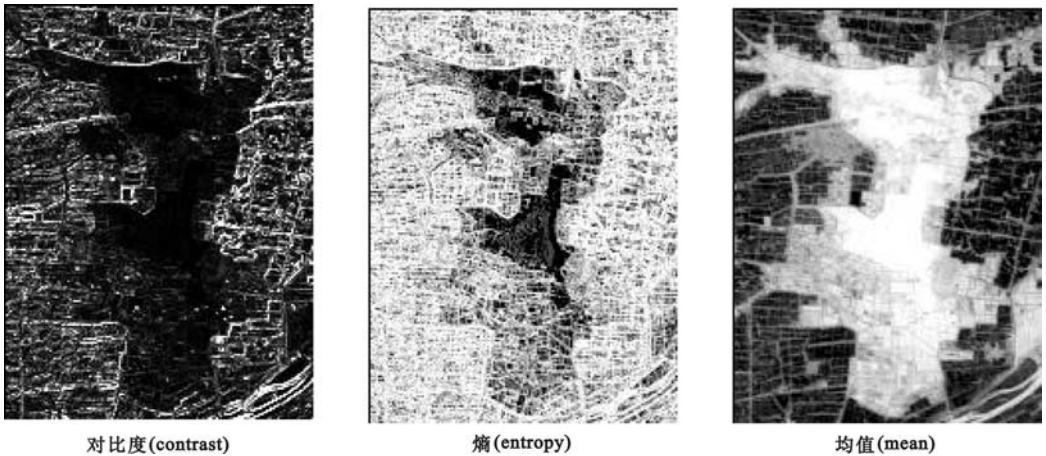


图2 研究区域影像纹理指标计算结果

Fig. 2 The computed result of the image texture index in study area

出, 对于对比度和熵值指标三种地物的值有一些交叉的地方, 这表明对这两种指标值采用阈值法分割图像无法有效区分这三类地物, 而从图 5 中看出: 三类地物采样点的纹理均值图显示三类地物的值基本互不相交, 固利用纹理量化指标均值的阈值分割可以简单有效的把三类地物区分开来. 因此我们选择均值作为提取围网养殖区的纹理指标, 经过对三类地物采样对比分析均值的大小, 最后确定纹理指标均值 49 和 38 作为区分自然水体、围网养殖区和其它的阈值.

#### 3.2 决策树方法提取结果

白马湖围网养殖区提取结果如图 6 所示, 湖泊围网养殖区为浅蓝色, 湖泊自然水面为深蓝色, 其它为浅赭色. 图 7 为白马湖区域中巴资源卫星 O2 星第 4 波段、第 2 波段和第 1 波段假彩色合成图, 图中红色区域

为农作物或植被,青色为人工建筑(道路、房子等),浅青色、蓝色和深蓝色区域为水体.

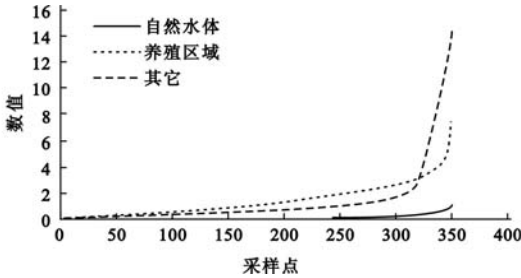


图3 三种不同地物采样点纹理对比度值图

Fig. 3 Three different objects sample point texture contrast value

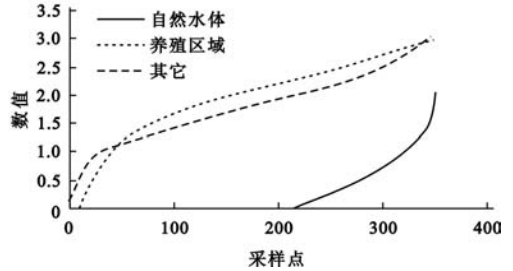


图4 三种不同地物采样点纹理熵值图

Fig. 4 Three different objects sample point texture entropy value

### 3.3 精度评价

将3.2提取结果与原始影像叠加分析,采用目视判读检验其提取效果,认为该方法的提取效果较好.为客观评价白马湖围网养殖区提取结果精度,由ENVI4.0软件随机选取了98个样本,建立混淆矩阵,见表2,围网养殖区提取精度为88.5%,自然水面精度为77.8%,其它精度为91.1%,总精度为86.7%,Kappa系数79.3%.

误分几类像元的分析原因如下:(1)把线性的水体误分为围网养殖区了,线性水体由于受周围地物的影响,纹理指标值与围网养殖区相近;(2)围网养殖区误分为自然水体,原因是比较大的围网养殖区内,水面纹理均一与自然水体相似;(3)其它类型地物和围网养殖区误分,主要原因是白马湖围网种植区和围网养殖区交错在一起,在边缘地方出现了误判.

图5 三种不同地物采样点纹理均值图

Fig. 5 Three different objects sample point texture mean value

表2 白马湖区域基于纹理特征的分类图精度表

Tab. 2 The accuracy table of Baima lake area based on image information classification

地面实况	图像分类结果				
	围网养殖区	自然水面	其他	总和	精度(%)
围网养殖区	23	1	2	26	88.5
自然水面	3	21	3	27	77.8
其他	3	1	41	45	91.1
总和	29	23	46	98	

## 4 结论与讨论

综上所述,采用NDWI阈值法与纹理特征辅助分类方法可以综合利用影像中的光谱信息与纹理信息,能较好地对湖泊围网养殖区信息进行自动提取,是中巴资源卫星影像或其它中高分辨率卫星影像提取湖泊围网养殖区的快速、有效的方法.研究采用的湖泊围网养殖提取体系中,应用NDWI阈值法提取水体及少部分光谱特征与之相似的人工建筑物和裸土;纹理量化指标均值指标能够较好地反映自然水体、围网养殖区

和其它地物内部结构的异质性,是识别光谱特征混淆的用地类型重要依据,决策树分类方法将 NDWI 与统计指标均值阈值进行逐步细化,能够较准确的提取出围网养殖区域.精度评价结果表明:以上方法简单可靠,提取精度较高.



图 6 白马湖围网养殖区提取结果图

Fig. 6 The extraction result figure of Baima lake enclosure culture area



图 7 白马湖 CBERS-2 数据 421 波段合成图

Fig. 7 Composed image of 421 band of Baima lake CBERS-2 data

但是基于纹理信息提取湖泊围网养殖区域也存在一些问题,值得注意和加以改进,如(1)对单一大块围网养殖区,区内水域难以和自然水体区别开来;(2)线状的河流水体易于被提取为围网养殖区.可以考虑的改进方法是对研究区域采用多光谱监督或非监督分类,把纹理信息作为辅助分类层;另外对分类结果矢量化,计算分类各图斑的形状指数,用以区分线性和非线性目标.因此在这几方面可以进行深一步的研究.

## 5 参考文献

- [1] 沈芳, 匡定波. 太湖流域典型中小湖群水资源利用及动态变化的遥感调查与分析. 遥感学报, 2003, 7(3): 221-226.
- [2] 杨英宝, 江南, 殷立琼等. 东太湖湖泊面积及围网养殖动态变化的遥感监测. 湖泊科学, 2005, 17(2): 133-138.
- [3] 王苏民, 窦鸿身主编. 中国湖泊志. 北京: 科学出版社, 1998: 288.
- [4] 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.
- [5] Swain P H, Davis S M. 朱振福, 童寿彬, 张晋等译. 遥感定量方法. 北京: 科学出版社, 1984: 145.
- [6] Ohanian P P, Dubes R C. Performance Evaluation for Four Classes of Texture Features. *Pattern Recognition*, 1992, 25: 819-833.
- [7] Haralick R M, Shanmugan K, Dinstein I. Texture Features for Image Classification. *IEEE Transactions On Systems, Man, and Cybernetics*, 1973, 3(6): 610-621.
- [8] 赵萍, 冯学智, 林广发. SPOT 卫星影像居民地信息自动提取的决策树方法研究. 遥感学报, 2003, 7(4): 309-315.