

# DEM 模型生成效率分析

叶海建<sup>①</sup>

(中国农业大学计算机网络中心)

**摘要** 对建立 DEM 模型的一般原理和影响 DEM 生成效率的因素进行了分析,指出改进 DEM 生成算法效率的关键在于改进离散数据存储结构和在其中的查找算法。

**关键词** DEM; 模型; 算法; 效率分析

**分类号** TP 311.52

## Efficiency Analysis of Building DEM

Ye Haijian

(Computer and Network Center, CAU)

**Abstract** The general theory and method of building DEM and the factors influencing the efficiency of building DEM are presented and analyzed. The key of improving the efficiency of the building algorithm is to improve the storage structure of dispersed sampling data and the search algorithm.

**Key words** DEM; model; algorithm; efficiency analysis

数字高程模型(DEM)是建立数字地面模型(DTM),最终生成三维电子地图和电子沙盘的基础。它不仅可以非常直观的方式展示一个地区的地形、地貌,而且也为各种地形特征的定量分析和不同类型专题图的自动绘制,提供了基本数据。

由于地形本身的复杂性和非解析性,试图用某种代数公式或曲面拟合方法对某一区域地形进行整体的数学描述是十分困难的,因此通常的做法是:对该区域进行网格划分,利用已经获得的三维空间离散数据,经过插值计算,求得网格交点的高程值,并以此作为该区域地面造型的依据,从而建立该区域的数字高程模型。

### 1 建立 DEM 的一般原理

#### 1.1 离散高程值的采样

离散高程值的获取可通过实地测量、卫星照片或航拍照片立体成像判读、数字化仪或扫描仪输入地图等多种方法实现。最常用的方法是利用等高线地形图获取离散高程值来建立区域的数字高程模型,这样比较简单、快速和可靠。需要注意的是,离散高程值的采样范围应该比实际建模区域大些,如图 1 所示。

#### 1.2 网格线交点高程值的计算

采样得到的高程数据是离散随机分布的,无法直接利用其生成地面模型,使用起来极为不

收稿日期:1999-07-01

①叶海建,北京清华东路 17 号 中国农业大学(东校区)147 信箱,100083

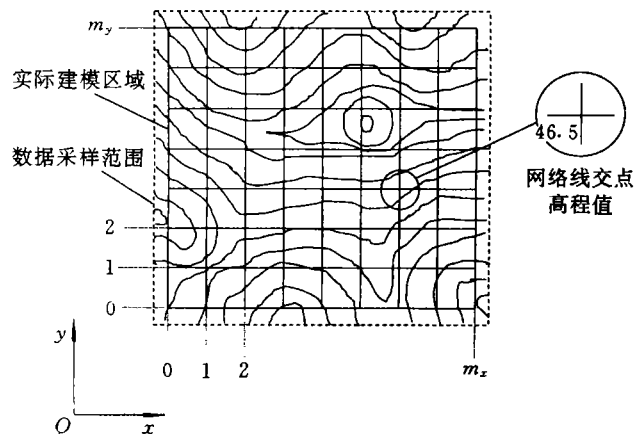


图1 网格化等高线地形图

便,因此一般需要将离散的高程数据网格化,形成二维阵列分布,图1中的网格即为数字高程模型的网格形式,其中的网格线交点含有相应的高程值。

图2为离散高程数据对网格线交点影响的示意图。 $P_{ij}(x_{ij}, y_{ij}, z_{ij})$ 为 $x-y$ 平面上第 $i$ 行第 $j$ 列的网格线交点, $z_{ij}$ 为 $P_{ij}$ 点的高程值。采样区采样得到的离散高程数据集合为

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

式中: $D_k$ ——第 $k$ 个离散高程数据;

$z_k$ —— $k$ 点高程值;

$n$ ——离散数据总数。

将采样得到的离散高程数据网格化的基本思想是: $P_{ij}$ 的高程值 $z_{ij}$ 是由于 $D_k$ 的高程值 $z_k$ 影响的结果<sup>[1]</sup>。显然,每个 $z_k$ 对 $z_{ij}$ 的影响是不同的,离 $P_{ij}$ 愈近,则影响愈大,反之则愈小。这种方法称为距离加权平均法<sup>[1,2]</sup>。 $z_{ij}$ 的计算公式如下:

$$z_{ij} = \begin{cases} \frac{1}{n} \sum_{k=0}^n w_k z_k & \text{当离散点 } D_k (k=1, 2, \dots, n) \text{ 与 } P_{ij} \text{ 不重合时} \\ z_k & \text{当离散点 } D_k \text{ 与 } P_{ij} \text{ 重合时} \end{cases} \quad (1)$$

式中 $w_k$ 为 $k$ 点权值,点 $D_k$ 到 $P_{ij}$ 距离 $d_k$ 的函数, $w_k = f(d_k)$ 。 $w_k$ 的确定对计算 $z_{ij}$ 的效率和精度有直接影响,通常取

$$w_k = \frac{1}{d_k^u}$$

式中 $u$ 为一常数,一般取 $u=2$ <sup>[1]</sup>。

### 1.3 离散数据个数 $n$ 的确定

从理论上讲,应将离散数据总数代入式(1)计算,但在实际计算中,采样得到的离散高程数

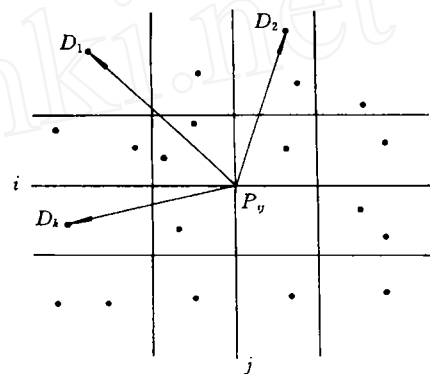


图2 离散高程数据对网格线交点的影响

据量  $n$  通常很大,而距离网格线交点  $P_{ij}$  较远的离散数据对  $P_{ij}$  点高程值  $z_{ij}$  的影响已经很小,因此,为提高运算效率,只取与  $P_{ij}$  点相距较近的若干个离散数据  $n'$ ,一般  $n' \ll n$ ; 故式(1)改为

$$z_{ij} = \begin{cases} \frac{1}{n'} \sum_{k=1}^{n'} w_k z_k & \text{当离散点 } D_k (k=1, 2, \dots, n') \text{ 与 } P_{ij} \text{ 不重合时} \\ z_k & \text{当离散点 } D_k \text{ 与 } P_{ij} \text{ 重合时} \end{cases} \quad (2)$$

通常采用搜索圆的方法<sup>[1,2]</sup>来确定  $n'$ ,即以  $P_{ij}$  点为圆心,  $r$  为半径确定一个搜索圆。凡落入搜索圆内的离散数据  $D_k$ , 均可代入式(2)计算。由于采样区内的离散数据分布不均匀,每个搜索圆内的  $n'$  值相差很大,因此,对代入式(2)计算的离散数据个数  $n'$  进行规定:一般取 4~10。当  $n' < 4$  时,应增大  $r$ ; 当  $n' > 10$  时,则应减小  $r$ 。

在计算  $z_{ij}$  时,还应考虑搜索圆内离散数据的分布情况,以避免数据“跑偏”影响 DEM 的精度。处理的一般原则是:在以  $P_{ij}$  为原点的直角坐标系内,每个象限内的离散数据个数至少为 1 个,否则应调整搜索圆半径  $r$ 。

## 2 建立 DEM 的效率分析

如图 1,  $x-y$  为水平面,  $z$  方向为高程。  $x$  方向网格数为  $m_x$ ,  $y$  方向网格数为  $m_y$ , 网格线交点总数为  $(m_x+1) \times (m_y+1)$ 。

按本文 1 中提出的方法计算所有网格线交点高程值的时间复杂度<sup>[3]</sup>

$$T(n, m_x, m_y) = \sum_{i=0}^{m_y} \sum_{j=0}^{m_x} T_{O_{ij}}(f(n)) \quad (3)$$

式中:  $T_{O_{ij}}(f(n))$  为第  $i$  行第  $j$  列网格线交点高程值的时间复杂度。

由式(3)可以看出,建立 DEM 的效率与网格的划分和计算每一网格线交点高程值的算法效率有关。

### 2.1 网格划分对建立 DEM 的影响

在一般情况下,  $m_x$  与  $m_y$  同数量级。当网格划分在  $x$  方向和  $y$  方向增加 10 倍时,网格线交点增加到原来的 100 倍,DEM 精度提高了,但建立 DEM 的时间消耗也将增加为原来的 100 倍,同时占用的内存空间也以百倍递增;相反,若减少网格划分数量,可以大大提高建立 DEM 的效率,但 DEM 的精度将受到影响:因此,在精度和运算效率之间应有一个合理的取舍。一般网格间距在 5 m 左右,精度已足够,根据建模区的长宽即可确定实际网格数<sup>[4]</sup>。

### 2.2 计算一个网格线交点高程值的效率对建立 DEM 的影响

由式(3)可知,  $T(n, m_x, m_y)$  与  $T_{O_{ij}}(f(n))$  成正比。在  $m_x$  和  $m_y$  确定的情况下,建立 DEM 的效率主要取决于计算每一个网格线交点高程值的时间复杂度  $T_{O_{ij}}(f(n))$ 。由式(3)可得

$$T(n) = \sum_{i=0}^{m_y} \sum_{j=0}^{m_x} T_{O_{ij}}(f(n))$$

当按式(2)计算  $z_{ij}$  时,其时间复杂度  $T_{O_{ij}}(f(n))$  主要取决于在已知的离散数据集  $D$  中查找满足计算条件(即落入搜索圆内)的数据的时间消耗,因此,离散数据集的数据结构和在其上的查找算法是影响  $T_{O_{ij}}(f(n))$  的关键。

最简单,也是最传统的做法是:将  $D$  以线性结构存放。由于  $n$  通常很大,并且不同情况下  $n$

的差异也很大,因此,一般以线性链表作为存储结构。计算一个  $z_{ij}$  时,按顺序查找遍历一次  $D$ , 则

$$T_{O_{ij}}(f(n)) = T_{O_{ij}}(n)$$

不失一般性,设计算  $z_{ij}$  的平均时间复杂度为  $T_{O'}(n)$ , 则有

$$T(n) = (m_x + 1) \times (m_y + 1) \times T_{O'}(n)$$

若设  $m_x = 1\ 000, m_y = 1\ 000, n = 10\ 000$ , 则  $T(n) \approx 10^{10}$ , 显然效率极差。

### 3 结 论

由上述分析可以看出,DEM 模型的生成效率受到网格划分和计算一个网格线交点高程值的效率两方面的影响。大多数情况下,由于精度方面的考虑,网格划分是预先确定的,因此,要提高 DEM 模型的生成效率,必须提高计算每一个网格线交点高程值的效率,其中,改进  $D$  的存储结构和在  $D$  中的查找算法则是关键。

### 参 考 文 献

- 1 胡友元,黄杏元主编. 计算机地图制图. 北京:测绘出版社,1987. 282~284
- 2 王永明,林行刚. 一种快速 DEM 生成算法. 计算机应用与软件,1998,15(4):28~30
- 3 严蔚敏,吴伟民编著. 数据结构. 北京:清华大学出版社,1997. 13~17
- 4 王来生,鞠时光,郭铁雄,等编著. 大比例尺地形图机助绘图算法及程序. 北京:测绘出版社,1993. 22~25