

基于 ORDB 的 Web GIS 空间数据管理研究

A Study on the Management of Spatial Data of Web GIS Based on ORDB

刘长岐 甘国辉

(中国科学院地理科学与资源研究所 北京 100101)

进入 20 世纪 90 年代以来, 随着网络技术的发展, 原来的基于单机模式的 GIS 已经无法满足用户对共享地理信息的需要, 越来越多的人期通过 Internet 获取地理信息, 基于网络的 GIS 顺势而生。由于大多数网络 GIS 的客户端采用了 WWW 协议, 因此学术界把类似的地理信息系统称为 Web GIS。其主要目的就是在网络环境下实现地理信息的存储、处理、分析与显示, 使用户能通过通用浏览器 (IE 或 NS) 访问并获取地理信息^[1]。

一、Web GIS 的特点和空间数据可视化的过程

Web GIS 作为当前 GIS 领域的研究热点和未来的发展方向, 相较于传统的 GIS, 有以下几个特点^[2]: (1) 作为 GIS 在广域网环境下的一种应用, Web GIS 是 GIS 与分布式计算技术以及 Internet 技术相融合的产物; (2) 客户端具有平台独立性, 即不论客户端的机器采用哪种操作系统, 只要是支持通用浏览

(2) 蛋白质 原子力显微镜在膜蛋白、游离的蛋白质分子及结晶的蛋白质等 3 个方面都有重要应用。例如, 可以用原子力显微镜研究蛋白质结晶的动态过程, 最大的优点是蛋白质晶体在母液中的生长过程, 可以在几乎不受外界干扰下用图像的方式记录下来, 用于研究大分子晶体生长的机制及一些相关现象。

(3) 多糖 由于多糖分子往往带有支链, 分子的均一性及线性不如 DNA 和蛋白质好, 相对而言得到的图像的分辨率差一些, 但是近年来这方面的发展很快, 取得了很多进展。具体表现为: 用原子力显微镜观测多糖分子的高级结构; 用 AFM 观测二维多糖网络结构, 并研究了多糖浓度及几种离子在不同浓度下对凝胶网络形成的影响; 用 AFM 直接观察以纤维素微纤维素为主体的植物细胞壁。

近来 Kuznetso 等用 AFM 与计算机模拟技术相结合的方法来研究蛋白质晶体的空间结构, 这为研究 DNA 等生物大分子的空间三维结构提供了重要启示。尽管 AFM 的分辨率受到探针尖端曲率半径的制约, 但由于其良好的性能, 无疑将会在生物大分子结构研究中发挥重要作用。

5. 原子力显微镜在其他研究领域的应用简述

(1) 在高分子领域的应用 原子力显微镜由于发明不久即应用于高分子领域, 弥补了扫描隧道显微镜不能观测非导电样品的缺欠, 因而受到重视, 应用范围也不断扩展。最近几年, 原子力显微镜的

应用已由对聚合物表面几何形貌的观测发展到深入研究高分子的纳米级结构和表面性能等新领域, 并由此导出了若干新概念和新方法。

(2) 在粘土矿物学研究中的应用 研究粘土矿物表面组成与表面结构的手段和方法多种多样, 但由于粘土矿物特殊的晶体结构, 这些手段和方法都有一定的局限性和不足。原子力显微镜能够分辨粘土矿物硅氧四面体片上的六方环结构及八面体片的羟基团, 是研究粘土矿物表面反应、表面改性及表面溶蚀作用的重要手段。

(3) 在生命科学领域的应用 原子力显微镜在研究生物大分子结构中的应用前面已经提到。此外, AFM 用于“透明质酸”、“髓鞘质”的研究也有了新的进展; 用于纳米水平研究染色体的超级结构时, 为下一级结构的研究打开了缺口。另外, 还用于研究活细胞的力学响应, 这为用物理方法制服病毒提供了可能性。就医学研究而言, 在一些领域内已经初步具备了进入临床病理学检验的基本条件。

综上所述, 原子力显微镜是功能强大的重要工具, 具有很多优点, 已经显示出它的巨大作用, 并且随着制造技术的不断完善和提高, 随着人们对它的了解的加深, 原子力显微镜必将不断拓宽应用领域, 显示出新的强大功能, 从而在很大程度上促进多个领域的研究进展。

(删去全部 25 篇英文、11 篇中文参考文献目录)

(责任编辑 王宏章)

器 (IE 或 NS), 即可访问服务器并获取所需的地理信息; (3) 服务器端的空间数据具有分布性, 即 Web GIS 的空间数据可能分布在 Web 上的多个服务器上, 根据用户提交的请求来决定从哪个服务器上获取信息。

当前学术界对 Web GIS 的研究主要集中在空间数据结构与模型、空间数据传输协议、分布式策略等方面。而 Web GIS 的原型产品, 其构造方法也不外乎文件共享、CGI(Common Gateway Interface) 方法、GIS 插件模式以及基于 COM/DCOM(运行于 Windows/Win NT 平台) 或 CORBA(运行于 UNIX 平台) 的组件技术等^[3]。但笔者认为, 作为当前业界研究重点的 Web GIS, 其核心问题依然是如何合理、高效地管理和传输空间信息, 基于此, 对 Web GIS 的空间数据管理技术进行研究, 就具有重要的现实意义。

根据 Allan Doyle 提出的在因特网上访问异构空间数据源的分步骤服务模型, Web GIS 空间数据的可视化要经过以下 4 个过程^[4]: (1) 从空间数据源中选择出要显示的地理实体数据; (2) 把选择出来的地理实体数据组合生成一个显示元素的序列; (3) 将显示元素按顺序生成最终要显示的地图结果; (4) 将准备好的地图送往显示设备进行最终显示。其中第一个过程即选择空间数据应该是后续操作的前提, 是 Web GIS 空间数据管理重点研究和解决的内容, 能否有效解决这一问题, 是事关 Web GIS 系统优劣的关键。而如何合理高效地从空间数据库中选出需要的空间数据, 则主要取决于 Web GIS 底层空间数据采用何种数据库进行管理。

下面本文将首先讨论 Web GIS 底层空间数据管理可采用的 3 种模式, 通过比较, 提出适合 Web GIS 底层空间数据管理的最优模式。

二、Web GIS 空间数据管理的 3 种模式

早期的 Web GIS, 其实现技术包括文件共享(局域网环境下通过将文件共享或目录映射实现网络操作)、CGI(通用网关接口, Common Gateway Interface) 模式(Client/Server 架构下, 客户端通过 CGI 动态访问 GIS 数据服务器, 获得静态的 HTML 处理文档) 以及 GIS 插件技术等。这些技术实现的 Web GIS 产品, 通常只能提供一些简单的功能, 如图形显示和查询等, 不能进行复杂的空间数据分析。随着分布式对象技术的出现, 建立在此基础上的构件式 Web GIS 成为当前主流的 Web GIS 解决方案。但不论是早期的 Web GIS, 还是当前的主流 Web GIS, 其实质仍然是通过 Web GIS 实现对空间数据的操作, 其空间数据的有效管理依然是在开发 Web GIS 产品前首先要考虑的一个问题。换言之, Web GIS 的网络性主要体现在数据访问方式上, 在底层的空间数据

管理上和单机 GIS 没有本质区别^[5]。根据数据库技术的发展现状, 当前的主流数据库产品主要有关系型、面向对象型和对象—关系型 3 种^[6], 相应地, Web GIS 空间数据管理的模式也有 3 种, 下面分别对这 3 种模式进行分析。

1. 纯关系型模式 关系型数据库具有坚实的数学基础, 经过 30 多年的发展, 已经无可置疑地成为当前数据库产品中最为成熟的。现在流行的数据库产品, 如 Oracle、SQL Server、DB2 等, 均建立在关系型理论的基础之上。关系型数据库管理系统(RDBMS) 对数据的管理是基于简单的二维表模式来实现的。如果只是管理属性数据, 它可以完成相当复杂的关系操作, 而且在数据完整性、安全性、独立性、高效的数据访问以及减少数据冗余度方面具有明显的优势。但对于空间数据的管理, 纯关系型模式则存在着先天的不足。首先, 空间数据不同于简单的属性数据, 它包含有坐标等空间信息, 如对一个简单的点状实体, 就需要存储其 X 坐标、Y 坐标甚至是 Z 坐标。要存储结构如此复杂的空间数据, 就必须设计复杂的 E—R 模型, 并在数据库中存放大量的二维表, 这将无限增大系统的容量和复杂性; 其次, 对于纯关系型数据库而言, 空间数据和属性数据通常不能放在一张二维表里, 割裂了逻辑上应为一个整体的地理信息, 难以保证数据的一致性。由于这些缺点, 虽然学术界对纯关系模式的 Web GIS 空间数据管理作了一些尝试, 但最终没有成为主流。

2. 面向对象模式 面向对象模式的核心是采用面向对象数据库(OODB) 管理空间数据。面向对象数据库是面向对象思想与数据库管理系统相融合的产物, 它引入了程序设计语言中的对象和类的概念。在面向对象的程序设计语言中, 我们知道, 对象是其自身所具有的状态特征及可以对这些状态施加的操作结合在一起所构成的独立实体, 类是对一组客观对象的抽象。在 OODB 中, 其管理的数据称为对象, 而实际操作则针对类来进行。由于对象本身的特性, 使 OODB 能够有效描述现实世界中的客观实体, 其语义表达能力强于纯关系型数据库; 同时, 由于 OODB 引入了对象和类的概念, 使得它具有包括抽象、继承、封装等一系列面向对象特征。这些特征使得 OODB 不仅能管理对象的状态, 还可以管理对象的行为。因此, OODB 具有 RDB 所不具备的对数据进行抽象描述的能力。正是因为 OODB 能够对数据进行抽象, 使得 OODB 的结构与行为是一致关联的, 因而 OODB 具有表达、演绎、推理知识的能力。由于 OODB 具有上述明显的特征, 理论而言, 它可以比 RDB 更为有效地管理空间数据。但是, 由于 OODB 的语义数据模型比较复杂且缺乏坚实的数学理论做基础, 使得业界对面向对象

的数据库系统难以形成一个统一的、大家公认的数据模型以及面向对象的数据库语言。因此,可以说 OODB 尚未成熟, 尽管已经对 OODB 进行了有益的探索, 也有了成型的面向对象的数据库管理系统^[7], 但作为非主流的数据库产品, OODB 还很难真正用于空间数据的管理。

3. 对象—关系模式 对象—关系模式是上述两种模式之间的一种折中方案, 主要是通过扩充关系型数据库的某些功能, 改进包括支持自定义数据类型、变长字段以及对象、方法、引用等面向对象的特性, 使其支持面向对象的数据模型, 这就是所谓的对象—关系型数据库 (ORDB)。当前一些大的数据库产品, 如 Oracle 8i、DB2—5 以上的产品都是这种系统。其优点在于既可充分利用成熟的关系型数据库强大的数据管理功能和可操作性, 又兼具面向对象数据库的对象和类的概念及数据模型的扩充, 可以很好地实现空间数据的管理, 其应该是当前 Web GIS 产品底层数据库管理的首先解决方案。

三、基于 ORDB 的 Web GIS 空间数据管理

1. 空间数据建模 空间数据具有多源性、多语义性、多时空性、多尺度等特点, 因此, 设计和建立能正确表达空间数据的模型是合理、有效管理空间数据的基础。尽管世界上一些发达国家的研究机构, 如国际标准化组织 (ISO) 的 TC211 小组、欧洲标准化组织 (EN/ TC28)、美国联邦地理数据委员会 (FGDC) 以及 OPEN GIS 联盟 (OGC) 等, 一直致力于空间数据的标准化研究, 但到目前为止, 还没有一种能够被广泛采用的描述空间对象的数据模型^[8]。

Web GIS 的网络性主要体现在数据访问方式上, 在底层的数据建模、数据存储方面和单机 GIS 并无本质区别。对基于 ORDB 进行底层数据管理的 Web GIS 而言, 用面向对象的观点来看, 用户可以将地理空间数据抽象为空间实体、图层、地图 3 个层次。其中空间实体是指具有一定几何形态和空间分布特征的事物和现象, 图层则是一定空间范围内具有相同属性的空间实体组成的有机集合, 而地图则是为表达更现实的语义而将图层按一定条件进行的合理组合。空间实体作为地理空间数据的基本元素, 具有两个最基本的特征, 一是属性要素, 一是空间要素^[9]。对空间实体的属性要素, 可以采用二维关系表的形式直接在数据库中存储和管理, 对空间实体的空间要素, 可通过定义长二进制字段以矢量方式按点、线、面的关系存储, 这样即可保证属性要素和空间要素的一致性。

2. 空间数据分布式计算 Web GIS 空间数据管理的一个很大的特点是空间信息的处理由服务器和客户端分担, 其中服务器端主要负责数据的组

织、管理和数据提取等功能, 而客户端则主要提供用户的数据分析功能并在一定的条件下向服务器发送数据请求。基于 ORDB 的 Web GIS 比之传统的 GIS, 其空间数据和属性数据统一存储在数据库中, 数据量更大, 数据表达更为复杂, 且数据分布更为分散, 因此, 如何提高空间数据的访问效率显得尤为重要。一个比较合理的解决方案是: 将消耗时间短而数据抽取率高的操作分布于客户端 (如不相关、相等、相邻等空间查询操作), 将空间运算复杂而数据抽取率低的操作分布在服务器端。这样既可以充分利用服务器的高速运算能力, 又可减小网络流量。

在基于 ORDB 管理底层数据的 Web GIS 中, 采用行之有效的空间索引技术对提高空间数据的访问效率也起着关键作用。所谓空间索引, 就是指依据空间实体的位置和形状或空间实体之间的某种空间关系, 按一定顺序排列的一种数据结构, 其中包含空间实体的概要信息如对象的标识、外接矩形及指向空间实体数据的指针。常见的空间索引一般是自上而下、逐级划分空间的各种树结构空间索引, 比较典型的有 R_树、R+_树等^[10]。

另外, Web GIS 空间数据访问经常遇到的一个问题是: 一个用户在客户端向服务器发出一个数据查询请求, 而其下一个查询请求所需的数据可能仍在上一个请求查询的范围之内, 为了加快系统的响应速度, 减少网络流量, 可以借鉴计算机硬件中的缓存处理技术^[11], 在服务器端通过直接在内存中建立与磁盘文件结构类似的内存数据库, 用于存储临时空间数据, 提高数据的访问效率; 在客户端, 通过在本节点建立数据缓存, 暂时存放刚刚从服务器获得的空間数据, 当再次访问这些数据时, 即可直接从本地缓存中读取, 这样不但可以减小网络传输流量, 而且可以明显提高数据的访问速度。

四、结论

随着遥感、GPS、全站仪数字化成图等空间数据采集手段的不断提高, GIS 空间数据库的容量呈几何倍数增大。目前, 对于一个中等城市级的 GIS 系统而言, 其数据量一般可达到 TB 级, 如我国已经建成的 1: 25 万地形要素数据库的数据量即达 8GB, 而正在建设的 1: 5 万地形要素数据库将更大^[12]。合理、高效地管理海量空间数据一直是 GIS 的研究核心和发展瓶颈。Web GIS 作为 GIS 在 Internet 上的应用, 是 GIS 走向社会化和大众化的有效途径, 也是 GIS 发展的方向之一。

Web GIS 底层的空间数据管理与单机 GIS 并无本质区别。目前许多大型商业数据库都支持对空间数据的管理, 如 Oracle 在 Oracle 8i 以后的版本中推

超弦——大自然的“琴弦” The Superstring—The “String” of Nature

桂起权¹ 贺天平²

(武汉大学哲学系, 教授¹ 武汉 430072; 山西大学哲学系, 博士生²)

一、弦理论的基本思路概要

记得有一位哲学家说过, “你要知道梨子……, 就要……。而你若要知道原子的组织同性质, 就得实行物理学和化学的实验”。现代科学告诉我们, 包括梨子在内的世界万物都由原子所组成, 而原子又由电子与夸克所构成。然而, 根据最新的超弦理论, 所有这些粒子实际上又都可以分析为在超微小尺度上振动着的一根闭合的弦。也就是说, 一切物质, 作为其终极实在的不是点状的粒子, 而只是细线状的小环。被人们认作粒子的, 实质上都像一根无限纤细的橡皮筋, 或似一根振荡、跳动的丝线。在超弦理论中, 把这种作为宇宙万物结构与功能的最小单元称作弦。图 1 即为此示意图——一个普通的梨子, 通过逐级分解(并放大), 可以显示出越来越小的结构组分, 最终归结为处在一定振动方式上的弦。

科学家始终以追寻各种自然力的统一和物质基本组成的统一为崇高理想。正因为如此, 爱因斯坦才付出了整个后半生的精力艰难地探索统一场论。只是当时时机、条件还不够成熟, 人们对 4 种基本相互作用了解得太少, 强力与弱力都尚未登上科学的历史舞台。爱因斯坦是过早地走到了他的时代的前头。不过在 50 多年之后, 爱因斯坦的统一理论

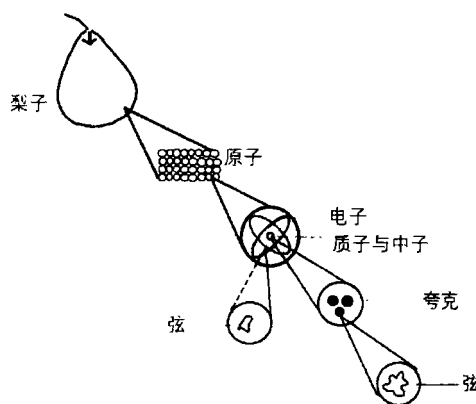


图 1 一个普通梨子的内部结构组分, 可以从原子一直分解到电子、夸克和弦

的理想又重新成为注意力的焦点。现在越来越多的学者趋于相信, 弦理论最有可能提供囊括一切自然力和物质的解释框架。例如, 弦理论认为, 现已观测到的基本粒子性质(见表 1)和 4 种基本的相互作用的表现方式(见表 2)无非是在表明, 作为宇宙最小结构与功能单元的基本弦存在着多种不同的振动模式。

弦理论的研究, 使人不能不想起古希腊毕达哥

出了 Oracle 8i Spatial, 提供了对象—关系数据模型, 在关系型数据库基础上引入面向对象技术, 以关系型数据库管理属性数据, 以面向对象技术处理地理空间数据, 实现了属性数据和空间数据的一体化管理, 提高了地理空间数据的运算效率。因此, 采用对象—关系型数据库 (ORDB) 管理海量空间数据应该是 Web GIS 较为理想的数据库管理平台。

参考文献

- [1] 方裕, 周成虎等. 第四代 GIS 软件研究, 中国图像图形学报, 2001, 6(A) (9), 817~ 823
- [2] 赵需生, 杨崇俊. Web GIS 的设计与实现, 中国图像图形学报, 2000, 5(A) (1), 76~ 79
- [3] 罗英伟, 汪小林等. 计算机辅助设计与图形学学报, 2002, 14(8), 790~ 795
- [4] 赵需生, 赵鲁等. <http://www.digitalearth.net.cn/>

- [5] 伍健. 分布式构件化 Web GIS 空间数据管理研究. 北京大学硕士论文, 2000-6
- [6] 丁力, 汪小林等. 中国图像图形学报, 2001, 6(A), (11), 76~ 79, 1101~ 1106
- [7] 冯玉才编著. 数据库系统基础(第二版). 武汉: 华中理工大学出版社, 1993
- [8] 周成虎, 李军. 地球空间元数据研究. 地球科学, Vol. 25(6), 579~ 585
- [9] 陈述彭, 鲁学军, 周成虎编著. 地理信息系统导论, 北京: 科学出版社, 1999
- [10] 龚健雅著. 整体 GIS 数据组织与处理方法. 武汉: 武汉测绘科技大学出版社, 1993
- [11] 李慕华. 分布式 Web GIS 构件化的研究和实现. 北京大学硕士论文, 2002-6
- [12] 常燕卿. 大型 GIS 空间数据组织方法初探. 遥感信息, 2000(2), 28~ 31 (责任编辑 王宏章)