

文章编号: 1007-6301 (2000) 03-0203-09

地球空间数据集成研究概况

李 军^{1,2}, 费川云²

(1. 中国科学院遥感应用研究所, 北京 100083; 2. 中国科学院地理科学与资源研究所, 北京 100101)

摘要: 数字地球空间数据的集成研究及应用始于本世纪 60 年代, 地理信息系统的出现及应用、多元数据的使用推动了地球空间数据集成研究及其应用。本文首先对地球空间数据概念进行描述, 然后对地球空间数据集成研究状况及存在的问题进行了详细的分析。结合数据集成需求及存在的问题, 论文分析了数据集成的发展方向。

关 键 词: 地球空间数据; 数据集成; 无缝数据集

中图分类号: N941 文献标识码: A

遥感及其它数据获取技术的应用使地球空间数据量迅速增加, 内容和形式的多元化已成为地球空间数据存在的特征; 需解决的问题复杂化、涉及的内容领域多样化, 要求使用多种数据源, 所以在应用项目中, 数据集成使用是不可避免的。地球空间数据有自己的特点, 对它的集成研究应从对这类数据的基本认识开始。

1 地球空间数据

数据与信息有直接关系。信息是物质存在和运动形式的一般描述 (周荫清, 1993), 是客观世界在人脑中的反映; 数据是由数字、字母和符号等表达的信息 (陈学谔, 1996), 是描述事物的状况、行为及其结果的有特定含义的字符或数字 (Martin, 1992)。两者有一定的区别和联系。数据指由观测者记录下来, 并存贮到计算机中的数值, 而信息则是数据使用者理解的这些数值的意思 (David J. Maguire, 1989), 如数字形式的全球平均气温值是数据, 而这些数据数值表达的冷暖状况则是信息。

地球空间数据是数据的一种类型。关于地球空间数据有几类概念: A 地理数据 (Geographical data (E. E. Shiryaev, 1987) 或 Geographic data (Harlan, 1995)), 强调数据与地理位置有关, 或者称为地理信息 (Geographic Information) (I. D. H. Shepherd, 1991), 泛指地球空间数据; B GIS 数据 (GIS data) (Susan, 1996) 或地球科学数据 (Earth-sciences data or Geoscientific data) (Graeme F., 1995), 指地球科学研究领域里, 在 GIS 环境下生成、使用的数据; C 空间数据 (Spatial data) (崔伟宏, 1995) 或地球空间数据 (Geospatial data) (FGDC, 1994), 强调数据与地球参考的空间位置有关。综合这些有不同

收稿日期: 2000-07; 修订日期: 2000-08

基金项目: “九五” 重中之重攻关项目 (96-B02-02-02) “重大自然灾害评估系统研究” 资助

作者简介: 李军 (1968-), 男, 河北大名县人, 博士后。1998 年在中科院地理研究所获得地图学与地理信息系统专业博士学位, 主要从事地学数据基础研究及地理信息系统应用基础研究。

侧重意义的概念, 我们认为地球空间数据 (Geo-spatial Data) 指的是与地球参考空间 (二维或三维) 位置有关的、表达与地理客观世界中各种实体和过程状态属性的数据。该概念强调地球空间数据的内涵是地学过程或特征, 其基本特征是空间位置。从认知的角度, 地球空间数据从来源到数据可分为三个阶段 (图 1)。客观世界中的过程和现象被人们认识后产生形式为文字、声波、电磁波、图象的类数据, 或者由数据获取设备直接形成客观世界的写真, 然后类数据表达为计算机能接收有特定含义的数字、字母或符号并进入计算机世界, 便成为数据。它与通常意义上的数据资料有一定的区别。

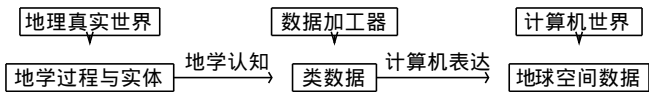


图 1 地球空间数据形成三个阶段示意图

Fig. 1 The three generation stages of geo-spatial data

地球空间数据具有分布式特征、多尺度特征、空间拓扑特征、数据量大、获取费用高、共享效益高、涉及专业部门多等特征。数据的最小单元是数据中的记录, 它记录每一个地学过程或地理特征的时空特征和属性, 多个同类的地理特征记录组织到一起则成为数据集。它是数据在数据库中存贮、处理、传输和输出的基本单元, 平时我们所说的地球空间数据其实是地球空间数据集, 如华北平原 1 50 万土地利用数据, 是含有若干土地利用记录、一定综合层次的数据集合。若干个数据集往往以不同的主线特征 (如区域、自然要素、社会经济要素或某种应用目的) 有机地在物理空间上或逻辑上组织到一起形成数据库。多元地球空间数据的使用则需要数据集成。

2 地球空间数据集成

集成 (integration) 的意思是指通过结合分散的部分形成一个有机整体, 地球空间数据集成的说法很多, 根据其侧重点可分如下几类 (I. D. H. Shepherd, 1991): ① GIS 功能观点认为数据集成是地理信息系统的基本功能, 主要指由原数据层经过缓冲、叠加、获取、添加等操作获得新数据集的过程; ② 简单组织转化观点认为数据集成是数据层的简单再组织, 即在同一软件环境中栅格和矢量数据之间的内部转化或同一简单系统中把不同来源的地理数据 (如: 地图、摄影测量数据、实地勘测数据、遥感数据等) 组织到一起; ③ 过程观点认为地球空间数据集成是在一致的拓扑空间框架中地球表面描述的建立或使同一个地理信息系统中的不同数据集彼此之间兼容的过程; ④ 关联观点认为数据集成是属性数据和空间数据的关联, 如 ESRI (1990) 认为数据集成是在数据表达或模型中空间和属性数据的内部关联; David Martin (1993) 认为数据集成不是简单地把不同来源的地球空间数据合并到一起, 还应该包括普通数据集的重建模过程, 以提高集成的理论价值。

从形式上, 数据集成是不同来源、格式、特点性质的地球空间数据逻辑上或物理上的有机集中, 有机是指数据集成时充分考虑了数据的属性、时间和空间特征、数据自身及其表达的地理特征和过程的准确性。

由此, 我们认为地球空间数据集成是对数据形式特征 (如格式、单位、分辨率、精度

等)和内部特征(特征、属性、内容等)作全部或部分的调整、转化、合成、分解等操作,其目的是形成充分兼容的数据集(库)(图2)。地球空间数据集成分为3个层次:概念层、逻辑层和物理层,涉及的内容分别为:数据集成的方式方法等概念模型、数据集成的逻辑表达和数据集成的具体实现。

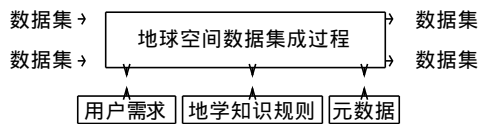


图 2 地球空间数据集成机理框架结构示意图

Fig.2 Construction diagram of geo-spatial data integration

3 地球空间数据集成目标

对于已存在的各类地球空间数据,数据用户对数据的空间范围有较强的随意性,如在使用卫星影像数据时往往会遇到研究区跨两景、多景的情况,数据集成的目的之一是从固定区域的数据集中高效获取用户随意确定的空间范围内的数据集。

数据集成的目标可以简单地表达为建立无缝数据集(库)。数据集(库)无缝表现在数据的空间、时间和属性上的无间断连续性(图3)。空间无缝指地理特征在不同数据集空间范围连续性;时间无缝指地学过程允许范围内的时间不间断;属性无缝指属性类别、层次的不间断特征。数据尺度已作为地球空间数据更根本的一个属性融合到了数据的空间、时间和属性中。数据集成即是寻找数据集之间连续性的表达方式,它表现为两个方面:不同尺度数据之间的集成和相同尺度数据之间的集成。不同尺度同种要素数据反映的是该地学要素过程在不同大小空间上表现的规律,其集成是使数据集之间不间断并能自然过渡,即形成全尺度的地球空间数据(或部分连续尺度);在相同尺度之间则主要是确定该尺度上表达某地学过程详细程度的标准,然后使在空间上邻接的地学特征能在物理上或逻辑上连接起来,对数据使用者不出现间断。

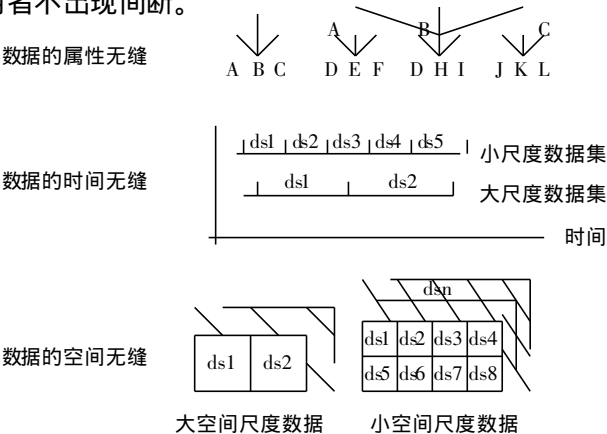


图 3 空间、属性、时间无缝数据集(库)的表现示意图

Fig. 3 Seamless integration of geo-spatial dataset (database) in the spatial, attribute and temporal dimensions

4 地球空间数据集成的研究状况及问题

地球空间数据集成是在应用基础上逐渐发展的,对其理论方法的研究处在逐渐成熟中。目前,在地球空间数据集成的应用和研究中仍存在许多问题需要深入探讨。

4.1 数据集成的基础研究

数据、集成软件及规则是数据集成的3个必备基础条件。数据是集成的对象;软件是可以处理空间特征、属性特征及其之间关联的通用或专题GIS软件,或是为数据集成专门设计的软件,它们可以实现集成的大多数操作;集成规则是进行数据集成的依据。数据集成基础研究包括数据集成机理及集成过程中诸多专题问题的研究,如:地球空间数据表达、误差传递及数据质量控制、多尺度数据处理等。

(1) 数据集成机理研究。数据集成机理是在集成的各个环节中处理各类问题的理论、方法及规则。它是以地学认知为前导,以地学规律、推理为主体内容的。数据集成的研究及应用层次取决于集成各领域的专题研究的进展。地球空间数据质量控制、地球空间数据误差传递、地球空间数据表达、地学元数据、数据交换等研究使地球空间数据集成理论和技术日渐成熟;地理认知理论的提出和发展为地球空间数据的表达、数据综合处理、定性的地球空间数据的定量化处理等提供了理论依据。

(2) 误差传递研究及数据质量评价与控制。关于误差形式、传递、消减方法及不确定性的描述表达,已有大量的文献进行论述(Bolstad, 1990; Goolchild, 1992; Fisher, 1993; Heuvelink, 1993; Howard, 1995; Philippe, 1997),并形成了适于某些应用领域的方法,但地球空间数据集成中的误差不是简单的图形问题,它与地学过程本身是密不可分的。

(3) 数据多尺度。从传统地图中继承下来的尺度特性在地球空间数据中依然存在,因而数据用户很难摆脱许多地图特征对地球空间数据的影响,在数据中尺度概念往往用数据精度表示,但基于数据自身的精度与尺度的描述则没有明确的标准,虽然许多学者已对该问题进行了深入的研究(Dan Lee, 1993; David 1994; C. B. Jones, 1996),但仍无法确定在一定精度上要用多少个中间点构成的线来描述一条河流,一条河流小到什么程度便不必表示出来这类问题。因而尺度问题在一定程度上只能用类似于制图概括的方法来处理。

(4) 地球空间数据表达。数据表达指如何用计算机的方法来表达地理客观世界的等级、层次特征及多种性质的地学过程和现象,其关键问题是计算机如何识别处理不同层次数据之间的联系(E E Shiryaev, 1987),从这种意义上说数据表达也是数据多尺度处理的内容,其表达的难点是模糊边界的描述(Andrej, 1997)、空间数据分类、组织、数据结构(Simon W., 1994)等。

(5) 其它问题。数据集成还涉及到数据集成价值、分布式地球空间数据库的使用、权限、收费等问题(David Rhind, 1992),这些方面也需要有可操作的标准和方法。

4.2 地球空间数据集成方法与应用

地球空间数据集成是基于地学内容、知识和规律的,在集成中对数据处理有两种性质:一是数据外部形式协调处理,其标志是数据空间特征相对位置、特征数量、属性的构成及层次不发生变化;二是数据特征内容的变化,即集成数据参与运算,空间特征、属性内容、时间特征尺度等或多或少发生了变化,或生成了新的数据集。集成方法是面向不同应用的。

Len Seligman (1996) 等人把数据集成系统分为在已有的系统中做新的界面、在不同数据源之间传递统一的访问请求、数据在结构松散的互操作系统中传输、数据仓库方式和数据移动等类型。每一种集成中都要用到诸如: 组成系统描述、界面描述、参考定义、语意相关性、转换功能模块库、访问控制和义务等。地球空间数据由于来源不同其参考体系及各种参数存在着很大差异, 如何使之匹配起来, 需经一系列的转换、一致化操作等过程。对集成方法和应用的研究可分为以下几个方面:

(1) 数据集成中的数据组织。鉴于地球空间数据的分布式特征及潜在的可视化表现力, 利用多媒体技术, 对地球空间数据进行可视化管理是实际可行的, 也是集成数据的未来组织形式。H. T. Uitermark (1996) 等人论述了以 OPGIS 和超图等为载体实现不同数据库中数据的动态调用方法; Wen-Syan Li (1996) 说明了地球空间数据与非地球空间数据集成的关系和复杂性, 并给出了各数据库之间动态集成机制与方法; Piotr Jankowski (1995) 讨论了在系统集成中基于公共数据库或数据动态交换机制的数据集成方法; Timothy L. Nyerges (1989) 就地球空间数据集成的结构和步骤方法进行了讨论。

(2) 集成中的多数据集叠加分析。集成中数据的叠加属于拓扑叠加, 其主要目的是根据数据内容之间的相关关系, 利用属性逻辑运算形成新的数据集。目前这种操作仍多采用基于栅格数据模型的叠加, Eugene A (1992) 给出了基于区域属性值滤波的栅格数据集成方法; David Martin (1993) 等人讨论了基于栅格 GIS 集成社会经济和自然环境数据应用中的问题、思路和方法。

(3) 数据转换。数据转换包括格式、属性分类等内容, 考察转换效果的主要标志是数据损失尽可能少, 其中研究最多的是数据在不同数据格式转换中的问题。N. M. Mattikalli (1995) 等人讨论了一种栅格数据到矢量数据的转换方法, 其做法是先将栅格影像图转成 Lattice 文件, 该文件中以对应网格的中心点的值来表示网格的值, 然后再转成 Grid 格式, 最后转换成矢量数据; Atsuyuki (1997) 等人提出用点到多边形 (Point-in-polygons) 方法, 由不规则区域的属性值转换成规则区域的值; Joseph. M (1990) 等人就集成中特征边界的提取、平滑及数据处理后的拓扑重建提供了理论研究及具体方法。

(4) 遥感数据与 GIS 数据的集成。遥感是地球空间数据最直接、时效性最强的来源形式, 其关键是如何把遥感数据与 GIS 数据结合起来, 这方面已进行了大量的研究。Wieslaw (1993) 论述了遥感与 GIS 数据在土地利用变化中的结合; A. L. O'Neill (1993) 等人说明了遥感与 GIS 数据集成在火灾监测及火灾损失评估中的应用方法; Susan Ringrose (1996) 讨论了遥感与 GIS 数据集成在地表植被变化检测中的使用方法; J. C. Hinton (1996) 对遥感与 GIS 数据集成的发展、技术及对软件的需求及其在环境主题应用中的作用进行了说明。

(5) GPS 数据与地球空间数据的集成。这两类数据的集成主要是用 GPS 修正已有的地球空间数据, 其原理和方法相对较简单。Wolfgang Irsen (1996) 讨论了用 GPS 更新地球空间数据的方法。

4.3 地球空间数据集成中存在的问题

在地球空间数据集成领域虽然做了大量的研究, 并且数据集成在许多的地球空间数据应用项目中得到了广泛的使用, 但对地球空间数据集成的研究仍有一些问题:

(1) 已有的集成研究多侧重于具体数据集集成方法的探讨及集成中某些专题内容的研

究。如误差传递、数据表达等,对地球空间数据集成的一般性问题,如集成的体系、方法论、规则、依据等讨论不多。没有建立起集成的理论体系,这在一定程度上限制了数据用户对数据集成的认识、理解及应用。

(2) 对地球空间数据集成具体专题的研究主要是从计算机的角度进行的。关于地球空间数据对地学过程的表达及数据要素相关性在数据中的表现讨论不多,数据质量、尺度及其关系评价处理方法仍停留在地图标准的层次上,没有建立起基于地球空间数据自身特征的质量和尺度表达、转换标准。数据集成中地学过程的认知表达、基于地学规则、知识的数据处理、基于地学规则的集成结果表达等环节的研究均不能脱离地学特征而独立存在。

(3) 数据集成中元数据的使用机制、方法研究不足。目前已形成诸多的地学元数据标准,但多数标准较冗长难以操作,并且多数把对数据集的说明及对数据特征的说明混在一起,对使用和管理都有一定的限制。我们认为元数据的内容应从地球空间数据集的共性开始,逐渐深入,对元数据的使用是从用户层出发,分析对数据集了解的层次性。

4.4 地球空间数据集成的发展趋势

根据地球空间数据集成的研究状况及集成应用项目的特征,我们认为地球空间数据集集成有以下趋势:

(1) 网络化。数据的分布式特征及项目需求数据的多元化,使集成应用项目涉及的部门、内容越来越复杂,要求在集成中能快速使用物理上分布于各个节点的数据,各类网络的建设为地球空间数据的网络化集成提供了条件。

(2) 集成机理、规范标准研究。网络、计算机及数据库技术只是为地球空间数据集集成提供了可能性,而其集成的真正实现与地球空间数据的自身特征分不开,而有关地球空间数据表达、组织、抽象等问题远没有形成有效的方法。因而基于认知科学、地学的集成机理、集成规则标准、普遍意义的集成方法等仍将是数据集成研究的主流。

(3) 集成知识规则的专家系统化。集成中用到的诸多知识规则不可能让数据用户全部掌握,而数据集成应用中又离不开这些规则,因而如何将各类集成中的知识规则转化成数据用户可操作的专家系统必将是数据集成研究的另一个方向。

参考文献:

- [1] O'Neill A L, Head L M, Marthick J K. Integrating remote sensing and spatial analysis techniques to compare Aboriginal and pastoral fire patterns in the East Kimberley, Australia[J]. *Applied Geography*, 1993, **13**: 67 ~ 85.
- [2] Andrej Vckovski. Digital Representation of Continuous random Fields, Geographic Information Research, Edited by Massimo Craglia & Helen Couclelis[M]. Taylor & Francis, 1997. 382 ~ 396.
- [3] Atsuyuki Okabe and Yukio Sadahiro. Variation in count data transferred from a set of irregular zones to a set of regular zones through the point-in-polygon method[J]. *Int. J. Geographical Information Sciences*, 1997, **11**(1): 93 ~ 106.
- [4] Bolstad P V, Gessler P, Lillesand T M. Positional uncertainty in manually digitized map data[J]. *Int. J. Geographical Information Systems*, 1990, **4**: 399 ~ 412.
- [5] Jones C B, Kidner D B, Luo L Q et al. Database design for multi-scale spatial information system[J]. *Int. J. Geographical Information Systems*, 1996, **10**: 901 ~ 920.
- [6] 陈述彭. 地理信息系统的基础研究——地球信息科学[J]. 地球信息——科学技术产业, 1997(3): 11 ~ 20.
- [7] 陈学诠, 陈洪亮. 数据库原理与工程应用[M]. 北京: 中国科技大学出版社, 1996. 3.
- [8] 崔伟宏. 空间数据结构研究[M]. 北京: 中国科学技术出版社, 1995. 17 ~ 22.

- [9] Dan Lee. An ideal means for multiple GIS data transformation and represents[J]. *Proceedings of the Third Workshop on Geographical Information System, Beijing*, 1993, 1: 66 ~ 79.
- [10] Kinder D B, Jones C B. A Deductive object——oriented GIS for handling multiple representations, *Advances in GIS: Proceedings of the Sixth international Symposium on Spatial data Handling*[A]. In: Waugh T C, Healey R G (ed.). London Taylor and Francis, 1994. 882 ~ 900.
- [11] Maguire D J. Computer in geography Longman scientific & Technical[M], 1989. 178 ~ 180.
- [12] David Martin, Ian Bracken. The integration and socioeconomic and physical resource data for applied land management information systems[J]. *Applied Geography*, 1993, 13: 45 ~ 53.
- [13] David Rhind. Data accessing, charging and copyright and their implications for geographical information systems[J]. *Int. J. Geographical Information Systems*, 1992, 6(1): 13 ~ 30.
- [14] Shiryaev E E. Computers and the Representation of Geographical data[M]. John Wiley & Sons, 1987. 1 ~ 11.
- [15] Eugene A Fosnight. Data integration through region-based nominal filtering[J]. *Int. J. Geographical Information Systems*, 1992, 6(5): 469 ~ 478.
- [16] Elaline Matthews. Global Geographical Databases for Modeling trace Gas Fluxes[J]. *Int. J. Geographical Information Systems*, 1993, 7(2): 125 ~ 142.
- [17] FGDC. Federal Geographic Data Committee[M]. Content Standards for Digital Geo- spatial Metadata, 1994.
- [18] Fisher P E. Algorithm and implementation uncertainty in viewshed analysis[J]. *Int. J. Geographical Information Systems*, 1993, 7: 331 ~ 347.
- [19] 弗·特普费尔 著. 江安宁 译. 制图综合[M]. 北京: 测绘出版社, 1982. 41 ~ 69.
- [20] Goodchild M K, Guoqing S, Shiren Y. Development and test of an error model for categorical data[J]. *Int. J. Geographical Information Systems*, 1992, 6: 87 ~ 104.
- [21] Bonham-Carter G F. Geographic Information Systems for Geoscientists: Modeling with GIS[M]. Pergamon, 1995. 20 ~ 50.
- [22] Onsrud H J, Gerard Rushton. Edited Sharing Geographic Information[M]. The State University of New Jersey, 1995. 1 ~ 7.
- [23] Heuvelink G B M, Burrough P A. Error propagation in cartographic modeling using Boolean logical and continuous classification[J]. *Int. J. Geographical Information Systems*, 1993, 7: 231 ~ 246.
- [24] Howard Veregin. Developing and testing of an error propagation model for GIS overlay operation[J]. *Int. J. Geographical Information Systems*, 1995, 9: 595 ~ 619.
- [25] Uitermark H T, Dutch Cadastre, Apeldoorn. The integration of geographic databases: Realizing geodata interoperability through the hypermap metaphor and a mediator architecture[A]. In: Geographical Information, Second Joint European Conference & Exhibition on Geographical Information Barcelona[C]. Spain 1996. ISO Press, 1996. 92 ~ 95.
- [26] Shepherd I D H. Information integration and GIS in Geographical Information Systems: Principles and Application Vol. 2[M]. In: Maguire D J, Goodchild M F, Rhind D W (ed.). Longman, London, 1991. 337 ~ 360.
- [27] Hinton J C. GIS and remote sensing integration for environmental applications[J]. *Int. J. Geographical Information Systems*, 1996, 10(7): 877 ~ 890.
- [28] Joseph M Piwowar et al. Integration of spatial data in vector and raster formats in GIS environment[J]. *Int. J. Geographical Information Systems*, 1990, 4: 429 ~ 444.
- [29] Len Seligman, Arnon Rosenthal. A Metadata Resource to Promote Data Integration[A]. In: First IEEE Metadata Conference April 16- 18, 1996[C]. NOAA Auditorium, Silver Spring, Maryland.
- [30] Mattikalli M, Devereux B J, Richards K S. Integration of remotely sensed satellite images with a geographical information system[J]. *Computers & Geosciences*, 1995, 21(8): 947 ~ 956.
- [31] Mattikalli N N, Devereux B J, Richards K S. Integration of remotely sensed satellite images with a geographical information system[J]. *Computers & Geosciences*, 1995, 21(8): 947 ~ 956.
- [32] 牛文元. 理论地理学[M]. 北京: 商务印书馆, 1992. 589 ~ 640.

- [33] Philippe Lagacherie, Susan Holmes. Addressing geographical data errors in a classification tree for soil unit prediction[J]. *Int. J. Geographical Information Systems*, 1997, **11**: 183 ~ 198.
- [34] Piotr Jankowski. Integrating geographical information systems and multiple criteria decision-making methods[J]. *Int. J. Geographical Information Systems*, 1995, **9**(3): 251 ~ 273.
- [35] Howding S W. 3D Geoscience Modeling—Computer Techniques for Geological Characterization[M], Springer, 1994. 71 ~ 86.
- [36] Susan Ringrose et al. The use of Integrated Remotely Sensed and Gisdata to determine Causes of Vegetation Cover Change in southern Botswana[J]. *Applied geography*, 1996, **16**(3): 226 ~ 247.
- [37] Nyerges T L. Schema integration analysis for the development of GIS databases[J]. *Int. J. Geographical Information Systems*, 1989, **3**(2): 153 ~ 183.
- [38] Uitermark T. The Integration of Geographic Databases realising Geodata Interoperability through the Hypermap Metaphor and a Mediator architecture[A]. In: *Geographical Information Second Joint European Conference & Exhibition on Geographical Information Barcelona*[C], Spain, 1996. ISO Press.
- [39] UCGIS. Research Priorities for Geographic Information Science[J]. *Cartography and Geographic Information System*, 1996, **23**(3): 115 ~ 127.
- [40] Li Wen-syan, Chris Clifton. Dynamic Integration in Multidatabase Systems [J]. *Journal of Database Management*, 1996, **7**(1): 28 ~ 40.
- [41] Michalak W Z. GIS in land use change analysis: integration of remotely sensed data into GIS[J]. *Applied Geography*, 1993, **13**: 28 ~ 44.
- [42] Wolfgang Irsen, Stefan Sandmann. Data Capture and Data Updating with GPS[A]. In: *Geographical Information Second Joint European Conference & Exhibition on Geographical Information Barcelona*[C]. Spain 1996. ISO Press. 1996. 45 ~ 52.
- [43] 周荫清 主编. 信息理论基础[M]. 北京: 北京航空航天大学出版社, 1993. 1.
- [44] 李军, 周成虎. 地球空间数据元数据标准初探[J]. *地球科学进展*, 1998, **17**(4): 55 ~ 63.

Overview of Study on Geo-spatial Data Integration

LI Jun^{1,2}, FEI Chuan-yun²

(1. Institute of Remote Sensing Application, CAS, Beijing 100083;

2. Institute of Geographical Sciences and Natural Resources, CAS, Beijing 100101)

Abstract: Application of Geographical Information Systems (GIS) and multi-resources geo-spatial data promotes the appearance of digital geo-spatial data integration in 1960's. But the weakness of study on geo-spatial data integration hinders its own development. As a overview on geo-spatial data integration, the paper focuses on fundamental concept, theory of data integration and problems in geo-spatial data integration research currently.

Geo-spatial data is one kind of data used to represent, describe geographical processes and geographical phenomena. From the point view of geographical cognition, geo-spatial data is abstraction of geographical entity controlled by geographical knowledge. Geo-spatial data integration, as argued by many researchers, is the process during which multi-resources and multi-scale geo-spatial data can be used in one uniform GIS software platform. The goals of geo-spatial data integration, put simply, are creating seamless

(including attribute seamless, temporal seamless and spatial seamless) dataset or database for certain application or general construction of geo-spatial database.

Study of geo-spatial data integration mainly focus on data integration mechanism, error transfer, data quality controlling, multi scale, and representation of geo-spatial data. Also many successful application of geo-spatial data integration were carried out, but there are still many shortcomings of the fundamental researches on geo-spatial data integration, such as integration principles, error propagation rules, multi-scale data integration and data quality evaluation and controlling. The main problems about geo-spatial data integration are short of general integration principles and rules, weak studying on geo-spatial metadata, ignoring the applications of geo-sciences principles and rules.

Based on the analysis on geo-spatial data integration, the authors list out some further research field on data integration such as network based data integration, data integration principles, metadata utility in data integration, geo-sciences expert knowledge system and their application in geo-spatial data integration.

Key words: geo-spatial data; data integration; seamless dataset