

RDSS 高程数据库的研究与设计

高 扬¹, 乌 萌²

(1 长安大学 地测学院, 陕西 西安 710054; 2 西安测绘研究所, 陕西 西安 710054)

摘 要: 针对 RDSS 系统对高程数据的应用需求, 采用了对高程数据按地形分类抽取和压缩方法, 设计了数据文件内数据的分块组织存储、四叉树索引和文件分级管理模型, 提出了双线性内插法、双三次内插法、移动曲面内插法相结合的数据拟合和提取方法。实验结果表明, 文中提出的高程数据库应用模型可以有效减少高程数据量, 控制精度损失, 并提高数据检索的速度, 满足 RDSS 系统的要求。

关键词: RDSS; 高程数据; 地形分类数据提取; 四叉树索引; 数据拟合

中图分类号: P208

文献标识码: A

文章编号: 1000-7180(2012)05-0161-04

A Researching on Elevation Data Model in RDSS

GAO Yang¹, WU Meng²

(1 College of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054, China;

2 Xi'an Surveying and Mapping Institute, Xi'an 710054, China)

Abstract: According to application requirement for elevation data in RDSS, this paper has presented a method which adopted terrain classified data extraction and compression for elevation data. A model with data block organization and storage in data files, quad-tree indexing and file classification management has been designed in this paper. A method of data fitting and extraction combining bilinear interpolation, bicubic interpolation and moving surface interpolation has been proposed. Experiment result illustrates that the elevation database application model referred in this paper will reduce the elevation data amount effectively, regulate the precision loss, speed up data searching and accord with the requirement of RDSS.

Key words: RDSS; elevation data; terrain classified data extraction; quad-tree indexing; data fitting

1 引言

采用 RDSS(卫星无线电测定业务)体制的卫星定位系统是由定位用户机以外的测量控制中心(MCC)测得卫星与用户间的距离,并由外部 MCC 系统确定用户位置,是一种非自主式定位或被动式定位。RDSS 定位采用三球交汇测量原理,为完成定位解算,需要使用高程数据进行迭代运算^[1-2],因此 MCC 需要预先存储高程数据。以 1:50 000 高程数据库为例,采用经纬网格结构,每个网格南北和东西间隔均为 1",每图幅高程文件形成 901×601 的格网,每个高程值占 2 个字节,则每个高程数据文件容量为 1083002 字节。可以看出,一个 RDSS 覆盖区域

的高程数据库是庞大的,如何高效地管理高程数据是一个值得研究的问题。

在 RDSS 系统中,每次定位解算都需要读取高程数据,高程数据的精度直接影响到定位解算的精度。由于 RDSS 覆盖区域高程数据量庞大,如何在满足数据精度的前提下,有效控制数据量,提高数据读取速度,是 RDSS 高程数据库设计的关键。本文针对 RDSS 的应用需求,研究一种能够实时准确地提供高程数据的管理模型。

2 高程数据库总体框架

RDSS 高程数据库系统从总体上可分为两大部分:高程数据库创建和高程数据库应用。其中高程数

数据库创建模块完成对原始高程数据的检测、压缩、数据组织与存储;高程数据库应用根据定位概略坐标完成高程数据的提取、数据解压、拟合运算和输出。其总体构成如图 1 所示,工作流程如图 2 所示。

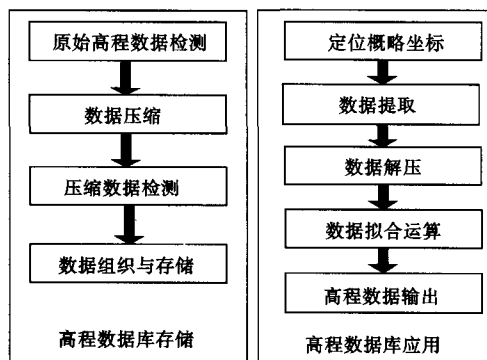


图 1 系统的总体设计

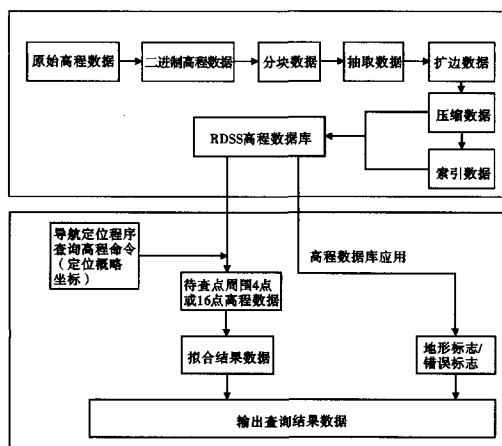


图 2 系统的软件数据流

3 高程数据的组织

为提高数据的读取速度,减少数据读取过程中的 IO 操作,高程数据需要装入定位计算机内存中,同时采用大容量内存锁定技术以确保高程数据不从内存换入换出。受内存容量的限制,高程数据需要预先压缩,减少数据量。本文根据地形特点,基于对高程数据按地形分类抽取和压缩方法,设计了数据文件内数据的分块组织、四叉树索引和文件分级存储结构,提高了数据的检索效率,满足系统的实时性要求。

3.1 高程数据的分类抽取和压缩

数据压缩方法很多,在高程数据压缩中常用的方法有整形量存储、差分映射及压缩编码等^[3]。在选择压缩方法时,在考虑尽量减少数据量的前提下,应保证压缩后的数据精度能够满足定位解算的精度要求,同时要求压缩和解压缩方法尽量简单,满足数据

读取的速度要求。

文献[3]中提出了一种按地形对高程数据分类抽取和压缩方法。其基本原理是将高程数据的图幅范围分为平原、丘陵、山地、海洋四种地形,其中平原和海洋高程数据密度可以降低较大,丘陵和山地高程的变化大,则相应区域的高程数据密度应适当降低或保持不变,密度确定的原则是保证定位解算的精度。对于不同地形采取不同的压缩策略:

(1)对于山地数据,保持原始存储密度,同时不进行压缩处理,以两个字节存储一个高程值;

(2)对于丘陵和平原数据,先进行抽取操作,将数据密度由 $1'' \times 1''$ 分别调整为 $2'' \times 2''$ 和 $6'' \times 6''$,然后采用直接和间接压缩法进行压缩处理。

该方法简单实用,能够有效减少数据量,并控制压缩时的精度损失,同时又能保证检索时数据解压效率,满足系统要求。其具体实现及试验分析见文献[4],这里不再赘述。

3.2 高程数据文件的分块组织

高程数据是以图幅为单位存储,每一图幅对应一个数据文件。 $1:50\,000$ 高程数据文件范围为 $10' \times 15'$,覆盖近 500 平方公里。这样大的区域内可能包含有多种地形,如果都根据最高密度进行数据抽取则不利于提高存储效率,而如果降低抽取密度则可能对地形剧变区域的高程精度造成损失。因此需要将数据文件按区域的分布位置和地形类别划分为小的数据块,每个数据块内部的数据抽取密度相同。

数据块大小的划分原则是十分重要的,直接影响到高程数据合理删减的程度。若划分得太大,则不可能反映地形的真实情况,或者造成精度损失,或者不利于对数据量的压缩,从理论上来说,划分得越小则数据密度确定得越合理,越有利于对数据的压缩,但如果数据块划分得太小,数据块数越多,数据块的检索将会变得复杂,不利于数据库数据的提取,给系统增加了负担。同时,也会使数据冗余增加较多。所以,适当确定数据块的大小,并建立高效的索引查找机制,可使数据合理抽取、压缩和快速提取得到兼顾。

为此,本文设计了基于多层四叉树空间按地形特征对数据分块的方法。其基本思想是:每个数据文件中数据块个数和每个数据块大小不固定,根据实际情况灵活划分。具体作法如下:

① 首先将图幅文件作为四叉树根节点,确定其分布区域内地形特征是否一致。若是则转③;否则,将图幅文件按空间分布位置将文件数据划分为四个

数据块,每块根据地形类别确定存储密度。

② 将每个数据块作为叶节点,并逐一进行判断,将内部地形特征不统一的数据块继续划分下去,直到每个数据块内部地形特征一致,从而就形成了一个基于多层二叉树的空间划分。

③ 利用文献[4]方法,对每个数据块内的数据按地形特征分别进行抽取和压缩。

④ 为每个数据块建立索引记录,每条索引记录主要包括该数据块在文件中的地址指针、该数据块的地形类别、数据密度等信息。索引记录结构如图3所示。

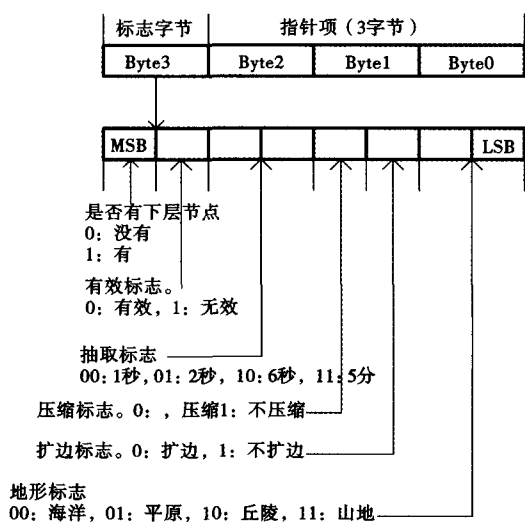


图3 索引记录结构图

另外,为了满足拟合需要,同时节省检索时间,每个数据块每边均向外扩一条边,增加一条边的数据。

如图4所示,1:50 000 高程数据文件包含有平原(A)、丘陵(B)、山地(C)三种地形。按照分布位置和地形特征将其划分为10个空间范围和密度不同的数据块,每个数据块为一规则矩形,对应二叉树的一个叶节点。A类数据块密度为 $6'' \times 6''$,B类数据块密度为 $2'' \times 2''$,C类数据块密度为 $1'' \times 1''$,第一层叶节点所对应的数据块空间范围为 $450'' \times 300''$,第二层叶节点所对应的数据块空间范围为 $225'' \times 150''$ 。

根据以上方法,每个高程文件根据区域按照地形特征划分为若干个块,并且根据地形标志抽取、扩边,形成新的块数据。在数据块内部,按照先纬度方向再经度方向(即先由南向北再由西向东)的顺序,依次压缩并且存放格网点上的高程数据,最后在文件中为每个数据块建立索引,这样通过索引就可以快速访问到对应的数据块。

3.3 高程数据文件的分级管理

RDSS覆盖区域广阔,高程数据文件数量众多,如果将区域内所有高程数据文件集中存放,可能增加操作系统负担,影响文件的检索效率。为此,本文采用了基于图幅比例尺三级管理体制:首先将RDSS覆盖区域划分为多个1:100万图幅数据集,再将每个1:100万数据集划分多个1:10万图幅数据集,最后将1:5万高程数据文件依次存入相应的1:10万数据集中。

采用上述方法对海量数据文件进行了合理的组织,在每个1:100万数据集中最多包含144个1:10万数据集,每个1:10万数据集最多包含4个1:5万高程数据文件,每一级对应的数据集数量不大,同时,使经纬度与高程数据文件存储路径之间建立了井然有序的连接,每个数据文件所在位置都可以根据定位点的概略坐标直接计算出来。在检索时,通过三次解算就可找到定位概略坐标所对应的高程数据文件。

建立的高程数据库模型如图5所示。

4 高程数据的提取

由于高程数据文件中高程数据是以栅格方式存储。大多数情况下,RDSS系统定位解算提供的定位点概略坐标没有落在高程数据文件中的栅格上,所以在提取定位点高程时,需要读取点位附近的高程数据,通过曲面拟合内插的方法来得到该点的高程值。曲面拟合的算法有很多,如移动曲面插法、线性内插法、双线性内插法、双三次内插法、有限元内插法、最小二乘配置法等。拟合算法的选择是非常重要的,研究表明,在某些特殊地貌,数学模型不同,计算精度上可能会产生一定的差异^[5],另外,有些算法过于复杂,系统占用率较高,速度较慢,不能满足定位系统的时间要求。

为此,本文采用了双线性内插法、双三次内插法、移动曲面内插法相结合的拟合方法,按地形特征选择拟合方法。具体如下:

(1)双线性多项式内插法能够保证相邻区域接边处的连续,不能保证光滑。但其计算量较小(只用四点参与拟合),比较适合在地势平缓的平原地区进行内插,并能收到良好效果。因此,在地势平缓的平原地区和海洋地区,采用双线性多项式内插,内插公式为

$$H_i = a_0 + a_1 x_i + a_2 y_i + a_3 x_i y_i \quad (1)$$

$$i = 1, 2, 3, 4$$

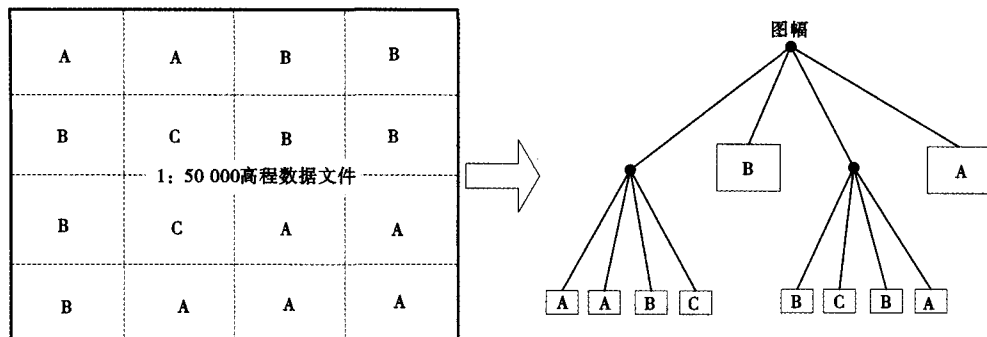


图4 高程数据文件依地形分块示意图

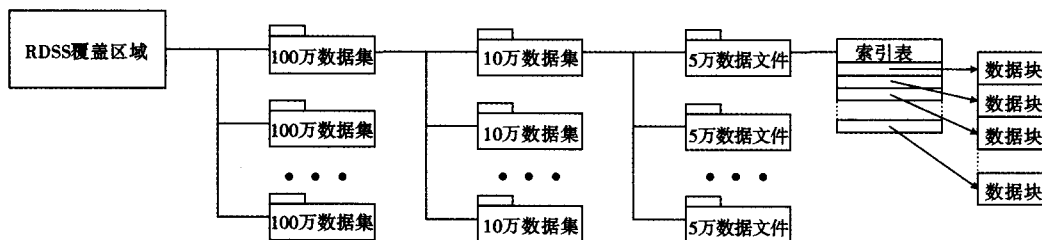


图5 高程数据库模型

其中 x_i, y_i 分别是内插点周围 16 个点的经、纬度坐标。

(2) 双三次拟合法能够保证相邻曲面之间的连续与光滑, 适合于连续变化的光滑地形, 如 U 字形地貌。因此, 在地形起伏不大的丘陵地区, 采用双三次内插, 内插公式为

$$H_i = A_0 + A_1 x_i + A_2 x_i^2 + A_3 x_i^3 + A_4 y_i + A_5 x_i y_i + A_6 x_i^2 y_i + A_7 x_i^3 y_i + A_8 y_i^2 + A_9 x_i y_i^2 + A_{10} x_i^2 y_i^2 + A_{11} x_i^3 y_i^2 + A_{12} y_i^3 + A_{13} x_i y_i^3 + A_{14} x_i^2 y_i^3 + A_{15} x_i^3 y_i^3 \quad (2)$$

$i=1, 2, 3, \dots, 16$

其中 x_i, y_i 分别是内插点周围 16 个点的经、纬度坐标。

(3) 移动曲面内插法是以待查点为中心, 定义一个局部函数去拟合周围的数据点, 不要求拟合面强行通过全部参考点, 具有较好的保凸性。同时, 它也要求数据点尽量密集, 使拟合面在整体上得到比较合理的支撑, 才能获得较好的插值质量。因此, 在地形起伏较大、格网点较密的山地区域, 采用移动曲面内插, 函数模型为:

$$H = Ax^2 + Bxy + Cy^2 + Dx + Ey + F \quad (3)$$

综上所述, 高程数据的提取流程如下:

① 根据查询高程申请所提供的定位点概略坐标, 检索高程数据文件图幅号。

② 根据四叉树索引检索定位点所在的数据块。

③ 提取并解压定位点周围的参与拟合的高程数据。

④ 根据定位点所在的地形标志确定拟合方法:
 λ 平原或海洋地区: 采用双线性多项式内插法;
 λ 丘陵地区: 采用双三次拟合法;
 λ 山地地区: 采用移动曲面内插法。

⑤ 拟合得到定位点高程并输出。

如图 6 所示。

5 实验与结论

本文对 125 幅 1:50 000 万高程数据按地形特征对数据分块组织和文件分级管理, 对数据进行了分类抽取和压缩, 并建立了四叉树索引。在 PC 机 (配置为 CPU: E8200; 内存: 2GDDR800) 上对建立的高程数据库按照本文方法进行了数据提取试验。试验随机选取了 800 个定位点, 根据概略坐标完成了参与拟合点的数据读取、拟合, 并将输出的高程值与原始值进行了比较。实验结果如表 1 所示。

表1 高程数据提取试验结果

原始数据量	压缩后数据量	实验点数	平均提取时间	中误差
135 MB	82 MB	800	0.35 ms	0.39 m

(下转第 169 页)

例2. 首先开发高风险高价值的用例,但是属于高风险高价值的用例有两个,可以按照这两个用例的价值高低决定先开发用例4,然后依次开发用例5、用例3、用例2、用例1。

3 结束语

本文提出了一种基于UML用例图中各个用例的实际情况来确定迭代策略. 用UML语言建立起来的用例模型能如实地反映出整个系统所要完成的功能. 利用用例的风险程度和价值决定整个敏捷开发过程中的迭代顺序,使高风险高价值的用例优先开发,而对高风险低价值的用例在开发过程中尽量避免,对提高整个系统的可靠性会有很大的帮助。

参考文献:

- [1] Wende Kong. Towards a formal and scalable approach for quantifying software reliability at early development stages[D]. Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park, 2009.

- [2] (美)柯恩,著. 敏捷估计与规划[M]. 宋锐,译. 北京:清华大学出版社,2007.
- [3] Ajith Reddy Guedem. Software architectural risk assessment[D]. Master Thesis submitted to the College of Engineering and Mineral Resources at West Virginia University, 2004.
- [4] Ahmed E Hassan. Architectural level risk assessment [D]. Dissertation submitted to the College of Engineering and Mineral Resources at West Virginia University, 2004.
- [5] Wensheng Hu, Zhouhui, yi Hong. A Method of FTA base on UML use case diagram[C]// ICRMS' [s, l]: IEEE, 2011: 757-759.

作者简介:

胡文生 男, (1969—), 博士生, 讲师. 研究方向为软件可靠性.

赵明 男, (1962—), 博士, 教授, 博士生导师. 研究方向为软件、硬件可靠性、应用统计.

杨剑峰 男, (1986—), 博士生. 研究方向为软件可靠性.

贾国荣 男, (1958—), 教授. 研究方向为电子商务.

(上接第164页)

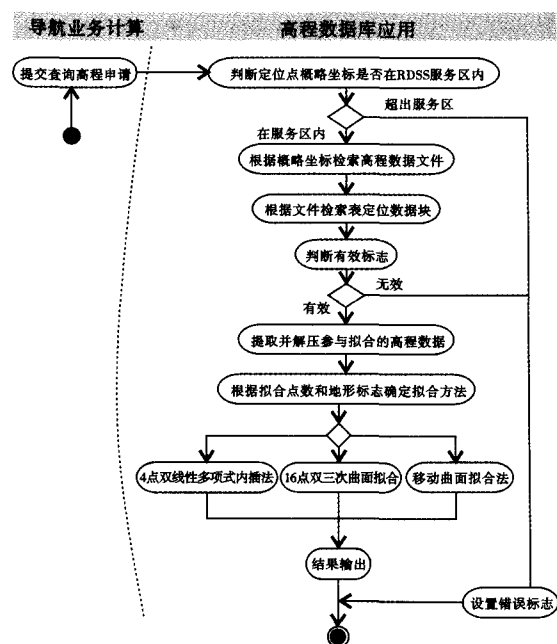


图6 高程数据的提取流程图

从实验结果看出, 本文提出的按地形特征对高程数据分块组织、四叉树索引和文件分级管理模型,

以及双线性内插法、双三次内插法、移动曲面内插法相结合的拟合方法, 可以有效减少高程数据量, 控制精度损失, 并提高数据检索的速度, 满足 RDSS 系统的要求。

参考文献:

- [1] 吴延忠, 李贵琦. 地球同步卫星定位[M]. 北京: 解放军出版社, 1992.
- [2] 谭述森. 卫星导航定位工程[M]. 北京: 国防工业出版社, 2007.
- [3] 陈建鹏. RDSS 高程数据抽取和压缩方法分析[J]. 测绘科学与工, 2009(1): 46-49.
- [4] 张祖勋. 数字摄影测量学[M]. 武汉: 武汉测绘大学出版社, 1996.
- [5] 高扬. 基于 DEM 应用内插算法的分析 [J]. 测绘科技, 1998(2): 23-27

作者简介:

高扬 男, (1969—), 博士研究生, 高级工程师. 研究方向为卫星导航及嵌入式 GIS.

乌萌 女, (1983—), 硕士, 工程师. 研究方向为卫星导航及车辆导航系统.