

# 一种用于空间数据多尺度表达的 R 树索引结构

邓红艳<sup>1),2)</sup> 武 芳<sup>1)</sup> 翟仁健<sup>1)</sup> 赵 倩<sup>3)</sup>

<sup>1)</sup>(信息工程大学测绘学院地图学与地理信息工程系 郑州 450052)

<sup>2)</sup>(中国科学院地理科学与资源研究所 北京 100101)

<sup>3)</sup>(空军指挥学院 北京 100097)

**摘 要** 针对现有 R 树无法支持空间数据多尺度表达的问题,提出了一种用于空间数据多尺度表达的 R 树变形索引结构:(1)允许空间对象出现在非叶结点上;(2)利用树的深度反映空间分辨率的变化,提供分辨率维的支持;(3)树的分支结构考虑对自动制图综合算法的支持.分析了该变形 R 树索引结构的空问数据多尺度查询过程,并着重对该索引结构生成算法中的约束条件、插入算法和分裂算法进行了讨论.针对相同数据源,使用该方法与基于四叉树的空问数据多尺度索引方法进行了对比实验,结果表明,该索引方法能有效检索多分辨率形式组织的空问数据,具有综合结果记忆功能,效率明显.

**关键词** 空间数据;多尺度表达;R 树;索引结构;地理信息系统

中图法分类号 TP391

DOI 号: 10.3724/SP.J.1016.2009.00177

## R-Tree Index Structure for Multi-Scale Representation of Spatial Data

DENG Hong-Yan<sup>1),2)</sup> WU Fang<sup>1)</sup> ZAI Ren-Jian<sup>1)</sup> ZHAO Qian<sup>3)</sup>

<sup>1)</sup>(Department of Cartology and Geographical Information Engineering, Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052)

<sup>2)</sup>(Institute of Geographic Sciences and Natural Resources Research, Beijing 100101)

<sup>3)</sup>(Airforce Command College, Beijing 100097)

**Abstract** Aiming at multi-scale representation of spatial data not supported by R-tree, a transmutation index structure of R-tree used for multi-scale representation of spatial data is proposed: (1) Spatial objects may appear at non-leaf nodes; (2) Depth of the tree is made use for reflecting change of spatial resolutions, and resolution dimensions is supported; (3) Structure of the tree's branches supports algorithms of automatic cartography generalization. Query of spatial datas among multiple scales of based on the R-tree's transmutation index structure is analyzed, and constraint conditions, insert algorithms and divide algorithms among creations of the index structure are emphasized. For the same data source, compared with multi-scale index based on quad-tree for spatial data, experiments prove that the transmutation structure can search spatial data organized by multi-resolution and memorize generalization results with high efficiency.

**Keywords** spatial data; multi-scale representation; R-tree; index structure; geographical information system

收稿日期:2005-12-31;最终修改稿收到日期:2008-04-17. 本课题得到国家自然科学基金(40671162,40671152)和国家“八六三”高技术研究发展计划项目基金(2007AA12Z211)资助. 邓红艳,女,1978年生,博士,讲师,主要研究方向为空间数据多尺度表达与地理信息系统. E-mail: denghongyan\_gis@163.com. 武 芳,女,1964年生,博士,教授,博士生导师,主要研究领域为空间数据多尺度表达与自动制图综合. 翟仁健,男,1980年生,博士研究生,讲师,主要研究方向为基于遗传算法的自动制图综合与空间数据库. 赵 倩,女,1963年生,博士,副教授,主要研究方向为数据库技术及应用.

## 1 引言

随着信息时代计算机技术的高速发展,地图学与计算机的综合产物——地理信息系统(Geographical Information System, GIS)也得到了迅速发展,其应用规模和需求不断扩大和增加.为解决用户对地理数据的不同应用需求和分析的需要,空间数据多尺度处理与表达已经成为当前地理信息科学领域一个亟待解决和前沿性问题<sup>[1]</sup>.

空间数据多尺度处理与表达涉及的问题有很多,但归纳起来主要有三个方面:首先是如何分割地理空间也就是如何划分多尺度序列子空间的问题,其中比较有代表性的是从空间认知的角度进行的研究<sup>[2-3]</sup>.其次就是如何生成子空间序列,同时动态建立并维护它们之间的空间关系,即空间多尺度数据的处理问题,不少学者从制图综合的角度对此进行了研究,并取得了一些阶段性成果,如基于小波理论的处理模型<sup>[4]</sup>、基于 Delaunay 图的处理模型<sup>[5]</sup>、基于遗传算法的处理模型<sup>[6]</sup>以及一些通用综合算法组合而成的处理模型等.最后是如何管理已生成的子空间序列,即空间数据多尺度表达数据索引结构的建立.这一问题不仅仅对于空间数据多尺度表达中的速度具有很大的影响,同时也决定了其分析和查询的速度,已经日益得到人们的重视.这一问题的基本解决思路是改进现有的空间数据索引方法以适合空间数据的多尺度表达,目前较有代表性的是基于四叉树的空间数据多尺度表达索引结构<sup>[7]</sup>.但是由于四叉树本身的限制,基于四叉树的索引在管理海量空间多尺度数据时效率太低,无法满足人们的需求.为解决这一问题,我们需要寻求新的用于空间数据多尺度表达的索引结构.

R 树<sup>[8-9]</sup>是目前 GIS 中应用非常广泛的一种空间数据索引结构,实践证明其对建立大数据量空间数据索引机制进行查询,相对于其它空间索引方法具有更加出色的表现.空间数据多尺度表达实质上也可以被看作是一种空间分辨率与浏览区域的组合查询,因此利用 R 树建立空间数据多尺度索引在需求上是完全合理的.但 R 树并未提供对多分辨率空间数据的索引方法,从而限制了它在多分辨率空间数据中的有效应用,目前也有少量这方面的研究工作<sup>[10-11]</sup>,但只是简单的探索性的研究,并未形成完整的解决方案.本文分析了传统 R 树进行空间数据多尺度表达的局限性,在此基础上提出了一种用于空间数

据多尺度表达的 R 树变形索引结构——SDMR 树(Spatial Data Multi-representation R Tree).

## 2 R 树用于空间数据多尺度表达问题的分析

R 树索引最早是由 Guttuman 在 1984 年提出的,经过不断的改进和优化,形成了一个以 R\* 树、Hilbert R 树、Shift-Optimal R 树等为代表性的 R 树类索引体系.由于 R 树实质是 B 树索引在空间维上的扩展,同时采用平行于数据空间轴的最小外接矩形(MBR)来代替复杂的空间对象建立索引,因此它不仅具有 B 树特有的动态平衡性,同时相对于其它索引方法而言,在处理空间数据尤其是海量空间数据时具有更大的灵活性和更高的效率.随着无缝 GIS 中海量空间数据多尺度表达的需要,人们利用 R 树来建立多分辨率数据索引的需求也越来越迫切.但 R 树本身并未提供对多分辨率空间数据的支持,因此直接将 R 树用于空间数据多尺度表达面临着许多问题,概括起来主要表现在 3 个方面<sup>[1,10]</sup>:

(1) 对象实体具有同等深度,在结构上不利于 GIS 多尺度表达.

R 树结构中,对象实体只能存在于 R 树的叶子结点中,具有相同的深度.而在 GIS 多尺度表达中,空间对象往往根据其分辨率的不同,形成树形分级结构.显然传统 R 树不具备描述这种多层结构的能力,无法直接用于空间数据的多尺度表达.

(2) 缺乏对空间对象分辨率维的有效支持.

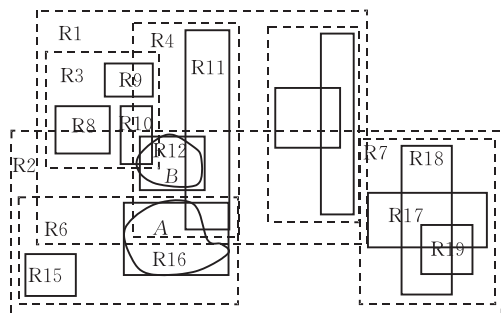
在空间数据多尺度表达过程中,分辨率维是空间对象是否显示的决定因子,整个多尺度显示过程的实质是分辨率维和空间维的联合查询过程.传统的 R 树索引是一种在空间维上的索引结构,缺乏对空间对象实体分辨率维的有效支持,因此无法直接用于空间数据的多尺度表达.

(3) R 树分支结构没有考虑自动综合的需求.

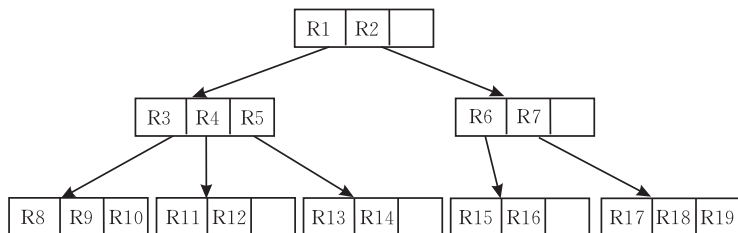
实现空间数据多尺度表达的一个关键因素就是自动综合算法的有效实施.现阶段,自动综合算法已不仅仅只对单个要素进行综合,同时还考虑了要素与要素之间的关系、要素与周围环境的语义关系等.传统 R 树索引建立过程中,对于语义关系、拓扑关系等综合辅助信息缺乏考虑,给自动综合实施带来了不便. R 树的经典结构如图 1 所示.图 1(a)中 A、B 表示需要进行合并的空间要素,利用传统 R 树生成算法被分到两个不同的分支中,如图 1(b)中的

R12 与 R16, 无论是从 A 找寻到 B 还是从 B 找寻到 A, 都需要进行另一分支的完全搜索, 这时效率不仅

没有提高, 相反还增加了算法的复杂度.



(a) R 树矩形分块结果



(b) R 树的结构

图 1 R 树结构

综上所述, 要将 R 树有效地用于空间数据多尺度表达, 需要对现有 R 树进行合理变形, 因此, 本文提出了一种用于空间数据多尺度表达的 SDMR 树索引结构.

### 3 SDMR 树

SDMR 树是基于 R 树的一种变形, 其基本思想是: (1) 借鉴 Oosterom 提出的在 Reactive 树<sup>[12]</sup>中引入对象实体重要度因子的思想, 在 R 树中引入分辨率维, 同时利用树的深度变化来反映空间数据多尺度表达中的分辨率变化; (2) 允许对象实体在较高的树层次上出现; (3) 在树的生成和操作算法中, 考虑空间对象的关系, 使得树的分支结果与实际地理特征相符合, 以便于采用综合算法. 下面对其关键问题进行讨论.

#### 3.1 SDMR 树的定义与多尺度查询过程

参照 Guttman 对于 R 树的定义形式<sup>[8-9]</sup>, 本文给出 SDMR 树的定义. 其中设  $M$  为结点中单元的最大数目,  $m$  ( $1 \leq m \leq M/2$ ) 为非根结点中单元个数的下限:

(1) 每个结点包括的单元个数介于  $m$  与  $M$  之间, 除非它同时是根结点.

(2) 根结点最少有两个孩子, 除非它同时是叶子结点.

(3) 允许空间对象出现在非叶结点上, 结点中单元的基本形式为  $(I, PointToChild, SpatialObjectID)$ , 如果  $PointToChild$  为空, 表示结点单元中存储的是空间对象信息, 其中  $I$  是空间对象的最小包含矩形 (MBR),  $SpatialObjectID$  指示的是空间对象, 简称空间对象单元; 如果  $PointToChild$  不为空, 表示该结点单元为树的分支结点, 其中  $I$  是包含

其所有子结点的 MBR,  $SpatialObjectID$  指示的是其子结点的综合结果 (如果  $SpatialObjectID$  为 0 表示其子结点综合结果没有产生及存储), 简称树枝结点单元.

(4) 树的深度表示空间数据多尺度表达的分辨率, 从根结点开始, 越往下分辨率越大.

(5) 每个非空间对象结点在地理区域划分上包括其子结点, 即综合操作不能在同一层次上跨结点操作.

根据 SDMR 树的定义可知, SDMR 树与传统 R 树的主要区别体现在第 (3)~(5) 点, 而这正好是传统 R 树被用来建立空间数据多尺度表达索引的局限性所在, 即是否允许空间对象出现在非叶结点上, 是否提供空间分辨率维的支持, 分支结构是否考虑对综合算法的支持. 正因为这些特征的存在, SDMR 树能够提供空间数据多尺度表达的视图查询.

假设已经将空间数据按照空间认知的方法分为  $n$  个子空间序列  $v_1, v_2, v_3, v_4, \dots, v_n$ , 即  $n$  种分辨率显示子视图 (子空间序列划分的问题属于空间数据多尺度表达中另一个重要问题, 本文将不予以讨论), 用 SDMR 树的纵向深度表示空间分辨率轴, 横向表示空间坐标轴, 则可以构成如图 2 所示空间数据多尺度查询坐标空间, 其中  $v_1, v_2, v_3, v_4, \dots, v_n$  分别用空间分辨率轴上一个坐标点进行刻画, 坐标点的位置由具有  $v_i$  (其中  $1 \leq i \leq n$ ) 分辨率空间对象最深结点决定. 当空间对象的显示分辨率为  $\{v \mid v_i < v \leq v_j\}$  (其中  $1 \leq i < j \leq n$ ), 则  $v_i, v_j$  之间某一结点单元可以直接指向空间对象. 当需要查询某一区域  $S$  在分辨率  $v_j$  的子空间视图时, 其实质就是在空间数据多尺度查询坐标空间中寻求符合条件坐标点, 即在区域查询上将依照 R 树查询方法, 在分辨率上查询小于等于  $v_j$  分辨率直接存储的空间要素和  $v_{j+1}$

的综合结果,具体步骤如下:

1. 子树的搜索. 从 SDMR 树的根结点  $T$  开始依次根据该结点中各单元情况进行判断: 如果是树枝结点单元则判断  $I$  与区域  $S$  的空间关系, 如果  $I$  与搜索区域  $S$  相交, 判断子树深度, 如果子树深度小于  $v_j$ , 则以该单元所指向的结点为子树的根结点, 重复上面的操作; 如果子树深度等于  $v_j$ , 则转至步 3. 如果是空间对象单元, 且子树深度小于等于  $v_j$ , 则调用步 2.
2. 空间对象结点判断. 判断空间对象与搜索区域  $S$  的空间位置关系, 如果空间对象落在搜索区域  $S$  内, 表明其满足搜索条件, 继续子树搜索.
3.  $v_j$  深度子树查询结果处理. 对于深度为  $v_j$  的子树, 由于 SDMR 树中的特征 3, 可以将其作为单独的综合处理区域, 其综合的尺度就是  $v_j$ , 综合结果可以是预先存储的, 也可以动态生成:
- 3.1. 如果结果是预先处理存储的结果则可以在树枝结点单元中通过  $SpatialObjectID$  直接获取, 然后调用步 2, 具体在数据存储中可以将综合存储的数据加以标识, 与原始

- 源数据进行区分.
- 3.2. 如果是动态生成, 其基本过程是: 将深度为  $v_j$  的子树独立出来, 分别构成单独的如图 3 所示的 SDMR 树  $ST_1, ST_2, ST_3, \dots, ST_k$  (其中  $k$  为查询到的子树的数量), 对于子树  $ST_i$  (其中  $1 \leq i \leq k$ ) 而言, 类似于步 1 找寻其子树, 若子树中存在空间对象目标或者综合结果则直接采用综合算法进行综合获取并存储, 同时修改相应的父结点中的  $SpatialObjectID$ , 该分支处理过程结束; 若是树枝结点单元, 不存在综合结果且深度小于  $v_{j+1}$ , 则调用步 1; 若是树枝结点单元并且深度等于  $v_{j+1}$ , 则进入下一分辨率的查询; 若存在  $v_{j+1}$  子树分支综合结果, 则依据综合因子  $v_j$  进行综合, 同时存储处理过的子树综合结果, 修改父结点中的  $SpatialObjectID$ ; 若不存在  $v_{j+1}$  子树综合结果, 则需要对  $v_{j+1}$  分辨率做类似于  $v_j$  的空间数据多尺度查询, 依此类推, 可以获取最高分辨率  $v_n$  上的查询结果, 然后通过不断综合处理获取各子树分支的综合结果, 逐步综合, 最后获取分辨率  $v_j$  上的视图空间, 如图 3 所示.

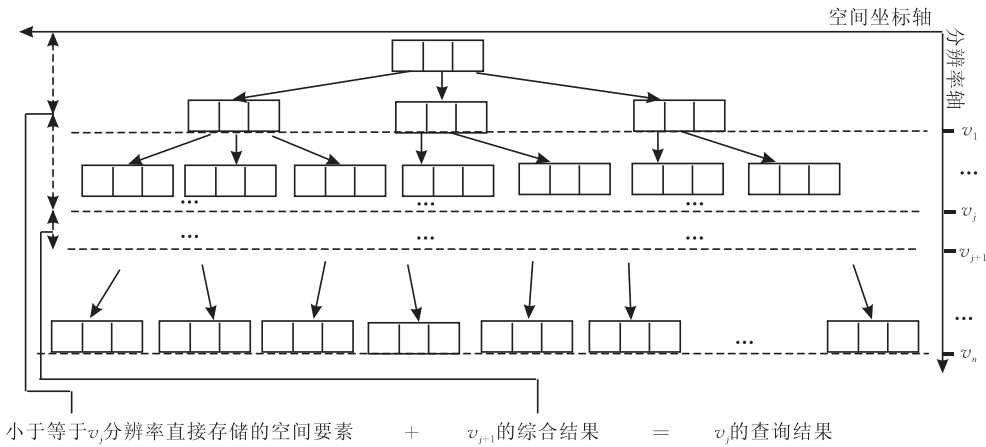


图 2 SDMR 树查询空间

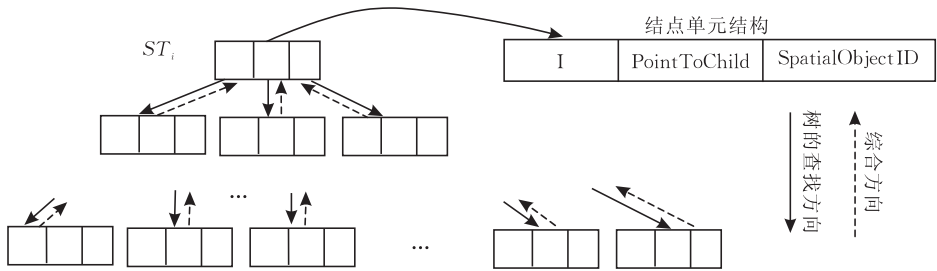


图 3  $ST_i$  子树查询过程

3.2 SDMR 树的生成算法

基于 SDMR 树的空间多尺度查询过程是以生成好的 SDMR 树为基础的,SDMR 树对空间多分辨率的支持决定了其生成过程不同于普通 R 树的生成过程,需要考虑一些特殊条件限制,其关键性问题包括两个方面:

(1) 约束条件生成

在空间数据多尺度表达过程中,境界、主要水系和高等级道路通常情况下是人们最先能看到的要素,同时也是制约其他要素正确综合和表达的关键因素,例如对于居民地的合并操作,不仅仅是不能进行跨道路合并,对于跨境界、跨水系的操作同样不合

理. 因此, 利用境界、主要水系和高等级道路形成的网络进行区域划分不仅符合地理特征而且也适合制图综合算法实施.

同时考虑到自动制图综合结果需保持空间分布要求, 本文结合 Agent 项目组提出的数据组织方式<sup>[13-14]</sup>, 提出了一种基于地理区域划分的约束条件, 其基本原理是: 首先对除境界、水系和道路网络以外的其它要素进行目标的自适应聚类, 以获取目标的基本空间分布和密度对比; 然后以最小位移限差的三倍为缓冲距离建立缓冲区, 探测目标间可能发生位移的区域, 将可能因为位移发生冲突的聚合类别进行合并; 最后利用境界、水系和道路网生成拓扑网络, 对经过聚类和缓冲区冲突探测处理的区域进行分割. 这样就构成了 3 个级别的约束范围: 约束 1 为分割后的单独聚类区域; 约束 2 为分割后的缓冲域; 约束 3 为境界、主要水系和道路网形成的区域. 将这 3 个级别的约束作为条件加入到传统的 R 树的插入和分裂算法中, 就形成了 SDMR 树的插入与分裂算法.

(2) SDMR 树的插入与分裂算法

SDMR 树的插入算法与传统 R 树的插入算法类似, 均包括 4 个步骤, 但由于加入了约束条件的考虑, 在具体实施过程中将有所不同:

1. 为新的空间对象寻找一个合适的结点位置.

1.1. 初始化. 令  $N$  为 SDMR 树的根结点.

1.2. 判断是否为合适的结点. 如果  $N$  的分辨率等于空间目标的显示分辨率, 则返回  $N$ .  $N$  即为要寻找的存储结点. 如果  $N$  的分辨率大于等于空间目标的显示分辨率, 则以空间目标为基础生成新的结点结构  $N$ , 同时返回  $N$ .

1.3. 选择合适的子树. 如果  $N$  不是合适的结点, 将空间对象依次与该结点的各单元进行约束条件的判断, 如果符合约束条件 1, 则令该单元为候选单元, 如果候选单元只有一个, 则令  $N$  为该单元所指的子结点, 从步 1.2 开始重复进行上面的步骤, 如果候选单元超过两个, 则按照 R 树中选择子树面积计算最小法则选择; 如果不符合约束条件 1, 则采用上述方法依次判断约束条件 2、约束条件 3; 如果以上 3 个

条件均不符合, 则按照 R 树中选择子树的面积计算最小法则选择合适的  $N$ .

2. 将新的空间对象记录到结点中

如果结点  $L$  中尚有空位(单元数小于  $M$ ), 则在  $L$  中记录新空间对象的 MBR 和 ID. 如果没有空位, 则将  $L$  结点中的单元(包括新的空间对象), 分为两个部分, 分别记录在结点  $L$  和新生成的结点  $LL$  中, 具体的结点分裂算法(以经典的平方耗费算法为基础)如下:

2.1. 为两类区域选择第一个矩形. 选择两个最不可能在同一类的矩形作为种子, 分别赋给两类, 种子的选取首先采用约束条件 3 进行判断, 不符合约束条件 3 的作为候选种子, 如果均符合约束条件 3, 则依次用约束条件 2、约束条件 1 进行判断选取候选种子; 如果约束条件 2、约束条件 3 均不满足, 则所有的矩形为候选种子; 如果候选种子为一对, 则将其作为种子; 如果候选种子超过一对, 则将候选种子采用经典平方耗费算法中的种子矩形方法确定种子.

2.2. 是否需要停止分配. 如果矩形集中的矩形均已分配完毕, 则分裂操作结束, 返回; 如果将矩形集中剩下的矩形都加入到某一类中, 才能使矩形的个数等于  $m$ , 则将所剩矩形都分配给该类, 并结束分裂操作, 返回.

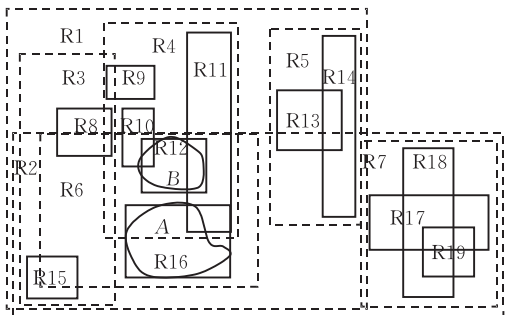
2.3. 分配矩形. 对矩形集中剩下的矩形进行分配, 选取其中的某一个矩形, 加入到其中一类, 基本判据是如果该类加入该矩形后, 满足约束条件 1, 而加入另一矩形不满足, 则加入该矩形; 如果该类加入该矩形后, 满足约束条件 1, 而加入另一矩形同样满足, 则采用平方耗费算法中的面积判断方法进行判断; 如果均不满足约束条件 1, 则采用约束条件 2、约束条件 3 进行与上相同类似的判断; 如果约束条件 2、约束条件 3 均不满足, 则直接采用平方耗费算法中的面积判断方法进行判断.

3. 调整树的结构.

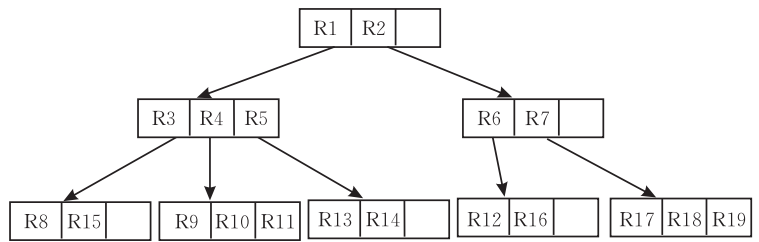
4. 生成新的根结点, 调整分辨率维的刻度.

如果由于孩子结点的分裂, 导致根结点中单元个数超过  $M$ , 则需将根结点分裂为两个结点, 并生成一个新结点作为根结点, 而原根结点分裂成的两个结点作为其子结点. 此时 SDMR 树增加 1 层, 相应的分辨率维刻度向下移动一个位置.

图 4 是 SDMR 树在初始过程中形成的矩形分块与树基本结构, 与图 1 相比, 需要进行合并的空间



(a) 矩形分块结果





要素  $A, B$  因为采用了约束条件的限制被分到同一结构分支中,有效解决了图 1 中由于  $A, B$  处于不同树结构分支而增加综合计算繁琐程度的矛盾.

### 4 实验结果与结论

本文采用 SDMR 树进行空间数据组织,针对某一区域进行比例尺 1:250000 至 1:500000 的空间数据多尺度显示实验(如图 5 所示). 其中图 5(a),为 1:250000 的实际比例尺显示结果,为了使结果看的清楚,图 5(b)~图 5(d)分别为相应比例尺下显示结果的放大图形. 观察图 5 的变化过程,特别是对于圆形标注出来的居民地区域,我们可以清楚的看到,对于图 5(a)与图 5(b),虽然比例尺变化不大,因为受图面显示区域大小的限制,图 5(b)的居民地显示细节不如图 5(a)详细,变化到图 5(c)与图 5(d),变化就更加明显. 这一过程正是空间数据的多尺度表达过程,充分说明 SDMR 树完全支持空间数据的多尺度表达.

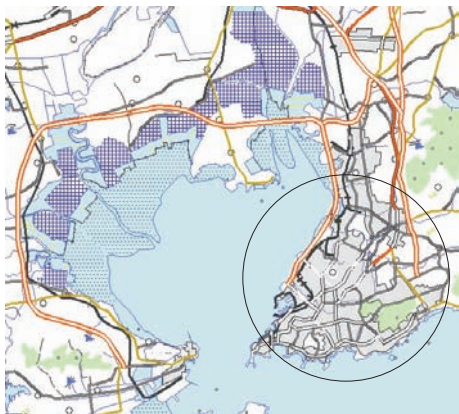
在此基础上进一步通过实验比较基于 SDMR 树的空间数据多尺度索引结构与基于四叉树的空间

数据多尺度索引结构的显示性能,实验数据依旧采用图 5(a)所示区域,采用同一综合算法库,依次实现空间数据 1:250000 到 1:500000 的多尺度显示功能. 表 1 是在不同显示比例尺下两种索引的不同耗费时间.

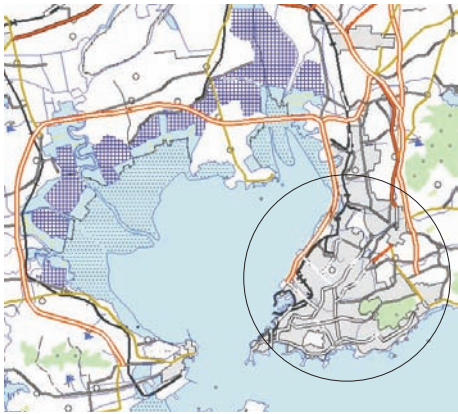
表 1 索引耗费时间对比

显示比例尺	基于 SDMR 树的 耗费时间/ms	基于四叉树的 耗费时间/ms
1:300000(第 1 次)	2967	3268
1:300000(第 2 次)	1322	3394
1:300000(第 3 次)	1245	3197
1:300000(第 4 次)	1304	3244
1:400000(第 1 次)	2176	3499
1:400000(第 2 次)	1298	3438
1:400000(第 3 次)	1289	3425
1:400000(第 4 次)	1273	3367
1:500000(第 1 次)	1833	3523
1:500000(第 2 次)	1255	3479
1:500000(第 3 次)	1312	3514
1:500000(第 4 次)	1236	3477

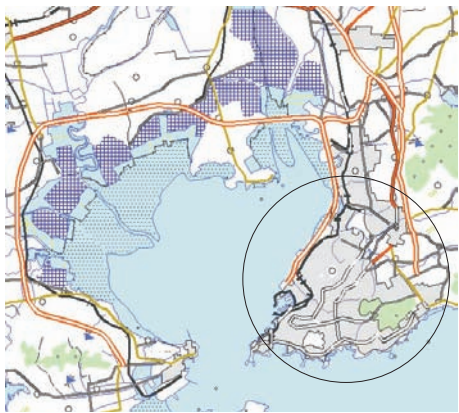
从表 1 纵向看,采用 SDMR 树的方法建立空间数据多尺度索引由于树中结点能够记录综合结果,因此下一次同等分辨率的查询时间将大大节省,同时在进行更低分辨率查询时,上一分辨率的综合结



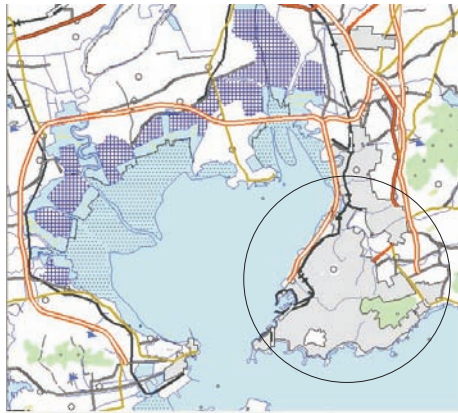
(a) 1:250000实际显示图



(b) 1:300000放大显示图



(c) 1:400000放大显示图



(d) 1:500000放大显示图

图 5 采用 SDMR 树的多尺度显示结果

果能够被利用,因此如果是按照从高分辨率到底分辨率的显示方向,显示时间也明显缩短.从表的横向结果进行比较,在同等分辨率的显示情况下,基于SDMR 树建立的索引方法由于在树的分支信息中本身就隐含了制图综合算法所需要的要素之间关系,因此在处理时间上要优于基于四叉树的方法.

由实验结果可知,SDMR 树不仅在显示效果上满足空间数据多尺度表达的要求,并且与传统四叉树相比具有更高效率.然而,要完全实现空间数据多尺度表达只研究其数据结构是不够的,研究更多高效率、高实用性的综合算法和模型也是解决这一问题的关键,这些都需要我们做进一步的工作.

参 考 文 献

[1] Mackaness William A, Ruas Anne, Sarjakoski L Tiina. Generalisation of Geographic Information: Cartographic Modelling and Applications. Amsterdam, Netherlands: Elsevier, 2007

[2] Vangenot Christelle. Multi-representation in spatial database using the MADS conceptual model//Proceedings of the International Cartographic Association Workshop on Generalization and Multi-Scale Representation. Leicester, England, 2004: 337-342

[3] Mark David M, Freksa Christian et al. Cognitive models of geographical space. Geographical Information Science, 1999, 13(8): 747-774

[4] Wu Fan. Research on multi-scale handling and representation of geospatial data[Ph. D. dissertation]. Wuhan University, Wuhan, 2002(in Chinese)  
(吴凡. 地理数据的多尺度处理与表示研究[博士学位论文]. 武汉大学, 武汉, 2002)

[5] Peng Hu, Qi Qing-Wen, Liu Zhao-Li. Progress in studies on automated generalization of spatial point cluster. IEEE International Geoscience and Remote Sensing Symposium, 2004, 13(8): 2841 -2844



**DENG Hong-Yan**, born in 1978, Ph. D. , lecturer. Her major research interests include multi-scale representation of spatial data and geographical information system.

**WU Fang**, born in 1964, Ph. D. , professor, Ph. D. supervisor. Her major research interests include multi-scale

[6] Deng Hong-Yan, Wu Fang, Qian Hai-Zhong, Hou Xuan. A generalization model of road networks based on genetic algorithm. Geomatics and Information Science of Wunan University, 2006, 31(2): 164-167(in Chinese)  
(邓红艳, 武芳, 钱海忠, 侯璇. 基于遗传算法的道路网综合模型. 武汉大学学报(信息科学版), 2006, 31(2): 164-167)

[7] Li Ai-Qin. Research on the seamless spatial database and its multi-scaling representation and processing[Ph. D. dissertation]. Wuhan University, Wuhan, 2001(in Chinese)  
(李爱勤. 无缝空间数据组织及其多比例尺表达与处理研究[博士学位论文]. 武汉大学, 武汉, 2001)

[8] Yan Wei-Min, Wu Wei-Min. Data Structure. 2nd Edition. Beijing: Tsinghua University Press, 1992(in Chinese)  
(严蔚敏, 吴伟民. 数据结构. 第2版. 北京: 清华大学出版社, 1992)

[9] Guttman A. R-tree: A dynamic index structure for spatial search//Proceedings of the ACM SIGMOD International Conference on Management of Data. Boston, US, 1984: 47-57

[10] Liu Chang, Pei Ji-Hong, Sun Hong-Yuan, Gong Xin. Multi-scale spatial index structure based on R-tree. Journal of Shenzhen University Science and Engineering, 2004, 21(3): 242-246(in Chinese)  
(刘畅, 裴继红, 孙宏元, 龚忻. 一种用于多分辨率空间数据的 R 树索引结构. 深圳大学学报(理工版), 2004, 21(3): 242-246)

[11] Chan Edward P F, Chow Kevin K W. On multi-scale display of geometric objects. Data & Knowledge Engineering, 2002, 40(1): 91-119

[12] Oosterom P V. The reactive-tree: A storage structure for a seamless scaleless geographic database//Proceedings of the Auto-Carto. Baltimore, US, 1991: 393-407

[13] Baeijs Christof, Demazeau Yves. SIGMA: Application of multi-agent system to cartographic generalization//Proceedings of the 7th European Workshop on Modelling. Eindhoven, Holand, 1996: 163-177

[14] Galanda Martain. Automated polygon generalization in a multi-agent system[Ph. D. dissertation]. Zurich University, Zurich, 2003

representation of spatial data and auto cartographical generalization.

**ZAI Ren-Jian**, born in 1980, Ph. D. candidate, lecturer. His major research interests include auto cartographical generalization based on genetic algorithms and spatial database.

**ZHAO Qian**, born in 1963, Ph. D. , associate professor. Her major research interests include database technology and application.

Background

The work is supported by the National Natural Science Foundation of China (grant Nos. 40671162, 40671152) and the High Technology Research and Development Program (863) of China under grant No. 2007AA12Z211. The objective of these projects is to improve stone technologies for further development of GIS. So far, some achievements have been made in fields of intelligent generalization algorithms and automatic controlling for intelligent cartography.

To meet growing requirements and applications expansion of GIS, especially requirements for application and analysis of spatial data, processing and multi-scale representation of spatial data has become a far-reach topic in geo-science field, and it is also a world-wide hard problem. To solve the problem, scientific structure of spatial data and cartography generalization are prerequisites. A key technique to the representation is how to manage sequences of sub-spaces, that is

how to build index structure for multi-scale representation of spatial data, because the technique will determine speed not only of representation but also query and analysis of the data. For the time being, solutions to the problem are improved structures on traditional index methods of spatial data, the best improved one is based on quad-tree, but it results in low efficiency for limitations of quad-tree.

In this paper, Spatial Data Multi-representation R-tree (SDMR), a transmutation index structure based on R-tree is proposed against limitations of traditional R-tree. Experiments prove that SDMR supports multi-scale representation of spatial data for all kinds of changes with higher efficiency than the improved one based on quad-tree, because the transmutation structure can memorize generalization results all through changes at the same scale and among different scales.