

基于空间信息多级网格的人口普查 数据采集系统的设计与实现

周俐俊¹ 朱欣焰¹ 邵振峰¹ 宋 珩¹

(1 武汉大学测绘遥感信息工程国家重点实验室, 武汉市珞喻路 129号, 430079)

摘要: 设计了人口普查多级网格(SIMGPC)的结构体系,重点探讨了网格划分、编码方法和网格数据库等关键技术,并实现了一个带有空间信息编码的人口普查数据PDA采集系统。

关键词: GIS; 空间信息多级网格; 人口普查多级网格
中图分类号: P208

人口普查是在国家统一规定的时间内,按统一的方法、项目、报表和标准时间对全国人口按地区逐户进行调查登记^[1]。目前,人口普查数据大多数是按行政区划统计的,而建设项目的规划、管理往往需要按流域等自然单元和自然资源与环境背景来分析,否则,就可能产生对人口地理分布现象的扭曲。例如我国西北地区的定居人口分布,相对集中于绿洲和沿交通干线地带,而且县域范围往往包括广袤的沙漠和戈壁,按行政区域统计的人口分布平均值或密度,几乎是毫无地理意义的^[2,3]。基于李德仁提出的空间信息多级网格(spatial information multi grid, SIMG)^[4,5],本文设计了一个人口普查多级网格(spatial information multi grid for population census, SIMGPC)的实验系统,该实验系统以不同的网格形状和不同层次的网格大小来开展人口普查和进行统计分析。以网格单元的统计分布,比较贴切地反映了高原、高山、荒漠、河谷以及绿洲、耕地等与人口分布的相互关系,提高了人口统计分析的地理精度,避免观念上的误导和生产规划、管理工作中的盲目性。通过网格的多层结构,可以把自然与人类活动的空间关系清楚地表达出来。

1 人口普查多级网格

人口普查多级网格(SIMGPC)是基于SIMG建立的一个带有空间信息编码的人口普查系统,

它以不同的网格形状和不同层次的网格大小来开展人口普查和进行统计分析。SIMGPC所需的网格划分的尺度根据实际地物的密集程度来确定,例如地物稀疏的地方(如沙漠)只需要高层网格,而地物密集的地方(如城市)同时具有从高层次到低层次的不同粗细网格。

SIMGPC实验系统的数据流包括以下两个方面:能接收PDA前端采集到的数据;从现有人口统计数据或者人口信息系统中转入的数据。当然,数据库中的数据还包括行政区划图、居民地层、水系层、人口统计数据等。基于SIMG的人口普查数据采集系统主要是通过带有GPS的手持PDA设备来采集人口普查数据,这些数据通过SIMGPC中划分的网格,统一存入SIMGPC本地网格数据库中,成为SIMGPC本地节点中的数据。根据实际地物的密集程度确定网格划分的尺度之后,通过PDA设备来采集低层次网格中的人口数据,高层网格中的数据是低层网格中的数据的汇总。通过SIMGPC的非完全树状结构,下级网格节点中的数据可以快速反映到上级网格节点中,能够快速地完成人口数据的统计和分析。

2 关键技术

2.1 网格划分

SIMG网格以固定的经差和纬差作为全球划分的基本层级,在基本层级的基础上,以四叉树方

式进行次级网格的划分。取经度 6° 、纬度 4° 为基本层级, 全球有 2 640 个网格, 能够保持网格的形状近似, 大体符合现有地图的分幅要求, 26 级划分以后可以表示地面 1 cm 大小。

2.2 网格编码

二叉树编码可以分为深度二叉树和线性二叉树, SIMG 网格不需要保存每个网格的深度信息, 宜采用线性二叉树进行编码。SIMG 网格编码分为三部分: 网格所在的基本级网格的行列号编码, 如 12.4 代表基本级网格中纵 12 横 4 的网格;

子网格的深度, 如 5 代表在基本网格后划分的第五层网格; 子网格在基本网格的子编码, 代表该网格在基本级网格中的 Morton 码, 使用基于自然数的 Morton 表示, 以 0 代表第一级的基本网格编码, 精度越高编码越长, 32 次划分需要占有 64 位的空间, 能够表达亚 mm 级的精度。如 12.4.5.7, 表示在基本级网格纵 12 横 4 的网格上, 进行了 5 级分割, 其编码为 00007。

2.3 网格数据库

SIMGPC 实验系统在 PC 端采用 Oracle9i 建立了网格数据库, 该网格数据库包括网格元表、PDA 数据采集表和网格人口统计表。网格元表用来记录网格的基本属性, 主要包括网格 ID、网格编码、网格四角点的 84 系经纬度坐标、网格中心点的 84 系经纬度坐标、中心点 84 系 3° 带高斯投影坐标、中心点 6° 带高斯投影坐标、与 54 系的转换参数、与 80 系的转换参数、网格的球面面积、平面面积、球面周长、平面周长、低纬边长、高纬边长、经线长等。

人口普查数据包括人口的位置信息和属性信息, 位置信息记录人口相对于网格中心点的相对坐标(即相对经度和相对纬度)和网格编码, 属性信息则是原有人口普查时需要采集的数据项, 主要包括身份证号码、姓名、性别、民族、出生日期、出生地点、籍贯、行政区划、门牌号、婚姻状况、文化程度等。PDA 数据采集表可根据实际需要采集的数据项进行扩充。

为了便于 PDA 采集的人口普查数据作为网格节点中的数据存入 SIMGPC 的网格数据库中, PDA 上采用 Pocket Access 数据库来存储采集的人口数据。Pocket Access 是与 PC 中的 Access 非常相似的数据库, 通过 ActiveSync 同步软件可以方便地将 Pocket Access 的“.cdb”文件与 PC 中的 Access 的“.mdb”文件相互转换和建立同步关系。然而, 在 PDA 中看不到 Pocket Access 数据库系统, 微软在 PDA 中只提供了 Pocket Ac-

cess 的数据库驱动, 而没有提供像 Pocket word、Pocket Excel 那样的 Pocket Access 数据库软件系统, 所以只能通过编程来实现对 Pocket Access 数据库的访问和操作。为了进行大量信息的输入, 并将同步后的 Access 数据库导入 SIMGPC 系统中的 Oracle 数据库中, 笔者在开发中采用了支持访问 Pocket Access 数据库的 ADOCE 数据库访问技术。由于 ADOCE 技术并没有相对应的 MFC 类库可以供开发者使用, 而直接调用 COM 库中的 ADOCE 操作函数又比较松散, 而且也很繁琐和复杂。因此, 笔者利用面向对象的技术重新对 COM 库中 ADOCE 操作函数进行了数据连接类和记录集类的封装, 以便在开发中使用。人口数据的录入、查询、修改、删除以及可视化显示等就是对 Pocket Access 数据库文件(*.cdb) 进行操作。

网格人口统计表结构与 PDA 数据采集表结构类似, 为了与网格元表连接, 添加了 MORTON 码等属性。

3 基于 SIMG 的人口普查数据采集系统设计与实现

3.1 系统体系结构

人口普查数据采集系统包括基本地图操作、网格划分及可视化、人口普查数据采集和 GPS 定位 4 个功能模块, 图 1 所示的系统体系结构包括了本系统的 4 个功能模块。从系统的逻辑结构来看, 本系统最底层为嵌入式操作系统 Windows CE, GPS 设备提供定位功能, 其上存储在本地的地图数据和人口普查数据, 系统的 4 个功能模块分别对地图数据和人口普查数据进行操作, 最上面一层是用户界面, 接受用户的操作并显示操作结果。

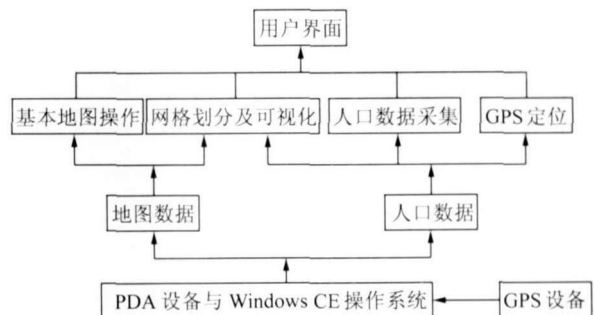


图 1 人口普查数据采集系统体系结构

Fig. 1 System Architecture of Population Census Data Collection System

3.2 主要功能设计

1) 基本地图操作功能, 包括地图显示、浏览、图层管理等。在 GIS 常见的地图显示、任意比例尺缩放、漫游、浏览全景图以及显示比例尺、指北针等方法的基础上, 利用图层管理、制图综合、地图符号图式化等方法, 使系统具有精简的内核和快速的浏览速度, 满足用户实际应用的需要。

2) 网格划分及可视化功能, 包括网格的生成、删除以及人口普查数据的可视化显示等。网格层位于地图图层之上, 网格划分的层级根据地图的比例尺来确定, 根据表 1 的网格划分精度分析, 实验系统采用的比例尺为 1: 500 的武汉大学测绘校区地图将划分到第 11 级网格。根据建立的空间索引、编码索引和专题索引, 可以使用网格空间位置或者网格编码查找网格对象, 在地图上显示网格所在位置以及网格中的人口信息, 还可以进行人口属性查询, 给定一个人的信息, 如姓名、身份证号码、门牌号等, 查询他所在的网格, 显示该记录点在地图上。通过 PDA 上生成的网格层和人口普查数据的可视化显示, 用户可以随时随地对人口数据进行查询和简单的统计分析。

3) 人口普查数据采集功能, 包括人口普查数据的入库、查询、修改、删除等。人口普查数据包含属性信息和位置信息, 属性信息即人口普查表中所列出的各个属性项, 位置信息包括网格编码、相对经度和相对纬度。通过 GPS 接收机获得定位点的 WGS84 经纬度坐标, 根据划分的网格, 可以得到该点所在网格的网格编码及其中心点坐标, 相对经度和相对纬度即是在该点的经纬度坐标与该点所在网格的网格中心点坐标的坐标相对量。PDA 采集的人口普查数据保存在 PDA 本地的 Pocket Access 数据库中, 通过与 PC 上 Access 数据库的相互转换或建立的同步关系, PDA 采集的人口数据将方便快捷地存入 SIMGPC 的网格数据库中。

4) GPS 功能, 主要提供 GPS 接收机的打开/关闭、基本参数的设置、获得定位点坐标、查看当前卫星状态和 GPS 数据等。根据 GPS 接收到的卫星定位数据, 对用户当前位置进行精确定位, 为用户提供人口普查数据的位置信息。

3.3 系统实现

本系统硬件采用 HP iPAQ Pocket PC 2210 掌上电脑; Pocket PC 2003 简体中文操作系统, Intel(R) PXA255 微处理器, 400MHz ARM, 32M ROM, 64M RAM, 内置 CF/SD 双扩展槽, 内置红外、蓝牙。同时, 集成 Rikaline GPS 6031 X7 型

蓝牙 GPS 模块。在整个系统的开发过程中, 基于 COM 模型的嵌入式 GIS 控件 hMap (Hand Map), 采用 Microsoft Embedded Visual C++ 4.0 进行开发。图 2 是生成的网格层界面; 图 3 是空间索引查询的结果, 用户在地图上感兴趣的区域使用点击笔点击, 即可得到该点所在的网格信息及其网格中的人口信息; 图 4 是编码索引查询的结果, 用户输入要查询的网格编码, 则可得该网格中的人口信息, 并将该网格和网格中的人口信息在地图上显示出来; 图 5 是专题索引查询的结果, 用户输入一个人的属性信息, 如姓名、身份证号码、门牌号等, 即可将此人在地图上显示出来。



图 2 网格生成界面
Fig. 2 Create Grid



图 3 空间索引查询结果
Fig. 3 The Result of Spatial Index Query

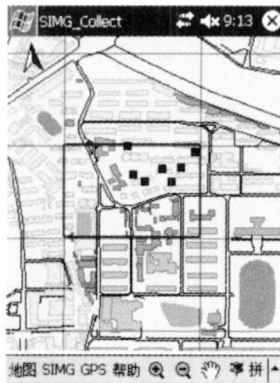


图 4 编码索引查询结果
Fig. 4 The Result of Code Index Query

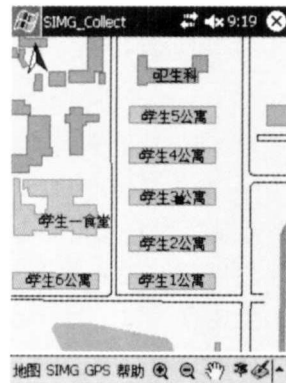


图 5 专题索引查询结果
Fig. 5 The Result of Special Index Query

SIMG 作为一种新的空间信息在计算机中的表示方法, 其最主要而且最有效地应用应当是国家和省市自然、经济、社会发展信息的及时获取、分析和在宏观决策中的应用。由于其数据结构与传统空间数据的差别, 在这类应用中 SIMG 将发挥其独特的优越性。

人口普查的成功不仅取决于人口资料的好坏,还取决于如何将人口数据和地理区域联系起来。基于 SIMG 的人口普查多级网格,有效的描述了人口的空间特性,能够为各级政府、部门社会提供人口地理信息的高水平综合性服务。SIMG-PC 实验系统的实现,为推动 SIMG 技术在国家和省市宏观决策走向实用作出尝试。采用带有 GPS 的手持 PDA 设备采集的人口普查数据,带有空间信息编码,可以方便快捷地导入 SIMGPC 系统的网格数据库中,确保了数据的准确程度和工作的效率,为人口普查的数据采集工作提供了一种新的尝试,北京市统计局计算中心预计在下一一次的人口普查工作中就可以应用手持 PDA 设备提高效率^[6]。

参 考 文 献

- [1] 唐棣, 刘德钦, 苏山舞. 关于建立中国人口地理信息系统的构想[J]. 测绘科学, 1997(4): 5-13
- [2] 陈述彭. 人口统计的时空分析[J]. 中国人口资源与环境, 2002, 12(4): 3-7
- [3] 金君, 印洁, 李成名, 等. 人口地理信息系统的设计与建设[J]. 工程勘察, 2002(3): 51-64
- [4] 李德仁, 朱欣焰, 龚健雅. 从数字地图到空间信息网

格——空间信息多级网格理论思考[J]. 武汉大学报·信息科学版, 2003, 28(6): 642-650

- [5] 李德仁, 邵振峰, 朱欣焰. 论空间信息多级网格及其典型应用[J]. 武汉大学报·信息科学版, 2004, 29(11): 945-950
- [6] 杨柳. 低端受冷落 高端寻突破 PDA 市场能否重塑? [OR/L]. http://www.enet.com.cn/enews/inforcenter/A20030701250726_1.html, 2003
- [7] Sahr K, White D. Discrete Global Grid Systems. Computing Science and Statistics. Interface Foundation of North America, Inc., Fairfax Station, VA., 1998
- [8] Dutton G. Universal Geospatial Data Exchange via Global Hierarchical Coordinates[C]. International Conference on Discrete Global Grids, Belmont, 2000
- [9] Goodchild M F, Yang Shiren. A Hierarchical Spatial Data Structure for Global Geographic Information Systems[J]. Computer Vision, Graphics and Image Processing, 1992, 54(1): 31-44
- [10] 傅曦, 齐宇. 嵌入式系统 Windows CE 开发技巧与实例[M]. 北京: 化学工业出版社, 2004

第一作者简介: 周俐俊, 硕士生。研究方向为 GIS 应用。

E-mail: zhoulijun@lmars.whu.edu.cn

Design and Implementation of Population Census Data Collection System Based SIMG

ZHOU Lijun¹ ZHOU Xinyan¹ SHAO Zhenfeng¹ SONG Heng¹

(1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: The core idea of Spatial Information Multirgrid (SIMG) is introduced and the structure of spatial information multirgrid for population census (SIMGPC) is designed. The key technologies are expatiated on the method to measure off grid, code and SIMG database. A population census data PDA collection system with spatial information code is designed and implemented.

Key words: GIS; spatial information multirgrid (SIMG); spatial information multirgrid for population census (SIMGPC)

About the first author: ZHOU Lijun, postgraduate, majors in GIS application

E-mail: zhoulijun@lmars.whu.edu.cn