

浙江省海宁市 TM 图像土地利用自动分类 精度评价方法的试验研究

朱光良

(浙江教育学院地理系 杭州 310012)

刘 南

(浙江省资源与环境信息系统重点实验室 杭州 310012)

摘 要 按照国家土管局土地利用现状分类标准,利用叠合的海宁市长安镇 1:1 万同时期土地详查电子地图的全域地面实况数据,在遥感图像处理常规分类方法的范畴内,采取最大程度利于提高分类精度的措施,所得的分类精度上限为 65%。研究表明,通常在试验区的小块局部中选择训练样本和选择评价分类精度的参考像元,对精度评价有显著影响,参考像元质量和数量的限制很可能是这些年来一些同类研究报道了较高分类精度结果的主要原因。

关键词 土地利用,遥感,地理信息系统,分类,精度评价

1 研究背景及问题的提出

现代中国面临的最严峻资源环境问题之一,是有效保护和合理开发利用土地资源;为此,首先必须有效地把握真实、准确和实时的土地利用现状数据。然而,由于种种原因,这个问题至今未能很好解决。经常性进行 80 年代至 90 年代初基于航测的土地详查是不现实的;利用卫星遥感进行调查方向正确,有巨大应用前景(例如今后应用米级遥感数据^[1]),但当前可经常利用的遥感数据源分辨率不够高,极大地限制了其土地分类调查的精度和实用性。

TM 数据是当前可能经常应用的主要遥感数据源。利用 TM 数据进行土地利用现状分类的精度不可能很高(特别在地块碎小的江南地区),对此,遥感界大多数同仁是有清醒认识的。

笔者在浙江省资源与环境信息系统重点实验室读研期间,利用该室建立的海宁市 1:1 万全市域基于 Arc/Info 格式的 GIS 数据库,作为 TM 数据进行土地利用现状自动分类精度研究的同时期的、全面而准确的地面实况,并采用能很好进行交叠的 ERDAS 遥感软件平台^[2]和 Arc/Info GIS 软件平台^[3],在常规

遥感分类方法范畴内,选取最优化的分类方法和措施,试图寻求利用 TM 数据进行土地利用现状自动分类能够达到的精度之上限。

2 试验区概况

海宁市位于浙江省北部、杭嘉湖平原的南缘,地势非常平坦,海拔在 4—8 m 之间。海宁市水系受杭嘉湖平原大水系控制,河港众多,水源丰富,属平原水网地区。经过长期的精耕细作,农田地块十分碎小。

本研究所采用的资料为包括海宁市在内的 1994 年 11 月 5 日的浙北 TM 影像,以及海宁市 1:1 万土地利用现状详查 GIS 数据库(1994 年秋变更版);研究工作试验区为海宁市长安镇。选取该镇的原因:第一,在海宁市各乡镇中,长安镇面积最大,总像元达 53155 个,有利于研究结果有较好的代表性;第二,在长安镇及周边地区可以找到分布比较均匀、质量较好的地面控制点(GCP);第三,长安镇境内有两条相交叉的、近直线的人工河,在 TM 影像上非常清楚,这有利于借以观察和分析 TM 影像与 GIS 矢量数据叠合的效果。

3 遥感数据几何精校正及其与土地详查矢量数据的叠合

3.1 遥感数据几何精校正

为了能有效利用1:1万土地详查矢量数据库作为地面实况来评价遥感数据的土地分类精度,首先要完成矢量数据库与遥感影像在同一坐标系下的几何配准。海宁市1:1万土地详查矢量数据库已采用北京大地坐标系(高斯-克吕格坐标 3° 带投影),因此,两种数据几何配准的主要工作,对遥感影像进行几何精校正,将其落实到高斯-克吕格坐标下。

为了提高几何精校正的精度,本研究采取扩大GCP选取范围的做法,即在长安镇四周选择了分布比较均匀的14个质量较高的控制点。这14个GCP一般在地形图上和TM影像上皆清楚、易于明断的道路、河流的交叉点。这里,“易于明断”主要指能相对好地确定GCP的图像坐标。

对14个GCP的图像坐标和高斯-克吕格坐标二者之间进行的多次拟合分析表明,采用其中9个GCP(仍然呈较均匀分布)的二次多项式拟合关系,其效果最佳。

$$U = a + bX + cY + dX^2 + eXY + fY^2 \quad (1)$$

$$V = a + bX + cY + dX^2 + eXY + fY^2 \quad (2)$$

这里 X 、 Y 是遥感影像上的图像坐标, U 、 V 是高斯-克吕格坐标。换言之,本次精校正采用基于9个分布于长安镇及其四周的GCP两套坐标间的二次多项式拟合关系,作为转换关系式,9个GCP的实际坐标与它在多项式拟合曲线上的对应值之间的差距(以像元为单位)的总均方差为

$$R = \left[\frac{1}{n} \sum_{i=1}^n (\Delta X_i^2 + \Delta Y_i^2) \right]^{1/2} = 0.22737 \quad (3)$$

这里 ΔX_i 、 ΔY_i 分别为第 i 个GCP坐标差距的 X 分量和 Y 分量, n 为GCP数目。长安镇TM图像像元实际平均尺度为28.5m,因此,上述用于精校正的9个GCP的均方差(RMS)相当于

$$RMS = 6.5m$$

3.2 几何精校正后的TM图像与1:1万详查数据库的叠合及精度分析

由于Arc/Info和ERDAS两个软件平台已有良好的连通性,本研究中经精校正到北京坐标系下的TM图像(ERDAS格式),很易于与Arc/Info格式的1:1万土地详查矢量数据进行叠合分析和处理。叠

合的结果呈现出非常好的吻合,主要的河渠、道路和其它明显地物,处处表现出相同的走势。

为了定量度量叠合精度,本研究设计了一种定量评价方法。其步骤是:(1)在试验区内选取不同方向、足够长度的线状地物;(2)利用线性或非线性回归方法,用一条矢量线去拟合线状地物在栅格影像上所表现出的跳跃、锯齿状像元串;(3)用这条拟合矢量线与1:1万矢量地图上的该线状地物拟合线进行比较,求取二者之间的最大、最小和平均偏离作为本研究叠合精度的定量度量。

本次实践中,长安镇两条贯穿全境的、大角度相交的人工河渠,即东北-西南向的左上塘河和西北-东南向的右上塘河,被选作叠合精度评价的线状地物。通过回归分析得出:左上塘河两条拟合中心线在图域内的最大距离为7.226m,最小距离为1.246m,平均距离为4.236m;右上塘河两条拟合中心线在图域内的最大距离为6.245m,最小距离为3.113m,平均距离为4.679m。

注意到左、右上塘河是大角度交叉的,这就避免了某一方向的偏差被掩盖未能测出的可能。由此可以得出结论,精校正后的TM图像与GIS矢量数据库相叠合的精度为4-5m(按平均值),或6-7m(按最大值),即叠合误差在1/4个像元或以内。

上述叠合精度评价看来是合理的。根据国家测绘局《1:1万比例尺地形图航空摄影测量内、外业规范》和《浙江省土地利用现状调查技术规范》的要求,1:1万土地详查矢量数据库的精度为5.8m(点位差);上面精校正GCP的均方差为6.5m,皆与上述叠合精度处于很相近的量级。当然,几何精校正的精度问题还有待进一步研究,上述均方差不一定能用来作为几何精校正精度的确切量度,但是,本文的主要目的不是作几何精度分析,而是在保证研究中的栅格与矢量两种数据能较好叠合的基础上,进行土地利用分类的精度评价。从这个角度,上述3种精度或误差的评价数据能很好地合拍,是有意义的。它表明,栅格与矢量数据的叠合误差在1/4个像元以内,1:1万土地详查矢量数据可以用来作为土地利用自动分类精度评价的地面实况数据。

4 分类方法及分类过程和结果

为了求得TM图像在江南平原水网农业区土地利用分类上所能达到的最高精度值,采用以下具体步骤:

4.1 预分类处理

为提高分类水平,分类前对研究区图像进行了各种增强处理,特别是通过多光谱空间的主成分变换(K-T变换)、缨帽变换(K-T变换)和非监督分类^[4](主要用 Isodata clustering),从多方面了解研究区 TM 数据的特点及其与地类实况的关联。

试验表明,对本例分类,还是 TM 原 7 个波段中的若干组合效果较好,其中以第 4、5 和 7 波段组合为最佳。这个结论与本室先期一些研究生的同类工作结论相同。本研究最后借以进行土地利用分类的特征是 TM₄、TM₅、TM₇。

4.2 分类体系的简化

TM 图像分类目标的粗细度对分类精度有很大的影响。为此,在土地详查分类体系的基础上,结合海宁市实际情况,本研究待分类的类别最后简化为 4 大类,即水域、园林地、建成区、耕地。

上述 4 大类所含细类如表 1。其中 116 地类呈水面状态,因而归入水域。城镇和道路甚难区分,合为一类有利于提高分类精度。其次,海宁市的园林地绝大部分为桑园,成分较为单一,也有利于提高分类精度。此外,本例图像为晚稻收割时图像,耕地类中占大部分的前两类也较为相近。总之,本研究把分类目标简化为上述 4 类,是十分有利于提高分类精度的措施。

表 1 分类关系表

Table 1 Relation table of classification

水域	园林地	建成区	耕地
116 变更养鱼	211 柑桔园	51 城镇	111 平坡水田
71 河流水面	212 其它果园	52 农村居民	14 旱地
74 坑塘水面	22 桑园	53 独立工厂	15 菜地
761 海涂	311 用材林	61 铁路	
77 沟渠	314 竹林	62 公路	
	36 苗圃	63 农村道路	
		78 水工建筑	

4.3 全域训练和以多重统计特征表达地类的分类方法

使分类精度尽可能攀高的另两个措施是:

全域训练

本研究为了提高训练水平,采取了一项极限措施,即进行全域训练;训练区没有地域限制,允许在

全图域内任何地点寻求最好的训练样本。这就大大地增加了训练较高质量样本的可能性。除此而外,本次训练还采取了两条措施。一条是利用增强或变换后的图像来帮助选择样本;另一条是借相应的土地详查矢量数据来确定训练样本。二者的目的都是为了提高样本训练的质量。

利用多重统计特征来表征地类

在通常的监督分类中,每一特定地类只给出一个统计特征(signature,表达该类在谱空间的范围和均值位置等)。但在本例中,4 个待定类是非常粗略的大类,各还含若干细类。因此,欲从一批训练样本总结出能统一表征某一大类的统计特征,是不可能的。事实上,在训练样本的过程中的确发现,同属于一个大类有着明显不同特点的、统计特征差异显著的样本群,一般是大类中亚类的反映。若将这些样本群合并起来作为大类的统一样本群并提取统一的 signature,那么,这个统计特征一方面不可能很好表征该大类的分布特点(在谱空间的位置、范围等);另一方面训练样本很可能呈远离正态分布的双峰、多峰分布,这将与下一步最大似然分类的基本要求相抵触。

由此缘故,本次研究允许每一大类含多重统计特征,即可以具有多个 signatures。在接下利用最大似然分类器进行像元归类时,像元所归的类,是 signature 或统计特征。最后,凡是归于某大类下属 signature 的像元,将计入该大类。

实践中,各大类所含统计特征或 signature 数分别为:水域 7 个,分别为钱塘江水域不同水深区域 3 个,河流 2 个,养鱼塘、其它池塘 2 个;城镇道路区 7 个,分别为城镇、农村居民点 3 个,铁路 1 个,公路 1 个,独立工矿 1 个,水工建筑 1 个;耕地类 5 个,分别为水田 2 个,旱地 2 个,菜地 1 个;园林地类 4 个,分别为桑园 2 个,竹林 1 个,其它果园 1 个。

最后,为了提高分类准确率,在最大似然分类时还利用 GIS 库给出的各大地类比例指定先验概率。

4.4 全域检验和分类精度估价

至此,本研究已采取一切可能的措施求取常规分类方法范畴内可能达到的最佳分类结果,现在的任务是,如何实事求是地对这个结果给以客观、可靠的度量或评价。

常规的分类精度评价方法是,选一组已知地面实况的参考像元,即检验区,把分类结果与这组参考像元作比较,由此得出分类精度。但实际工作中,与

训练区一样,检验区通常也是有限的;而图域内不同局部的地况可能会有不少差异。这样,同样的分类结果可能因检验区的位置和选择的不同而得到完全不同的评价或度量。为了避免这样的偏差,本研究将利用所掌握的全图域实况,对所有分类像元进行逐个分析评价,或者说全域检验,以得到真实可信的分类精度。

全域检验的具体做法是,先把1:1万土地详查矢量数据转换成一个栅格文件,该栅格文件在几何上与遥感图像严格匹配(坐标和像元大小完全相同)。栅格文件的像元值由GIS库地类属性确定,并转化为可与分类图像比较的数据形式。例如,在分类图像中水域代码为1,那么,由GIS库转化而来的该栅格文件中,原地类为水域的像元的属性值,就取为1。转换中,栅格文件的每个像元在GIS矢量图形上不一定对应一个单纯的地类,这时采用面积较优法。换言之,当栅格文件的像元所对应的矢量图形的地类为单一类别时,以矢量图形的地类所对应的属性值给像元赋值;当栅格文件的像元所对应的矢量图形的地类为两个或两个以上类别时,则按矢量图形中地类面积大的地类的属性值给像元赋值。

如此转换而来的栅格文件,给出了评价上面所得到的分类结果的全图域参考像元。这个栅格文件以自动化的方式与遥感图像分类文件进行逐行逐列的全域比较,凡相应像元属性值相同者为正,反之为误,如此给出全域检验下分类结果的精度评价(表2,表3)。

表2 试验区分类混淆矩阵

Table 2 Error matrix of test area

归类数据	参 考 数 据				行总数
	水域	耕地	园林地	建成区	
水 域	2329	1521	564	1007	5421
耕 地	1299	20151	3554	844	25848
园林地	429	2832	4123	780	8164
建成区	1683	2965	2923	6151	13722
列总数	5740	27469	11164	8782	53155

以上分类精度总体偏低,究其原因,除混合像元占有相当大的比例外,还有同物异谱和异物同谱对分类精度的影响。例如,长安镇镇所在地的大片建成区在分类中被分到水域一类,尽管采取很多办法进行改善,但这种异物同谱现象还是局部存在。

表3 分类精度汇报表

Table 3 Accuracy report of test area

类 别	参考总数	归类总数	正确归类数	生产者精度/%	用户精度/%
水 域	5740	5421	2329	40.57	42.96
耕 地	27469	25848	20151	73.36	77.96
园林地	11164	8164	4123	36.93	50.50
建成区	8782	13722	6151	70.04	44.83
总 计	53155	53155	32754	/	/
总精度/%			61.620		

上述结果表明,传统方法范围内利用TM图像进行江南平原水网农业区土地利用分类的精度,最高只能达到60%或略多。当然,不同季节、时间和地点的TM图像效果不会相同,分类精度会有所差别。但是,本次研究的TM图像是若干晴好天气中的图像,而11月份一般又是较有利于土地利用分类的季节,可以认为,在传统方法范围内再提高也有限,取分类精度上限为65%,是合理的。

5 地面实况区选择对分类精度的影响

5.1 选取不同局部的检验区对分类精度的影响

如果上面分析的结果是正确的,那就意味着江南产粮区TM图像土地利用分类实际工作的正确率可能还不到60%,因为实际应用时不可能进行全域训练。这样,以往这类分类结果报道精度偏高的一个重要原因,可能来自面积有限的检验区位置或选择的影响。为此,在保持训练、分类完全相同的条件下,不采用全局检验,而是用不同的小块局部检验区,进行分类精度评价。具体做法是,在长安镇南部,中部和北部分别选取一个像元数在4000至6000左右的区域作为检验区,对前述分类结果进行精度评价。表4至表6中给出了以这3个检验区进行评价的结果。

上述结果说明,同样的分类结果(即前面给出的最佳分类结果),由于选择不同的检验区,得到了不同,甚至是明显不同的精度评价。当然,这样的结果也是在意料之中的。同一图像内必然存在区域差别,不同局部取为检验区,结果自然不同。在本例中,长安镇南部因有若干并行的河渠道路穿过,造成较多的混合像元。因此南部检验的精度评价最低。

检验区选择对分类精度评价的影响,还因检验

区样本数量的大小而变化。一般而言,样本数愈少,

表 4 长安镇南部检验区的分类精度汇总表

Table 4 Accuracy report for south reference area

类 别	参考总数	归类总数	正确归类数	生产者精度/%	用户精度/%
水 域	545	636	190	34.86	29.87
耕 地	2197	1937	1389	63.22	71.71
园林地	675	680	164	24.30	24.12
建成区	821	985	312	38.00	31.68
总 计	4238	4238	2055	/	/
总精度/%			48.49		

表 5 长安镇中部检验区分类精度汇总表

Table 5 Accuracy report for middle reference area

类 别	参考总数	归类总数	正确归类数	生产者精度/%	用户精度/%
水 域	538	626	344	63.94	54.95
耕 地	2144	1836	1607	74.95	87.53
园林地	789	763	497	62.99	65.14
建成区	515	761	386	75.95	50.72
总 计	3986	3986	2834	/	/
总精度/%			71.10		

表 6 长安镇北部检验区分类精度汇总表

Table 6 Accuracy report for north reference area

类 别	参考总数	归类总数	正确归类数	生产者精度/%	用户精度/%
水 域	672	516	196	29.17	37.98
耕 地	2858	2889	2109	73.79	73.00
园林地	1401	861	482	34.40	55.98
建成区	1195	1860	881	73.72	47.37
总 计	6126	6126	3668	/	/
总精度/%			59.88		

偶然性愈大。例如,当样本数仅为 400 个上下时,精度评价变化于 40% 至 80% 之间。当检验区样本像元更少时,变动还有所增大。

还有一个问题,不少文献建议随机选取检验像元可消除这种偏差,这固然不错。然而,若地面实况有限,则随机选取也只能在该局部以内,仍然不能改善大局。本研究虽然未采用随机地选择单个参考像元的方法,但是这并不影响工作的客观性。大块的检验区较随意选取,所有像元无一例外地参与检验,

这与在该区内随机挑一部分同样客观,甚至更好。事实上,在一个局部小区内随机挑选参考像元,终究还是局限在该局部,其客观性是无法与全局检验相比拟的。

5.2 选取不同局部的训练区对分类精度的影响

训练区选择对分类精度的影响,前人已有论述^[5]。为了估价选择小块不同局部训练区对分类精度的影响,本步骤的工作是,在长安镇北部、中部和南部分别选取数千个像元大小的训练区,采用同样的训练、分类方法进行最大似然分类,然后进行全域检验方式的精度评价。表 7—9 给出了分类精度评价结果。

表 7 长安镇南部训练区的分类精度汇总表

Table 7 Accuracy report for south training site

类 别	参考总数	归类总数	正确归类数	生产者精度/%	用户精度/%
水 域	5740	6335	2021	35.21	31.90
耕 地	27469	25617	18663	67.94	72.85
园林地	11164	8245	3207	28.73	38.90
建成区	8782	12958	4077	46.42	31.46
总 计	53155	53155	27968	/	/
总精度/%			52.616		

表 8 长安镇中部训练区的分类精度汇总表

Table 8 Accuracy report for middle training site

类 别	参考总数	归类总数	正确归类数	生产者精度/%	用户精度/%
水 域	5740	5464	2241	39.04	41.01
耕 地	27469	25984	19853	72.27	76.40
园林地	11164	8067	3879	34.75	48.08
建成区	8782	13640	6014	68.48	44.09
总 计	53155	53155	31987	/	/
总精度/%			60.182		

首先,这 3 个结果低于表 2 和表 3 其原因是这 3 个结果与表 2 和表 3 的检验方法相同(全域检验),差异仅来自训练区的不同;而表 2 和表 3 是全域训练的结果,即在整个图域任意选取最有利训练样本的结果。因此,其分类精度应当高于受区域限制的训练结果。其次,长安镇中部的结果与表 2 和表 3 差异不大,是因为中部地区是样本较好的地带。事实上,前面“全域训练”所选的样本,多数来自中部。

南部训练区的结果最差,是由于那里混合像元较多的缘故。长安镇北部的结果居中。实际上,北部训练区选择时有意避开长安镇城镇区,如果在长安镇以及附近进行训练,效果也不会好。

表9 长安镇北部训练区的分类精度汇总表

Table 9 Accuracy report for north training site

类别	参考总数	归类总数	正确归类数	生产者精度/%	用户精度/%
水域	5740	6220	2145	37.38	34.47
耕地	27469	24238	18266	66.50	75.36
园林地	11164	8693	3801	34.05	43.73
建成区	8782	14004	5547	63.16	39.61
总计	53155	53155	29759	/	/
总精度/%			55.985		

当训练区尺度减小到1000个像元以下时,训练区选择的影响将加剧,低质量训练区将导致分类精度降到50%、甚至40%以下。事实上,在长安镇南部混合像元较多的区域,若训练样本限定在数百个像元内选取,常常会找不到较满意的signature,在那种情况下,强行继续分类工作,其效果可想而知。

6 结论

本研究通过利用海宁市长安镇1:1万同时期土地详查矢量数据库作为地面实况数据,对TM图像在江南水网地区进行土地利用现状监测的分类精度极限进行研究,得到以下结论:

第1,在遥感图像处理常规分类方法范畴内,采取最大程度利于提高分类精度的措施,所得的分类精度上限为65%。

第2,检验区和训练区对分类精度评价的影响都是明显的,且这两种区域的尺度愈小,影响愈显

著。参考像元质量和数量的限制很可能是这些年来一些同类研究报道了较高分类精度结果的主要原因。

第3,如果以全域训练的分类结果为标准,检验区的不同导致分类精度的评价可能偏高,也可能偏低;而训练区的不同只能导致分类精度评价的降低。

考虑到江南丘陵河谷地区的地块细碎程度通常不亚于平原水网地区,上述结论意味着,在中国江南地区,利用TM图像进行土地利用现状的自动分类,其精度离土地详查工作的技术要求还有相当距离。当然,利用模糊分类、神经网络和专家系统等方法,可以一定程度地改善分类精度;但是,卫星遥感能实用于土地详查及变更的根本途径,看来还是采用更高分辨率的数据源,例如不久将可及的米级商用卫星遥感数据。

参考文献 (References)

- 1 W. Stoney. A New Space Race Is On. *GIS WORLD* 1998, 11(3): 44-46.
- 2 ERDAS. Inc. ERDAS Field Guild, Version 7.5. July, 1991.
- 3 ESRI. Inc. PC ARC/INFO starter KIT, PC ARC/INFO Data Conversion Version 3.4D Plus. 1992.
- 4 Peng Wanglu. Electronic Processing of Remote Sensing Data and the Geographic Information System. Beijing, Press of Teachers University of Beijing. 1991, 122-143, 149-158. (In Chinese) [彭望碌编著. 遥感数据的计算机处理与地理信息系统. 北京:北京师范大学出版社, 1991, 122-143, 149-158.]
- 5 Barnsley, M. S. Barr. Inferring Urban Land Use from Satellite Sensor Images Using Kernel-Based Spatial Radassification. *Photogrammetric Engineering & Remote Sensing*, 1996 62(8): 949-958.

作者简介

朱光良,男,1963年6月出生。1983年7月在北京师范大学天文系获学士学位,1997年7月在杭州大学地理系(浙江省资源与环境信息系统重点实验室)获硕士学位。现主要从事自然地理、遥感与地理信息系统教学、科研工作,已发表论文6篇。

The Experimental Research to Accuracy Assessment of Land-use Automatical Classification in Haining City of Zhejiang Province

ZHU Guang-Liang

(*Zhejiang Educational College, Hangzhou 310012*)

LIU Nan

(*Zhejiang Provincial GIS Laboratory, Hangzhou 310012*)

Abstract It is still not clear that how high the upper limit of land-use classification accuracy can be reached using TM imagery in southeast plain areas of China featuring mass irrigation networks and small tracts. In this paper a result for the upper limit, 65%, is presented based on the GIS layer (as ground truth) of the detailed land-use inventory map at 1:10000 scale being well overlaid with a TM image of the same period, according to the national land-use classification system by taking all measures that may make the accuracy higher within the framework of conventional image classification methods. The results show that the selection of training samples and reference pixels for accuracy assessment in a limited part of test area has significant influence on accuracy assessment. The limitation of quality and quantity of reference pixels is probably the main reason why some early researches of this kind reported higher classification accuracy.

Key words Land-use, Remote sensing, GIS, Classification, Accuracy assessment