

常用矢量化软件的精度比较研究

——以 R2v、Arcgis、Mapgis 为例

钟燕林, 田永中*, 徐永进, 吴文戡, 王林松, 陈军, 石永明 (西南大学地理科学学院, 重庆 400715)

摘要 在对专业矢量化软件 R2v 和自带矢量化模块的非专业软件 Arcgis、Mapgis 充分研究的基础上, 应用这 3 种软件分别对地形图和土地利用现状图进行矢量化, 结合第二次土地大调查, 提出了 R2v 结合 Arcgis 的矢量化方法, 以期将矢量化精度提升到更高水平。

关键词 空间数据; 矢量化; Arcgis; R2v; Mapgis; 措施; 方法

中图分类号 S126 **文献标识码** A **文章编号** 0517-6611(2008)36-16227-02

Comparative Study on the Accuracy of Common Vectorization Softwares

ZHONG Yan-lin et al (School of Geographic Sciences, Southwest University, Chongqing 400715)

Abstract Based on the full study on the professional vectorization software R2v and unprofessional software Arcgis and Mapgis with vectorization module, the topographic map and land use actuality map were digitized by using these three kinds of softwares. Combining with the second land use investigation, the vectorization method of combining R2v with Arcgis was put forward, so as to reach higher vectorization accuracy.

Key words Spatial data; Vectorization; Arcgis; R2v; Mapgis; Measures; Methods

空间数据是地理信息系统(GIS)的操作对象, 是 GIS 最基本和最重要的组成部分, 其数据质量的好坏, 直接影响空间数据建库的质量优劣, 甚至间接影响到国民经济和国防建设的发展^[1], 而地图的矢量数字化作为获取地理空间信息的重要手段是建立 GIS 数据库的基础工作之一, 也是建立 GIS 的“瓶颈”^[2]。因此, 研究空间数据的矢量化具有重要意义。

在空间数据矢量化研究中, 很多专家取得了丰硕的成果。胡晋山等于 2005 年应用 Epscan、GeoScan 这 2 种软件进行矢量化, 验证了当扫描分辨率大于 300 DPI 时, 矢量化精度提高不明显^[3]; 包忠聪等总结出 R2v 结合 ArcInfo 的方法可以更高效地应用于数字化地形图^[4]。笔者在前人的基础上, 用 R2v、Arcgis、Mapgis 这 3 种软件分别对地形图和土地利用现状图进行矢量化, 发现 R2v 结合 Arcgis 的矢量化方法能更好地提高矢量化的精度和速度。

1 矢量化方法与工具

1.1 数字化地图方法 矢量化是将栅格图像转换为矢量图像的过程, 现在一般都采用人机交互与自动跟踪相结合的方法在屏幕上完成栅格式数字图的矢量化, 故又称屏幕数字化。将纸质地图转成数字化地图有 3 种方法: 一是用数字化仪或计算机屏幕进行矢量化, 也称为手扶跟踪数字化; 二是扫描; 三是对地图中线状要素进行提取。经过多年的比较研究发现, 扫描矢量化比手扶跟踪数字化更好, 能够更好地避免由于疲劳或粗心引起的人为错误。扫描文件必须经过矢量化(即将栅格线转化为矢量线)的跟踪描绘过程, 才算完成数字化^[5]。随着地理信息系统产业的蓬勃发展, 基础地理信息质量控制日益受到广泛重视, 作为基础地理信息数据来源重要手段的地图扫描数字化, 由于速度快、精度高, 已成为地图数字化的主流^[6]。由此而产生的矢量化软件有很多, 市场上常见的矢量化专业软件有德国 Softelec 公司的 VPStudio, Able 公司的 R2v、澳大利亚的 Provec、德国的 VPmax、清华三维及武汉大学自主开发的软件。此外还有具有矢量化功能的

相关 GIS 制图软件, 如 Arcgis、Envi、Mapgis、SuperMap 等^[7]。

1.2 软件开发平台及工具简介 矢量化软件的种类虽多, 但矢量化的能力却参差不齐, 笔者在经过市场抽样调查统计并结合美国《CADALYST》杂志等评测的基础上, 对专业矢量化软件 R2v 和具有矢量化功能的 GIS 制图软件 Arcgis、Mapgis 的矢量化功能进行研究。

1.2.1 R2v 简介。Able 公司的 R2v 是一款销量名列世界前三甲的高级光栅图矢量化软件系统, 它价格适中, 能独立运行, 是地图和 GIS 领域非常好的工具。它可以处理多种格式的光栅(扫描)图像, 输出文件格式多种, 是一个可以用扫描光栅图像为背景的矢量编辑工具, 且具有良好的适应性、高精度度。

1.2.2 Mapgis 简介。Mapgis 是中国武汉地质大学开发的一套集数据处理、编辑、校正、矢量化、数据转换于一身的地理信息系统, 是中国国家科委唯一推荐的国产 GIS 优选平台。该系统是一个工具型地理信息系统, 具备完善的数据采集、处理、输出、建库、检索、分析等功能和良好的二次开发能力。

1.2.3 Arcgis 简介。Arcgis 是美国 ESRI 公司 20 世纪 70 年代后期推出的地理信息系统, 早期被称为 ArcInfo, 主要运行在 UNIX 系统下。20 世纪 90 年代推出全新的基于桌面的地理信息系统, 称为 Arcgis, 不仅可以实现地图矢量化, 还能进行制图、数据管理、地理分析、数据编辑等功能, 是一个集成的、可伸缩的系统。

2 矢量化分析

2.1 矢量化过程 该研究的 3 种软件矢量化过程为: 原图准备→扫描原图→打开图像→转换图像格式→新建图层→二值化处理→选择矢量化方法(交互式跟踪)→分层矢量化→保存和导出文件。

2.2 试验控制指标 鉴于试验用的图件空间要素较稀疏, 采用了扫描分辨率为 300 DPI 的某区纸质地形图(等高线)、土地利用现状图(1:10 000)(聚酯薄膜图), 用 3 种软件分别矢量化。试验误差控制要求在放大条件下进行数字化采集, 将数字化误差控制在 1 个像素以内。其余的控制指标如下: 采点密度 0.2 mm, 接合距离 0.02 mm, 悬挂距离 0.007 mm, 细化距离 0.007 mm, 纹理距离 0.01 mm。

作者简介 钟燕林(1984-), 女, 四川遂宁人, 硕士研究生, 研究方向: 国土资源管理与区域开发。* 通讯作者, 副教授。

收稿日期 2008-10-20

2.3 试验数据分析 试验所用地形图为纸质图,土地利用 现状图为聚酯薄膜图,试验数据如表 1~3 所示。

表 1 R2v 矢量化地形图和土地利用现状图点坐标差分布

Table 1 The distribution of point coordinate differences in the topographic map and land use actuality map digitized by R2v

坐标差分 布区间//mm Distribution interval of coordinate difference	R2v 矢量化地形图点坐标差分布 Distribution of point coordinate differences in the topographic map digitized by R2v				R2v 矢量化土地利用现状图点坐标差分布 Distribution of point coordinate differences in the land use actuality map digitized by R2v			
	Δx		Δy		Δx		Δy	
	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number
0~0.05	174	91	148	129	167	133	158	149
0.05~0.10	55	65	68	60	48	38	41	45
0.10~0.15	38	30	25	28	42	24	51	30
0.15~0.20	26	16	11	26	21	23	14	6
>0.20	3	2	4	1	2	2	2	4
总计 Total	296	204	256	244	280	220	266	234

表 2 Arcgis 矢量化地形图和土地利用现状图点坐标差分布

Table 2 The distribution of the point coordinate differences in the topographic map and land use actuality map digitized by Arcgis

坐标差分 布区间//mm Distribution interval of coordinate difference	Arcgis 矢量化地形图点坐标差分布 Distribution of point coordinate difference in the topographic map digitized by Arcgis				Arcgis 矢量化土地利用现状图点坐标差分布 Distribution of point coordinate difference in land use actuality map digitized by Arcgis			
	Δx		Δy		Δx		Δy	
	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number
0~0.05	165	123	127	139	146	129	133	116
0.05~0.10	55	45	59	59	70	47	62	75
0.10~0.15	34	28	64	14	36	26	75	16
0.15~0.20	26	19	17	12	28	14	5	13
>0.20	2	3	2	7	1	3	2	3
总计 Total	282	218	269	231	281	219	277	223

表 3 Mapgis 矢量化地形图和土地利用现状图点坐标差分布

Table 3 The distribution of point coordinate difference in the topographic map and land use actuality map digitized by Mapgis

坐标差分 布区间//mm Distribution interval of coordinate difference	Mapgis 矢量化地形图点坐标差分布 Distribution of point coordinate difference in the topographic map digitized by Mapgis				Mapgis 矢量化土地利用现状图点坐标差分布 Distribution of point coordinate difference in land use actuality map digitized by Mapgis			
	Δx		Δy		Δx		Δy	
	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number	(+) 的个数 (+) number	(-) 的个数 (-) number
0~0.05	149	124	126	126	158	124	135	145
0.05~0.10	46	49	56	62	58	38	35	45
0.10~0.15	33	29	67	20	25	46	56	45
0.15~0.20	36	28	17	15	24	14	18	10
>0.20	3	3	2	9	12	1	6	5
总计 Total	267	233	268	232	277	223	250	250

2.4 评价矢量化的标准 土地利用现状图精度的判断标准是平面位置监测点位中误差 M_s ;对于地形图,该研究不考虑原图的测量误差和高程误差,判断其矢量化过程中精度的标准 M 有 2 个:平面位置监测点位中误差 M_s 和高程精度 mb 。

$$S = \pm \sqrt{(x_i - x)^2 + (y_i - y)^2} \tag{1}$$

$$M_s = \pm \sqrt{\frac{\sum S_i^2}{n}} \tag{2}$$

$$m_b = \pm \sqrt{\frac{(H_i - h_i)^2}{n}} \tag{3}$$

式(1)中, S 为检测点坐标差, i 为平面检测点个数;式(2)中 M_s 为点位中误差;式(3)为高程检测的中误差计算公

式, H_i 为检测点的实测高程, h_i 为检测点的图上注记高程, n 为高程检测点个数^[8]。

2.5 试验结果 根据表 1~3 的试验结果,用公式(1)~(3)进行运算得到:土地利用现状图中 R2v、Arcgis、Mapgis 3 种软件的点位中位差分别为 0.116、0.121、0.125;地形图中 R2v、Arcgis、Mapgis 3 种软件的点位中位差分别为 0.115、0.131、0.133,高程精度分别为 0.224、0.268、0.283。

3 结论与讨论

从精度上来说, R2v 矢量化精度最高,中位点误差分别为 0.115 和 0.116;其次是 Arcgis,中位点误差分别为 0.131 和

2.2 不规则三角网 TIN 不规则三角网是通过不规则分布的数据点生成连续三角面来逼近地形表面。它是根据区域有限个点集将区域划分相连的三角面网络,区域中任意点落在三角面的顶点、边上或三角形内。如果点不在顶点上,该点的高程值通常通过线性插值的方法得到(在边上的点用边的 2 个顶点的高程,在三角形内的点则用 3 个顶点的高程)^[2]。ArcGIS Engine 也为我们提供了现成的接口函数 ISurface 中的函数 InterpolateShape(IGeometry * pShape, IGeometry * * pOutShape, VARIANT * pStepSize)或 GetProfile(IGeometry * pShape, IGeometry * * pOutShape, VARIANT * pStepSize)来实现。其中 pStepSize 为插值步长。插值步长对栅格 DEM 与 TIN 来说,其重要性有所不同。对 TIN 来说,插值步长不是很重要,在 ArcGIS Engine 的默认处理中,采用三角形顶点及三角边进行插值,结果已经能反映 TIN 数据的地形信息。同理,实际上,最终河道断面数据的精度还是由原始采样点的精度、密度与分布决定的。

2.3 等高线 等高线是地面上高程相等的点所连成的闭合曲线。原则上讲,等高线都应该是闭合的,但是在实际的出图中,等高线被打断成若干个目标,是多个独立的几何对象。利用河道断面线与等高线的相交来获取断面数据,主要基于的是河道断面线与等高线的交点,而不是指定点的高程值。通过 ArcGIS Engine 中 ITopologicalOperator 的函数 Intersect(IGeometry * Other, esriGeometryDimension resultDimension, IGeometry * * intersection)来实现,得到的是交点处的 x 、 y 坐标,同时获取此处的等高线高程值。最后通过断面点系列有序排列形成河道断面图。

利用河道栅格 DEM、不规则三角网 TIN 以及河道等高线数据都可以插值生成河道断面数据。但是不同方式下生成的河道断面数据是有差别的。相比较而言,不规则三角网 TIN 与等高线插值生成的河道断面数据更接近(图 3),栅格 DEM 方式获取的河道断面数据与上述 2 种方式相比,相差较

大。这个比较是基于同一时期同一比例尺数据而言的。

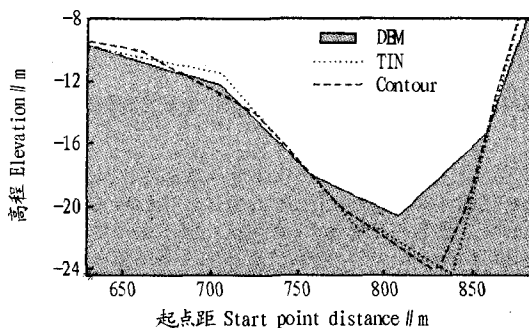


图3 栅格 DEM、TIN、等高线获取断面高程数据

Fig.3 The obtained elevation data based on grid DEM, TIN and contour

3 结论

该研究的河道断面生成与数据获取,解决了传统河床演变断面分析必须立足固定断面的局限。以河道地形测量数据为基础,利用 GIS 空间分析,可根据需求任意布设河道断面,河床演变的断面分析不再受到实测固定断面资料的局限。同时,利用河道栅格 DEM、不规则三角网 TIN 以及河道等高线数据都可以插值生成河道断面数据。但是不同方式下生成的河道断面数据是有差别的,可结合具体应用选择适合于分析的断面数据获取方式,从而提高数据的质量。

参考文献

- [1] 钱宁,张仁,周志德.河床演变学[M].北京:科学出版社,1987.
- [2] 车德福. GIS 中数字高程模型的建立与应用研究[D]. 南宁:广西大学,2004.
- [3] 杨明,梁国亭,赖瑞勋,等.基于曲面内插技术的数字地形生成原理及应用[J]. 水利学报,2007,38(2):221-225.
- [4] DENG X Z, ZHAN J Y, SU H B, et al. Simulation and analysis of land system structure changes in Huang-Huai-Hai plain area[J]. Agricultural Science & Technology, 2007, 8(3-4):45-52.
- [5] 王耀军,朱长青,王志伟.基于 Coons 曲面的规则格网 DEM 表面模型[J]. 测绘学报,2008,37(2):217-222.

(上接第 16228 页)

0.121; Mapgis 的中位点误差分别为 0.133 和 0.125。

从适用环境上来说, R2v 对地形图和土地利用现状图同时适用,但 R2v 中没有手形漫游工具,不能为空间对象赋予属性,自动跟踪时线划不够圆润。

Arcgis 在矢量化土地利用现状图和地形图时的速度略高于 Mapgis,两者不仅具有矢量化模块,还具有数据编辑、输出、建库、检索、分析等功能,使用范围相对比 R2v 更广。同时, Arcgis 功能强大,矢量化运行平台庞大,对硬件的要求比 Mapgis 对硬件的要求高,价格比 Mapgis 昂贵,对于成本预算高、精度要求较高的项目很适合; Mapgis 是国产软件,对硬件的要求低,性价比远高于国外软件,不需要庞大运行平台,可靠性和实用性强,对于低成本且不要求绝对精度的项目非常合适。

本着提高矢量化速度和保证矢量化精度的原则,结合第二次土地调查的实践,笔者在前人基础上总结得出:在矢量化过程中使用 R2v 进行半自动化跟踪数字化并保存为 R2v 工程文件,在此基础上,若资金和硬件配备充足,将数字化的

图层以 shp 格式导出,用 Arcgis 进行后期编辑,完成属性数据输入的方法不仅可以提高矢量化效率,还能降低人为因素产生的误差,能更快更高质量地实现地图数字化。若资金和硬件配备不够完善,必须在成本较低的情况下完成数字化,可以采用 R2v 结合 Mapgis 的方法进行。这 2 种方法不仅能更好地利用 R2v 强大的数字化优势,更能充分发挥 Arcgis 和 Mapgis 强大的后期编辑功能,真正做到提高生产效率,保证数字化产品的质量。

参考文献

- [1] 黄杏元,马劲松.地理信息系统概论[M].北京:高等教育出版社,2003.
- [2] 蓝悦明,杨晓梅.手工数字化地图误差的分布检验[J]. 测绘通报,2003(4).
- [3] 胡晋山,康建荣.地图扫描数字化误差分析及控制[J]. 测绘科学,2005,30(2):108-109.
- [4] 包忠聪,岳彩荣,何超.基于 R2v 与 Arcinfo 的地形图数字化方法[J]. 林业调查规划,2006,31(5):29-32.
- [5] CHANG K T. Introduction to geographic information systems[Z]. The McGraw-Hill Companies Inc, 2003:64.
- [6] 胡晋山,马明栋,李博.地形图扫描数字化质量控制[J]. 测绘通报,2004(12):53-55.
- [7] 徐祖顺. gis 入门与提高[M]. 重庆:重庆大学出版社,2001.