

Internet 环境下土地空间数据组织与表达

张立亭^{1,2}, 周世健², 祝国瑞¹

(1. 武汉大学资源与环境科学学院, 武汉 430079; 2. 东华理工学院测量系, 抚州 344000)

摘要: 为了较好地实现网上土地空间数据共享, 提高土地空间数据网上传输速度, 该文研究了土地空间数据组织与表达方法, 分析了土地空间数据的特点和用户需求, 并从数据分层表达和数据格式方面进行研究, 通过实验, 得出了一套客观的数据。经过对数据及其图件的对比, 分析表明应根据用户的需要选择数据存储格式, 当用户只浏览图件而不需要交互操作时, 尽量采用 jpg 格式或 gif 格式存储图件; 当需要为用户提供交互式操作图件时, 可以用 Map Info 的 tab 格式或 ESR I 的 shp 格式存储图件。

关键词: 土地信息; 数据分层; 数据组织

中图分类号: F311

文献标识码: B

文章编号: 1002-6819(2005)10-0174-04

张立亭, 周世健, 祝国瑞 Internet 环境下土地空间数据组织与表达[J]. 农业工程学报, 2005, 21(10): 174-177.

Zhang Liting, Zhou Shijian, Zhu Guorui Land spatial data organization and expression based on internet[J]. Transactions of the CSAE, 2005, 21(10): 174-177. (in Chinese with English abstract)

0 引言

土地作为人类生存和发展的物质基础, 表明了人类的活动离不开土地。国外研究表明: 大约有 80% 的政府决策与土地有关, 而且, 许多土地利用决策欠科学在很大程度上也是缘于土地的地理数据资料不完整^[8], 因此, 最大限度地共享土地信息是社会发展的必然要求。Internet/Intranet 技术的出现与应用的普及, 为土地空间数据的共享提供了可能。但是, 由于土地空间数据的数据量巨大, 造成网上传输的困难, 尽管网络技术和 WebGIS 技术在不断发展, 带宽、网速都有了很大的提高, 但是仍然满足不了土地空间海量数据的快速传输的要求。况且, WebGIS 尚处于发展初期, 与传统的 GIS 相比, 大多数只能提供简单的空间信息发布和查询功能^[7], 同时, 由于目前 WebGIS 仍然存在着诸多不足^[9], 诸如: 支持单一数据源而非多数据源的获取, 因而不能满足日益增长的信息交互的需求; 数据集中非分布式存放与管理不利于信息的共享与更新; 缺乏丰富的空间信息表现手法以及与网络资源特别是借助浏览器表达的能力等等。这些问题也为土地空间数据的共享带来了困难。尽管, 国家制订相关的数据交换标准^[3, 6], 但是在共享效果方面仍然存在许多需要解决的问题。为此, 要解决这些问题, 实现土地信息资源的共享, 就必须对土地空间数据进行科学合理地组织与表达。而且, 对于土地信息系统(LIS)应用来说, 用户关心的不仅是文字内容, 而且还关心动态图形或数据, 不同的用户对空间数据也有不同的需求重点和兴趣。如果不加区别地将土地空间数据安全要素图件传输给用户, 势必造成大量的数据冗余, 不能很好地保证网络快速传输的速度和用户的用图要求。为了既要满足用户的用图要求, 又要保证网络传输速度, 本文将针对土地空间数据的特点和用户的要求, 仅就土地空间数据的分层组织进行探讨。

1 土地空间数据的特点与用户分析

1.1 土地空间数据的特点

土地信息是指表征土地系统诸要素的数量、质量、分布特征、相互联系和变化规律的数字、文字、图像和图形等的总称^[1]。它反映了有关土地实体的性质、特征和变化状态, 是对表达土地特征与现象之间关系的土地数据的解释。在土地信息系统中, 土地数据可分为三种类型: 空间特征数据(定位数据)、时间属性数据(尺度数据)和专题属性数据(非定位数据)^[1]。土地信息属于空间信息, 其位置的识别是与数据联系在一起的, 这是土地信息区别于其他类型信息的最显著的标志。空间数据是有关空间位置、专题特征及时间信息的符号记录。因此土地空间数据是反映有关土地信息的空间位置。

2 用户分析

网上土地信息用户大致可以分为政府和分散的个人用户两类。对于政府来说, 各部门需要通过网络传输大量的土地空间信息和属性信息, 用于政府工作决策; 土地管理部门可以通过网络获取相关各政府部门的资料和图件, 及时更新土地数据, 用于土地利用规划决策; 而个人用户则可从网上浏览到所需要的有关土地方面的空间和属性信息, 以及土地政策等, 如土地市场、土地拍卖等相关的空间位置信息等。针对不同的用户需求, 应当有不同的图形数据组织表达方式, 这样可以减少网络传输的数据冗余, 保证传输速度。

1.2 空间数据存储格式

由于不同 GIS 厂商对 GIS 空间数据的理解和定义不同, 其空间数据的存储格式也不尽相同, 因此, 出现了多种空间数据格式。在这些众多的数据格式中, 由于数据结构不同, 数据量也必然不同。目前, 通用 GIS 软件空间数据矢量交换格式有: ESR I E00 格式(e00); ESR I Shape 格式(shp); Map Info 交换格式(mif/tab); AutoCAD 的 dxf 格式(dwg); 地球空间数据交换格式(vct)。通用空间栅格数据格式有 tif、gif、jpg、bmp 等^[4]。为了更好地适合网上空间数据传输的需要, 我们既要保证网上的传输速度, 又要保证土地空间数据的共享, 因此, 在存储土地空间数据时, 要有便于数据交换通用的数据格式。本文在研究土地信息的分层处理时, 将依据以上通用数据格式进行研究。

3 土地空间数据的分层

土地空间数据的表示方法有矢量数据和栅格数据, 前者位置明显, 属性隐含; 而后者属性明显, 位置隐含。他们各有优势, 都是有效地表示土地信息的方法。但是, 不管是矢量表示还是栅

收稿日期: 2005-04-21 修订日期: 2005-07-04

基金项目: 国家自然科学基金项目资助(40171082); 江西省教育厅科技项目资助(赣教技[2005]208), 江西省数字国土重点实验室开放基金资助(DLLJ200509)

作者简介: 张立亭, 男(1966-), 武汉大学资源与环境科学学院, 博士生, 东华理工学院测量系副教授, 主要从事地图制图学与地理信息工程专业的理论与应用研究。江西抚州 东华理工学院测量系, 344000。 Email: ltzhang@ecit.edu.cn

格表示, 土地空间数据都是海量的。为了适于网上土地信息的传输的要求, 我们对土地空间数据进行分层组织。具体方法如下。

3.1 矢量数据的分层

土地信息的矢量图在数字化过程中, 将数字化以后的土地信息图按不同的性质和类型分成多个层层叠加的透明层, 每一透明层称为一个图层。数字图件是按图层来组织的, 不同的图层包含了整个地图的不同的主题内容。如土地利用现状图, 按类分层, 第一层包含行政区域图, 第二层包含耕地, 第三层包含园地等等, 层层叠加, 即构成一幅完整的土地利用现状图(如图 1)。对于各种图件来说, 按图层组织更具有可视性、可编辑性、可选择性和可自动标注性。将按主题内容分层的土地信息图件, 存储在不同的服务器数据库中。当用户通过网络向服务器发出请求时, 服务器可根据用户的要求将用户需要的主题层发回给用户, 如果用户需要多个主题层的内容, 服务器亦可将多个图层进行叠加传输给用户。从而减少不必要的数据传输, 减轻服务器和网络的负担。

3.2 栅格数据的分层

栅格数据结构尽管具有数据量大的缺点, 但是其结构简单, 空间数据的叠加与组合十分方便, 便于进行空间分析, 因此, 仍然可用来表示土地空间数据。可用栅格数据表示土地信息的图件有很多种, 如土地利用现状图、土地定级估价图、土地利用规划图等。

在用栅格数据结构表示土地空间信息时, 为适合网上数据传输的需要, 减少不必要的数据传输量, 可对土地信息图件作分层管理, 即将彩色地图扫描成栅格数据格式存储, 或者将矢量图转换成栅格数据格式存储, 在存储栅格数据的时候, 我们将分别按地图的颜色层扫描彩色地图或由矢量图直接转存为栅格图, 按统一的格式存储数据。

4 实例及分析

根据上面的方法, 我们将深圳龙港区土地利用图(如图 1)按矢量格式和栅格格式进行分层。在土地利用现状图中, 不同的颜色表示不同用途的土地类型。分层时根据不同的颜色按照土地利用分类分别提取耕地、园地、林地、疏林地、交通用地、城镇居民点及交通、农村居民点用地、灌木林、水域和裸岩石砾地等。如图 2 是, 分层时把城镇及交通放在一层, 是想比较线带面图形和交通纯线状图形在数据量上的区别。栅格数据以 .jpg 格式、.gif 格式、.tif 格式、.png 格式和 .bmp 格式存储; 矢量数据以 .cdt、.dwg、.tab、.shp、.e00 等格式存储, 得出的数据量如表 1、表 2 所示。

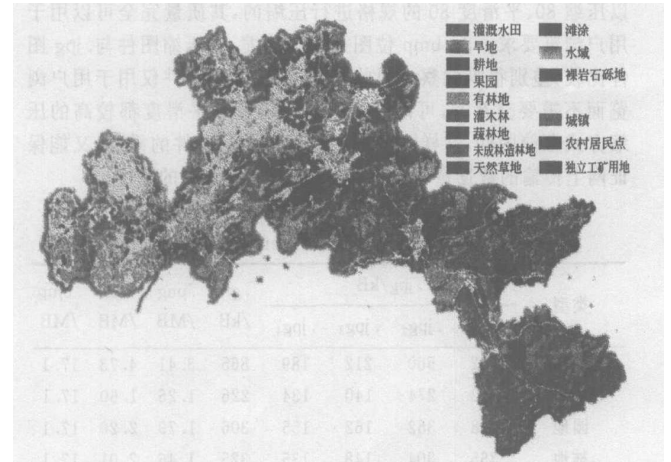


图 1 龙港区土地利用图

Fig 1 Land use map of Longgang district

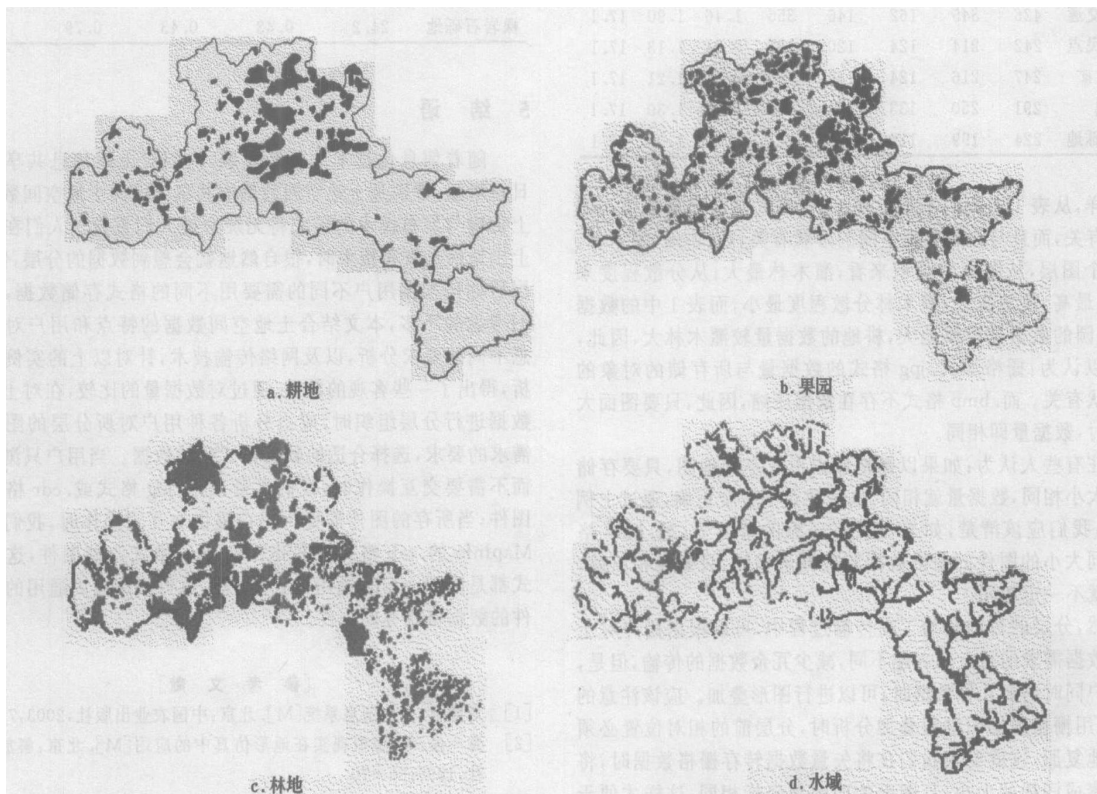


图 2 土地利用分层图

Fig 2 Delamination maps of different land use patterns

表 1 中, .jpg, .bmp 两种格式, 都是在 cdr 的基础上转存的栅格数据. jpg 格式和 bmp 格式都是 RGB 色彩(24 位元)^[2], 转存图件大小都是以宽 3000 像素、高 2000 像素规格存储的, 解析度 300dpi. 其中, .jpg 格式又分别以四种压缩方式存储, 即压缩 10、平滑度 10(记 .jpg₁); 压缩 10、平滑度 80(记 .jpg₂); 压缩 80、平滑度 10(记 .jpg₃); 压缩 80、平滑度 80(记 .jpg₄). 从数据量的大小可以看出, 在 jpg 格式的压缩过程中, 压缩值和平滑度都与数据量都有关系. 尽管 jpg 使用一种有损压缩算法, 无损压缩算法能在解压后准确再现压缩前的图象, 而有损压缩则牺牲了一部分的图象信息来达到较高的压缩率, 但是这种损失很小以至于人们很难察觉. 从图件的质量来说, 完全能够满足用户用图的需要, 以上 jpg 的四种格式存储的图件的效果差异不大, 图 2 就是以压缩 80、平滑度 80 的规格进行压缩的, 其质量完全可以用于用户浏览要求. 而 bmp 位图是无损压缩, 其压缩图件与 jpg 图件比较, 差别很难察觉. 因此, 当土地信息的图件仅用于用户浏览而不需要交互时, 可以尽量选取压缩值和平滑度都较高的压缩方式存储图件, 这样既可以保证用户浏览图件的需要, 又能保证网上传输的数据量尽量小, 从而保证网络传输的速度.

表 1 不同栅格格式的数据量

类型	.jpg/kB				.gif/kB	.png/MB	.tif/MB	.bmp/MB
	.jpg ₁	.jpg ₂	.jpg ₃	.jpg ₄				
土地利用	702	560	212	189	865	3.41	4.73	17.1
耕地	312	274	140	134	226	1.26	1.60	17.1
园地	418	362	162	155	306	1.79	2.20	17.1
林地	365	304	148	135	325	1.46	2.01	17.1
灌木林	296	254	132	125	390	1.31	1.85	17.1
疏林地	299	258	135	129	203	1.14	1.50	17.1
交通	300	256	138	129	182	1.01	1.29	17.1
城镇及交通	426	345	162	146	355	1.46	1.90	17.1
农村居民点	242	214	124	120	137	0.84	1.13	17.1
独立工矿	247	216	124	119	143	0.91	1.21	17.1
水域	291	250	133	124	165	1.01	1.30	17.1
裸岩石砾地	224	199	119	115	121	0.79	1.06	17.1

同样, 从表 1 中我们不难发现, .jpg 格式的数据量不仅与图面大小有关, 而且与存储数据空间的形状有关, 如: 耕地、灌木林、果园三个图层, 从图斑的面积来看, 灌木林最大; 从分散程度来看, 果园最高, 耕地次之, 灌木林分散程度最小, 而表 1 中的数据量显示, 果园的数据量较耕地大, 耕地的数据量较灌木林大, 因此, 我们可以认为: 栅格形式 .jpg 格式的数据量与所存储的对象的形状有关. 而 bmp 格式不存在数据压缩, 因此, 只要图面大小相同时, 数据量即相同.

以往有些人认为: 如果以栅格数据形式存储地图, 只要存储图件的大小相同, 数据量就相同, 这种观点不完全正确, 通过上例的比较, 我们应该清楚, 如果在没有压缩的情况下(如 bmp 格式), 相同大小的图件的栅格数据其数据量相同, 只要存在压缩, 数据量就不一定相同.

当然, 分层的目的是为了在传输过程中, 可以根据用户对土地空间数据需求的重点和兴趣不同, 减少冗余数据的传输, 但是, 如果用户同时需要几种数据时, 可以进行图形叠加. 需要注意的是, 在利用栅格数据图进行叠加分析时, 分层前的相对位置必须能很好地复原, 这就要求我们在将矢量数据转存栅格数据时, 将图面设置成同样的大小, 且图面的西南角坐标相同, 这样才便于栅格数据的叠加.

表 2 中, 列出了 5 种较为通用的矢量格式的数据量, 这些数据量大小也存在较大的区别, 如 cdr 格式的数据量明显地较其

他格式的数据量小, 是因为, .cdr 格式不带属性格式, 而其余的三种矢量数据都带有属性格式, 从数据结构来说, 其复杂程度远远高于 cdr 格式, 因此, 如果用户不需要属性数据时, 也可以用 cdr 格式传输. 同时, ESR I 的 e00 格式和 ESR I 的 .shp 格式的数据量明显大于其他格式的数据量, 主要是由于其数据结构不同^[5]. 再说, 从土地的利用类型与数据量的关系来说, 图中土地利用图、城镇居民点图层的数据量较其他用地类型的数据量大, 原因是这两种图中, 属性数据较其他类型复杂. 至于, 图中图面的形状与数据量没有太多的关系, 如, 同是以线状形式为主的交通用地层与水域用地层, 交通用地除了要记录线路的名称外, 线路的等级、路面宽度、路面材料、路基的宽度等也是必须标明的, 故其属性当然比水域用地的属性复杂, 数据量也较水域用地的数据量大. 因此, 可以认为, 矢量图的数据量的大小与空间数据的属性有关.

表 2 不同矢量格式数据量比较

Table 2 Comparison of data quantity of different vector formats

类型	.cdr/kB	.dwg/MB	.tab/MB	.shp/MB	.e00/MB
土地利用	865	13.10	13.40	24.1	92.9
耕地	47.5	0.67	0.67	1.22	3.39
园地	83	1.09	1.08	2.04	5.61
林地	102	1.36	1.33	2.55	7.90
灌木林	56.9	0.81	0.83	1.49	4.22
疏林地	53	0.78	0.78	1.45	4.01
交通	119	1.19	1.19	2.20	10.20
城镇及交通	426	7.49	7.72	13.50	43.50
农村居民点	36.6	0.54	0.54	0.99	2.71
独立工矿	35.7	0.56	0.55	1.02	2.81
水域	118	1.18	1.03	1.88	5.37
裸岩石砾地	24.2	0.43	0.43	0.79	2.15

5 结 语

随着信息化技术的不断发展, 人们对土地信息共享的要求日益增强, 特别是土地空间数据的共享. 因此土地空间数据的网上传输与发布技术的研究将无疑会被人们重视. 人们在研究网上土地信息发布技术时, 很自然地就会想到数据的分层, 但是, 分层后如何根据用户不同的需要用不同的格式存储数据, 研究者们考虑的不多, 本文结合土地空间数据的特点和用户对土地信息不同的要求分析, 以及网络传输技术, 针对以上的实例进行分析, 得出了一些客观的数据, 通过对数据量的比较, 在对土地空间数据进行分层组织时, 应当分析各种用户对所分层的图件对象需求的要求, 选择合适的数据格式存储数据. 当用户只浏览图件而不需要交互操作时, 我们尽量采用 jpg 格式或 cdr 格式存储图件; 当所存的图件需要为用户提供交互式操作时, 我们可以用 Map Info 的 tab 格式或 ESR I 的 .shp 格式存储图件, 这两种格式都是国际上 GIS 通用格式. 当然, 也可以用国内通用的 GIS 软件的数据格式存储图件.

[参 考 文 献]

- [1] 刘耀林. 土地信息系统[M]. 北京: 中国农业出版社, 2003, 7, 1-2
- [2] 高俊, 等. 虚拟现实在地形仿真中的应用[M]. 北京: 解放军出版社, 1999: 66-70
- [3] 中国可持续发展信息共享系统的研究开发项目管理办公室. 中国可持续发展信息共享系统的研究开发集中共享数据上交技术规定[EB/OL]. <http://www.acca21.org.cn/projet/xinxigongxiang/sjsj.htm>, 2003.5

- [4] Gregory K Wallace. The JPEG still picture compression standard [EB/OL]. <http://www.gischina.com/maindoc/sinchin/gisforum/format/jpg.pdf>
- [5] ESRI. ESRI shapefile technical description [EB/OL]. <http://www.gischina.com/maindoc/sinchin/gisforum/format/shapefile.pdf>, 1998
- [6] 中国可持续发展信息共享系统的研究开发项目管理办公室. 中国可持续发展信息共享数据库数据字典内容标准 [EB/OL]. <http://www.sdinfo.net.cn/ngcc/sdinfo/document/stdic.htm>.
- [7] Hall W, Davis H, Hutchings G. Rethinking hypemedia: the microcosm approach [M]. Boston: Kuwer Academic Publishers, 1996
- [8] Government of Jamaica. Land information management systems [EB/OL]. 1997 http://www.spatialvision.com/LAND_INFORMATION_SYSTEMS.doc
- [9] 韩海洋, 等. Internet 下多数据源、超媒体空间信息的分布式调度与管理 [EB/OL]. <http://www.e2map.com>, 2003

Land spatial data organization and expression based on internet

Zhang Liting^{1,2}, Zhou Shijian², Zhu Guorui¹

(1. School of Resources and Environmental Science, Wuhan University, Wuhan 430079, China;

2. Department of Survey, East China Institute of Technology, Fuzhou, Jiangxi 344000, China)

Abstract Land spatial data organization and expression method was studied in order to realize the on-line land spatial data shares and enhance the land spatial data on-line transmission speed. The characteristics of the land spatial data and the customer's need were analyzed, moreover the data delimiting expression and the data format aspect were studied, through experiment one set of objective data was obtained. After contrasting and analyzing the data and the map, the data storage format according to user's need should be chosen. When the user only browses the map but does not need alternately to operate it, the .JPG format or the .GIF format is recommended to use to store graphic document. When the interactive operation graphic document for the user is needed, the .TAB format of the MapInfo or the .SHP format of the ESRI is recommended to use to store graphic document.

Key words land information; data delimitation; data organization