

网络地图渐进式传输中的粒度控制与顺序控制

艾廷华

(武汉大学资源与环境科学学院,地理信息系统教育部重点实验室,武汉 430072)

摘要 网络地图渐进式传输中数据粒度剖分与传输顺序是该过程控制的两个基本问题,数据粒度过于精细,会获得高分辨率的动画效果,但付出的代价是数据量增大;传输顺序则反应了适应不同领域用户需求从重要到次要数据流导出的先后次序。通过空间数据网上传输与其他视频、动画、图像数据传输的对比,提出了空间数据粒度控制的3个层次划分,即要素级、目标级、几何细节级,通过多个实例应用分析讨论了粒度剖分精细与数据量压缩的技术途径。关于传输顺序建立了数据逻辑组织的空间尺度与语义层次组合矩阵,通过数据元素线性扫描的矩阵行、列优先不同顺序,提出空间尺度优先与语义层次优先的传输过程,讨论了两种传输过程的适用范围。

关键词 渐进式传输 地图综合 流媒体 网络 GIS 空间数据粒度

中图法分类号: P208 文献标识码: A 文章编号: 1006-8961(2009)06-0999-08

Granularity and Order Issues in Progressive Transfer of Web Map

AI Ting-hua

(School of Resource and Environment Sciences, Wuhan University, Wuhan 430072)

Abstract Data granularity and transmission order are two basic questions when transferring steaming maps progressively over web. The map data with high granularity is able to result in high resolution animation but the data volume increases greatly. The transmission order reflects the data stream sequence from coarse to refined at different domains. This study investigates the hierarchical structure of map data organization and presents a granularity classification with three levels, namely feature level, object level and geometric detail level. Through example analysis, the study discusses the application of three granularity partitioning in progressive transmission as well as the strategies to reduce data volume. For transfer order analysis, a mat of semantic and scale is built to describe the map data set and two kinds of transmission order is offered to apply in different progressive transfer processes, namely the semantic priority order and the scale priority order.

Keywords progressive transfer, map generalization, streaming media, web GIS, spatial data granularity

1 引言

作为IT家族中的一员,网络地图在信息服务模式上越来越强调“以人为本”、“个性化”的技术特色,其表现为用户要求能深层次地参与系统的运作,而不是被动地迎合,希望能自主控制地图数据的传输与可视化。适应这一需求,流媒体渐进式传输日

益成为一种受欢迎的传输模式,将部分的控制权交给用户,通过边传输边使用,随时调整传输的粒度、进程。这一渐进式的传输模式在逐级精细化的传输过程中,当信息内容满足用户需求时可随时暂停传输,避免了数据完全下载之后再使用的时间浪费(某些用户可能只需要粗分辨率数据)。空间数据在流媒体传输中所表现出来的从概略到精细化的渐进式可视化,同时具有空间信息导航的功效,满足用户

基金项目:国家高技术研究发展计划(863)基金项目(2007AA12Z209,2007AA12Z346-5(1));国家科技支撑计划基金项目(2006BAB10B01-C);教育部新世纪优秀人才支持计划基金项目(NCET05-0623)

收稿日期:2008-12-03;**改回日期:** 2009-03-26

第一作者简介:艾廷华(1969~),男,教授。2000年于武汉测绘科技大学获工学博士学位。主要从事地图综合研究。E-mail: Tinghua_ai@tom.com

可能的从多角度、多视点由终端到服务器对空间信息的访问、认知、分析,并获得不同层次不同分辨率下的信息表达,增强地图网络服务的自适应性,避免地图数据从北到南、从西到东完整传输中用户只能等待的被动局面(目前大部分地图网站采用这一传统的传输显示模式)。Google Maps 即采用了按重要性差异逐渐叠加的渐进式显示方式,满足用户获取信息具有从概略到精细化的过程。

当前网上传输的数据中,视频、音频、图像等普遍采用了流媒体传输方式,而以矢量数据为主的网络地图要实现这一传输方式还面临诸多问题。Google Earth 中对遥感图像较好地运用了流媒体传输模式,但对矢量地图的传输在流媒体效果上还很不理想,表明矢量数据实现渐进式传输比栅格数据要困难,表现在数据粒度剖分与层次化组织、渐进式

传输中的控制、数据的叠加恢复等方面。视频、音频本身以时间尺度为控制因子将数据组织为线性结构,而矢量地图需要以空间分辨率、语义分辨率为控制因子将数据组织为多层次的非线性结构,将其映射为由时间尺度控制(在不同时刻输出不同信息内容)需要特殊的转换方法。对“非线性层次结构”的矢量地图实现渐进式传输,有一系列问题要解决,诸如传输“粒度”如何划分、数据容量如何压缩、跨尺度间空间关系的维护等。从图 1 的对比中可看出两种数据结构下实现渐进式传输的差别。与结构简单的图像数据相比,矢量数据包含多种几何形式、具有多种空间关系,难于采用图像的诸如“二叉树”层次结构、3 维形体的 LOD (Level of Detail) 层次结构来实现多尺度表达技术。

流媒体渐进式传输是空间数据多重表达在网络

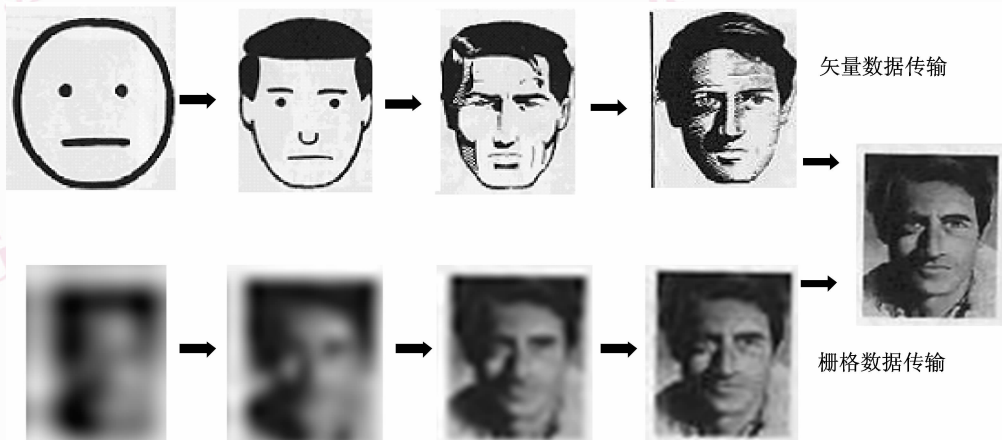


图 1 矢量数据结构与栅格数据结构的渐进式传输比较

Fig. 1 A comparison of progressive transfer between vector data and raster data for the same scene

传输环境中的一个应用,它与地图综合相关^[1],可以看作是地图综合的逆过程(即从粗到细)。要实现矢量地图流媒体传输,关键是基于“粒度”控制因子对地图数据建立逐级精细化的多层次表达模型,并在服务器端建立数据线性索引结构,通过线性检索输出得到数据的从粗到细的精细化表达。该问题的研究赢得广泛兴趣,Bertolotto 和 Egenhofer 最早提出地图数据渐进式传输的概念,并建立了初步的形式化模型,运用综合算子及其组合定义了实现这一流媒体传输形式的几种操作^[2]。Buttenfield 提出了服务器上对空间数据依据重要性的有序化组织策略,认为渐进式传输需要寻求有效的树结构描述细节的层次化^[3]。Han 和 Tao 则从技术实现角度设计了渐进式传输的服务器/终端概念框架,制定了数据

流组织中如何按表达粒度对数据分割“打包”的原则^[4]。Mackness 对网上矢量数据传输的分辨率进行了讨论,建立了传输中数据综合抽象度与分辨率的函数关系^[5]。

对于支持矢量地图渐进式传输的数据结构的研究,一方面该领域对过去的综合模型及数据结构进行剖析,筛选出适用于网络传输的数据结构;另一方面,针对网络传输的特点设计研究新的综合模型与数据结构。Oosterom 提出了 On-the-fly-generalization 的概念^[6],和网络传输的需求正好吻合,后来依据这一思想研制了面向渐进式传输的多边形剖分树模型 GAP-tree (generalized area partitioning) 以及逐级精细化剖分的操作算法^[6-7]。可以服务于这种传输的数据结构还包括 Ballard 早期建立的用于曲线逐

步分解的 Strip-tree^[8],以及 Oosterom 在 Douglas 曲线化简算法基础上建立的平衡二叉树 BLG-tree^[6]。在最近几年,计算机图形学中的 LOD 技术应用到 GIS 数据组织中,通过尺度变换产生了良好的效果。LOD 绘制技术是在粗糙图形表达基础上逐级添加高分辨率的信息内容,达到由“模糊”向“精细化”过渡的可视化效果^[9]。

从数据流程上分析,渐进式传输涉及 3 个环节的技术问题:(1)在服务器端建立多尺度空间数据组织,基于特定的数据模型获得不同尺度的流方式数据集的动态导出;(2)从服务器到终端传输中的流方式控制,包括在线式综合技术的嵌入,以及数据流在“粒度”上的控制;(3)到达终端后多尺度数据的集成及一致性空间关系的实时维护。本文试图对第 2 个问题渐进式传输的过程控制进行研究,包括网络地图渐进式传输中数据粒度剖分与传输顺序控制。通过空间数据网上传输与其他视频、动画、图像数据传输的对比,提出了空间数据粒度划分的 3 个层次,即要素级、目标级、几何特征级。关于传输顺序建立了数据逻辑组织的空间尺度与语义层次组合矩阵,提出两种控制优先的传输顺序。

2 数据粒度控制

网络地图渐进式传输可看作是数据表达从空间尺度到时间尺度的映射转换,在空间尺度控制下的不同层次的数据细节映射到时间轴一帧一帧逐级传输,在时间域上的一个快照对应于在某个空间尺度上的一个表达,在空间尺度要获得较高分辨率的表达,则将要等待较长的传输时间。该映射过程使得我们可以将视频、动画技术的流媒体传输、数据压缩策略借用到矢量地图的流媒体渐进式传输技术中。

视频技术的目标是通过小数据量实现逼真、清晰的动画效果,该技术领域研究的一个关键问题就是解决高保真数据压缩。类似地,空间数据渐进式传输也面临着同样的问题,一方面要获得连续的尺度表达效果,另一方面又不能让数据量无限增大。为解决该矛盾,就要选择合适的数据传输粒度,粒度指空间数据在尺度域变化中的最小分辨单位,在渐进式传输中对应每一步传输的数据体。从空间数据多尺度表达的角度,粒度划分则是对多尺度表达“多”到什么程度的回答。传统技术上,国家序列比

例尺地形图对同一地区采用诸如 1:10 000, 1:50 000, 1:100 000, 1:250 000 等尺度的表达,按照该划分在服务器端存储这些版本的数据库,传输时按要求输出适宜比例尺的数据,也可以获得渐进式传输逐步精细化表达效果,但该变化不连续,是离散的跳跃式变化,说明数据粒度划分过粗。

2.1 三级粒度划分

在视频技术中粒度指一秒钟存储多少“帧”,粒度划分精细,才能获得连续的动感。对地图数据而言,要实现粒度精细的渐进式传输,就需要服务器端能导出粒度精细的不同尺度表达数据。当比例尺发生微小变化 $d(s)$ 时,数据表达在什么层次上发生变化是首先要解决的问题。显然,多尺度空间数据库不可能也没有必要存储无限小粒度的数据表达,比例尺的变化还是离散的、跳跃的,但划分的粒度应适合于具体的应用需求。

GIS 数据可看作是在 3 个层次上的结构组织,即要素层、目标层、几何细节层。要素是指具有相同语义特征的目标集;目标是具有独立地理意义的表达实体,是构成要素的基本单位;几何细节是几何表达上的划分的结构体,是构成目标的基本单位,如构成河流目标的“弯曲”特征,构成面状目标的三角形剖分单元。根据该数据组织层次的划分,网络地图渐进式传输的数据粒度也可划分 3 个层次:要素级、目标级和几何特征细节级。

2.2 要素级与目标级粒度

显然,要素级粒度划分最粗糙,每一步传输显示一个要素层的数据体,相邻两步间的数据跨度较大,终端用户通过数据“层”定制实现要素级的渐进式传输。要素级粒度划分及渐进式传输主要面向特定的应用领域,根据任务目标区分主体要素层、背景要素层,犹如专题地图的表达。

目标级数据传输粒度的划分,表现为根据地理目标的重要性进行排序,传输时逐个导出目标,该重要性度量可根据其几何特征的显著性(面积大小、长度、宽度等)、语义特征重要性(如高等级道路 vs 次要道路)等特征确定。目标集在排好序的线性结构上与表达尺度可建立映射关系,在地图综合技术中通过“选取”算子实现,建立目标传输的“资格比例尺”,当尺度变化到某一刻时,够资格的目标便显示出来,变化的对象以完整的目标图形出现或消失。目标级上的数据粒度划分与多尺度表达可以有两种数据组织,一是横向向上并列关系组织目标实体,按重

要性次序建立线性结构;二是纵向上叠置关系组织目标实体,通过间接地空间组合表达目标。

第 1 种情形由于是并列关系,渐进式传输中后输出的目标不影响前输出的目标,只是总体上增加了重要性次一级的新目标。例如对河系的传输,通

过河流长度、Horton 码、汇水面积等因子对河流重要性排序,传输时逐级输出主河流、2 级河流、3 级河流等^[10],如图 2 所示。对建筑物居民地传输按其面积大小排序逐级输出建筑物多边形,即可获得从总体到细节、从重要到次要的渐进式传输效果。

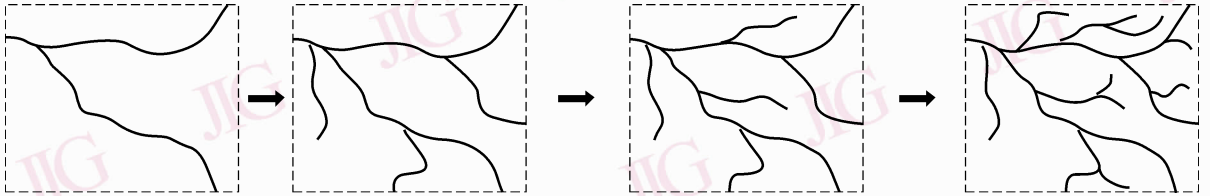


图 2 基于河流长度、Horton 码、汇水面积综合条件排序,渐进式传输河流数据

Fig. 2 The progressive transfer of catchment data based on the river order considering length, Horton code and watershed area

第 2 种情形,渐进式传输中后输出的目标不仅作为新目标出现,同时通过纵向叠置改变前输出目标的特征。GAP-tree 数据结构^[6-7]即属该情形,树的父节点对应的多边形包含其子节点对应的多边形,建立 GAP-tree 线性索引结构后,当前节点与上一个节点的关系为“子父关系”或“兄弟关系”,渐进式传输时子节点对应多边形输出后通过叠置差运算将改变父节点对应多边形的面积,获得更精细化的表达效果。图 3 为扩展 GAP-tree 示例,图 4 为与该数据结构对应的土地利用多边形渐进式传输过程^[7]。

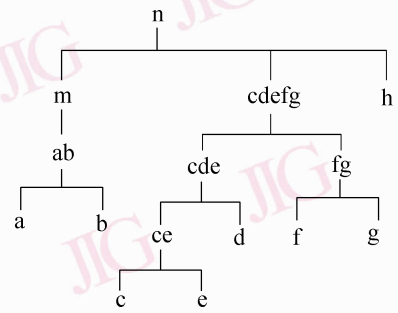


图 3 扩展 GAP-tree 示例

Fig. 3 An example of extended GAP-tree

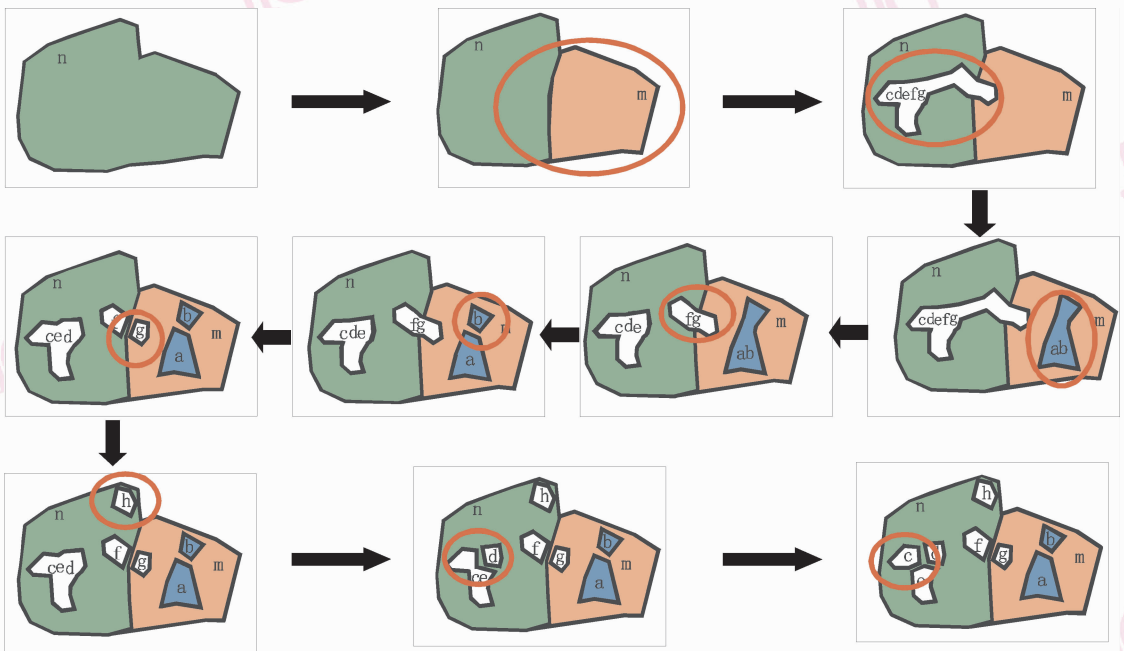


图 4 扩展 GAP-tree 数据结构支持下的土地利用多边形渐进式传输

Fig. 4 The progressive transfer of land use data supported by extended GAP-tree

2.3 几何特征级粒度

几何特征是构成目标表达的基本单位,传输粒度定位到几何特征级上,可获得目标表达的动态演变过程,从简单的线、面演变为复杂的高精确度的线、面。渐进式传输达到该层次的粒度划分才是真正的连续式动画效果,几何特征演变从粗到细是地图综合算子“化简”的逆过程。如果化简算法能连续地导出微小精度差异的表达,可以用该算法来进行几何特征级粒度剖分。

Douglas-Peucker DP 算法是曲线化简的著名算法,把算法执行过程的中间结果按尺度特征详尽地记录下来,建立线性索引即可获得线目标的渐进式传输。DP 算法对曲线的化简是通过矢高的判断依次删除距离基线远的点,最彻底的化简结果是两端点的连接。如图 5 所示,每删除一个点相当于用一个三角形的一边去置换另外两边,反过来,逐级还原化简前原始表达对应为逐级用三角形某两边置换第 3 边。这样对曲线的从粗糙到精细的拟合表达为一系列三角形的累积及其累积过程的边的置换,这里每一个累积的三角形为相邻传输间的差异。对于图 5,将所有变化三角形累积到一起,得到曲线的逐级精确表达为

$$f(s_i) = f(s_0) + \Delta_A + \Delta_B + \Delta_C + \Delta_D + \Delta_E + \Delta_F + \Delta_G$$

其中, s_i 为特定尺度, $f(s_i)$ 为与该尺度对应的表达。截取表达式中的部分项,得到不同尺度下的曲线拟合表达,遗漏的部分项为截断误差,图 5 中从上到下曲线的几种拟合表达依次为

$$f(s_1) = f(s_0) + \Delta_A$$

$$f(s_2) = f(s_0) + \Delta_A + \Delta_B + \Delta_C$$

$$f(s_3) = f(s_0) + \Delta_A + \Delta_B + \Delta_C + \Delta_D + \Delta_E$$

$$f(s_4) = f(s_0) + \Delta_A + \Delta_B + \Delta_C + \Delta_D + \Delta_E + \Delta_F + \Delta_G$$

从 $f(s_1)$ 到 $f(s_4)$, 表达的分辨率越来越高,按照顺序三角形传输获得高粒度的渐进式传输效果。图 6 是基于该原理对实现曲线渐进式传输截取的几个快照。

这样通过对经典化简算法改进获得了线目标高粒度的层次剖分,用于网络地图几何特征级粒度上的渐进式传输,其他综合算法可做类似改进,作者前期研究建立的多边形图壳层次化剖分及变化累积模型^[11]可用于面目标的渐进式传输。

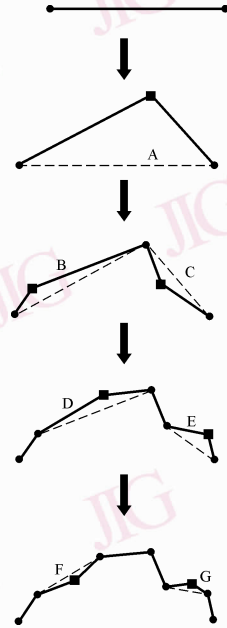


图 5 DP 算法化简表现为序列三角形边的置换

Fig. 5 Triangles accumulation to represent curve at multi-scales based on DP algorithm

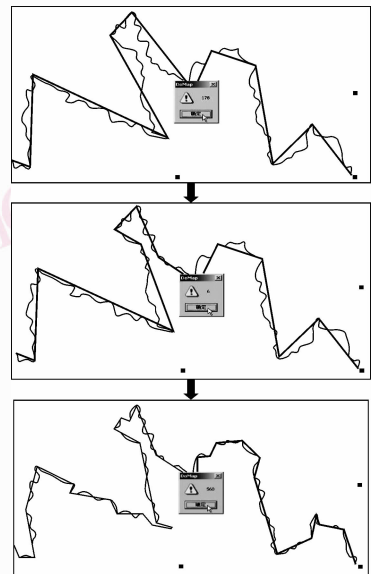


图 6 DP 算法支持下几何特征级粒度传输级可视化演变过程

Fig. 6 The progressive transfer and visualization of curve object with geometric detail granularity

在服务器端数据预组织及渐进式传输中如何进行数据粒度划分与传输控制,取决于应用终端用户

感兴趣信息内容的层次结构。比如,用户感兴趣的是河流流域的网络结构,划分的粒度达到河流目标层即可,没有必要将河流的弯曲逐步演变出来,如果用户感兴趣的是河流的分布,则要深入到弯曲层次,逐步地将河流的曲线表达由概略到详细的过程演变出来。一般地,对数据粒度的划分有如下的原则:发生变化的对象层次比感兴趣的信息内容的层次低一级。

3 传输顺序

渐进式传输中数据按照重要性次序输出,重要的目标或目标表达的重要特征优先输出。决定数据重要性意义的因素包括空间尺度和语义特征。在序列比例尺地形图上,重要的、几何特征显著的信息才

能在小比例尺地图上得以表达,因此表达尺度可以作为目标重要与否的衡量标准。另一方面是否重要还取决于用户的需求,即语义特征的倾向性,面积很小的房屋因为其独特的历史文化意义,使得其重要性突出,相反面积较大的房屋由于没有语义特征上独特性,并不能将其归为重要建筑物类而优先表达。

尺度与语义两者相互独立,某一尺度下包含有各种专题语义的目标,某一语义特征下的目标可以有不同尺度特征的数据表达。多尺度数据库表达的目标可以置入“尺度”×“语义”的 2 维矩阵中,如图 7 所示。该矩阵的行表示某一语义特征目标(如水系类、道路类等)在不同空间尺度下的多重表达,矩阵的列表示某特定尺度(如 1:10 000,1:50 000 等)下包含的多种语义特征目标。矩阵的任意元素 a_{ij} 表示在语义类 i 尺度 j 下的一个表达。

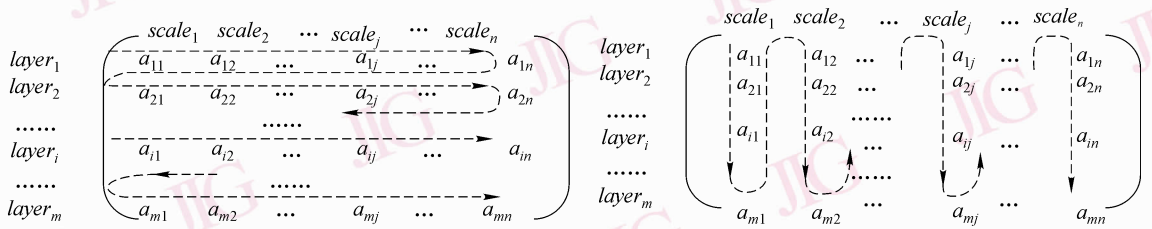


图 7 渐进式传输的两种顺序:语义占优与尺度占优

Fig. 7 Two transfer orders in mat organization: semantic priority and scale priority

基于多尺度空间数据库的渐进式传输,需要对上述矩阵元素建立线性索引结构,传输时扫描该线性结构逐个输出,基于用户视图截取线性结构上的部分数据集实现某尺度、某语义特征的表达。对矩阵结构下的目标建立线性索引有“按行占优”和“按列占优”两种方式,分别对应“语义占优”和“尺度占优”,如上图所示。“语义占优”的多尺度表达体现为某语义层目标在各种尺度下表达传输完成后方可跳到另一语义层目标中,类似于专题地图中的背景层与主题层的区分,这种表达顾及了用户的倾向性。“尺度占优”的线性结构体现为按尺度特征建立目标的顺序表,对语义特征的考虑是平均化的,不特别倾向于某一专题语义的目标,类似于序列比例尺普通地图的表达,六大要素类型平均考虑,仅由尺度特征决定目标是否表达。图 8 通过实例表示了两种传输顺序渐进式传输效果,图 8 完整的数据包含水系、道路、居民地 3 种语义要素的不同尺度表达,这些目标可以根据河流等级、长度、面积大小等指标排序建立线性结构,假定用户最关心居民地的分布。语义

占优的渐进式传输按居民地、道路、水系顺序传输,即使次要的居民地也比高等级道路、一级河流优先传输。尺度占优的渐进式传输顺序为面积大的居民地、等级高的道路、级别高的河流、面积小的居民地、次等级的道路、低级别的河流……该顺序反映了尺度控制效应。

渐进式传输的目的在于提供一种自适应用户需求的表达,通过过程化逐步显示,给用户预留参与控制的机会,将用户想要的内容优先输出。为了应对多用户的需求,并预见其感兴趣的内容,可以对用户分类或按特定任务目标设计渐进式传输的语义顺序。例如对于导航地图的传输就优先传输道路、附属设施加油站、地标建筑物等。上述矩阵对语义层次和尺度分辨率的剖分仍然是离散化的,即按该传输顺序显示用户仍然会获得明显的跳跃感受,如果要获得无缝尺度下的连续显示感受,可以提高剖分的精细度,逼近连续尺度变换,或者在两离散尺度间通过内插获得中间尺度的过渡性表达。

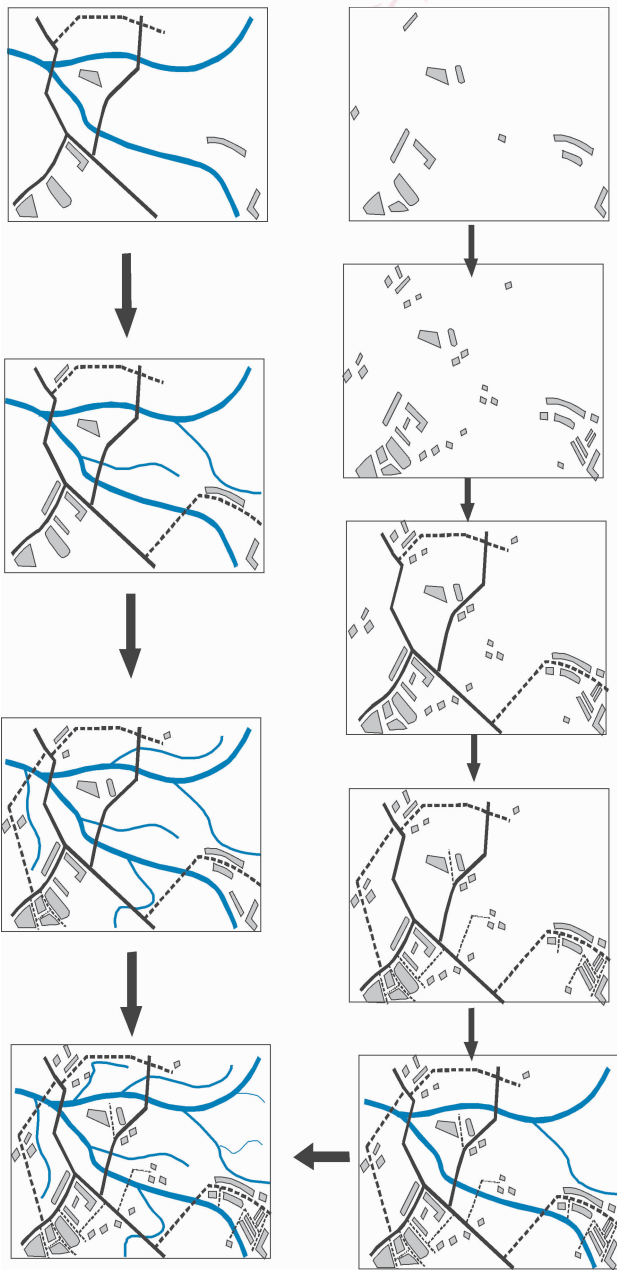


图8 尺度占优与语义占优控制下的两种传输顺序比较

Fig. 8 Two kinds of transfer order respectively corresponding to scale priority and semantic priority

4 结论

网络地图的渐进式传输涉及服务器数据组织、中间传输过程控制、终端数据显示等多个技术环节,对传输过程控制中的数据粒度和传输顺序问题进行了初步研究,在概念层次上对粒度剖分的原则及尺度、语义占优的传输顺序控制进行了讨论。从传统

的地图设计角度,渐进式传输属于地图内容与形式表达在动态过程中的体现,但地图“动”起来后便产生一系列新问题要研究,传输的粒度与顺序控制问题即为其一。

多尺度表达模型的建立和线性结构组织是矢量地图渐进式传输的关键,根据不同传输粒度的划分,在目标级和要素级粒度层次上的多尺度表达比较简单,而在几何细节层次上的多尺度表达,需要寻求具有尺度敏感性的地图综合算子支持。所谓尺度敏感性指尺度发生微小的变化综合算法就可派生出不同的表达,综合算法能在一定尺度范围内动态地输出数据流,需要关注综合过程,而不是像传统地图综合技术那样只关注综合结果。例如地图数据从1:1万到1:10万的综合,应当展示比例尺10倍变化的动态过程,而不是只考虑目标比例尺的结果。目前寻求几何特征级上的多尺度表达及渐进式传输的应用是一个迫切需要解决的问题。

矢量数据渐进式传输的实现与地图综合技术密切相关,传输的进程可看作是地图综合过程的逆过程,在地图综合领域出现了连续式综合、在线式综合、按需制图(on-demand mapping)等新问题的研究^[12-13]。当这些技术还不成熟,还难于获得高粒度的渐进式数据传输效应及自适应用户需求的个性化传输顺序时,可以采用折中的方案,如采用多个比例尺版本的数据存储、预定制多个候选传输顺序、在线处理与离线尺度划分相结合等技术路线,初步实现渐进式传输。

参考文献 (References)

- 1 Yang B, Purves R, Weibel R. Efficient transmission of vector data over the internet [J]. International Journal of Geographical Information Science, 2007, 21(2):215-237.
- 2 Bertolotto M, Egenhofer M. Progressive transmission of vector map data over the world wide web [J]. GeoInformatica, 2001, 5(4): 345-373.
- 3 Buttenfield B P. Transmitting vector geospatial data across the internet [A]. In:Proceedings of GIScience 2002[C], Berlin, 2002, 2478: 51-64.
- 4 Han H, Tao V, Wu H. Progressive vector data transmission[A]. In: Proceedings of 6th Association of Geographic Information Laboratory Europe[C], Lyon, France, 2003: 103-113.
- 5 Mackaness W A, Elizabeth Glover. Dynamic generalisation from single detailed database to support web based Interaction [A]. In: Proceedings of the 19th International Cartography Association[C], Ottawa, 1999:1175-1183.

- 6 Oosterom P Van. Reactive Data Structure for Geographic Information Systems[M]. Oxford: Oxford University Press, 1994.
- 7 Ai T, Oosterom P van. GAP-tree extensions based on skeletons[A]. In: Proceedings of the 10th Advances in Spatial Data Handling[C], Berlin: Springer-Verlag, 2002:501-514.
- 8 Ballard D. Strip trees: a hierarchical representation for curves[J]. Communication of the Association for Computing Machinery, 1981, 14(2): 310-321.
- 9 Weibel R, Dutton G. Generalising spatial data and dealing with multiple representations[A]. In: Geographical Information Systems: Principles, Techniques, Management and Applications[M], 2 Ed, Cambridge, 1999:125-155.
- 10 Ai T, Liu Y, Chen J. The hierarchical watershed partitioning and data simplification of river network[A]. In: Proceedings of the 12th Advances in Spatial Data Handling[C], Vienna: Springer-Verlag, 2006:617-632.
- 11 Ai T, Li Z, Liu Y. Progressive transmission of vector data based on changes accumulation model [A]. In: Proceedings of the 11th Advances in Spatial Data Handling[C], Berlin: Springer-Verlag, 2004:85-96.
- 12 Cecconi A. Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping [D]. Zurich, Switzerland: University of Zurich, 2003.
- 13 Sester M, Brenner M. Continuous generalization for visualization on small mobile devices[A]. In: Proceedings of the 11th Advances in Spatial Data Handling[C], Berlin: Springer-Verlag, 2004:355-368.