

微机 GIS 系统软件设计的原理与方法*

王 峰 任伏虎

(北京大学地理系)

谢 昆 青

(北京城市规划设计研究院)

1987年12月10日收稿

摘 要

本文介绍了北京大学遥感技术应用研究所(国家遥感中心技术培训部)于1988年4月推出的地理信息系统通用软件 PURSIS 的设计原理与方法。由于它是在 IBMPC/XT, AT 及系列兼容机上研制的,因此便于推广使用。为了克服微机在速度和容量上的种种限制, PURSIS 系统采用了一些较复杂的技术手段,使得与其它系统相比,具有如下明显特点:(1)利用软件技术,为图形的输入和编辑创立了一个良好的交互式工作环境,使用户能方便地输入和修改各种地图。(2)采用压缩编码技术存贮图形和图像数据,从而大大减轻了微机系统的空间负担。(3)设计了高效率的数据格式转换算法,使矢量数据能快速转换为网格数据。(4)设计了 COLOR 136 软件,将廉价的 COLOR-400 卡支持的 16 种颜色扩充为 136 种颜色,成功地解决了专题图件的显示问题。(5)解决了图形数据和属性数据的联合操作问题,使用户能在图的逻辑概念上进行多种操作。此外,数字地形模型和其它数据在操作时的联系甚为紧密。(6)提供了一些常用的应用工具,如专家权重、叠加分类、专家命题等模型。使用户在这些模型支持下解决一些各自的专业问题。(7)解决了 FORTRAN 语言、编译 DBASE III 和 8086 宏汇编语言之间的相互调用问题,使系统设计灵活多变,并能充分利用微机的系统功能。

关键词 地理信息系统 系统软件 应用模型

地理信息系统(GIS)已成为近年来地学、测绘学、遥感技术与计算机科学相结合的新的发展方向。它正日益受到世界各国、各级政府部门的普遍重视和支持。从国际上十余年的发展情况来看,以某个学科或问题为对象的专题信息系统具有较高的使用价值,小型或微型计算机地理信息系统的建立,充分表现出其强大的生命力。本文介绍 PURSIS (Peking University Remote Sensing Information System) 软件的研究与设计。

一、系统设计的基本原理

GIS 计算机系统软件应具有数据输入、存贮、操作、处理、输出和系统维护等功能。我们把 GIS 软件分成两大部分:一是通用软件;二是应用软件。前者是 GIS 软件设计的基本内容。它是各类专业用户共享的软件资源,如数据输入、数据库管理、数据操作和输出等。

* 本项工作在承继成教授的指导下进行,参加工作的还有柳林等,在此表示感谢。

后者则是在通用软件支持下根据解决某一专题问题的特定模型而设计的应用程序。图 1 示出了上述关系。

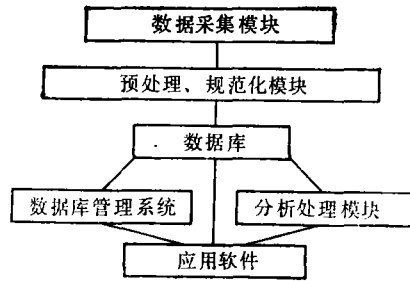


图 1 系统中各模块间的关系
Fig. 1 Relation among the modules

二、系统设计语言和基本硬件配置

(一) 系统设计语言

进行 GIS 软件设计(应用软件除外),应采用能涉及计算机底层的计算机语言,如汇编语言、C 语言等,只有这样,才能充分开发计算机的系统功能。PURSIS 软件采用汇编语言、FORTRAN 语言和 DBASE III-Clipper 语言的混合编程技术。FORTRAN 程序主要用于数值运算、逻辑控制等方面;当需要进行计算机中断处理或系统功能调用时,由 FORTRAN 主程序调用汇编子程序来实现。关于非图像数据(属性数据)的管理和操作,利用编译 DBASE III-Clipper 语言进行控制。

GIS 软件提交给用户的工作方式(用户界面)大致可分为两种:菜单工作方式和语言工作方式。考虑设计和使用上的方便,PURSIS 采用了菜单工作方式。

(二) 系统的基本硬件配置

- (1) 主机 IBM PC XT, AT 及其兼容机, 640KB 内存。
- (2) 大于 30MB 硬盘存贮空间。
- (3) 高分辨率彩色/图形显示器及控制卡 (COLOR-400 卡或 EGA 卡)。
- (4) 宽行打印机。
- (5) 数字化仪(如 TRUE GRID 8000 系列, DT4100 等)。
- (6) 绘图仪(如 DMP 系列或 DXY800 系列等)。
- (7) 8087 或 80287 协处理芯片。
- (8) 若需要处理汉字时,应配置支持 COLOR-400 卡或 EGA 卡的汉卡。

三、系统采用的数据格式

GIS 要求的数据大致可分为三类:影像数据;图形数据;属性数据。由于影像和图形

数据的数据量十分巨大,在微机上建立 GIS 必然受到存贮空间的限制,所以应该采用压缩编码技术来减轻系统的空间负担。这里介绍 PURSIS 对影像和图形数据采用的三种数据格式。

(一) 网格数据

也称直接编码数据。目前我们主要按这种格式存贮遥感影像和地面高程模型 (DEM) 数据。影像数据通过与图像处理系统通讯获取; DEM 数据根据数字等高线插值产生。上述遥感影像数据,对 PURSIS 而言,通常是指图像处理系统处理后得到的遥感分类图件。

(二) 多边形数据

在 PURSIS 中,多边形数据是数字化仪输入斑块状图形(如土壤类型图、土地利用图、行政区域图等)的产品。根据图形的拓扑关系,依次对每个多边形以及组成该多边形的弧段进行编码,并记录组成弧段上连续点的坐标。其数据文件由弧段编码文件、多边形编码文件、弧段坐标文件、弧段连接成多边形的索引文件这四个部分组成,从而形成多边形数据结构。多边形数据虽然能极大地节省存贮空间,但在操作网格数据、设计大部分图形间的运算过程中较为复杂,操作效率也往往因此而降低。所以,我们仅仅把它作为 PURSIS 系统的数据格式之一。它在 PURSIS 系统中的作用是,把矢量数据向网格数据转换,作为图形编辑时的基本数据,作为图形输出的数据。

(三) 游程长度编码数据

游程长度编码 (Run Length Code 缩写为 RLC) 数据,是一种压缩编码数据(详见附 1)。对于有较少灰度要求的图像(如各种数字地图),RLC 是较为实用的方法之一。它可以在多边形数据或网格数据的基础上产生。在 PURSIS 中全部的图形操作均以 RLC 数据为基础。

四、PURSIS 系统软件功能

系统按分层结构设计,最高一层描述总功能,其它层次则一层比一层具体地实现这些功能。反映在用户工作界面上,是一级级的菜单选择项。

(一) 数据采集模块

(1) 地图数字化。主机与数字化仪通过 RS-232 异步通讯接口实现数据的传输。从数字化仪得到的数据按矢量结构存贮。

(2) 图形编辑。PURSIS 的交互式图形编辑具有以下功能:点、线的删除;合点;加点点;加线;矢量数据到网格数据的相互转换;矢量图的拼接。

(3) 与图像处理系统通讯获得记录在 CCT 磁带上的遥感影像数据。

(4) 从键盘输入属性数据,由编译 DBASE III 建立属性数据库。

(二) 数据预处理和规范化模块

数据预处理包括图像数据的几何校正；同一地区或同一等级图间的归化比例尺和配准；数字化仪坐标与计算机图形坐标的转换；由数字地形图插值产生地面高程模型 DEM (图版 I 图 2)。

数据规范化的目的在于使全部数据能够按系统所要求的格式实现存贮，如：图形数据统一按游程长度编码的格式存贮；影像数据和地面高程模型数据统一按网格格式存贮；进行编辑或绘图仪输出的图形数据按矢量格式存贮；属性数据按 DBASE 所要求的格式存贮等。为了实现数据的流通，必须建立各种数据格式间的转换。我们把满足上述格式且能进行格式间转换的数据称为 PURSIS 系统的规范数据，符合数据规范也称为数据入库。

(三) 图像数据库系统

图像数据库是 GIS 的核心，它管理着图形、影像和属性数据。图像数据库的基本功

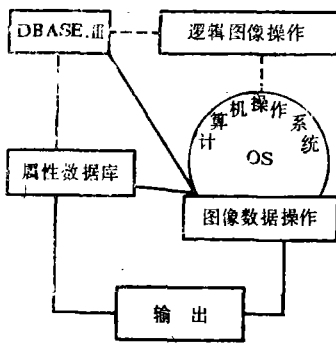


图 3 图像数据库系统

Fig. 3 Picture data base system

能是对数据进行查询、操作、修改、更新、增加、删除。PURSIS 图像数据库的原理框图见图 3。

图像(影像和图形)数据的控制 and 操作，目前还是利用计算机操作系统的文件管理来实现的。为了方便 GIS 用户，在 DBASE 管理系统中设计的查询和操作，虽然是对与图像有关的非图像数据进行的，但在系统内部，对图像数据进行的相应操作也随之自动完成。这种方式称为逻辑图像操作。这样我们就沟通了编译型 DBASE 语言——Clipper, FORTRAN 和 8086 汇编语言之间的相互调用，并用这三种语言的混合编程设计出图像数据库系统。作为大众数据库的 DBASE 编程，本文从略。下面介绍在操作系统支持下的图像

数据操作软件。

1. COLOR-136 彩色图像显示软件

PURSIS 将原装图形卡换成 COLOR 400 卡(分辨率可达 640×400 像素)后，每幅图像可以显示出 16 种颜色。尽管如此，仍不能满足显示地学分类图件的需要。我们将颜色编码重新组合(组合方案共有 C_{16}^6 种)，设计了 COLOR-136 彩色图像显示软件(详图附 2)。该软件将 COLOR 400 卡支持的 16 种颜色扩充到 136 种，从而较好地解决了在微机以廉价的硬件满足显示各种分类图件和地图的需要。

2. 属性数据库

属性数据库存储与图像有关的各类非几何属性信息，通过类别码与图形数据相联系。属性数据库的初始结构由数据采集模块在图形数字化的同时建立，用户可以视任务的要

求对属性数据库结构进行编辑修改, 录入属性数据。非几何属性逻辑操作的结果可以反映到图像数据上, 也可利用软件设置的光标, 对屏幕图像进行查询, 查询的内容包括图像坐标、颜色码、类别码、图幅尺寸、图像比例尺和属性数据库中的全部数据项, 为图像的几何校正、综合分析、获取信息等提供了有效的手段。

3. 图像操作

PURSIS 图像操作内容包括: 单要素或多要素提取; 图像平移、旋转、缩小、放大、剪取、游动和开窗处理; 图形拼接; 调配图斑颜色; 多幅图几何叠置; 多幅图加权叠置; 图像的数学变换。

4. 绘图仪输出

5. 系统维护

(1) 安全性保护措施

安全性的基本原则是防止无关人员获取其不应知道的数据。通常采取的措施有: 检查用机权限; 存取控制; 密码变换等。

(2) 完整性保护措施

完整性的破坏是指数据失真或因错误而造成数据间的矛盾。完整性保护措施是由设计人员制定一系列约束条件, 当数据超出约束时, 自动提出让操作人员进行检查和修正。

(3) 系统恢复

系统恢复一般借助于系统副本等手段实现。

(4) 系统重构

系统经过长期使用后, 由于种种原因, 系统性能将愈来愈差, 如空间利用率、检索和操纵效率越来越低等。为了改变这种状态, 必须重新组织系统。

(四) 分析处理模块和模型实现工具

分析处理模块为 PURSIS 用户提供一套数学工具。其中包括求最大值、最小值、数学期望、方差、多元回归、直方图分析、直方图均衡、线性拉伸、数值量化、专题图加权等等。这一系列数学运算程序可以作用于图像数据, 也可以作用于属性数据。

此外, 还提供了依据命题逻辑设计模型的实现工具。下面简述其原理。

设有一个为农业规划部门提供种植水稻的自然条件参考模型。其条件为:

(1) 积温 C 高于 3200°C ; (2) 降雨量 $R > 800$ 毫米, 或当降雨量 < 800 毫米时, 有灌溉条件 W ; (3) 坡度 $S < 8^{\circ}$; (4) 无霜期 F 大于 200 天。

上述条件要求模型中有 5 个参变量(用专题图表示), 即积温分布图、降雨量图、水利条件图、坡度图和无霜期分布图。上述图件经数字化后存入计算机中。根据条件, 写成命题逻辑的表达式为:

$$P \rightarrow ((C > 3200) \wedge ((R \geq 800) \vee ((R < 800) \wedge W)) \wedge (S < 8) \wedge (F > 200))$$

式中 P 表示为“可种植水稻”; “ \rightarrow ”表示“如果…则…”; “ \wedge ”表示“并且”; “ \vee ”表示“或者”。

PURSIS 为用户提供的模型实现工具, 就是把类似于上式的命题表达式自动翻译成

计算机程序。所以,用户只要写出解决自己问题的命题表达式就可以很快求解。

五、建立专题信息系统

以上阐述的是 PURSIS 的基本内容,它也是微机 GIS 通用功能软件部分。任何一个信息系统都离不开这类软件的支持,但一个信息系统的存在还必须有应用软件。应用软件是面向专业问题的。不同的应用软件与通用软件组合,就构成不同的专题信息系统。建立专题信息系统大致可分成两步:首先,在通用软件的支持下,把收集的各种数据输入计算机,并按系统要求的格式进行预处理,完成数据入库。第二是根据专业模型由用户设计应用软件,在此之前,往往要对数据库中的数据进行一些查询、操作或处理,这一系列工作都可通过调用通用软件功能实现。应用软件与系统的接口方面,则应按系统的要求存取系统数据库中的数据。

作为 PURSIS 软件的应用实例之一,我们利用专家评分模型产生了河南省冬小麦生长环境适宜性评价图(图版 I 图 4)。

六、结 束 语

我们对 PURSIS 系统做了如下指标测试:

(1) 矢量数据到网格数据的转换,采用了我们自己设计的一种全新算法——边界代数填充法。转换时间为 1—5 分钟。

(2) 游程长度编码文件与网格数据文件的相互转换,根据图像复杂度不同,约花费 10—60 秒。

(3) 图像显示为即时响应,完成一幅图像的显示约 1 分钟。

(4) 全部图像操作均可在 1—3 分钟内完成。

(5) 图像查询、属性数据库操作均为即时响应。

(6) 绘图仪输出,根据图件的复杂度不同,约花费 1—2 小时。

关于游程长度编码文件相对于网格数据文件的压缩比,统计结果表明在 10 倍左右。正是因为采用了游程编码的压缩技术,使得一部分操作的时间花费大为节省。

一个计算机软件的正确性、完整性、一致性等问题,通常不能用数学方法加以证明,它只能在长期的运行中不断改进和完善。

附 1. 游程长度编码

图像上的一个扫描行是整数的序列 x_1, x_2, \dots, x_n 。把取相同值的连续的若干像元称为一个游程,则序列可分成 K 个游程 ($K \leq N$)。从此,把序列 x_1, x_2, \dots, x_n 映射成整数对 (g_i, l_i) 的序列 ($i = 1, \dots, K$)。每个整数对描述一个游程, g_i 表示第 i 个游程的取值; l_i 表示该游程的长度(像元个数)。这样,一条扫描线上 N 个像元的信息只要用 K 个整数对就能保持信息的表达。这种压缩方法称为 RLC 方法。可见,RLC 方法对专题图进行数据压缩是十分有效的。

附 2. 微机图像显示软件 COLOR-136

COLOR-136 (简称 C136) 软件以美国 SIGMA DESIGNS 公司生产的 COLOR 400 高分辨率彩色图形控制板作为硬件支持。COLOR 400 板在高分辨 (640 × 400) 图形模式下, 每个像素的颜色由四个调色平面决定。为了满足显示专题分类图件的需要, 设计了 COLOR-136 软件, 它把原来的 16 种颜色扩充到 136 种组合色。我们称这种组合原理为交叉点排列原理, 即每一组合色由两种原色组合而成, 虽然有 C_2^{16} 种组合方案, 加上原有的 16 种颜色, 共得到 136 种颜色。例如为了得到普蓝色的视觉效应, 我们可以用蓝色 (色码=1) 和绿色 (色码=2) 进行交叉排列 (图 5)。此时, 寄存器的设置为: $BL = 15$; $AH = 12$; $DX =$ 像元的行号 (L); $CX =$ 像元的列号 (R); $AL =$ 像元的色码 (C)。利用基本输入输出 (BIOS) 指令系统的中断 ($INT IOH$) 功能, 将在屏幕的第 L 行, 第 R 列处, 按照色码为 C 的颜色显示。上述过程遍历屏幕上所有点后, 就可以得到一幅彩色图像。

1	2	1	2
2	1	2	1
1	2	1	2
2	1	2	1

图 5 交叉排列原理
Fig. 5 Principle of the cross

Principle And Method For GIS Software Design On Microcomputer

Wang Zheng Ren Fuhu

(Department of Geography Peking University)

Xie Kunqing

(Urban Planning and Designing Institute of Beijing)

Abstract

This paper will brief you on establishment of one kind of GIS system—PURSIS (Peking University Remote Sensing Information System) on microcomputer, IBM PC/XT, AT, and simply explain its basic characters, functions for data processing, digital data formats, picture data base system and expert system. Based on the idea of modular programming, we should organically combine many thematic GIS, each of them becoming a subsystem. The data collected and stored by GIS can be shared by any subsystem. The main system is composed of some share processing modules, which is the share software resource by the users. In this paper, we also introduce some techniques for the organization of the system, basic conception of the expert system, data encoding, color display for the picture on the IBM PC, etc.

Key words Geographic Information Systems System Software Application Models

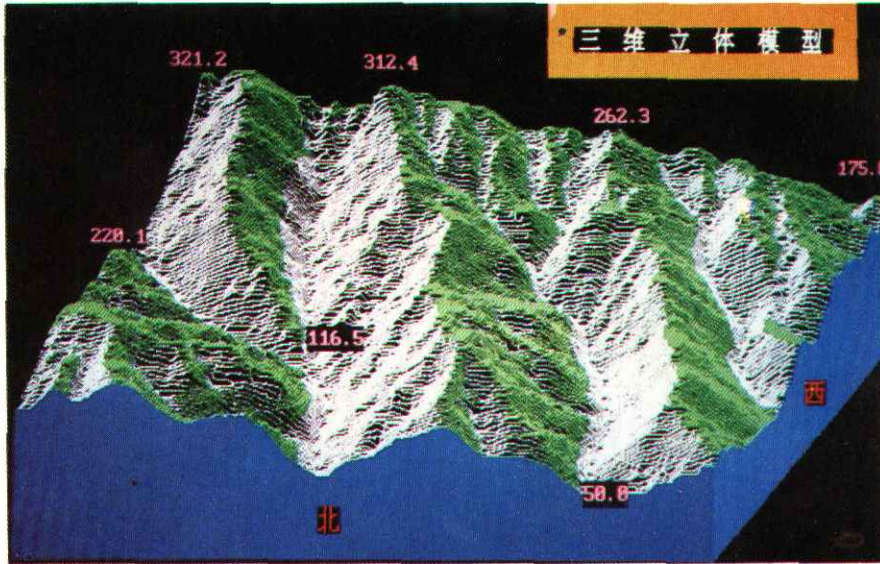


图 2 DEM的三维立体模型显示

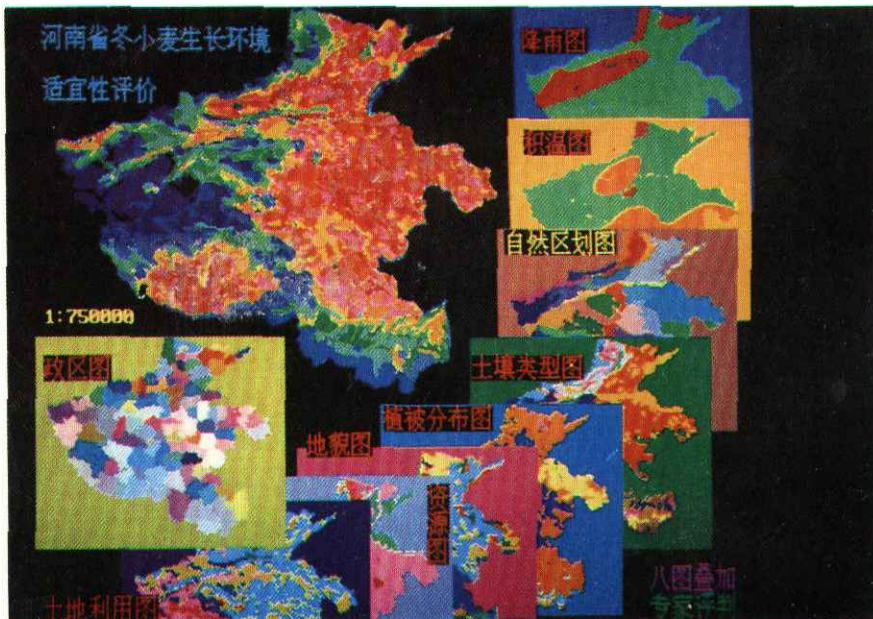


图 4 河南省冬小麦生长环境适宜性评价图