

基于对应分析的湖泊围网养殖范围提取

王 静^{1,2}, 高俊峰¹

(1 中国科学院 南京地理与湖泊研究所, 江苏 南京 210008 2 中国科学院 研究生院, 北京 100039)

摘 要: 为了提高基于遥感数据的湖泊围网养殖区提取精度, 提出了基于对应分析的湖泊围网养殖区的提取方法, 并以 ETM+ 卫星遥感数据为基础, 将其应用于太湖围网养殖区的提取。研究结果表明: (1) 基于对应分析的方法具有较高的提取精度, 围网养殖区的提取精度达到了 95.59%, 总体精度为 91.14%, Kappa 指数为 0.86; (2) 提取结果表明, 2001 年太湖围网养殖区面积为 111.76 km², 占湖泊总面积的 76.29%。实验表明, 基于对应分析的方法是遥感影像湖泊围网养殖区提取的一种新的有效方法。

关键词: 对应分析; 纹理分析; 决策树分类; 围网养殖区; 太湖

中图分类号: TP751/TP79 **文献标识码:** A

1 引 言

湖泊网围养殖的开展, 极大地促进了我国湖泊渔业的发展^[1]。地处长江三角洲经济圈的太湖流域, 伴随着经济的快速发展和产业结构的调整, 湖泊围网养殖迅速发展, 养殖面积不断扩大, 超密度网围水产养殖已成为这一流域中小型湖群水资源过度开发利用的一个重要特征^[2]。同时, 围网养殖也带来一系列负面影响, 其过度密集的围网养殖正在严重破坏着湖泊生态, 直接影响到湖泊蓄水供水、水质净化、洪水调蓄和气候调节等多种社会和生态功能的发挥^[1]。科学规划湖泊围网养殖区, 合理利用湖泊水面, 严格控制湖泊围网养殖面积的扩大, 以及对其进行遥感动态检测显得尤为重要, 湖泊围网养殖区范围和面积的确定作为遥感动态监测的一项重要内容, 对其进行研究具有重要意义。

湖泊围网养殖区的遥感提取作为一个崭新的课题, 目前已经受到研究者的重视, 相继开展了一系列的研究^[2-5], 在他们的研究当中, 主要侧重于对两个时段湖泊围网养殖区遥感分类结果变化的分析, 而对遥感提取的方法研究的相对较少。为了提高湖泊围网养殖区的提取精度, 我们提出了一种基于对应分析的湖泊围网养殖区提取方法。虽然

对应分析在生物和统计领域得到了广泛的认同和应用, 但在遥感领域的应用仍是相对较新的^[6,7], 正在积极的探索当中。1999 年, Carr 和 Matanavi 利用对应分析对 SPOT 影像进行了处理, 并与主成分分析和因子分析的结果做了比较^[6]; 同年, Xie 和 Xia 利用对应分析对 TM 影像进行处理, 从而估计棉花的种植面积^[8]; 刘庆生等对内蒙古自治区大青山地区色尔腾山的 TM 影像进行了对应分析, 进而区分各类岩石^[9]; 2006 年, Cakir 等将对应分析应用到土地利用/覆被的变化检测中并取得了优于主成分分析的结果^[7]; 2007 年, Gao 等基于对应分析对复杂地形区 ETM+ 影像进行了土地利用/覆被分类研究, 并取得了较好的结果^[10]。本文运用该方法对太湖围网养殖区进行提取, 并与基于主成分分析等其他方法的提取结果进行精度比较。

2 数据与方法

2.1 研究区概况及数据获取

2.1.1 研究区概况

太湖又名西涌子湖, 地理位置是 31°29′—31°42′N, 119°44′—119°53′E, 地处江苏省常州市西南 13km 处的太湖上游, 跨无锡、常州两市 (图 1)。

收稿日期: 2007-05-08 修订日期: 2007-10-10

基金项目: 国家科技部科技支撑计划 (编号: 200700000353) 和江苏省科技厅太湖水污染治理科技专项 (编号: BS2007162)。

作者简介: 王 静 (1979—), 女, 硕士研究生, 就读于中国科学院南京地理与湖泊研究所。主要从事水环境遥感、水资源以及水环境等方面的研究。发表论文 1 篇。Email: w_j75@163.com

漏湖由泻湖演变而成, 是太湖流域第二大湖泊, 总面积 146km^2 。漏湖湖面形状犹如一长茄形, 长约 22.1km , 最大宽 9.0km , 平均宽 6.63km , 湖盆地形平坦, 无显著起伏, 平均湖底高程 1.63m , 平均水深 1.26m ^[1]。漏湖是中国最早开展网围养殖的湖泊之

一, 自 1984 年开始网围养殖试验, 至 21 世纪初形成网围养殖高峰, 从网围养殖品种和分布来看, 漏湖形成了 3 个不同的养殖区即南部的网围河蟹养殖区, 中部的资源增殖保护区和北部的网围养鱼区^[1, 2]。



图 1 漏湖在太湖流域的位置及其 ETM+ 影像假彩色 (5 4 3) 合成图

Fig. 1 Location of Gehu lake in Taihu Basin and its Landsat-7 ETM+ band 5, 4, 3 false color image

2.1.2 数据获取

本研究所使用的数据主要有 Landsat-7 ETM+ 遥感影像、Google Earth 上截取的高分辨率漏湖影像、1:1 万漏湖利用现状图和实地调查数据。选取的 Landsat-7 ETM+ 遥感影像的获取时间为 2001-07-26, 轨道号为 path119/row38, 来源于马里兰大学, 数据记录格式为 GEOTIFF, 产品级别为正射校正。结合漏湖分布范围, 从 Landsat-7 ETM+ 影像中截取研究区 (图 1), 共计 $770\text{行} \times 508\text{列}$ 像元。图 1 为由漏湖 ETM+ 波段 5(R)、4(G)、3(B) 合成的假彩色影像, 从中可以看出, 漏湖中自然湖泊水面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。

2.2 研究方法及技术流程

2.2.1 对应分析原理

对应分析是在因子分析的基础上发展起来的, 又称 $R-Q$ 型因子分析, 与主成分分析主要研究变量之间的相互关系不同, 对应分析既研究变量之间的关系, 还研究样品之间的关系, 不仅如此, 它还能在同一个直角坐标系内同时表达出变量与样品两者

之间的相互关系^[11]。对遥感影像进行对应分析主要包括 5 个计算步骤^[6, 7, 9, 12]:

(1) 将遥感影像转换成二维矩阵 X (其中行代表波段 (p), 列代表像元 (N)), 并对数据进行标度和统一量纲。

$$p_{ij} = \frac{1}{T} X = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{p1} & p_{p2} & \cdots & p_{pN} \end{bmatrix} \quad (1)$$

式中, $T = \sum_{i=1}^p \sum_{j=1}^N x_{ij} p_{ij} = \frac{x_{ij}}{T}$, $p_i = \sum_{j=1}^N p_{ij} p_j = \sum_{i=1}^p p_{ij}$ 即 T 为矩阵 X 的数据值总和; p_{ij} 为每个元素出现的频率; p_i 为波段 i 的边缘概率; p_j 为像元 j 的边缘概率。

(2) 计算矩阵 Z 。

$$Z = \frac{p_{ij} - p_i p_j}{\sqrt{p_i p_j}} \quad (2)$$

(3) 作 R 型因子分析。首先计算矩阵 Z 的方差-协方差矩阵 $A = Z \times Z'$ 的特征值 λ_i ($i=1, 2, \dots, m$) 和特征向量 U_i ($i=1, 2, \dots, m$), 然后计算 R 型因子载荷矩阵 F 。

$$F = \begin{bmatrix} u_{11} \sqrt{\lambda_1} & u_{12} \sqrt{\lambda_2} & \cdots & u_{1m} \sqrt{\lambda_m} \\ u_{21} \sqrt{\lambda_1} & u_{22} \sqrt{\lambda_2} & \cdots & u_{2m} \sqrt{\lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{p1} \sqrt{\lambda_1} & u_{p2} \sqrt{\lambda_2} & \cdots & u_{pm} \sqrt{\lambda_m} \end{bmatrix} \quad (3)$$

F 为 R 型因子的载荷矩阵。

(4) 作 Q 型因子分析。对应于步骤 (3) 所得的特征值 λ_i ($i=1, 2, \dots, m$), 关于矩阵 $B = Z' \times Z$ 的特征向量 V_i ($i=1, 2, \dots, m$) 为 $V_i = Z' U_i$; 计算 Q 型因子载荷矩阵 G :

$$G = \begin{bmatrix} v_{11} \sqrt{\lambda_1} & v_{12} \sqrt{\lambda_2} & \cdots & v_{1m} \sqrt{\lambda_m} \\ v_{21} \sqrt{\lambda_1} & v_{22} \sqrt{\lambda_2} & \cdots & v_{2m} \sqrt{\lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ v_{p1} \sqrt{\lambda_1} & v_{p2} \sqrt{\lambda_2} & \cdots & v_{pm} \sqrt{\lambda_m} \end{bmatrix} \quad (4)$$

(5) 将矩阵 G 转换成影像, 即为对应分析结果图像。

2.2.2 技术流程

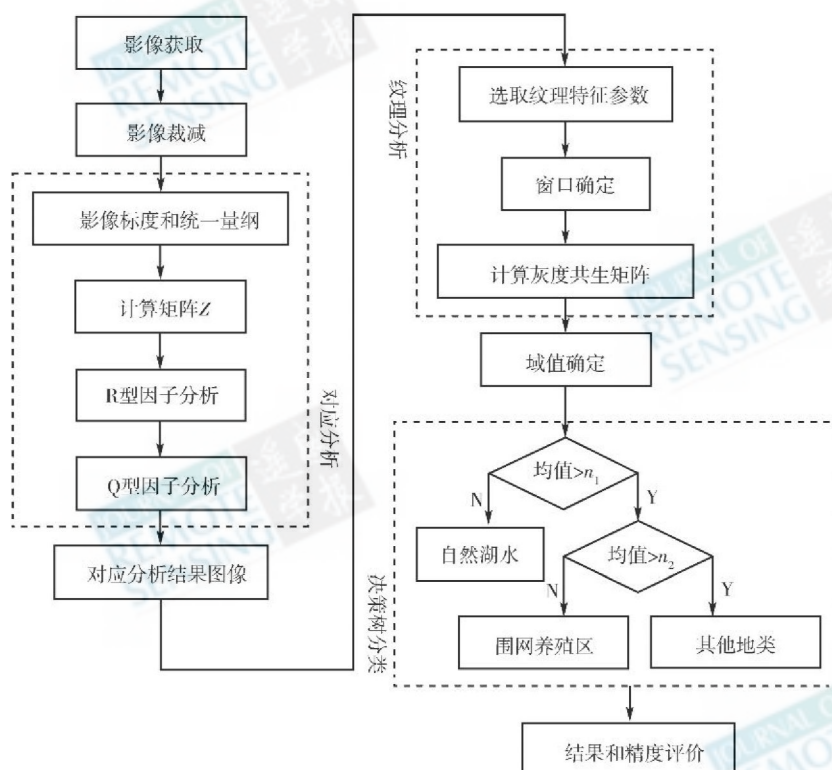


图 2 湖泊围网养殖区提取算法流程图

Fig. 2 Flowchart to extract the enclosure culture area in lake

围网养殖区提取的技术流程主要包括 4 步, 即影像预处理、对应分析、纹理分析和决策树分类 (图 2)。由于获得的 Landsat-7 ETM+ 遥感影像产品级别为正射校正, 所以本研究中影像预处理主要是根据研究区范围来裁减获得研究区 ETM+ 影像。然后利用 2.2.1 节介绍的对应分析方法与计算步骤对裁减获得的研究区 ETM+ 影像进行对应分析, 并获得对应分析结果图像。对其进行纹理分析, 图像的灰度共生矩阵作为统计纹理分析方法已被理论证明并且实验显示它在纹理确定上是一个很好的方法, 广泛用于将灰度值转化为纹理信息^[13], 故本文采用的纹理分析方法是统计方法中的灰度共生矩阵, 初选的几个常用纹理特征统计量主要为均值、变化量、对比度和熵值, 在此基础上进一步确定用于围网养殖区提取的纹理特征统计量, 并进行纹理分析窗口大小的确定, 进一步计算获得用于自然湖泊水面、围网养殖区和其他地类等 3 种地类分类提取的最佳阈值。根据最佳阈值, 采用决策树分类方法分类提取湖泊围网养殖区, 并对分类提取结果进行精度评价。

3 结果分析

3.1 基于对应分析法的提取结果

3.1.1 对应分析结果

影像经过对应分析,生成的 6 个成分中各分量信息内容呈逐渐递减趋势,前两个成分包含了原始影像 95.76% 的信息。其中第一成分主要综合了 4、5、1 波段的信息,包含了原始影像 92.20% 的信息,从而有利于在此基础上对其进行后续分析,进行围网养殖区的提取。第一成分中,波段 4 的贡献最大,载荷达 0.62,主要原因是该波段处于水体强吸

收区;其次是波段 5,载荷为 0.56,主要是该波段位于水的吸收带 (1.4μ, 1.9μ) 之间,受两个吸收带的控制,对水的敏感性较强的缘故。第一主成分图上,水体是黑色,围网是低灰阶值,高灰阶值的为其他地类;第二主成分包含了原始影像 3.59% 的信息,主要综合了 4、5 波段的信息,其载荷均超过了 0.5;第三成分主要综合了 1、3、2 波段的信息;第四成分主要综合了 3、2 波段的信息;第五成分主要综合了 7、5 波段的信息;第六成分主要综合了 1、3、2 波段的信息,但 4 个成分总共仅包含原始影像信息的 4.24%,其中第六成分所占的比例几乎为 0。影像对应分析后的特征向量矩阵和特征值如表 1。

表 1 对应分析后的特征向量矩阵和特征值

Table 1 Eigenvector matrix and eigenvalues of the image after correspondence analysis

	第一成分	第二成分	第三成分	第四成分	第五成分	第六成分
波段 1	0.413772	0.151646	-0.67856	-0.01367	0.175953	0.572768
波段 2	0.225109	-0.30401	0.468373	-0.63226	-0.05407	0.409789
波段 3	0.209068	0.231341	0.521415	0.67176	0.023924	0.456877
波段 4	-0.62339	0.646099	0.016128	-0.24995	-0.12783	0.330221
波段 5	-0.55655	-0.51608	-0.06591	0.189953	0.538176	0.281286
波段 7	-0.18987	-0.38377	-0.20905	0.224141	-0.81214	0.32727
特征值	1.43×10^{-7}	5.57×10^{-9}	3.21×10^{-9}	1.86×10^{-9}	1.49×10^{-9}	3.17×10^{-21}
贡献率	9.22×10^{-1}	3.59×10^{-2}	2.07×10^{-2}	1.20×10^{-2}	9.64×10^{-3}	2.05×10^{-14}
累计贡献率	9.22×10^{-1}	9.58×10^{-1}	9.78×10^{-1}	9.90×10^{-1}	1.00	1.00

3.1.2 纹理分析结果

(1) 纹理特征统计量的确定。首先,将 ETM+ 遥感影像对应分析后的第一主成分作为计算灰度共生矩阵的初始数据,使用 ENVI 4.2 软件计算灰度共生矩阵以及均值、变化量、对比度和熵 4 个纹理特征统计量。其次,对该 4 个纹理指标分析结果图进行样点采集,自然湖泊水面、围网养殖区以及其他地类 3 种地类分别采集约 300 个样点,并分别统计这些样点的 4 个纹理信息特征统计量在不同窗口 (3×3, 5×5, 7×7, 9×9, 11×11, 13×13, 15×15) 时的值。试验结果表明,对于变化量、对比度以及熵值 3 个指标,自然湖泊水面、围网养殖区以及其他地类等 3 种地类的值交叉现象比较严重,不利于 3 种地类的分类提取;而对于均值这一特征统计量,除在 3×3 窗口时有少量重叠外,在其他窗口 (5×5, 7×7, 9×9, 11×11, 13×13, 15×15) 时均无重叠现象,有利于该 3 种地类的分类提取,故选取均值这一纹理特征统计量作为进一步分析的依据。

(2) 纹理分析窗口大小的确定。纹理分析窗

口的确定主要是针对不同的纹理参数和不同地物,对不同大小的窗口进行纹理特征值计算,如果特征值从某一窗口开始趋于稳定,则此窗口大小符合此研究对象。由于本文的主要研究对象是围网养殖区,所以,只须计算不同窗口下围网养殖区的纹理特征均值,试验结果表明 (图 3),当窗口像素达到 9×9 时,其值开始趋于稳定,因此,选择 9×9 窗口做进一步分析。

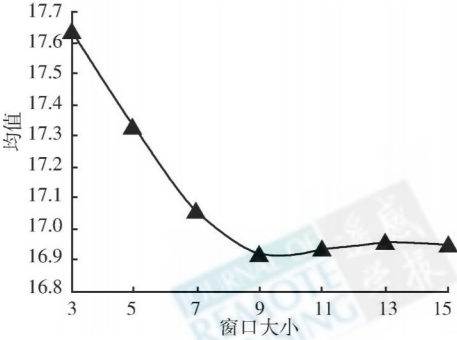


图 3 窗口大小与纹理特征均值关系图

Fig. 3 Relationship between texture window and texture index mean

3.1.3 阈值确定

运用 9×9 纹理分析窗口, 计算得到基于均值纹理特征统计量的影像纹理分析结果, 并根据对自然湖泊水面、围网养殖区以及其他地类等 3 种地类的采集样点, 计算得到该 3 种地类的样本特征值统计表 (表 2), 则阈值 n_1 , n_2 的取值范围分别为 $n_1 \in [10.2, 11.0]$, $n_2 \in [22.9, 29.2]$ 。

进而通过测试阈值和分类精度之间的关系确定最佳阈值, 分类精度最高时的阈值即为最佳阈值。为了提高测试速度和分类精度, 具体操作过程中主要采用“两步法”^[14]来实现。第一步即分别确定 10.2, 22.9 为 n_1 , n_2 的初始阈值, 以 0.1 为步长, 逐步测试阈值与精度之间的关系, 找出能产生高精度的阈值确定为“初阈值”; 第二步再以 0.01 为步长在“初阈值”周围进一步测试, 得到最佳阈值 $n_1 = 10.29$, $n_2 = 29.20$ 。

表 2 9×9 窗口影像采样点纹理特征均值统计表
Table 2 The statistics of texture index mean for three kinds of samples in 9×9 window

地类	最小值	最大值
自然湖泊水面	8.4	10.2
围网养殖区	11.0	22.9
其他地类	29.2	49.3

3.1.4 决策树分类结果

根据计算获得的最佳阈值 $n_1 = 10.29$, $n_2 = 29.20$ 运用决策树分类方法对漏湖围网养殖区进行提取。首先根据“均值 < 10.29 ”将自然湖泊水面提取出来, 然后再依据“均值 < 29.20 ”将围网养殖区提取出来, 得到分类结果图 (图 4)。并计算得漏湖围网养殖区面积为 111.76km^2 , 占湖泊总面积的 76.29% ; 漏湖自然湖泊水面面积为 31.18km^2 , 占湖泊总面积的 21.28% ; 其他地类面积为 3.56km^2 , 占湖泊总面积的 2.43% , 主要为建筑物以及出露的洲滩。

3.1.5 精度评价

通过在影像上随机选取验证点数据作为样本, 利用 ENVI 4.2 软件中的精度分析功能对分类后的数据进行精度评价, 计算得到影像分类精度评价结果 (表 3)。本次分类的总体精度为 91.14% , Kappa 系数为 0.86 说明分类方案的总体分类质量较好。但是, 仍有错分和漏分现象, 如围网养殖区在分类过程中出现较大的错分现象, 错分误差为 13.41% , 其他地类在分类过程中出现了较大漏分现象, 漏分误差为 21.43% 。

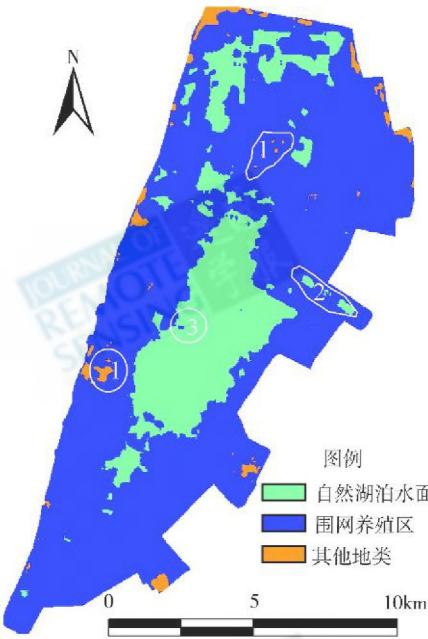


图 4 漏湖围网养殖区提取结果图

Fig. 4 The extraction result of enclosure culture area in Gehu lake

表 3 漏湖 ETM + 影像分类精度

Table 3 Precision of the classification for the Gehu lake ETM + image

地类	错分误差 /%	漏分误差 /%	精度 /%
自然水面	3.77	8.93	91.07
围网养殖区	13.41	4.05	95.95
其他地类	4.35	21.43	78.57

总体精度 = 91.14% , Kappa 系数 = 0.86

3.2 基于对应分析法与其他方法的分类效果比较

为了进一步评价基于对应分析法 (分类法 1) 围网养殖区的提取精度, 我们把该方法与其他 3 种方法进行了比较。第一种是基于主成分分析法 (分类法 2), 该法主要是利用主成分分析法替代对应分析法, 与分类法 1 不同的是, 在分析过程中, 采用了前 2 个主成分, 其累计贡献率为 96.81% ; 第二种分类方法是在分类法 2 的基础上, 引入经过纹理分析后的波段 4 (分类法 3) 做进一步分析, 以提高分类精度; 最后一种分类方法是直接利用原始波段进行纹理分析, 再进行决策树分类 (分类法 4), 而不做对应分析或主成分分析。

利用分类法 2、分类法 3 和分类法 4 进行分类, 围网养殖区提取精度分别是 78.38% , 91.89% 和 82.43% ; 总体精度分别是 82.17% , 89.17% 和

83.54%; Kappa系数分别是 0.71、0.82和 0.73。通过比较发现,在围网养殖区提取精度和总体精度上,分类法 1 较分类法 2 分别高 17.57 和 8.97 个百分点,分类法 1 较分类法 3 分别高 4.06 和 1.97 个百分点,分类法 1 较分类法 4 分别高 13.52 和 7.6 个百分点,分类法 1 表现出最好的分类精度,而且分类法 1 仅用了第一成分,在操作上较其他 3 种方法简便。由此可见,在同样条件下,基于对应分析法的精度高于其他 3 种分类方法,而且可操作性更强。图 5 表示了 4 种不同分类方法的分类效果。

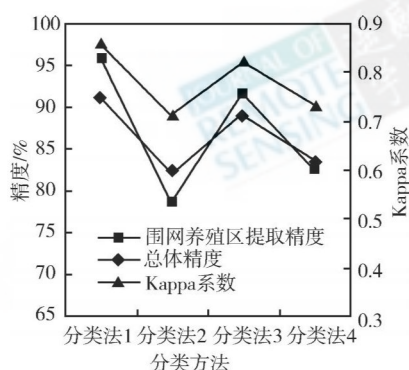


图 5 基于 4 种不同分类方法的精度比较

Fig.5 Comparison of the classification precision based on 4 different methods

3.3 漏湖围网养殖面积变化情况分析

漏湖自 1984 年开始网围养殖试验,当时网围养殖面积仅占湖泊总面积的 0.1% 左右;1986 年起,养殖面积逐年扩大,至 1994 年网围面积达 23.18 km²,占湖泊总面积的 15.8%;1997 年网围养殖区域的面积达 64.13 km²;1998 年以后网围养殖区域的面积更是达到 113.2 km²,占湖泊总面积的 78.8%^[1,2]。1998 年相比 1984 年,网围养殖区面积增加了 113.05 km²,年平均增长约 8.08 km²,这主要是受政策影响和经济利益驱动造成的,在“七五”(1986—1990 年)和“八五”(1991—1995 年)期间曾作为农业部重点攻关对象进行了网围养殖技术的研究^[1]。网围养殖区的迅速扩大,加剧了湖泊富营养化过程,导致漏湖水体生态环境日益恶化,对网围养殖进行控制显得尤为重要,2001 年,漏湖围网养殖区面积减少为 111.76 km²,占湖泊总面积的 76.29%,相比 1998 年减少 1.44 km²,漏湖围网养殖区迅速增长的态势得到了有效遏制。图 6 表示了漏湖围网养殖区面积变化情况。

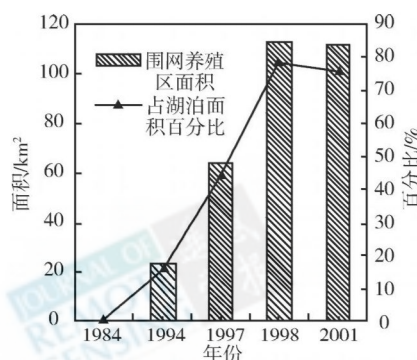


图 6 漏湖围网养殖区面积变化情况

Fig.6 Change chart of enclosure culture area in Gehu lake

4 结论与讨论

(1) 本研究提出了一种基于对应分析的湖泊围网养殖区提取新方法,并应用该方法对漏湖围网养殖区进行了提取。提取结果表明,2001 年漏湖围网养殖区面积为 111.76 km²,占湖泊总面积的 76.29%;精度检验结果表明,基于对应分析的提取方法的总体精度为 91.14%,其中,围网养殖区的提取精度达到了 95.95%,Kappa 系数 0.86 表明利用对应分析法可以快速有效地提取湖泊围网养殖区的范围和面积。

(2) 湖泊围网养殖区提取的高精度是与对应分析自身的优越性分不开的,因为对应分析既研究各波段(波段 1—5、7)之间的关系,还研究像元之间的关系,不仅如此,它还能在同一个直角坐标系内同时表达出波段与像元两者之间的相互关系。正是由于在分析过程中,它不仅考虑了影像波段特征属性及其相互关系,同时也考虑了像元特征之间的相互关系,较单纯考虑波段之间相互关系的主成分分析有所改进,有利于提取精度的提高。本研究应用该方法对漏湖围网养殖区进行了提取,并取得了较好的效果,因此认为如果能获得高质量的 TM/ETM+ 遥感图像并能进行严格有效的图像预处理,该方法也可以应用到对其他湖泊的围网养殖区的提取。

(3) 利用对应分析对漏湖围网养殖区进行提取,虽然获得了较高的提取精度,但仍然存在一定的误分现象,例如图 4 中标注有“1、2、3”序号的白色范围线内的部分。范围线 1 内的橙色小块实际地类是围网养殖区,但却被误分成其他地类,通过查看高分辨率的遥感影像发现此处的围网中有大

量成片的水生植物,因此改变了围网养殖区本来的光谱特征,出现了同谱异物的现象;范围线 2 内的浅蓝色小块是围网养殖区却被误分成自然湖泊水面了,这主要是因为较大的围网养殖区内,水面纹理均一与自然水体相似等原因造成的;范围 3 内的深蓝色小块是自然湖泊水面被误分为围网养殖区。在实际的应用中需要具体分析提取结果,以剔除误分区域。

(4)本研究将对应分析应用到 ETM+ 影像,探讨了围网养殖区提取的可行性,并与基于主成分分析等其他 3 种方法的提取结果进行了比较,证明了对应分析在太湖围网养殖区提取上的优越性,如果将该方法应用到 IKONS、QUICKBird 等高空间分辨率遥感影像上将有助于进一步提高围网养殖区的提取精度,但这有待于更进一步的实验。

(5)对应分析相比主成分分析等其他方法具有一定优势,而当前其应用研究仅局限于岩石分类、棉花种植面积估计、土地利用/覆被分类、土地变化检测等以及湖泊围网养殖区提取等少数几个方面,研究案例很少,其应用范围仍有待做进一步的深入研究。

参考文献 (References)

- [1] Lin Q G, Kong Y J, Chen L Q, *et al*. Effect of Pen Aquaculture on Community Structure and Species Diversity of Zoo Benthos in Gehu Lake [J]. *Chinese Journal of Applied & Environmental Biology*, 2005, 11(5): 566—570 [刘其根, 孔优佳, 陈立桥等. 网围养殖对太湖底栖动物群落组成及物种多样性的影响 [J]. 应用与环境生物学报, 2005, 11(5): 566—570]
- [2] Shen F, Kuang D B. Remote Sensing Investigation and Analysis for Water Resources Utilization and Its Dynamic Change of Representing Mid-or Small Lake Groups in Taihu Drainage Area [J]. *Journal of Remote Sensing*, 2003, 7(3): 221—226 [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 [J]. 遥感学报, 2003, 7(3): 221—226]
- [3] Yang Y B, Jiang N, Yin L Q, *et al*. RS-based Dynamic Monitoring of Lake Area and Enclosure Culture in East Taihu Lake [J]. *Journal of Lake Science*, 2005, 17(2): 133—138 [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 [J]. 湖泊科学, 2005, 17(2): 133—138]
- [4] Li X G, Jiang N, Yang Y B, *et al*. Remote Sensing Investigation and Survey of Lake Reclamation and Enclosure Aquaculture in Lake Taihu [J]. *Transactions of Oceanology and Limnology*, 2006, (1): 93—99 [李新国, 江南, 杨英宝等. 太湖围湖利用与网围养殖的遥感调查与研究 [J]. 海洋湖沼通报, 2006, (1): 93—99]
- [5] Li J J, He L H, Dai J F, *et al*. Extract Enclosure Culture in Lake Based on Remote Sensing Image Texture Information [J]. *Journal of Lake Science*, 2006, 18(4): 337—342 [李俊杰, 何隆华, 戴锦芳等. 基于遥感影像纹理信息的湖泊围网养殖区提取 [J]. 湖泊科学, 2006, 18(4): 337—342]
- [6] Cakir H I, Khorram S, Nelson S A C. Correspondence Analysis for Detecting Land Cover Change [J]. *Remote Sensing of Environment*, 2006, 102: 306—317.
- [7] Carr J R, Matanawi K. Correspondence Analysis for Principal Components Transformation of Multispectral and Hyperspectral Digital Images [J]. *Photogrammetric Engineering and Remote Sensing*, 1999, 65(8): 909—914
- [8] Xie J, Xia D S. Correspondence Analysis (CA) in TM Image Composition of Cotton Estimation [A]. *Proceeding of SPIE* [C]. 1999, 3868: 455—459
- [9] Liu Q S, Lin Q Z, Wang Z G, *et al*. Study of Correspondence Analysis on Landsat TM [J]. *Journal of Image and Graphics*, 1999, 4(2): 161—165 [刘庆生, 蔺启忠, 王志刚等. TM 图象对应分析研究 [J]. 中国图象图形学报, 1999, 4(2): 161—165]
- [10] Gao Y N, Zhang W C, Wang J, *et al*. LULC Classification of Landsat-7 ETM+ Image from Rugged Terrain Using TC, CA and SOFM Neural Network [A]. *IGARSS2007* [C]. Barcelona, Spain, 23-27 July 2007.
- [11] Hu L P. The Modern Statistics and Application of SAS System [M]. Beijing: Military Medical Science Press, 2000 [胡良平. 现代统计学与 SAS 应用 [M]. 北京: 军事医学科学出版社, 2000]
- [12] Zhou C H, Luo J C, Yang X M, *et al*. Geographical Understanding and Analyses of Remotely Sensed Imagery [M]. Beijing: Science Press, 2001 [周成虎, 骆剑承, 杨晓梅等. 遥感影像地学理解与分析 [M]. 北京: 科学出版社, 2001]
- [13] Bo H, Ma F L, Jiao L C. Research on Computation of GLCM of Image Texture [J]. *Acta Electronica Sinica*, 2006, (1): 155—158 [薄华, 马缚龙, 焦李成. 图像纹理的灰度共生矩阵计算问题的分析 [J]. 电子学报, 2006, (1): 155—158]
- [14] Yang C F, Tamura M. The New Method for Detecting Change of the Landscape the Differencing Image PCA Method and Its Application in the Liaohe River Delta [J]. *Acta Geographica Sinica*, 2004, 59(4): 592—598 [杨翠芬, 田村正行. 差分主成分分析法在辽河三角洲景观变化中的应用 [J]. 地理学报, 2004, 59(4): 592—598]

Extraction of Enclosure Culture in Gehu Lake Based on Correspondence Analysis

WANG Jing^{1,2}, GAO Jun-feng¹

(1 *Nanjing Institute of Geography and Linnology, CAS, Jiangsu Nanjing 210008 China*

2 *Graduate School, CAS, Beijing 100039 China*)

Abstract Lakes have been facing a number of disadvantageous eco-environmental issues due to the aggravation of human activities and the global climate change. So lake dynamic monitoring becomes more and more important. The extraction of lake enclosure culture area as an important content for lake dynamic monitoring has been highly valued. However, few researches on extracting lake enclosure culture area have been conducted, and little attention has been paid to the extraction method, and the extraction precisions using existed methods were unfavorable. In order to improve the extraction precision of lake enclosure culture area, a new method using correspondence analysis (CA) to extract lake enclosure culture area is proposed in this paper. The CA method relies on squared deviations between pixel values and their expected values (joint probabilities), and it explores the relationship between not only the image bands but also the pixels. Although CA has routinely been an accepted technique among ecologists and statisticians, its application to remote sensing is still relatively new. It is the goal of this study to detail CA methodology as a means of achieving highly accurate and efficient lake enclosure culture extraction.

The methodology mainly includes four steps as follows. Firstly, the data is preprocessing. Secondly, the preprocessed data is transformed using the CA algorithm. Thirdly, the texture analysis is applied to the CA first component. Finally, the enclosure culture area is extracted by decision tree according to the most superexcellent threshold value. In this study, the Gehu lake is selected as a case to validate this method based on the multispectral data of Landsat-7 ETM+ acquired on July 26, 2001, which is obtained in a standardized orthorectified GeoTIFF format downloaded from internet. The Gehu Lake (31°29'—31°42'N, 119°44'—119°53'E) is located in the middle and lower reaches of the Yangtze River, which is the second largest lake in Taihu basin, and its water area is 146.0 km². It is one of the earliest lakes which utilized the enclosure culture to cultivate freshwater fish or shellfish. And the classification results are evaluated using the precision analysis function in ENVI 4.2.

The results are as follows: (1) the extraction accuracy of enclosure culture area is 95.59%, overall accuracy is 91.14%, with overall Kappa value of 0.86, and (2) the area of Gehu lake enclosure culture is 111.76 km², 76.29% of total Gehu lake area. Then the results are compared with those with the other three methods, namely, the method based on principal component analysis (PCA), the method combining the PCA first two components with band 4 and the method based on original bands without CA or PCA analysis. And the results of the comparison test showed that: (1) the overall accuracies for the three methods above mentioned are 82.17%, 89.17% and 83.54% respectively, and (2) the classification accuracies are obvious less than that based on CA transformation.

To sum up, the CA is proved to be a new powerful multivariate analysis technique for extracting enclosure culture area. But there is still some commission and omission phenomenon, such as the higher commission phenomenon in the classification of lake enclosure culture area with the commission error of 13.41%, as well as the higher omission phenomenon in the classification of the other category with omission of 21.43%. So it is necessary to improve the method to reduce the commission and omission error. And the further research should be done to test the CA method on other lakes for extraction of enclosure culture, and should also be focus on whether it can be suitable for application to high spatial resolution imagery, such as SPOT, KONOS and QUICKBIRD.

Key words correspondence analysis; texture analysis; decision tree classification; enclosure culture in lake; gehu lake