

一种视口相关的空间数据多尺度检索方法^{*})

陈 苹 吴秋云 景 宁

(国防科技大学电子科学与工程学院 长沙 410073)

摘 要 合理高效的数据对象选择策略是提高分布式空间数据检索效率的重要因素之一。基于人眼视觉特性,综合考虑了视口尺度对分布式空间数据检索的影响,提出了一种视口尺度相关的分布式空间数据检索方法,将视口尺度集成到分布式空间数据对象选择算法中,能根据视口尺度自适应地调整分布式空间数据对象的选择策略。实验结果表明,该方法在小视口尺度检索时能显著降低处理和传输的数据量,以较小代价快速得到查询结果的近似解。

关键词 多尺度空间数据库,数据检索,空间信息系统

A View-port Dependent Method for Multi-Scale Spatial Data Retrieval

CHEN Luo WU Qiu-Yun JING Ning

(College of Electronic Science and Engineering, National Univ. of Defense Technology, Changsha 410073)

Abstract Effective data selection strategy is one of the important issues of distributed spatial data retrieval. This paper proposes a view-port dependent method for distributed spatial data retrieval based on human vision features with fully considering the effects of view-port with different scale. By integrating the proposed model into the distributed spatial data retrieval algorithm, the method can adaptively tune the distributed data selection strategy according to the scale of current view-port. Experiment shows that the proposed method has favorable performance in small scale view-port query, and can rapidly obtain the approximate result with low cost.

Keywords Multi-scale spatial database, Data retrieval, Spatial information system

在地理信息系统和空间数据库领域中,分布式空间数据检索日益受到人们的重视。分布式空间数据检索的核心是分布式空间查询技术,它关注的是从地理分布的不同空间数据源中检索出满足查询要求的数据,并通过一定的空间运算,返回给查询提交者。分布式空间数据检索所包含的查询类型一般有空间范围查询、空间连接查询和空间聚集查询等,其代价除了通常的计算代价和 I/O 代价外,通信代价成为主要因素之一。由于空间信息获取和数据生产所固有的分布性,在面向大规模空间信息应用的背景下,分布式的空间数据检索必将成为主流的空间信息应用形式之一。

尺度特征是空间数据的固有特征之一,是指数据表达的空间范围的相对大小,不同尺度下空间数据所表达的信息密度有很大的差异,在不同的尺度层次上对空间数据进行访问与表达是现代空间信息系统的重要功能。

多尺度空间数据检索是该领域的一个新的研究问题。一般地,对空间数据的组织都是基于单一尺度,不论空间查询的范围如何,空间数据检索都面向同一细节层次进行处理,该方法不但难以从不同的角度对空间现象进行刻画,而且在进行大范围查询时会使得数据查询和信息检索的代价较大。我们注意到,空间数据库不但具有分布性的特征,而且一般都非常大,在单一的尺度层面上进行空间查询和处理显然不能满足实时检索的要求,因而需要建立一种有效的基于空间数据多尺度特征的数据结构与检索算法以支持分布式空间数据的快速检索,即对于给定的查询条件,利用这种索引结构与算法能够在小代价情况下快速得到查询结果的近似解,从而在整体

上提高基于反馈机制的空间数据检索性能。对于分布式空间数据的索引结构包括 M-R 树和 MC-R 树^[1,2],面向空间数据多尺度组织的数据结构则包括 Reactive 树, BLG 树, Multi-scale R 树、3D R 树^[3-5]等。除此之外,在空间数据库中进行多尺度数据组织的方法还有 z-value 索引^[6]等。

对于空间数据的多尺度检索问题,一般通过比较被检索空间对象的尺度和查询窗口的尺度实现,而与进行查询结果表达处理的视口无关。随着应用环境的不断发展,影响空间数据多尺度表达的因素已不仅仅局限于查询窗口的尺度,还和表达视口的尺度有关^[7],例如, PDA 上的空间信息表达能力就和高分辨率显示设备的表达能力相差较大。因此,面向已有的以及潜在的新型应用环境,有必要研究视口尺度对于多尺度空间信息检索的影响。为此,本文基于人眼视觉特性,综合考虑了视口尺度对分布式空间数据检索的影响,提出了一种视口尺度相关的分布式空间数据检索方法,该方法将视口尺度集成到分布式空间数据对象选择算法中。实验结果表明,本文所提出的方法能根据视口尺度自适应地调整分布式空间数据对象的选择策略,在小视口尺度检索时能显著降低处理和传输的数据量,在较小代价情况下快速得到查询结果的近似解。

1 分布式空间数据检索与多尺度特征

分布式空间数据检索的目标是根据提交的空间查询,从物理分布的不同空间数据源中检索出满足查询要求的数据,并通过一定的空间运算,返回给查询提交者。随着 Web 服务

^{*})基金项目:国家“八六三”高技术研究发展计划基金项目(2002AA134020, 2003AA135110)。陈 苹 博士研究生,研究方向为空间数据库、地理信息系统和 Web 信息检索;景 宁 教授,博导,研究方向为地理信息系统与数据库技术。

等新型分布式计算体系结构的发展,空间信息应用开始采用“服务”的分布式体系结构,以解决长期困扰空间信息应用领域中的互操作问题。典型的基于空间信息服务的分布式空间数据检索系统体系结构如图 1 所示。

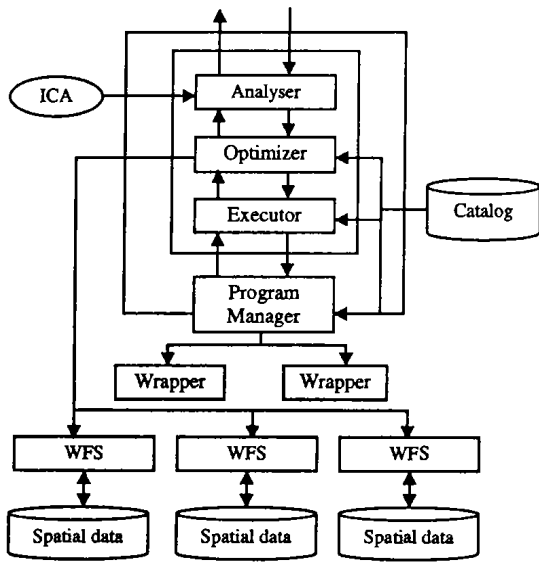


图 1 基于空间信息服务的分布式空间数据检索系统^[3]

在上述系统中,各分布的空间数据源通过 WFS 服务进行包装,数据中介则包括查询分析器、查询优化器与执行器,它根据分布式数据模式转换规则(ICA)和系统目录对提交的查询进行分析与优化,将查询分解为子查询,形成分布式查询计划,启动相应的空间数据服务执行子查询,并汇集各子查询结果,形成完整的查询结果返回给查询提交者。

在基于服务的分布式空间数据检索中,只有那些与检索范围相交的数据源才应该被访问和查询。一般选择算法对空间数据源抽取带有空间语义的元数据,建立空间索引,剪除与检索范围无关的数据源。设 D 为所有数据源的集合, Q 为针对全局模式提交的查询,一般的数据源选择算法在 D 中选择出与 Q 相关的一个数据源子集 $D' = \{D_1, D_2, \dots, D_N\}$,其中 N 为相关数据源个数,进行查询分解后得到面向各个数据源的子查询集合 $\{q_1, q_2, \dots, q_N\}$,其中 q_i 表示 Q 需要在数据源 i 上执行的局部查询, C_i 表示在数据源 i 处理局部查询的计算代价, S_i 表示局部查询在数据源 i 产生的结果集大小, α 表示在网络中单位数据的传输代价。则查询 Q 的总体代价估计 C_Q 为:

$$C_Q = \sum_{i=1}^N (C_i + \alpha S_i) = \sum_{i=1}^N C_i + \alpha \sum_{i=1}^N S_i \quad (1)$$

由上式可以看出,查询 Q 的总体代价由两部分构成:查询的计算代价和数据的传输代价。因此,查询分析器与优化器的主要目标是:

- 1) 减少参与查询的分布式空间数据源的数目,从而降低查询执行总体计算代价;
- 2) 减少查询执行过程中必须在网络上传输的数据量,从而降低通信代价。

为了实现上述目标,在查询分解和优化时一般采用下面两种思路:

- 1) 应用信息集成领域中的模式匹配和映射技术,剪除数据模式上与查询模式不匹配的数据源;
- 2) 应用分布式空间索引技术^[1,2],剪除在空间范围上与查

询空间范围不相交的数据源。

在基于服务的空间信息应用中,最终端的空间信息服务常常是 WMS,这时,空间数据的多尺度特征将对最后的结果表达产生重要的影响。应用空间数据多尺度特征,过滤对最终表达结果无影响的空间数据源或者数据对象,将能够进一步降低查询代价。

传统的对于空间数据多尺度的研究集中在制图综合领域,近年来,由于应用的推动,结合空间数据库技术进行空间数据多尺度检索的研究成为一个新的研究热点。Horhammer 等提出了将尺度作为空间索引的一部分的方法^[4],并通过实验说明了能够有效地减少查询的 I/O 代价。Zhou 等提出了构建空间尺度集成 3D R-tree 索引的方法^[5],能够有效地减少了总体计算代价;陆桑璐等提出在数据空间和对象之间建立一个映射机制(z-value)^[6],将多尺度的概化处理作为空间数据库预操作(pre-database),提高整体系统的查询性能。

上述研究中,均未考虑视口对于不同尺度下空间数据表达的影响。视口是指用于空间数据表达的可视化窗口。随着应用环境的不断发展,视口形式已经从单一的 CRT 显示器发展到不同类型、不同分辨率的显示终端,如高分辨率投影、PDA、手机等。影响空间数据多尺度表达的因素已不仅仅局限于查询窗口的尺度,还和表达视口的尺度有关^[7]。对于相同的空间数据集,在不同分辨率视口中的信息表达能力相差较大。因此,面向已有的以及潜在的新型应用环境,有必要研究视口尺度对于多尺度空间信息检索的影响。

2 视口相关的空间数据检索模型

2.1 基本思想

尺度是衡量空间对象某几何特征量大小的量,空间数据尺度的描述没有统一的定义,一般有两种基本的方法:一是人为指定不同空间对象的尺度,二是基于空间对象的几何属性来自动定义空间对象的尺度。在文[7]中,作者根据人眼的视觉特性和空间对象的几何特征给出了不同的空间对象尺度定义方法。尺度函数的定义可以扩展到高维空间,其具体内容对形式表述无影响。

视口指最终查询结果的显示窗口。图 2 表示了在保证视觉效果基本不变的情况下,一个空间数据集在不同尺度视口下的不同表达。其中(a)表示在大尺度情况下的表达,该结果被表达为最详细形式,(c)表示在小尺度情况下的表达方式,该结果被表达为最简化的形式,(b)介于二者之间。

1) 在视口中小尺度情况下,检索返回的结果集被简化表达,大量的细节信息对于最终显示结果没有贡献;

2) 在视口小尺度情况下,不但检索返回的结果集被简化表达,而且某些数据对象由于尺度太小而对于最终的表达结果没有贡献,如来自于数据源 H 和 I 的数据对象。

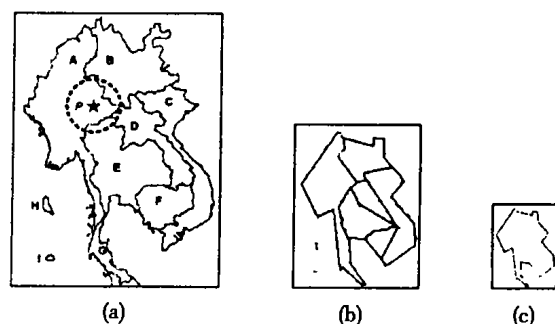


图 2 相同数据集在不同尺度视口下的表达

对于上述(b)、(c)两种情况,如果仍然向所有相关数据源发送完整的查询,则处理和传输对最终结果无贡献的数据将引起检索效率的降低。如果能够构造出空间对象表达方式与视口尺度之间的关系,则可以在给定查询窗口和视口尺度的条件下,计算某对象在该条件下对于最终结果的重要程度,若重要程度较小,则该对象可以在检索结果的最终表达中被忽略,从而能够过滤对于最终表达无贡献的空间数据对象,乃至空间数据源,提高整体检索性能。

2.2 模型描述

将视口尺度作为影响检索结果的因素之后,决定空间数据对象对最终结果贡献的因素主要有三个:数据对象的尺度,查询窗口的尺度和视口的尺度。基于这个认识,综合考虑人眼视觉特性,可以建立一个包含上述三个因素的数学模型,计算某数据对象对于最终结果的影响程度。

下面明确讨论中要用到的几个基本概念:

定义 1 视口(view port)指最终查询结果的显示窗口,记为 V ,比如 128×128 的显示窗口。

定义 2 查询窗口(Query window)指显示窗口所对应的实际地理空间范围,记为 W ,如 $10\text{km} \times 10\text{km}$ 的范围。

定义 3 尺度函数(Scale function)是计算某对象尺度的函数。记为 $SF(x)$,其中, x 为空间对象。尺度是衡量空间对象某几何特征量大小的量,本文中讨论的是二维空间对象,因此选择尺度函数为求面积函数。尺度函数的定义可以扩展到高维空间,其具体内容对形式表述无影响。

以往的研究认为,从纯几何角度考虑,影响空间数据对象在某一尺度下显示方式的因素主要是对象尺度与查询窗口尺度的相对比值,当该值处于不同的门限区间 T 时,对象将被进行相应的处理。其中,常用的门限包括:①对象过滤门限 T_f :用于判定对象是否应该被过滤;②对象简化门限 T_s :用于判定对象是否应该被简化表达。

上述的数学模型是一个二元函数,表示为:

$$T = f(SF(O), SF(Q)) \quad (2)$$

在给定 O 和 Q 后,该模型中门限区间是固定的,没有考虑到视口大小对空间对象表示的影响,即具有“视口无关”性,但在目前越来越多的应用中,视口大小对于空间对象表示的影响是不可忽略的。

通过对人眼某些视觉现象的观察,并结合视觉生理、心理学等方面的研究成果,发现:(1)人眼对图像边缘区信息的失真很敏感;(2)人眼对图像平滑区信息的失真比较敏感;(3)人眼对图像纹理区信息的失真不敏感^[10]。视口尺度对空间对象过滤门限的影响主要体现在以下几个方面:

1)门限的确定。对于包含相同数目空间对象的视口来说,当视口尺度扩大时,空间对象分布较稀疏,视口具有图像平滑区的视觉特征,空间对象的过滤对视觉效果的影响较大, T_f 及 T_s 应较小;当视口尺度缩小时,空间对象分布较密集,视口具有图像纹理区的视觉特征,空间对象的过滤对视觉效果的影响较小, T_f 及 T_s 应较大。因此,综合考虑人眼的视觉特性,在视口尺度较小的情况下,可以过滤更多的空间对象而对最终表现结果的无大影响。也就是说,空间对象的门限不再是固定不变的,而是随着视口的尺度大小而变化。

2)门限的变化率。门限的变化率也随着视口的大小改变,在保持视觉效果基本不变的前提下,视口尺度较大时,门限变化率较小;视口尺度较小时,门限变化率较大。

3)边界条件。根据对象显示的物理意义,当视口尺度趋向

于无穷大时,无论尺度多小的对象都要显示,即门限 $\lim_{SF(V) \rightarrow \infty} T_f = 0$;当视口尺度小于某一值 a 时,所有对象都被过滤(如视口尺度小于人眼分辨率时)。

由以上,构造过滤门限 T_f 的数学模型:

$$T_f(V) = \begin{cases} \frac{1}{\eta SF(V)} & SF(V) \geq a \\ 0 & a > SF(V) \geq 0 \end{cases} \quad (3)$$

其中 η 是视觉分辨力常数, η 越大,则分辨力越大,过滤的数据对象越少,反之,过滤的数据对象越多。相应的,可以定义简化门限 T_s 的数学模型。

将式(3)代入到式(2),可得隐函数如下:

$$T(V) = f(SF(O), SF(Q)) \\ \Rightarrow \eta f(SF(O), SF(Q)) SF(V) = 1 \quad (4)$$

由以上可以看出,将 $\eta SF(V)$ 引入到视口无关模型中,通过计算 $\eta f(SF(O), SF(Q)) SF(V)$ 与 1 的比值,可以确定该空间数据对象在不同尺度视口下的表现方式。

为讨论简单起见,我们设

$$f(SF(O), SF(Q)) = SF(O) / SF(Q) \quad (5)$$

并定义 $\eta f(SF(O), SF(Q)) SF(V)$ 为空间对象 O 的尺度选择性因子 SS (Scale Selectivity)。由式(3)和式(5)可得:

$$SS(O) = \begin{cases} \frac{\eta SF(O) \cdot SF(V)}{SF(Q)} & SF(V) \geq a \\ 0 & a > SF(V) \geq 0 \end{cases} \quad (6)$$

空间数据对象尺度选择性的取值决定了其表现方式和处理方式,在实际应用中,空间数据对象在不同尺度视口下具有多种细节表达层次,为讨论简单起见,本文只设定三个层次:不表达、简化表达、完全表达。则相应的尺度选择性门限为两个:

- 过滤门限:判定不表达与表达,记为 t_f
- 简化门限:判定简化表达与完全表达,记为 t_s

当 $SS(O) < t_f$ 时,该数据对象被过滤,生成查询计划时将不对该数据对象生成子查询;

当 $t_f \leq SS(O) < t_s$ 时,该数据对象将被简化表达,生成查询计划时将生成一个简化的子查询;

当 $SS(O) \geq t_s$ 时,该数据对象将被完全表达,将生成完整的子查询。

运用该模型,我们可以在分布式查询计划生成时,通过候选数据对象的尺度选择性指标,过滤掉那些对最终结果无影响的数据对象,并进而过滤掉在确定尺度意义下无关的数据源。

3 基于 MC-R-tree 的检索算法与代价分析

MC-R 树是面向分布式空间数据组织和并行处理的一种索引结构^[2]。在主站点(Master Site)上,MC-R 树的结构与标准 R 树类似,不同在于其叶子结点的结构为(MBR, client-site-id);在子站点(Client Site)上,MC-R 树的结构与标准 R 树相同。

一般地,基于 MC-R 树的分布式查询处理分为以下步骤:

1. 空间查询将首先被提交到主站点,从主站点 MC-R 树的根结点开始,逐级比较直至叶子结点。当某叶子结点命中之后,此空间查询将被分解为相应的子查询,并提交到该叶子结点所对应的子站点上去。

2. 子站点接收到子查询后,启动本地的查询处理器,应用本地的 MC-R 树索引完成查询,得到查询结果,并向主站

点返回查询结果。

3. 主站点收集所有子查询的结果,进行后续处理(如配准,叠置分析等)后返回给查询提交者。

在上述的处理过程中,如果在步骤1中加入尺度选择性模型,则能够在分发子查询之前将尺度选择性小于过滤门限的数据源过滤掉,从而减少参与查询执行的子站点数目;如果在步骤2中加入空间数据对象的尺度选择性模型,则能够在执行子查询时将尺度选择性小于过滤门限的数据对象过滤掉,并将尺度选择性小于简化门限的对象按照简化方式表达,从而减少通过网络传输返回给主站点的数据总量。

基于上述的思想,我们给出主站点和子站点上 MC-R 树尺度扩展的检索算法,它是一个 MC-R 树的深度优先搜索算法。

算法 1:主站点 MC-R 树尺度扩展检索算法 MCR-MS (MC-R tree Master Scale selection)

输入:R:当前 MC-R 树的根结点;Q:查询窗口;V:视口;
 t_s :简化门限; t_f :过滤门限

输出:result:命中的子站点集合

- (1)对于 R 中的每个子结点 E
- (2)令 I 等于 E 的 MBR 与 Q 的交
- (3)如果 I 不等于空,则
- (4)计算 SS(I)
- (5)如果 $SS(I) \geq t_s$,则
- (6)如果 E 是叶子结点,则将 E 对应的子站点 ID client-site-id 放入 result 中
- (7)如果 E 是非叶子结点,则将 E 设为当前 MC-R 树的根结点,以 E,Q,V, t_s , t_f 作为新参数递归执行本算法

算法 2:子站点 MC-R 树尺度扩展检索算法 MCR-CS (MC-R tree Client Scale selection)

输入:R:当前 MC-R 树的根结点;Q:查询窗口;V:视口;
 t_s :简化门限; t_f :过滤门限

输出:result_simp:简化表达的数据对象集合;result_full:完整表达的数据对象集合

- (1)对于 R 中的每个子结点 E
- (2)令 I 等于 E 的 MBR 与 Q 的交
- (3)如果 I 不等于空,则
- (4)计算 SS(I)
- (5)如果 E 是非叶子结点,则
 - (5.1)如果 $SS(I) \geq t_s$,则
 - (5.2)将 E 设为当前 MC-R 树的根结点,以 E,Q,V, t_s , t_f 作为新参数递归执行本算法
- (6)如果 E 是叶子结点,则
 - (6.1)如果 $SS(I) \geq t_s$,则将 E 对应的数据 ID 加入到 result_full 中
 - (6.3)如果 $t_s > SS(I) \geq t_f$,则将 E 对应的数据 ID 加入到 result_simp 中

由上述算法可以看出,如果搜索结点(树结点或者叶子结点)的空间范围与检索范围不相交或者交的尺度选择性小于过滤门限,则该结点及其子树都将被剪枝。通过这种尺度扩展的选择算法,既能剪除与检索范围不相关的数据对象,又能剪除尺度太小而对最终查询显示结果无贡献的数据对象,因此能够进一步减少参与全局查询的数据对象数目,提高分布式空间数据检索的整体查询效率。

设 N' 为进行完全表达处理的数据源数目, N'' 为进行简

化表达的数据源的数目,则有 $N' + N'' \leq N$,另外,令 d_i 表示在数据源 i 上进行简化计算的代价。则基于上述算法的查询总体代价估计 $C_{Q'}$ 为:

$$C_{Q'} = \sum_{i=1}^{N'} (C_i + \alpha S_i) + \sum_{j=1}^{N''} (C_j + \alpha S_j + d_j)$$

$$= \sum_{i=1}^{N'+N''} C_i + \sum_{j=1}^{N''} d_j + \alpha \left(\sum_{i=1}^{N'} S_i + \sum_{j=1}^{N''} S_j \right) \quad (7)$$

$C_{Q'}$ 同样由两部分构成,一是计算代价,包括进行查询计算的代价和进行简化计算的代价;二是传输代价。

与式(1)比较,当 $N'' = 0$ 时,有 $N' \leq N$,由式(7)易知 $C_{Q'} \leq C_Q$;当 $N'' > 0$ 时,有 $N' \leq N$, $\sum_{i=1}^{N'+N''} C_i \leq \sum_{i=1}^N C_i$,在网络速度相

对于 CPU 计算速度较慢的情况下,有:

$$\sum_{j=1}^{N''} d_j + \alpha \left(\sum_{i=1}^{N'} S_i + \sum_{j=1}^{N''} S_j \right) < \alpha \sum_{j=1}^{N''} S_j$$

从而 $C_{Q'} \leq C_Q$ 。

综合上述两种情况,可以得出:基于视口相关的多尺度数据检索方法的查询总体代价将小于一般方法的总体代价。

4 实验结果与讨论

为了验证本文提出方法的有效性,我们设计了一组实验,在模拟环境下验证视口相关检索方法的性能。我们在局域网中构造了实验环境,实验数据取自 DCW(Digital Chart of the World)数据集^[17]。我们选取其中包括中国和南亚部分的数据作为模拟数据集,因为该部分数据的几何特征较丰富且尺度变化较大。

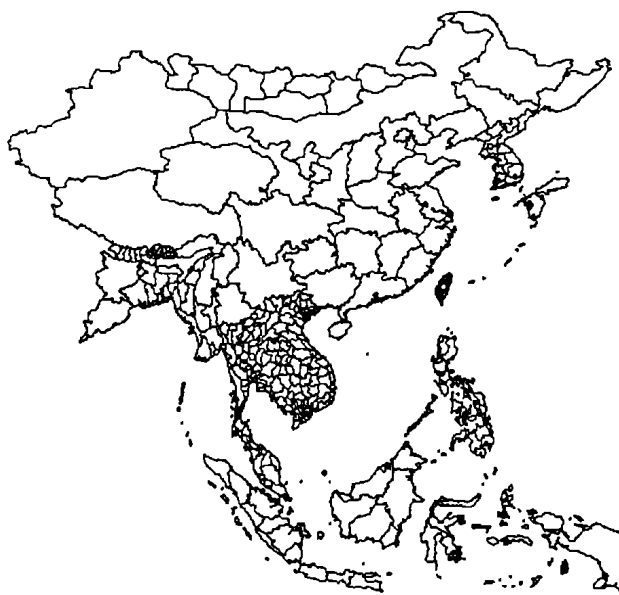


图 3 实验数据集

实验将其中每个行政区模拟为一个分布的数据源。共模拟 359 个分布的空间数据源。实验中采用的三种方法是:

- 1)基于 MC-R 树的视口相关检索方法,视口尺度为 100×100 ;简称 MCR-V1;
- 2)基于 MC-R 树的视口相关检索方法,视口尺度为 400×400 ;简称 MCR-V2;
- 3)基于 MC-R 树的视口相关检索方法,视口尺度为 1600×1600 ;简称 MCR-V3;

实验中采用的参数是: $\eta = 0.25$; $t_f = 1$, $t_s = 100$, $\alpha = 100$;
 对每一种方法,我们构造不同尺寸的查询窗口进行查询,

统计在该查询下各种方法所命中的必须远程访问的数据源个数,并记录相应的网络数据传输量。考虑到空间数据的分布并不均匀,在每一组查询测试中,我们将查询窗口放置在 10 个不同的位置进行测试,并将各结果的统计平均作为该组查询

的结果。

图 4(a)表示了不同尺度视口下查询窗口大小和访问的子站点数目之间的关系,图 4(b)表示了不同尺度视口下查询窗口大小和网络传输数据量之间的关系。

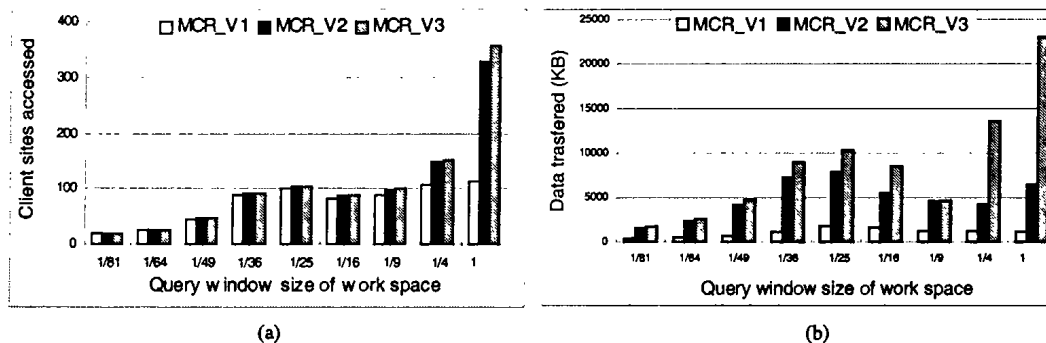


图 4 实验结果

可以看出,在不同查询窗口下, MCR_V1 平均需要访问的子站点数目最小,网络数据传输量也最少;MCR_V3 平均需要访问的子站点数目最大,网络数据传输量也最大。视口相关模型能够在保证视觉效果基本不变的前提下,面向不同尺度视口自适应地调整分布式空间数据源的选择策略和网络数据传输量,从而使系统整体检索性能得到提高。而在视口无关的模型中,检索策略不会随着视口尺度的变化而变化,要么在满足大视口检索性能的前提下而导致小视口检索时冗余数据较多,要么在满足小视口检索性能的前提下而降低大视口检索时的视觉质量。

价情况下快速得到查询结果的近似解。该方法带来的不足是,由于简化计算的影响,被简化数据区域的表现质量有所下降,对此,我们将进一步研究减少视觉质量损失的方法。

目前,该策略与方法已被初步应用到面向城市空间信息应用的原型系统中(图 5),并期望在面向 Web 的分布式空间信息服务(如 WMS, WCS, WFS)中得到广泛的应用。

参考文献

- 1 Koudas N, Faloutsos C, Kamel I. Declustering spatial databases on a multi-computer architecture. In: Proc. of Extended Database Technologies (EDBT), 1996
- 2 Schnitzer B, Leutenegger S T. Master-Client R-trees: A New Parallel R-tree Architecture
- 3 Oosterom V P, Schenkelaars V. The development of an Interactive multi-scale GIS[J]. Intl. Journal of Geographic Information Systems, 1995. 489~507
- 4 Chan E P F, Ueda K. Efficient query result retrieval over the Web. In: Seventh Intl. Conf. on Parallel and Distributed Systems, 2000. Proc. ,2000. 161~170
- 5 Zhou S, Jones C B. Design and Implementation of Multi-scale Databases. In: C S Jensen et al. Eds. SSTD 2001, Springer-Verlag Berlin Heidelberg, 2001. 365~384
- 6 陆桑璐,周晓方,陈贵海,谢立. 空间信息检索及其数据库概化技术. 软件学报, 2002, 13(8): 1534~1539
- 7 Zhou S, Jones C B. A Multi-representation Spatial Data Model. In: Proc. of SSTD 2003 (LNCS 2750). Greece: Springer, 2003. 394~411
- 8 Boucelma O, Essid M, Lacroix Z. A WFS-based mediation system for GIS interoperability. In: Proc. of ACM GIS 2002, McLean: ACM Press, 2002. 23~28
- 9 Horhammer M, Freeston M. Spatial indexing with a scale dimension. In: 6th Intl. Symposium on Spatial Databases, SSD'99, Hong Kong, Springer, 1999. 52~71
- 10 Wang X Y, Yang H Y. A fast image coding algorithm based on human visual system[J]. Journal of Software, 2003, 14(11): 1964~1970
- 11 Digital Chart of the World data server[DB/OL]. <http://www.maproom.psu.edu/dcw/>

另外,我们注意到在查询窗口较大时,小尺度视口检索的性能改善幅度较大,而在查询窗口较小时,三者差别不大。这是由于在查询窗口较大和视口尺度较小时,更多的子站点由于尺度选择性小于门限而被过滤了,而在查询窗口较小时,过滤的数据源比较少。



图 5 城市空间信息服务原型系统

结束语 本文综合考虑了分布式空间数据检索中人眼视觉特性与查询窗口、显示视口的关系,建立了一个计算分布式空间数据对象对于检索结果的尺度选择性模型,提出了一种分布式空间数据对象的选择方法。该模型和方法被集成到基于分布式空间索引结构的对象选择算法中,能够根据查询窗口的尺度和显示视口的尺度对分布的空间数据源以及数据对象进行动态选择,进一步降低了分布式空间数据集的获取代价。实验结果表明,本文所提出的方法能够根据视口尺度自适应地调整分布式空间数据对象的选择策略,在小视口尺度检索时能够显著降低处理和传输的数据量,能够在较小代