

基于全球剖分框架的多源空间信息区位关联与综合表达方法

杨宇博¹ 程承旗² 郝继刚³ 陆楠¹

(北京大学遥感与地理信息系统研究所 北京 100871)¹ (北京大学工学院 北京 100871)²

(北京遥感信息研究所 北京 100192)³

摘要 针对全球海量多源空间信息之间关联关系建立困难的问题,提出了基于全球剖分框架的多源空间信息区位关联模型解决方案。首先,采用 GeoSOT 剖分框架对多源空间信息进行统一组织与管理;然后,根据提出的“对象-面片-数据”空间信息三层组织机理概念模型,设计并实现了基于全球剖分框架的多源空间信息区位关联模型;最后,进行了多源空间信息区位关联模型的综合表达应用实验。实验结果表明:本模型可以有效地建立起地理对象与多源空间信息之间的区位关联关系,可有效提升海量多源空间信息的空间关联与表达应用效能。

关键词 地球剖分,GeoSOT 剖分框架,多源空间信息,区位关联,综合表达

中图分类号 P208 文献标识码 A

Area-location Association and Comprehensive Expression of Multi-source Spatial Information Based on Global Subdivision Framework

YANG Yu-bo¹ CHENG Cheng-qi² HAO Ji-gang³ LU Nan¹

(Institute of Remote Sensing and GIS, Peking University, Beijing 100871, China)¹

(College of Engineering, Peking University, Beijing 100871, China)² (Institute of Remote Sensing Information, Beijing 100192, China)³

Abstract Traditional GIS lacks the ability to establish the relationship between geographic object and its related global massive multi-source spatial information. GeoSOT subdivision framework was used to organize and manage multi-source spatial information in a unified manner, and "Object-Mesh-Data" three-layer spatial information organizational mechanism was put forward. Based on this mechanism, multi-source spatial information area-location association model was designed to support association and query of spatial information. In order to validate this model, multi-source spatial information comprehensive expression experiment was carried out. The experiment result proves the effectiveness of this model, which can improve the application performance of massive multi-source spatial information.

Keywords Global subdivision, GeoSOT subdivision framework, Multi-source spatial information, Area-location association, Comprehensive expression

1 引言

在 GIS 系统中,反映同一空间事物的信息以多种形式存在,导致了空间信息多源性的产生。空间信息的多源性主要表现在获取手段多源性、存储格式异构性、多时空与多尺度性、多语义性、分布式特性等方面^[1]。在现有空间信息管理模式下,若要查询某一区域或地理对象相关的空间信息,需要将其空间范围按照不同数据的组织基准与坐标系统进行转换,才能在特定数据源中进行查询操作。当数据源增多时,这一过程愈加繁琐,降低了空间信息的使用效率。

在上述背景下,空间信息的多源性导致地理对象与不同数据间的关联关系难以建立,而现有的数据转换与空间数据互操作方法都不能从根本上解决此问题,需要发展一种全新的空间数据模型^[2]来支持多源空间信息的高效集成与综合运

用。全球剖分框架既符合计算机对空间信息离散化处理的需求,又摒弃了地图投影的束缚^[3],为统一组织与管理全球海量、多源、多尺度空间信息提供了一种新的思路,有利于建立地理对象与多源空间信息间的关联关系^[4,5]。

2 研究背景

全球剖分框架主要研究如何将球面剖分为等面积和等形状的层次状面片,具有基于全球、多分辨率、支持尺度变换、空间位置分布均匀、融合空间索引机制等优点,国内外学者提出了多种全球剖分框架^[6-9]。

本文采用的 GeoSOT 剖分框架^[10]属于等经纬度的四叉树剖分网格体系,其核心思想是将度、分、秒三级网格进行逻辑上的外延(见图 1),使得网格编码具有整数特征,易于计算与储存。其 0 级网格定义为以赤道与本初子午线交点为中心

到稿日期:2012-07-06 返修日期:2012-10-15 本文受国家 973 计划项目(61399)资助。

杨宇博(1985-),男,博士,主要研究方向为地理信息系统研究与工程应用,E-mail:yyb@pku.edu.cn(通信作者);程承旗(1961-),男,博士,教授,主要研究方向为地理信息系统工程;郝继刚(1978-),男,博士,主要研究方向为遥感影像处理;陆楠(1984-),女,博士,主要研究方向为地理信息系统研究与工程应用。

点的 $512^{\circ} \times 512^{\circ}$ 方格, 依次向下四分至 9 级网格, 大小为 $1^{\circ} \times 1^{\circ}$ (其中部分网格没有地理含义, 不再向下划分, 下同), 10 级网格将起始数值空间大小由 $60'$ (即 1°) 外延到 $64'$ 后, 继续向下四分至 15 级网格, 大小为 $1' \times 1'$, “秒”级网格剖分方式参照分级网格, 秒以下 22—32 级网格严格按照四分方法划分和编码。由此形成完整的 32 层 GeoSOT 剖分框架, 其二进制一维编码如图 2 所示。

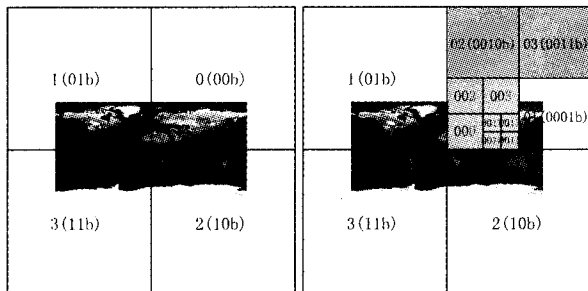


图 1 GeoSOT 剖分网格 1—4 级剖分方案

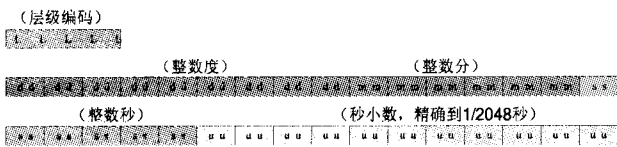


图 2 GeoSOT 剖分框架 32 层二进制一维编码

基于 GeoSOT 框架组织全球多源空间信息, 使得各类信息具有了统一的组织基准, 并且 GeoSOT 剖分框架所具有的多尺度层次特性, 有利于多源空间信息的区位关联, 特别是在针对不同尺度间多源信息的区位关联操作方面, 相比于传统方法具有较大优势。

3 基于 GeoSOT 的多源空间信息区位关联模型

3.1 模型的基本原理与架构

多元空间信息的区位关联, 是指根据空间信息特有的区域与位置属性特征, 在统一组织框架下, 依据一定的规则与方法, 通过空间信息间的共有区位特征, 建立起多源空间信息与地理对象间的关联关系。

在 GeoSOT 全球剖分框架下, 以剖分面片为基本单元对数据进行分块标识操作, 使多源空间数据具有了统一的组织基准。同时, 将地理对象所在区域也使用剖分面片的组合表示^[1], 即使用剖分面片作为地理对象与空间数据的“中间层”, 形成“对象-面片-数据”的空间信息三层组织机理, 以替代传统模式下“对象-数据”的两层组织方式, 三者间的关系如图 3 所示。

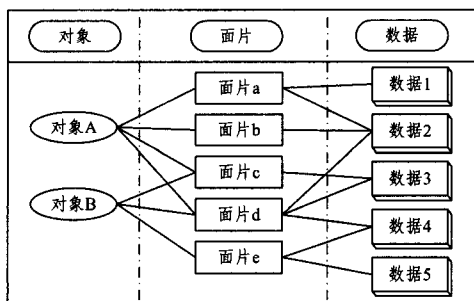


图 3 空间信息的三层组织机理

基于“对象-面片-数据”的空间信息三层组织机理可以提升空间数据的检索和查询效率, 进而提高数据的应用效能。

以此为基础, 可以实现基于全球剖分框架的多源空间信息区位关联, 模型架构如图 4 所示。

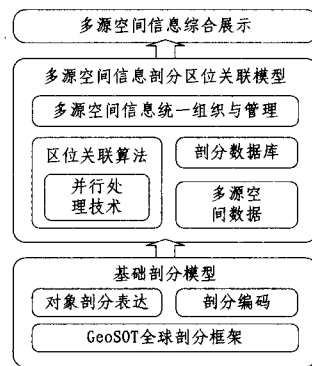


图 4 多源空间信息剖分区位关联模型架构

从图 4 中可以看出, 多源空间信息区位关联模型的基础剖分模型包括 GeoSOT 全球剖分框架、对象剖分表达及剖分编码等 3 个部分; 在此基础上设计的区位关联模型, 其数据来源为剖分数据库和多源空间数据, 区位关联算法采用并行处理技术, 可以实现多源空间信息的统一组织与管理, 从而最终完成多源空间信息综合展示。

3.2 区位关联的具体步骤及算法

在剖分体系下进行区位关联操作, 需要使用 GeoSOT 框架对多源空间数据进行统一划分与组织。对于影像数据、数字高程数据等空间数据, 划分方式分为物理划分与逻辑划分两种: 物理划分指对数据文件进行物理分块划分, 同时生成面片金字塔, 使影像数据完全按照 GeoSOT 面片进行存储与组织; 逻辑划分指不改变数据的存储结构, 从逻辑上对其进行多层划分, 生成划分结果对应的索引文件, 记录每个逻辑面片的起始点在影像文件中的像素坐标值。对于文字、图片、视频等非传统意义上的空间数据, 在 GeoSOT 剖分框架下, 也可以根据其描述或者获取点的地理位置信息, 对其赋予剖分编码。

上述划分操作使多源空间数据具备了区位关联的数据标识基础, 具体的区位关联操作包括以下步骤(算法流程见图 5): 首先, 获取某一区域对应的编码集合; 然后, 逐一将集合中每个编码与编码库中已有编码进行比较操作, 若没有找到相同编码, 则返回空值, 若找到相同编码, 则在该区域的关联数据中记录相同编码对应的数据源信息。

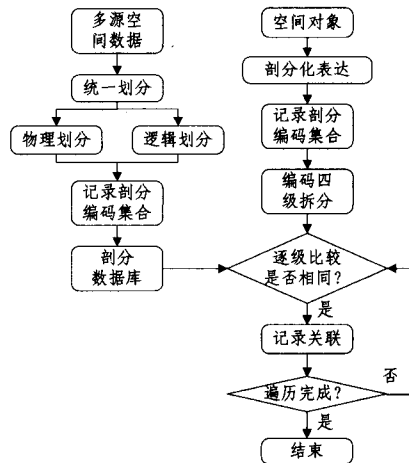


图 5 剖分区位关联算法流程图

在设计区位编码与编码库中已有编码进行比较的具体方法时, 考虑到编码库中记录条目数可能较多的原因, 将编码分

为度、分、秒、秒以下 4 段(划分方式见图 2)。按照由度级编码段到秒以下编码段的次序分别与剖分数据库中的数据源剖分编码集合进行对比,任何一级编码段不同则都不需继续向下一级编码段比较,具有相同编码段则继续向下一级编码段比较,以此提高对比效率。

4 区位关联与综合表达实验

传统的多源空间信息关联方法需要在数据库中记录目标信息,并根据目标标识 ID(一般为按记录先后而定的顺序码)对不同表中的目标相关信息进行关联。本文基于剖分区位关联模型,可以直接通过剖分区位标识进行多源空间信息之间的关联操作,减少了维护关系型数据库中目标信息的工作量,同时通过并行处理技术可以有效提高信息关联效率。以此为基础,本文实现了多源空间信息的综合表达方法,并开发出原型验证系统。

图 6 概括了地理对象相关联多源空间信息的综合表达流程。基于剖分区位关联模型,使用区位标识编码库中的编码在多源空间数据库中提取多源空间数据,并在基于 WorldWind 的剖分影像数据球平台上完成综合表达实验。

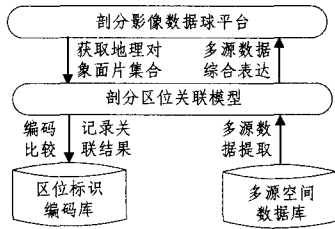


图 6 多源空间信息综合表达流程图

在实验中,随机选取某区域内 100 个互不重叠的地理对象,并选取该区域内影像数据、DEM 数据以及部分地理对象相关文字、图像数据,对上述数据物理划分并编码,进行区位关联实验。图 7 是对北京大学理科教学楼进行综合表达的系统界面。

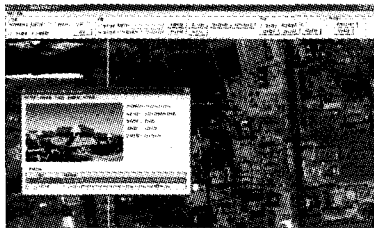


图 7 综合表达系统界面

表 1 区位关联计算时间对比

对象数	剖分面片层级	面片总数	线程数	平均关联时间/ms	并行加速比
100	16	152	1	0.57	3.563
			4	0.16	
100	17	483	1	1.98	3.600
			4	0.55	
100	18	1537	1	6.48	3.620
			4	1.79	
100	19	4995	1	19.50	3.631
			4	5.37	
100	20	16833	1	68.82	3.663
			4	18.79	

为了体现出算法的并行特性,结合 OpenMP 并行处理技术^[12]实现区位关联模型在算法层级的并行,并记录关联操作

的时间,对比并行前后的结果。从表 1 中可以看出,通过在不同层数据中使用并行技术,当设置为 4 线程时可以使关联效率提升 3.6 倍左右;同时随着剖分层级的升高,并行算法的加速比逐步升高。

为了得到更直观的分析结果,将面片个数与平均关联时间分别取 log,并作出不同剖分层级并行前后平均关联时间的对比分析图,如图 8 所示。从图 8 中可以看出,平均关联时间随着面片数量的增加而增加,通过回归分析可以推测出平均关联时间与面片数量基本呈正比关系。也就是说,选择更高精度层的面片进行关联操作会使关联时间逐步增加,因此,如何选择合适的剖分层级进行关联操作尤为重要。

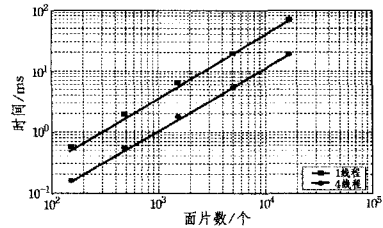


图 8 计算时间对比分析图

结束语 海量多源空间信息在 GIS 系统中的应用面临着新的挑战。本文提出的基于 GeoSOT 剖分框架的多源空间信息区位关联模型充分考虑了各类型空间信息的特点,提高了空间信息的应用效能,并具备较高的区位关联效率。关于增大剖分层级对关联效率的影响问题,还需要进一步研究如何根据不同数据来源与不同的用户需求,选择合理的剖分面片层级进行区位关联操作。

参考文献

- [1] 李德仁,邵振峰,朱欣焰.论空间信息多级格网及其典型应用[J].武汉大学学报:信息科学版,2004,29(11):945-950
- [2] 王家耀.军事地理信息系统(MGIS)在现代化战争中的作用及其发展[J].信息工程大学学报,2000,1(4):102-105
- [3] 童晓冲.空间信息剖分组织的全球离散格网理论与方法[D].郑州:解放军信息工程大学,2010
- [4] 张雪伍,苏奋振,石忆邵,等.空间关联规则挖掘研究进展[J].地理科学进展,2007,26(6):119-128
- [5] 马荣华,马小冬,蒲英霞.从 GIS 数据库中挖掘空间关联规则研究[J].遥感学报,2005,9(6):733-741
- [6] Ottoson P, Hauska H. Ellipsoidal Quadrees for Indexing of Global Geographical Data[J]. Int. J. Geographical Information Science, 2002, 16(3): 213-226
- [7] Goodchild M. Discrete Global Grids for Digital Earth[C]// Proceedings of International Conference on Discrete Global Grids. Santa Barbara; [s. n.], 2000
- [8] 李德仁,肖志峰,朱欣焰,等.空间信息多级网格的划分方法及编码研究[J].测绘学报,2006,35(1):52-56
- [9] 程承旗,关丽.基于地图分幅拓展的全球剖分模型及其地址编码研究[J].测绘学报,2010,39(3):295-302
- [10] 程承旗,任伏虎,濮国梁,等.空间信息剖分组织导论[M].北京:科学出版社,2012
- [11] 杨宇博,程承旗,宋树华.面向地理对象多尺度表达的剖分编码方法研究[J].地理与地理信息科学,2010,26(5):33-36
- [12] Barbara C, Gabriele J, Ruud V D P. Using OpenMP: Portable Shared Memory Parallel Programming[M]. The MIT Press, 2007