

多源数据更新地理空间数据库的理论及关键技术

Theory and Key Technology of Updating Geospatial Database Based on Multi-resource Spatial Data

张 锦/ZHANG Jin^{1,2}, 董晓媛/DONG Xiao-yuan¹, 金雁中/JIN Yan-zhong¹

1. 太原理工大学测绘科学与技术系,太原 030024

2. 同济大学交通运输学院,上海 200092

1. Department of Geomatics, Taiyuan University of Technology, Taiyuan 030024, China

2. School of Transportation Engineering, Tongji University, Shanghai 200092, China

[摘要] 地理空间数据整合与更新是目前国际地理信息科学界的前沿课题。数字合并计算、多尺度空间数据集计算环境是有效的空间数据整合实现技术,多源空间数据统一资源信息服务平台是实现整合空间信息服务的有效途径,利用卫星遥感影像数据和应用 GIS 空间目标综合技术更新空间数据库是有效的空间数据库更新技术途径。论文对空间数据整合和更新的理论及技术进行了较为全面的分析,指出地理空间数据库整合增强并丰富了空间数据库的服务能力,地理空间数据库更新提升了地理空间信息服务的质量和价值。

[关键词] 地理信息系统,空间数据整合和更新,数字合并,自动综合,变化检测

[中图分类号] TP311.131

[文献标识码] A

[文章编号] 1000-7857(2005)08-0071-04

Abstract : Integration and updating of geospatial data is a key problem in Geomatics. Conflation and integration computer environment of multi-scale spatial data are valid implement technology of spatial data integration. The unified information service platform of multi-resource spatial data is an effective means to get spatial information service. Using satellite image data and applying GIS spatial objects generalization technology to update spatial database are effective approaches. In this paper, the theory and technology of spatial data integration and updating are throughly analyzed, and we also present that integration of geospatial data enhances the service ability of spatial database, and the updating upgrade the value of geospatial information service.

Key Words: GIS, integration and updating of spatial data, conflation, auto-map generalization, change detection

CLC Number: TP311.131

Document Code: A

Article ID: 1000-7857(2005)08-0071-04

1 引言

地理空间数据整合与更新是目前国际地理信息科学界的前沿课题,在空间数据库建设起步较早、应用水平较高的发达国家更是一个倍受关注的科学问题。由于我国的地理空间数据库的建设起步晚,许多重要的数据库刚刚完成或正在建设中,整合和更新的理论、方法及技术体系的研究还很薄弱,迫切需要进行科学攻关。在地理空间数据库更新数据源和更新方法方面,遥感数据和遥感技术在数据更新中已越来越重要,利用大比例尺地图对小比例尺地图进行级联更新是一种重要的数据更新途径,地理空间数据更新过程中应注意整合利用行业数据资料。主数据库增量数据建模、遥感影像与矢量地图自动配准、重要地

理要素变化信息自动提取、多尺度数据库联机更新的自动综合方法,以及客户数据库更新方法等地理空间数据持续更新框架问题,是这一研究领域的关键技术和主要研究方向^[1]。

2 多源空间数据集理论

2.1 多源空间数据特征或特性分析

空间数据作为数据的一类,具有数据的一般特性;由于在获取方式、表示和管理方面的显著技术特色,空间数据又具有一些非常鲜明的特性。一般意义上的空间数据包括图形数据(或几何数据)和属性数据。图形数据或几何数据又可按其形态或获取与存储方式的差异分为矢量数据、DEM(数字高程模型)格网或 TIN(不规则三

角网)数据,以及栅格数据。属性数据的分类比较复杂,笼统地可分为一般属性数据和专题属性数据;基本的细分类别为几何类型信息、分类分级信息、数量特征信息、质量特征信息和名称信息等^[2]。空间数据的一般特性可概括为抽样性、概括性、多态性和空间性^[3]。随着空间数据获取与集成技术的发展,空间数据的多尺度性(或多分辨率特性)、多源性、时态特性、客观性、相关性等成为空间数据分析与应用中必须予以深入研究的数据特性。

空间数据的多源性一方面表现为空间数据获取手段的多样性,如测量、调查、统计汇总、现有资料数字化、影像数据获取等,另一方面对于不同的空间数据采集方法和管理系统,空间数据也具有不同的存

收稿日期:2005-06-03

基金项目:山西省自然科学基金资助项目(20031039)、中国博士后基金资助项目(20040350496)

作者简介:张 锦,男,山西省太原市迎泽西大街新矿院路 18 号太原理工大学测绘系主任,教授,主要研究方向为地理信息系统模型理论、空间数据整合与更新、3S 集成技术和应用;E-mail: zjgis@sina.com

储、交换格式。不同的数据获取方式或数据表达方式,其数据格式存在较大的差异,如影像数据和矢量数据、DEM与图表等。这些问题无疑给多源空间数据的集成应用带来技术难度,但是,多源空间数据的集成应用为空间数据整合和更新、空间信息服务也带来了无可比拟的数据资源和信息资源优势。

2.2 多源空间数据库的整合理论和技术

2.2.1 数字合并

数字合并(Conflation)是从不同的数据源合并为一个新的、最优的数据集的算法过程,“新”和“优”表现在空间和属性两个方面。数字合并算法的关键技术是如何自动探测候选的匹配地物的特征要素。利用数字合并算法可实现将一组弧段与其它不同精度的弧段匹配,匹配对应关系是一对多或多对多的关系。数字合并同时可将合并前弧段的属性传递给合并后的对象,传递过程并不是简单的属性数据重叠,而是通过一种智能算法计算出合并后对象匹配的属性的过程(如图1所示)。

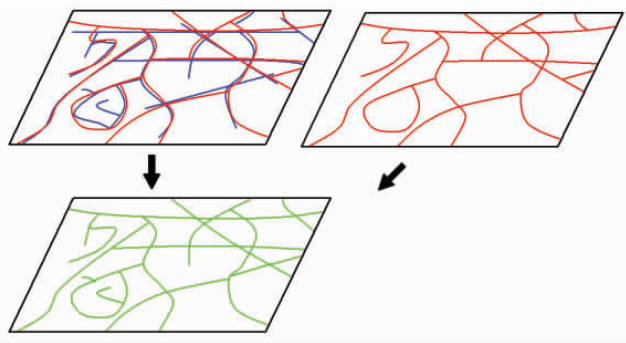


图1 数字合并算法应用示例

数字合并算法^[4,9]为多源数据分析、整合应用提供了有效方法,它分为地图合并计算、基于点的合并计算、基于线的合并计算和基于规则的合并计算。通常的数字合并算法包含三步迭代过程:特征匹配、聚焦区域对象重新排列、定位和属性冲突解决。线性数字合并的算法过程主要为:探测同态对象、同态对象匹配。数据库匹配^[6,7]对数字合并算法提出了更高的要求,目的是综合利用数据库信息资源,提供多源数据匹配的方法模型。

2.2.2 多尺度空间数据库集成管理和计算环境

多尺度空间数据库集成管理和计算环境为多源空间数据整合和计算提供了支持环境。

多尺度空间数据主要包括DLG(数字线划图)、DEM(数字高程模型)和DOM(正射影像图)数据。由于这些数据的获取方式、表现内容存在很大差异,所以,在建立多尺度空间数据的组织模型时,这些数据应分别予以考虑。又由于相当多的应用需要将上述3D数据(DLG、DEM、DOM)集成应用,所以,建立的多尺度数据模型应能提供集成数据的管理、计算与表现能力。Jones C B提出了多尺度空间数据库的概念模型GEOSYSSEY,并将多尺度空间数据库分为内涵数据库(Intensional Database)和外延数据库(Extensional Database)两部分。内涵数据库主要由两部分构成:一是规则数据库(包括:空间推理、时态推理和属性推理规则数据库),二是过程或算法(几何算法和模型综合算法)。外延数据库由对象目录(包括:Real-World Object Directory)和元数据库构成。对象目录的内容为语义、空间和时态数据库,元数据库的内容为分类层次和质量属性。在该模型基础上的多尺度空间数据管理和计算的实现策略是:利用组合适当的元数据的多分辨率数据结构递进层次存取物理存储的空间对象;从显式存储的关系中直接或间接演绎确定空间关系,如不存在显式存储的关系,则应用几何算法确定几何对象表现时所需的空间关系;如不存在目标尺度匹配的

几何对象,则应用综合算法进行GIS几何目标自动综合,计算生成所需尺度的空间对象^[8]。

多尺度工程下的工作区数据对应多比例尺的空间数据集,如1:5万的工程对应的工作区数据包含1:5万、1:2.5万和1:1万等比例尺的地形图数据,也就是说这时的工作区起到局部工程数据管理的作用。显然,如果在工作区中引入多比例尺地形图数据索引操作可解决这一问题,可是这将使得新的模型在很大程度上不兼容传统的单分辨率工作区模型。解决这一问题的方法是,仍然在工程文件中建立索引操作,也就是在传统的单分辨率工程文件的索引基础上附加多比例尺索引操作,具体操作方法如下。

工程数据管理和索引文件从1:100万比例尺的地形图数据开始,按照国家基本比例尺图形图分幅编号标准,逐级建立多比例尺数据管理层次索引,每个确定的比例尺地形图数据集对应传统GIS系统相同的工程、工作区内容。当然,也可自定义地形图多比例尺映射表,并按照映射表的规则建立工程索引。另外,由于不同比例尺的空间数据可能对应着不同的地图投影方法和不同的分带,以及影像、DEM数据在边缘带的跨带操作等,所以,还需定义、管理空间数据的投影转换或换带计算。按照这一方法,最终的规范化集成多分辨率空间数据组织和管理对象模型见图2。

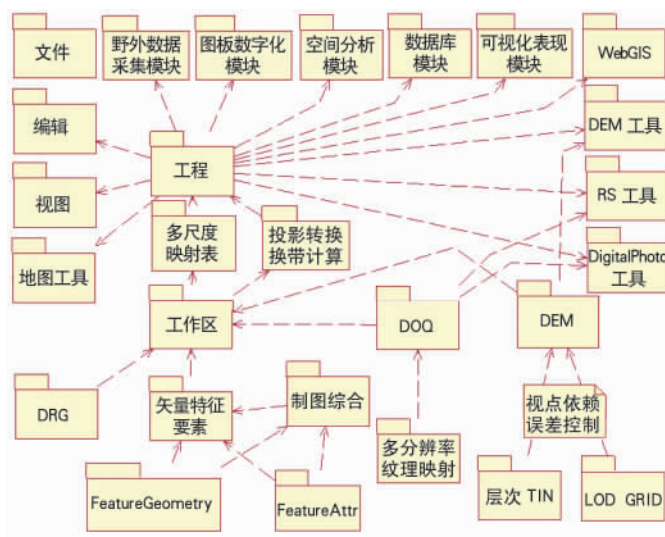


图2 规范化集成多分辨率空间数据组织和管理对象模型^[9]

2.2.3 多源空间数据库集成模式和统一资源信息服务平台

空间数据格式、语义描述的不统一使得空间数据的集成管理和应用变得非常复杂,甚至不可行。直接数据格式转换中的语义丢失和数据版权问题、数据互操作模式的效率、支持软件系统和相适应的GIS系统体系架构问题,以及直接数据访问的非交换数据格式不公开等问题,使得这一领域的进展与应用的需求存在较大的差距。目前广泛采用的方式还是对数据进行格式转换,辅以后转换后的二次编辑。

如何在动态、异构组织间实现协同的资源共享以及协同解决某一问题?寻找新的支撑平台,构建统一的资源信息服务平台是解决这一问题的可行之路。网格技术为空间信息资源共享、交换和互操作提供了新的技术平台,图3为网格的服务环境示意图^[10],图4为OGSA(开放网格系统架构)体系架构图。

根据5层沙漏结构中各个组成部分与共享资源的距离,将对共享资源进行操作、管理和使用的功能分散在5个层次。越向下层就越接近于物理的共享资源,因此,该层与特定资源相关的成分就比较多;越向上层就越感觉不到共享资源的细节特征,也就是说上层是更加抽象资源的表示,因此,就不需要关心与底层资

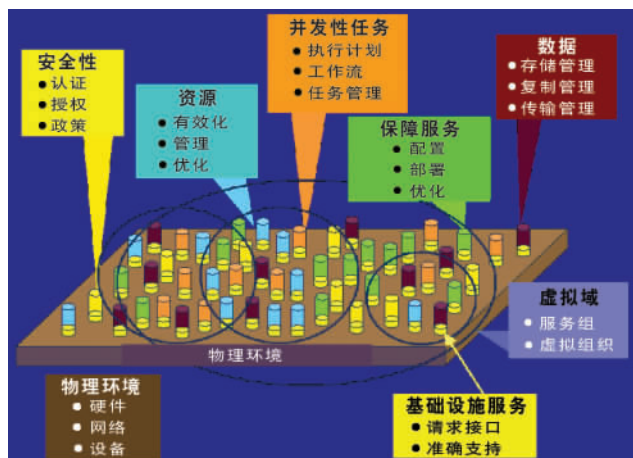


图3 网格服务环境示意图

源相关的具体实现问题。

3 空间数据库更新的关键技术

3.1 利用卫星遥感影像数据更新空间数据

在空间数据库更新的过程中,为了达到快速更新的目的,需要对新旧数据源进行快速变化检测并提取变化特征对象。变化检测可以通过数据的对比、匹配来检测变化区域、变化特征对象。典型的方法有以下4种。①模板匹配法:通过模板与影像场中被模板覆盖区域的相关性度量计算搜索变化区域。②微分纠正法:以全自动方式获取密集同名点对并作为控制点,由密集同名点对构成密集三角网(小面元),利用小三角面元进行微分纠正,以实现影像的精确配准。③数学形态学方法:是研究数字影像形态结构与快速并行处理方法的理论,将大量复杂的图像处理运算用基本的位移和逻辑运算组合来描述和实现。利用数学形态学方法可实现图像的增强、分割、边缘检测、结构分析、骨架化、组分分析等,算法便于进行并行处理和硬件实现,从而提高影像处理的计算效率。④最小二乘影像法:可以充分利用影像窗口内的信息灵活地引入各种已知参数和条件进行整体平差计算,使影像匹配可达到1/10甚至1/100像素的高精度。

对于遥感影像数据,有效的几何校正方法是实现变化检测和特征提取的基础。通过已有地形图、实地丈量 and 遥感图像判读等方法和技术手段精确标定控制点和检查点的实地和图像位置,并应用RTK GPS(实时动态卫星定位系统)精确测量控制点和检查点点位,完成控制点的实地、图像标定和点位测量。几何校正模型可采用二次多项式模型、RPC(有理多项式系数)+平均高程+二次多项式模型、RPC+DEM+二次多项式模型和Rubber Sheeting(橡皮纸)模型进行卫星遥感图像几何校正精度的对比分析研究。如精度不满足要求,则可通过改变控制点分布、增加控制点的数量等措施来改进几何校正的精度。所需增加的控制点一般不需要再进行GPS实测,可充分利用待更新的地图资料。建议的方法为:选择待更新数字地形图非变化部分且在遥感图像上又是明显特征点的那些点作为新增的控制点,并在数字环境下精确捕捉坐标。对精确校正后的遥感影像进行地物边界跟踪,构建拓扑关系、注记和制图,完成地形图的更新。图5为某区域控制点分布、布置图。图6为通过实地丈量测距离在影像图上精确确定图像控制点位置的测量方法。图7为更新后的地形图与影像图叠置对照。

3.2 应用GIS空间目标综合技术更新空间数据库的理论和关键技术

综合计算是地图学的核心研究方向之一。通过对地图制图本质理论的分析研究,提取地貌和地物的形态结构并模型化,给出初始状态下空间目标变换为新的状态(新的尺度)空间目标的结构化变化理论,在此基础上再进行移位、符号化和夸大处理^[11,12]。

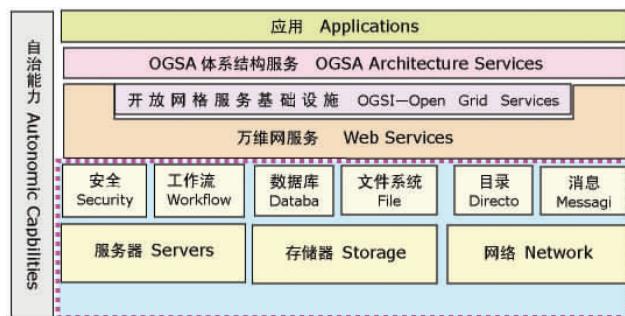


图4 OGSA 的主要架构

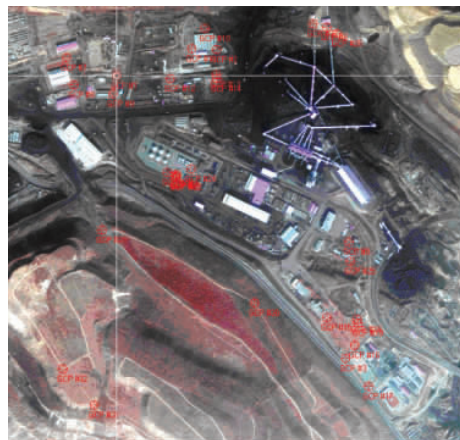


图5 控制点分布图



图6 图像控制点位置测量方法

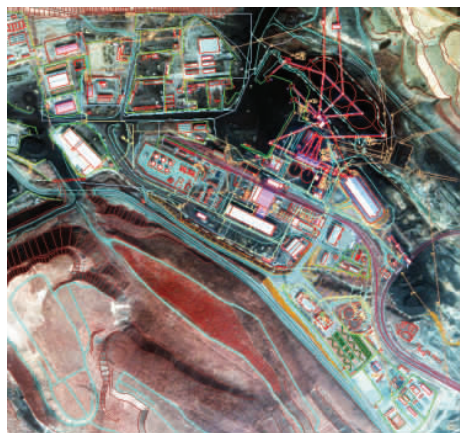


图7 更新以后的地形图与影像图叠置对照

综合计算又是 GIS 发展与应用的重点研究方向。运用数据结构理论和地图制图综合理论,通过对空间目标的分割、标识,建立尺度变化时空间目标的综合模型和算法^[13]。伴随 GIS 的发展,地图综合已逐步演变为 GIS 空间目标综合。GIS 空间数据库相对于一般的地图数据文件,其空间目标的属性和空间关系(有的需事后重建),以及规范化的空间数据模型、数据存储结构等,是综合过程可以利用的重要知识源,结构化、规范化的空间数据便于应用综合模型进行自动综合计算。

怎样将图像、矢量和表面综合算法建立联系,从而产生集成的多源数据环境下的空间目标自动综合算法 (Images, vector and surface generalization—how can it be linked)?这是 ISPRS / ICA 2002 Workshop Session 3(国际摄影测量和遥感学会/国际制图协会,2002 年学术讨论会)的讨论议题,讨论认为影像分析和矢量数据综合这两个领域是紧密相关的,因为它们处理的是相同的现实世界对象。

经过多年的研究实践,该领域已基本形成了综合规则、单目标综合算法、多目标综合算法、局部综合算法、全局综合求解综合算法等地图综合或空间目标综合的理论和体系。但是,这些综合算法基本上是以围绕单源空间数据进行的,目标是从一种尺度的空间数据通过综合算法以尽可能少的人机干预完成综合过程,最终综合生成目标尺度的空间数据集。这时如果引入多源、多尺度数据集,包括多尺度 DLG 数据、多分辨率 DOM 或遥感影像数据和多分辨率 DEM 数据,一方面使综合过程多了一些参考(系),另一方面可从多尺度 DLG 和影像数据获得邻近目标尺度空间对象的结构特征、边界、形态等信息。基于这些信息,应用相适应的算法可探测典型化、移位和夸大化的空间对象,最终使得综合过程变得容易,自动综合变得可行。多数据源同样给问题的解决带来更多的知识,对象关联的属性、多数据源集成数据挖掘中产生的知识对于综合同样是重要的、有益的。

通过空间目标综合技术建立多尺度空间数据库,同时完成由多源空间数据、大比例尺空间数据自动更新小比例尺空间数据库,是空间数据库更新和整合的根本性技术。但是,这一问题的解决目前仍面临极大的挑战,尚存在诸如算法问题、语义问题、冲突问题、规则应用问题、地图美学问题等,需要我们深入研究和进一步完善。

4 结论

地理空间数据整合与更新是目前国际地理信息科学界的前沿课题。本文对空间数据整合和更新的理论和技术进行了较为全面的分析,通过实例说明多源数据集成整合和更新空间数据库的技术途径及应用价值。地理空间数据库整合增强和丰富了空间数据库的服务能力,地理空间数据库更新提升了地理空间信息服务的质量和增值。

参考文献(References)

- [1] 中国科协空间数据整合与更新青年科学家论坛简报[EB/OL],2004, http://ngcc.sbsm.gov.cn/laboratory/chinese/news/xiaoxi15_Presentation/简报.pdf
- [2] 毋河海.地图数据库系统[M].北京:测绘出版社,1996
- [3] 郭仁忠.空间分析[M].武汉:武汉测绘科技大学出版社,1997
- [4] Filin S, Y Doytsher. Linear approach to map conflation: Matching of polylines[J]. *Surveying and Land Information Systems*, 1999, 59(2): 107~114
- [5] Filin S, Y Doytsher. The detection of corresponding objects in a linear-based map conflation [J]. *Surveying and Land Information Systems*, 2000, 60(2): 117~28
- [6] Devogele T, Parent C, Spaccapietra S. On spatial database integration [J]. *International Journal of Geographical Information Science*, 1998, 12(4): 335~352
- [7] Walter V, D Fritsch. Matching spatial datasets: A statistical approach [J]. *International Journal of Geographical Information Science*, 1999, 13(5): 445~73
- [8] Jones C B, et al. Database Design for a Multi-Scale Spatial Information System [J]. *International Journal of Geographical Information System*, 1996, 10(8): 901~920
- [9] 张锦. 多分辨率空间数据模型理论和实现技术研究[M]. 北京:测绘出版社,2004
- [10] Jay Unger, Matt Haynos. A visual tour of Open Grid Services Architecture. IBM Developer Works China—Grid computing[EB/OL], <http://www-900.ibm.com/developerWorks/cn/grid/gr-visual/index.shtml>. 2003
- [11] 毋河海,等. 地图水系与地貌自动综合研究专辑[J]. 武汉测绘科技大学学报, 1995, 20(增刊)
- [12] 王桥, 毋河海. 地图信息的分形描述与自动综合研究[M]. 武汉: 武汉测绘科技大学出版社, 1998
- [13] Oosterom P Van, Vincent Schenkelaars. The Development of an Interactive Multi-scale GIS[J]. *International Journal of Geographical Information System*, 1995, 9(5): 489~507

(责任编辑 黄永明)

·科技动态·

2005 年 7 月国内报刊重要科技文章篇目辑览

(报纸名称后面的数字是发表日期,期刊名称后面的数字是发表期数,N 代表报纸,J 代表期刊)

爱因斯坦和同行审稿制度的一次冲突 (刘寄星)物理[J] 7
有理逼近的一些最新进展(Ⅱ)——倒数逼近的研究综述 (虞旦盛,等)数学进展[J] 3
科学家“纠正”生化和细胞生物学教科书传统认识 (李 斌)新华社[N] 1
航天、宇航和光障——航天对基础科学的影响 (宋 健)科技日报[N] 5
夏朝立国前后的气候突变与中华文明的诞生 (王绍武)气候变化研究进展[J] 1
昆虫拍翼方式的非定常流动物理再探讨 (余永亮,等)力学学报[J] 3
为纯科学呼吁(节选) (亨利·奥古斯特·罗兰[美])中国基础科学[J] 3
大学生受电脑电磁辐射的研究 (胡炎弟,等)安全与环境学报[J] 3
中国科技界向爱因斯坦学什么 (范良藻)科学时报[N] 5
我国电子废弃物立法的困局与出路 (韩立琳,等)环境保护[J] 5
中国科学家对生命起源多个难题做出统一解释 (赵 路)科学时报[N] 6
生物医药关键技术发展趋势 (王 萍)中国生物工程杂志[J] 6
固体表面动态现象的研究进展 (赵跃智,等)材料导报[J] 7
生物制药业对美国经济的贡献 (美国 Milken 研究所)Biotech World [J] 6
当代科学影响哲学的几个重要基础理论问题 (邬 焜)自然辩证法研究[J] 7
国家科技进步奖励制度改革的若干建议 (赵春明,等)科学与科学技术管理[J] 7
科研成果应具备的要素 (王 选)光明日报[N] 7
新能源离我们有多远 (杜祥琬)光明日报[N] 7
我国科学数据库资源现状堪忧 (刘英楠)科学时报[N] 8

心管性痴呆中药治疗的实验研究新进展 (沈颖婕,等)中国中药杂志[J] 10
提高创新能力,推动自主创新 (路甬祥)求是[J] 13
从曙光十年的发展看自主创新 (李国杰)科技日报[N] 12
走进果蝇的“认知世界” (郭爱克)科学时报[N] 13
全球变暖、海水变淡、两极失衡 (迈克尔·席贝尔)参考消息[N] 1
哲学与医学的交融性思维 (麻天祥)科学时报[N] 4
扭转乾坤的科学报告——《科学——没有止境的前沿》发表 60 年述评 (苏开源)科学时报[N] 6
提高科技人员待遇不应靠“项目”经费 (苏纪兰)光明日报[N] 14
中国传统海洋文化中的自然观和科学观 (宗正海)光明日报[N] 14
国家实力根本源于基础研究 (马大猷)科学时报[N] 15
CO₂的地质存储——地质学的新课题 (周 蒂)自然科学进展[J] 7
双链 RNA 技术与植物病毒研究 (张晓婷,等)云南农业大学学报[J] 4
差异蛋白组学研究技术新进展 (孙 薇,等)化学通报[J] 6
喜马拉雅冰川融化潜藏巨大威胁 (理查德·苏埃德·帕里)参考消息[N] 13
探索未知是科研的快乐与动力 (朱日祥)科学时报[N] 18
学术品位是重要的衡量标准 (刘 兵)科学时报[N] 18
2 个可激系统的螺旋波同步行为(袁国勇,等)河北师范大学学报[J] 4
(吴晓晴 辑)