

# 基于四叉树结构的DEM生成算法

叶海建

安乐

(中国农业大学计算机网络中心) (北京特种车辆研究所)

**摘要** 提出了一种以四叉树结构存储采样数据的方法,给出了建立四叉树结构的算法和基于四叉树结构建立DEM模型的实例,结果表明DEM模型的生成效率得到显著提高,尤其是在数据量大的情况下,效果非常明显。

**关键词** 算法; 四叉树; 数据结构; DEM; 模型

**分类号** TP 311.52

## A Fast Quadtree-based on Algorithm of Building DEM

Ye Haijian

An Le

(Computer and Network Center, CAU) (Beijing Special Vehicle Institute)

**Abstract** The method of Storing Sampling with a quadtree structure is presented. The efficiency of the quadtree-based algorithm of building DEM is remarkably raised. The more much sampling data and more much grided points, the more remarkable is the raising of the efficiency of building DEM. Moreover, the algorithm of constructing quadtree and an example of building DEM on the base of quadtree structure are presented.

**Key words** algorithm; quadtree; data structure; DEM; model

矩形网格的数字高程模型(DEM)是由一系列排列整齐的、具有不同高程值的网格点构成的。这些网格点的高程值由采样数据点的高程值插值计算得到。由于地形的复杂性和数据量的庞大,快速确定网格点高程值所需的采样数据成为提高DEM模型生成效率的关键<sup>[1]</sup>。为此,笔者提出一种基于四叉树存储结构的DEM生成算法。

### 1 四叉树存储结构的基本思想

树型结构是一种非线性结构,它在查找算法中有其独特的优势<sup>[2]</sup>。若将采样数据以某种形式划分,形成四叉树,缩小查找区域,减小查找数据量,将会提高查找效率。

建立采样数据四叉树结构的基本思想<sup>[3]</sup>是:首先定阈值,检查区域内采样数据个数是否超过阈值;若超过,则将该区域4等分,然后对4个子分区进行同样的处理,直到所有子分区内的采样数据个数均不超过阈值为止。具体操作为:设 $n_q$ 为预先给定的阈值,若采样区域内采样数据个数 $n > n_q$ ,则将采样区4等分,得到4个子区,每个子区内的采样数据个数为 $n_i$  ( $i=0, 1, \dots, 3$ , 且 $n_i \neq 0$ ,  $n_i = n$ );若4个子区中存在 $n_i > n_q$ 的子区,则将其再4等分,直到所有子区的采样数据个数 $n_i \leq n_q$  ( $i=0, 1, \dots, 3$ )为止,由此得到的树即为采样数据的四叉树结构。在此四叉树中,只有叶子结点含有采样数据。

收稿日期: 2000-01-17

叶海建,北京清华东路17号中国农业大学(东校区)147信箱,100083

如图1所示,在采样区域内随机分布着一系列采样数据点,当取 $n_q=10$ 时,其采样数据的划分如图所示。图2为图1中采样数据的四叉树表示,“○”为四叉树结点,“□”表示对应结点(即子区)所含的采样数据。在子区内,由于数据量已控制在阈值范围内,因此,为简便起见,一般以线性链表存放。

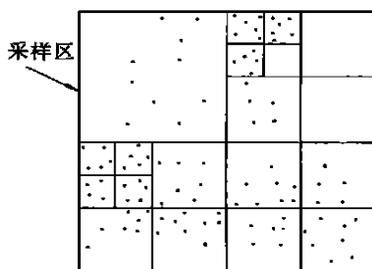


图1 采样区内采样数据的划分

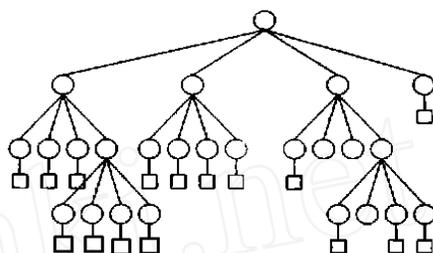


图2 采样数据的四叉树表示

## 2 建立采样数据四叉树结构的算法

设采样区域内采样得到的数据集为

$$D = \{D_k(x_k, y_k, z_k) \mid k = 1, 2, \dots, n\}$$

式中: $D_k$ 为第 $k$ 个采样数据; $x_k, y_k, z_k$ 为 $k$ 点的三维坐标, $z_k$ 为高程值; $n$ 为采样数据总数。现将 $D$ 划分,建立采样数据四叉树结构,其算法可分为2个部分。

I:

- 1) 确定采样区域 $A$ ,该区域长为 $l$ ,宽为 $b$ ;
- 2) 若 $A$ 内采样数据个数 $n=0$ ,则执行6);
- 3) 若 $n>0$ ,则生成根结点,将所有采样数据移入根结点;
- 4) 若 $n > n_q$ ( $n_q$ 为阈值,预先给定),则执行6);
- 5) 若 $n > n_q$ ,执行II;
- 6) 结束。

II(结点分裂算法):

- 1) 将当前结点区域 $A$ 4等分为4个子区 $A_i$ ( $i=0,1,\dots,3$ ),生成4个孩子结点;
- 2) 依次计算落入各子区内的采样数据个数 $n_i$ ( $i=0,1,\dots,3$ ),将父结点中的采样数据分别移入相应孩子结点;
- 3) 依次检查 $A_i$ 内的 $n_i$ ( $i=0,1,\dots,3$ ),若 $n_i > n_q$ ,则递归执行II;
- 4) 直到所有结点均满足 $n_i \leq n_q$ ;
- 5) 结束。

## 3 阈值的选取

采样数据四叉树结构的建立,实际上是将采样数据所属区域按采样数据个数不超过 $n_q$ 的原则,不断对其进行4等分的过程;因此,四叉树的深度和形态与采样数据总数 $n$ 、阈值 $n_q$ 和采样数据的分布有关。

对于不同地形, 其采样数据的总数  $n$  和采样数据的分布是不同的, 因此, 可以对四叉树深度进行调节的只有  $n_q$ 。显然,  $n_q$  愈大, 四叉树深度愈小, 遍历四叉树算法效率高, 但结点内搜索效率降低; 反之,  $n_q$  愈小, 四叉树深度愈大, 遍历四叉树算法效率低, 但结点内搜索效率较高。所以,  $n_q$  的取值应合理, 一般以四叉树的深度不超过 10 为宜。

### 4 网格线交点高程值的计算

由文献[1]可知, 第  $i$  行第  $j$  列的网格线交点  $p_{ij}$  的高程值  $z_{ij}$  由下式计算

$$z_{ij} = \begin{cases} \frac{1}{n} \sum_{k=1}^n W_k z_k & \text{采样点 } D_k (k = 1, 2, \dots, n) \text{ 与 } p_{ij} \text{ 不重合} \\ z_k & \text{采样点 } D_k \text{ 与 } p_{ij} \text{ 重合} \end{cases}$$

式中:  $n$  为离  $p_{ij}$  点最近的采样数据个数,  $z_k$  为其中第  $k$  个采样数据点的高程值,  $w_k$  为  $k$  点权值<sup>[1]</sup>。

DEM 模型网格线交点高程值的计算从左下角开始, 先按行计算, 后按列计算。

由于采样数据四叉树是按采样区域建立的, 而网格是按DEM 建模区划分的, 采样区比建模区大, 因此, 前者的 4 分区线与后者的网格线一般不重合, 如图 3 所示。现在的问题是, 计算  $p_{ij}$  点高程值  $z_{ij}$  时, 如何在  $D$  中查找到离  $p_{ij}$  点最近的  $n$  个采样数据? 通常的做法是利用以  $p_{ij}$  点为圆心的搜索圆来确定  $n$  个采样数据。由于采样数据已经通过 4 分区划分到不同的子区, 存储在四叉树结构中, 并且每个子区的采样数据个数均不超过  $n_q$ , 因此, 搜索采样数据时, 只要搜索被搜索圆覆盖和与搜索圆相交的子区即可, 这样可以大大减小搜索范围和搜索数据量, 从而提高搜索算法的效率。

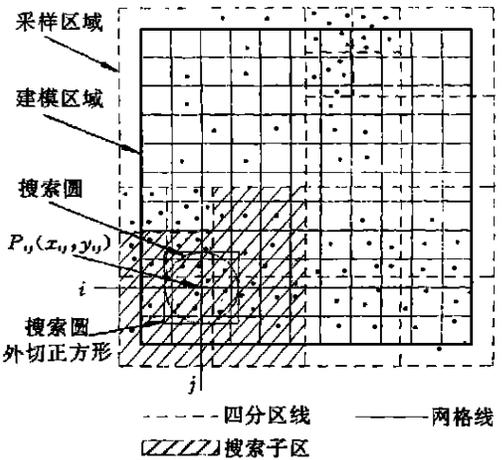


图 3 搜索区域的确定

另外, 由于搜索圆的边界是圆的, 不易确定与其相交的子区, 因此, 确定搜索子区时, 以搜索圆的外切正方形所覆盖的子区和与其相交的子区作为搜索范围, 如图 3 所示。外切正方形确定的搜索范围可能会比搜索圆确定的搜索范围大一些, 但不会增加很大, 故不会影响总的效率。只需遍历四叉树即可快速确定搜索区域, 从而确定搜索数据范围, 进而确定  $n$ 。

### 5 实例

图 4 为华东某地的等高线图<sup>[3]</sup>, 面积大约为  $10\ 700\text{m} \times 9\ 260\text{m}$ 。数字化后得到的采样数据总数  $n = 24\ 620$ 。现在要建立该区域的DEM 模型, 硬件平台为 Pentium III450MHz, 内存 128MB, 硬盘空间足够大; 软件环境为 Windows NT 4.0, Visual C++ 5.0。表 1 示出在不同的网格间距情况下, 采样数据分别以四叉树和线性表 2 种结构存储时, 生成DEM 模型所消耗

的时间对比情况。图5为经计算后建立的DEM地面模型。

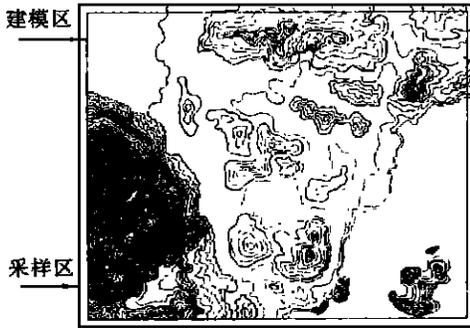


图4 等高线图

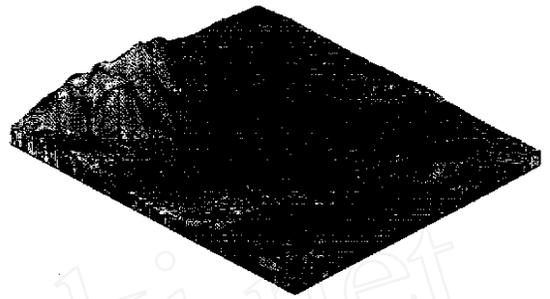


图5 DEM模型

表1 采样数据以不同结构存储时生成DEM模型所需时间

m in

存储结构	网格交点数/个			
	3 990	159 159	992 817	3 967 273
四叉树	0.500	1.584	14.286	61.668
线性表	9.200	41.428	517.780	2 272.500

## 6 结束语

采用基于四叉树存储结构的DEM生成算法,效率得到显著提高,尤其是在数据量大的情况下,效果更加明显。DEM模型生成算法效率的提高,有利于地形数字化类软件的实用化和普及,该算法已嵌入笔者开发的“电子沙盘制作系统”的软件中,效果良好。

## 参 考 文 献

- 1 叶海建 DEM模型生成算法的效率分析. 中国农业大学学报, 1994, 4(6): 81~ 84
- 2 严蔚敏, 吴伟民编著. 数据结构. 北京: 清华大学出版社, 1997. 13~ 17
- 3 安 乐. 电子沙盘中DEM/DTM生成算法和模型建立算法的研究及其系统的开发: [学位论文]. 北京: 中国农业大学, 1999