MAPGIS 中扫描矢量化及其相关技术*

刘修国　谢　忠

(中国地质大学信息工程学院,武汉 430074)

摘　要　针对扫描矢量化过程不同阶段存在的扫描图象可视化、线条有效提取、变形数据校正等问题,提出或发展了解决相应问题的方案. 相应方案在 MAPGIS 基础平台的 MAPRVS 智能矢量化系统中得到精巧实现,并取得了令人满意的应用效果.

关键词　地理信息系统,栅格矢量化,扫描图象可视化,线条提取,变形数据校正.
中图法分类号　TP391

第一作者简介　刘修国,男,副教授,1969 年生,1991 年毕业于中国地质大学计算机系,获学士学位,现主要从事 GIS 方面的科研及应用开发工作.

地理图件的扫描矢量化,是解决地理图件数字化瓶颈难题的必然途径. 由于地理图件固有的幅面大、曲线关系复杂、线型种类繁多、符号注记遍布等特点,决定了地理图件扫描矢量化的难度[1]. 如何在地理图件矢量化过程中很好地解决扫描栅格数据的快速可视化,有效地提取关系复杂的线条,并确保结果矢量数据的质量满足要求,是地理图件矢量化软件系统必须妥善解决的问题.

针对地理图件矢量化过程中的问题,MAPRVS 智能矢量化系统发展了最久未用淘汰技术. 采用空间相关最久未用淘汰策略调度四叉树线性编码[1]的栅格分块数据,有效地实现了扫描栅格图象数据的快速可视化,发展了线条跟踪提取技术[2]. 采用线条轮廓跟踪夹比取中矢量化算法,极大地提高了线条识别的精度和效率. 利用控制点 Delaunay 三角剖分[3]借助仿射变换校准变形区域的方法,确保了结果矢量数据的质量.

1　栅格图象数据的快速可视化技术

1.1　图象数据四分标准化

内存容量有限、运行速度较慢是微型机的不足之处,而扫描栅格图象数据却具有数据量大、格式多样化的特性. 要在微机硬件平台上快捷高效地实现对扫描图象数据的矢量化,系统首先必须将栅格图象数据分块组织并转化为系统内部统一的标准格式,同时通过对分块数据施行编码以允许系统随机存取图象数据的各个分块. 基于这一考虑,MAPRVS 矢量化系统采用了四分标准化方案.

首先,根据内存容量、硬盘存取速度,系统自动将栅格图象数据标准化为 256 ×256 或 512 ×512 等可变大小的数据分块;然后对各分块数据施行线性四分树编码,并将二维图象分块数据映射到线性编码的一维连续空间中,从而实现对图象分块数据的快速定位和随机存取. 试验表明,对 A0 幅面的地理底图以 600 DPI 的分辨率扫描,数据量高达几十兆. 按 512 ×512 大小分块组织数据,系统在一般微机上也能够很好地实现对扫描栅格数据的快速可视化和处理.

1.2　空间相关最久未用淘汰策略

尽管系统采用了四分标准化方案将二维分块数据映射到一维空间上随机存取,但若对分块的扫描栅格数据,每次只调入其中一个分块到内存中,这不仅没有充分利用微机的有限内存,而且系统为实现处理不得不来回调入不同的图象分块数据,这会极大地降低系统的处理效率. 为此 MAPRVS 矢量化系统在对图象数据四分标准化的基础上,采用了空间相关最久未用淘汰策略以优化系统对分块数据的存取,提高了系统的处理性能.

空间相关最久未用淘汰策略,其实质是利用内存虚拟存取技术. 实现时,系统在内存中开辟若干个分块大小的虚拟缓冲区,以同时容纳多个分块数据. 开始时系统调入当前及其周边的多个分块数据,然

1998 年 3 月 10 日收稿.
*国家"九五"重中之重科技攻关项目(No. 96 - B02 - 03)资助.

© 1994-2008 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

后系统按一定优先顺序(如处理方向)将待处理的分块数据调入内存;当没有可利用的空闲块缓冲区时,系统便根据空间相关性加权计算最久未用系数来确定待淘汰的数据分块,并确定待调入的数据分块. 采用空间相关最久未用淘汰策略,较好地解决了分块数据在内外存间的有效交换.

2 矢量化技术和算法

2.1 传统的细化追踪矢量化技术及算法

为了提高扫描矢量化的自动化程度,系统在比较了传统的边界追踪、模板、中心线及剥皮等经典细化算法的基础上,选择了算法相对简单、直观且细化结果也较理想的逐层剥皮细化算法[1]. 逐层剥皮细化的基本思想是:先找到一个位于线划影象边缘上的象元,接着以此象元为中心,按一定顺序检测其 8 个邻域的灰度值,以确定本中心象元是否应置为 0,并找到与本中心象元相邻的边缘象元以继续追踪剥皮. 该算法的优点是只对边缘象元进行处理,因此细化效率高. 在对细化提取的骨架线矢量化时,系统会自动考虑细化引入的断线,在追踪矢量化过程中采取部分容错限差,以提高细化矢量化的线条质量.

实践表明,对于图形信息单一、线条清晰的扫描栅格图象数据,传统的细化追踪矢量化技术具有一定的长处. 它不仅可以得到较好的矢量化结果,更主要的是可提高扫描输入矢量化的自动化程度,解决栅格线划影象向矢量数据自动转换的问题.

2.2 非细化、交互式追踪矢量化技术及算法

相对于细化追踪矢量化,非细化追踪矢量化和交互式矢量化功能两者都采用的是两边同步行走夹比取中矢量化算法. 同传统的细化追踪矢量化技术相比,算法不必先对栅格线划影象细化提取中心骨架线,而是直接对栅格线划影象追踪两边边界从而加以矢量化. 该算法的基本思想是:通常情况下(见图 1a),系统在栅格线划影象两边同步行走,并取其中心线作为线划影象线条的矢量线;在遇到毛刺及各线的接头时(见图 1b),系统按一定的限差控制栅格线条两边的走向和走步. 若在一定的限差允许范围内可以通过,系统则继续夹比取中矢量化该线条;否则在非细化追踪矢量化功能中,系统终止该线条的矢量化而开始另一线条的矢量化. 在交互式矢量化算法中,系统会提请用户选择以引导系统可进行的下一个执行动作.

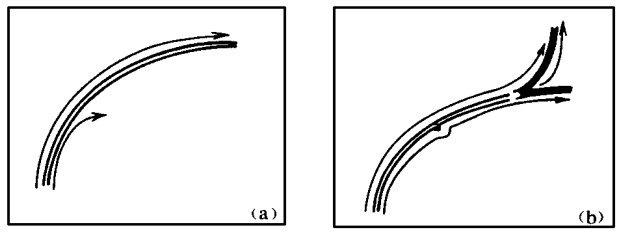


图 1 两边同步行走夹比取中矢量化算法原理

Fig. 1 The principle of line contour tracking vectorization algorithm

比较而言,非细化追踪矢量化算法较细化追踪矢量化算法,可以避免细化引入断线等缺点,同时也具有较高的效率. 而交互式矢量化算法由于提供了退点、加点、导向、选择等交互功能,可有效地选取必需的图形信息矢量化,这不仅提高了系统矢量化操作的灵活性,也可保证矢量化结果数据的质量.

3 矢量数据抽稀算法和校正

3.1 矢量数据抽稀算法

扫描矢量化产生的矢量图形数据,随着扫描分辨率的提高其矢量化生成的矢量数据量也越大;但大量的数据点是冗余的,必须进行数据压缩以提取必要的数据点. 由于道格拉斯 - 普克(Douglas-Poiker)矢量数据抽稀算法既可保持原线条的特征,又可根据制图精度要求确定抽稀限差等优点[4],因而在 MAPRVS 矢量化系统中精巧地实现了该算法.

该算法实现抽稀过程是:先将一条曲线首尾点虚连一条直线,求其余各点到该直线的距离,选其最大者与规定的临界值相比较,若小于临界值,则将直线两端间各点全部舍去,否则将离该直线距离最大的点保留,并将原线条分成两部分,对每部分线条再实施该抽稀过程,直到结束. 抽稀结果点数随选取限差临界值的增大而减少,应用时应根据精度要求来确定抽稀限差临界值,以获得最好的结果.

3.2 矢量数据局部非线性校正

在扫描矢量化处理中,通常会遇到由于图纸伸缩、扫描走纸等引起的数据非线性变形. 这类变形不能通过简单的平移、旋转或比例等线性处理来校正,而必须采用非线性局部处理来校正. MAPRVS 矢量化系统提供了对矢量结果数据进行局部仿射变换的校正功能.

仿射变换具有直线变换后仍为直线,平行线变

© 1994-2008 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

换后仍为平行线,并保持简单的长度比,而在不同方向上变换后长度比会发生变化等特点.实际应用中较多采用能保持边界拓扑一致性的分块双仿射四点变换,它是较简单的一次变换,其变换表达式为

$$X_n = A_1 \times X + A_2 \times Y + A_3,$$

$$Y_n = B_1 \times X + B_2 \times Y + B_3.$$

式中 3 对待定系数,只要给定不在同一直线上的 3 对对应点的坐标即可求得.

在 MAPRVS 扫描矢量化系统中,系统利用变换控制点将变形区域 Delaunay 三角剖分化[3],然后对每一三角剖分实行仿射变换.该方法的特点是:由变换控制点形成的 Delaunay 三角剖分较均匀地覆盖变形区域;而对每一个三角形的仿射变换只利用三角形的 3 个顶点条件,保证两相邻接三角形公共边界上的点其变换的象是唯一的.这样,在 Delaunay 三角剖分覆盖的变形区域内任一点都能变换到相应的位置,从而实现以小区域内简单变换模型解决大区域内复杂变形的校准.

4　结语

MAPRVS 扫描矢量化系统由于采用了分块处理、类线性四分树编码方案组织栅格数据存储,并采用局部空间相关最久未用淘汰策略调度虚拟存取技术,加之合理的内部数据结构,较好地解决了微机内存少、扫描输入数据量大的矛盾.系统在扫描栅格图象数据的高效处理、快速可视化、分块调度、交互矢量化、矢量数据压缩及变形失真非线性校正处理等方面具有明显特色.

与先进实用的扫描矢量化技术的有机结合,使得 MAPRVS 扫描矢量化系统以其方便灵活、实用高效的性能赢得了用户.已有数个大型综合信息管理系统使用 MAPRVS 对多达上千幅的地理底图数字化.实际工程应用表明,使用 MAPRVS 扫描矢量化系统对数字化大数量地理底图具有用时少、精度高、易拼接的明显优点,是建立优质的大型地理信息管理系统地理图件数字化的有效工具.

参　考　文　献

1　徐庆荣,杜道生,黄伟等.计算机地图制图原理.武汉:武汉测绘科技大学出版社,1992. 58

2　谭建荣.基于图形约束的工程图扫描图象直线整体识别方法.计算机学报,1994, 17(8): 561～569

3　柯正谊,何建帮,池天河.数字地面模型.北京:中国科学技术出版社,1993. 127～130

4　季民河.以非欧几里得方法对 Douglas - Poiker 线性地理要素简化算法在保持空间变化特征能力方面的评价.见:林辉主编.CPGIS 第一次年会论文集 ——地理信息系统的发展与前景.北京:科学出版社,1993. 162

RASTER-TO-VECTOR RELATIONAL TECHNIQUE IN MAPGIS

Liu Xiuguo　Xie Zhong

(Faculty of Information Engineering, China University of Geosciences, Wuhan 430074)

Abstract　This paper analyzes the key factor of the problems existing in the process of raster-to-vector, such as scanning image visualization, extracted line vectorization, distorted data correction etc.. It gives a new method to solve these problems. The new method is implemented in MAPRVS, which is a raster-to-vector subsystem of MAPGIS.

Key words　geographic information system, raster-to-vector, scanning image visualization, extracted line, distorted data correction.

© 1994-2008 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net