

# 基于平面不规则三角网的 DEM 数据压缩与质量分析

刘春<sup>1,2)</sup> 吴杭彬<sup>1)</sup>

<sup>1)</sup>(同济大学测量与国土信息工程系, 上海 200092) <sup>2)</sup>(现代工程测量国家测绘局重点实验室, 上海 200092)

**摘要** 针对构建的不规则三角网,依据相邻三角形法线向量之间的夹角为判断依据,给出一种基于平面不规则三角网(TIN)的DEM数据压缩方法。在详细推导了压缩准则的基础上,给出了压缩判断阈值的确定方法。为了说明该压缩方法能顾及地形特征,对压缩方法的可靠性作详细介绍,并说明了该方法的可行性。

**关键词** 平面不规则三角网 数据压缩 DEM

中图分类号: P237.3 文献标识码: A 文章编号: 1006-8961(2007)05-0836-05

## Data Compress of DEM Based on Plane Triangulation Irregular Network

LIU Chun<sup>1,2)</sup>, WU Hang-bin<sup>1)</sup>

<sup>1)</sup>(Department of Survey and Geo-Informatics, Tongji University, Shanghai, 200092)

<sup>2)</sup>(Key Laboratory of Advanced Engineering Surveying of SBSM, Shanghai 200092)

**Abstract** A new method of data compression of Digital Elevation Model (DEM) which is based on the Triangle Irregular Network (TIN) is presented in this paper. In the method, the angle between triangle normal vectors is used as the measure for identifying points which should be removed, in which way the original DEM data can be compressed while the data storage space is reduced and the successive disposal speed will increase under the precondition of minimizing the loss of model precision. Furthermore, the reliability of the data compression method is analyzed to show the feasibility of the method.

**Keywords** TIN (triangulation irregular network), data compression, DEM (digital elevation model)

## 1 引言

DEM(digital elevation model)是地形的一种表现形式,它是选取定义坐标场中的已知X、Y、Z坐标值的离散点序列对地面的一个简单的统计表示。人们通过对环境的数据获取、表面建模以及其他的一些相关技术建立感兴趣对象的DSM(digital surface model)模型。描述地形的DEM主要数据来源有影像,地形图以及其他一些数据源。获得数据之后,可以根据规

则和不规则的空间插值方法建立。DEM在测绘和遥感、农林规划、土木和水利工程、军事领域、地学分析以及地理信息系统等各个领域获得了广泛深入的研究和普遍应用<sup>[1]</sup>。DEM通常用TIN(triangle irregular network)和GRID两种数据结构表示。

然而对于大范围的DEM,不同细节层次的表达对于DEM数据量的要求受到软件和硬件的限制。因此探讨针对DEM的数据压缩是目前DEM研究的内容之一。数据压缩本是一种计算机数据处理技术,在计算机地图制图和地理信息系统中,通常利用

基金项目:国家自然科学基金项目(40501061);“基础地理信息与数字化技术”山东省重点开放实验室课题(SD040211)

收稿日期:2005-12-26; 改回日期:2006-03-14

第一作者简介:刘春(1973~),男,副教授。2000年12月获同济大学大地测量与测量工程博士学位,2001年11月~2004年3月于同济大学海洋科学博士后流动站从事博士后工作。主要研究方向为GIS空间数据不确定性理论、LIDAR数据处理和应用,已于国内外刊物发表论文50余篇。E-mail: liuchun@mail.tongji.edu.cn

数据压缩技术减少数据量。目前已有不少针对线状地物和面状地物的矢量数据进行压缩的方法。而对 DEM 的数据进行压缩,主要用于减少存储数据量,提高后续处理速度。一般在下列两种情况下进行数据压缩:①是在 DEM 原始数据中,存在一定程度的数据冗余,影响 DEM 的应用处理速度;②是当通过大尺度的 DEM 派生中小尺度的 DEM 时,其地形表达精度要求降低<sup>[2]</sup>。

文献[2]提出 DEM 数据压缩,DEM 地形压缩误差( $E_p$ )的概念,在此基础上提出一种以 DEM 地形压缩误差影响为选取资格的数据压缩方法。利用该方法可以为 DEM 的数字综合提供一种技术手段,减小 DEM 的存储数据量,提高 DEM 后续处理速度。文献[3]提出一种小波多尺度分析的方法,建立 DEM 网格数据综合小波变换模型,并提出通过 DEM 网格数据综合来实现等高线的制图综合。文献[4]采用倾角 Morton 编码(slopping angle morton coding, SMC)探讨了对数字高程模型格网数据进行综合简化的问题,文献[5]则从稠密格点形成的 GRID 网格集中,通过选取“重点集”生成描述地面起伏的三角形剖分网来进行综合简化。

可见现有对 DEM 的数据压缩研究大多针对网格数据,而网格数据往往已经通过空间内插获得,在内插过程中有内插误差的引入。此外针对网格数据的压缩方法很难考虑地形特征,而压缩的可靠性是保留地形原有特征。因此本文根据原始数据构建的不规则三角网,给出对 DEM 数据进行压缩的方法,并对其压缩的可靠性展开讨论。

## 2 基于不规则三角网的 DEM 数据压缩

### 2.1 DEM 不规则三角网建立

在数字高程模型中,不规则三角网通过从离散分布的数据点生成的连续三角面来逼近地形表面。在所有可能的三角网中,狄洛尼(Delaunay)三角网在地形拟合方面表现尤为出色。狄洛尼三角网为相互连接且互不重叠的三角形的集合,每一个三角形的外接圆内不包含其他的点。狄洛尼三角网构网方法有分而治之的算法,数据点逐次插入算法和三角网生长算法<sup>[1]</sup>。

### 2.2 压缩模型

对离散数据进行狄洛尼三角形构网后,DEM 可以表达成每个三角形(面)的 3 个顶点的链表。地面模型已经建立,对这一模型数据的压缩必然会导

致模型精度的降低。压缩过程中必须是基于这个已经建好的三角网模型,使得删除点对模型的影响最小。显然,平坦区域上的点构成的三角网包含的信息量最少,同时这些三角形构成面的法线向量也就越趋近于平行,也就是说法线向量的夹角越小。利用三角网的这一性质,可以对三角网进行进一步压缩,而使损失的模型精度最少。

如果过点  $p$  的三角形有  $n$  个,那么两两三角面之间的法线向量夹角有  $C_n^2$  个,求出这  $C_n^2$  个夹角的最大值,并且当这个最大值小于设定的阈值时,则这个点就删除,否则予以保留。

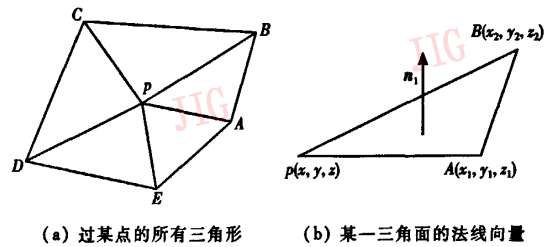


图 1 平面法向量示意图  
Fig. 1 Sketch map of plane normal vectors

如图 1 所示,  $p, A, B$  3 个点构成一个三角形,  $n_1$  为其法向量,显然在模型坐标空间可以得到向量  $\vec{pA}$  和向量  $\vec{pB}$ :

$$\begin{aligned} \vec{pA} &= (x_1 - x, y_1 - y, z_1 - z) \\ \vec{pB} &= (x_2 - x, y_2 - y, z_2 - z) \end{aligned}$$

则

$$\begin{aligned} n_1 &= \vec{pA} \times \vec{pB} \\ &= (x_1 - x, y_1 - y, z_1 - z) \times \\ &\quad (x_2 - x, y_2 - y, z_2 - z) \\ &= \begin{vmatrix} i & j & k \\ x_1 - x & y_1 - y & z_1 - z \\ x_2 - x & y_2 - y & z_2 - z \end{vmatrix} \end{aligned} \quad (1)$$

其中,  $i, j, k$  为  $X, Y, Z$  方向的单位向量。令

$$n_1 = (dx_1, dy_1, dz_1)$$

则可以得到

$$\begin{aligned} dx_1 &= (y_1 - y)(z_2 - z) - (y_2 - y)(z_1 - z) \\ dy_1 &= -(x_1 - x)(z_2 - z) + (x_2 - x)(z_1 - z) \\ dz_1 &= (x_1 - x)(y_2 - y) - (x_2 - x)(y_1 - y) \end{aligned} \quad (2)$$

同理可以得到过  $p$  点的另外一个三角面的法向量为

$$n_2 = (dx_2, dy_2, dz_2)$$

将向量  $n_1$  和  $n_2$  平移至同一顶点, 并作向量  $n = n_2 - n_1$  (如图 2 所示), 显然角  $\theta$  即是两向量之间的夹角, 也就是两个三角面之间的夹角。显然,

$$n = n_2 - n_1 = (dx_2 - dx_1, dy_2 - dy_1, dz_2 - dz_1)$$

同时, 令

$$l_1 = \|n_1\|, l_2 = \|n_2\|, l = \|n\|$$

于是可以得到:

$$\cos\theta = \frac{l_1^2 + l_2^2 - l^2}{2l_1l_2}$$

即

$$\theta = \arccos\left(\frac{l_1^2 + l_2^2 - l^2}{2l_1l_2}\right) \quad (3)$$

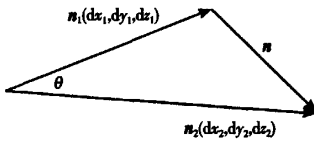


图 2 向量夹角示意图

Fig. 2 A angle between vectors scheme

对于地面模型内的每个点都可以求出  $C_n^n$  个  $\theta$  角 (其中  $n$  为包含该点的三角形个数), 求出其夹角最大值。如果最大  $\theta$  角小于设定的阈值  $T$ , 那么说明包含这个点的所有三角形在设定的范围  $T$  内近似趋于一个平面, 该点可以被删除; 否则, 说明包含这个点的所有三角形至少有两个存在较大的差异, 该点不能被删除。

需要注意的是, 由于向量叉乘满足右手法则, 因此, 求解法向量时应注意所有三角形的顶点应该按照一定的顺序 (例如: 逆时针) 给出。

### 2.3 阈值选取的依据

在以上模型中, 需要确定阈值  $T$ 。  $T$  值的大小直接决定了压缩效率和压缩误差。压缩效率和压缩

误差也存在一定的关系, 一般来说, 压缩比越大, 压缩误差也越大。因此阈值的选取应该根据地面模型允许的压缩误差决定。由于真实的地面并不知道, 所以一般选取一些能近似表达模型误差的量  $S$  作为评价依据, 用来描述压缩误差。  $S$  的选取可以采用不同的准则, 例如可以根据 DEM 采样点高程差的方差等来选择。

设压缩前 DEM 的网格采样点的高程为  $h_i^{(0)}$ , 压缩后对应点的高程为  $h_i^{(1)}$ , 其中,  $i = 1 \sim n$ ,  $n$  为采样点的个数。那么

$$S = \frac{1}{n} \sum_{i=1}^n (h_i^{(0)} - h_i^{(1)})^2 \quad (4)$$

自适应压缩过程的流程如下: (设允许压缩误差  $e$ , 压缩误差与允许压缩误差的允许差值  $\varepsilon$ , 阈值的变化步长  $\Delta T$  均已给定)

(1) 选取大概的阈值  $T$ ;

(2) 对 TIN 用阈值  $T$  进行压缩, 并对压缩前后的 DEM 进行网格采样, 利用 (4) 式计算压缩误差  $S$ ;

(3) 进行  $S$  与  $e$  的大小判断:

1)  $|S - e| > \varepsilon$ :

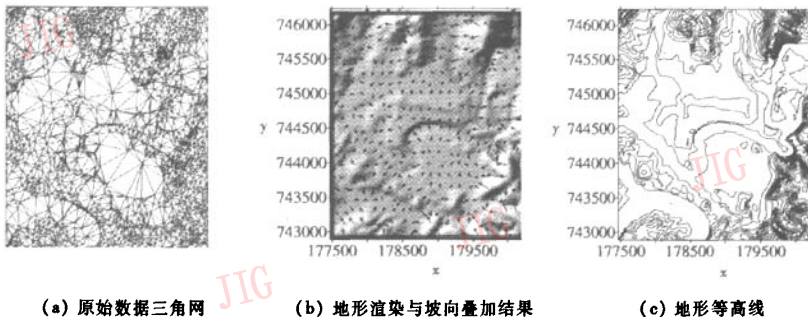
①  $S - e > \varepsilon$ , 则缩小阈值: 若  $T = T - \Delta T$ , 则转入步骤 2;

②  $S - e < \varepsilon$ , 则放大阈值: 若  $T = T + \Delta T$ , 则转入步骤 2;

2) 退出循环,  $T$  即为该允许误差  $e$  下的最佳阈值。

### 3 数据压缩的质量分析

选取某试验区的高程采样点, 进行狄洛厄构网并进行分析。原始数据包括 2 672 个采样点, 构网后三角形个数为 5 192, 原始数据构网结果如图 3 所示。



(a) 原始数据三角网

(b) 地形渲染与坡向叠加结果

(c) 地形等高线

图 3 原始数据构网结果

Fig. 3 Delaunay Triangle results of raw data

对建立的三角网模型用上述方法压缩,并分别取 5°,7.5°,10°,12.5°为阈值,压缩结果和重新构网的结果如图 4 所示。

表 1 给出了各种阈值下压缩结果的数值。从表 1 可以看出,随着阈值的增大,点压缩比  $R_{point}$  和三角形压缩比  $R_{triangle}$  迅速增大。其计算式如下:

$$R_{point} = \frac{N_0 - N_1}{N_0} \times 100\% \quad (5)$$

其中,  $N_0$  为原有点数,  $N_1$  为保留点数。

$$R_{triangle} = \frac{M_0 - M_1}{M_0} \times 100\% \quad (6)$$

其中,  $M_0$  为原有三角形数目,  $M_1$  为压缩后三角形个数。

表 1 各种阈值下的压缩结果

Tab. 1 Compression results from different threshold value

阈值	删除点数	点压缩比 (%)	删除三角形数	三角形压缩比 (%)
5	30	1.12	60	1.16
7.5	105	3.93	210	4.04
10	247	9.17	494	9.51
12.5	462	17.29	924	17.80

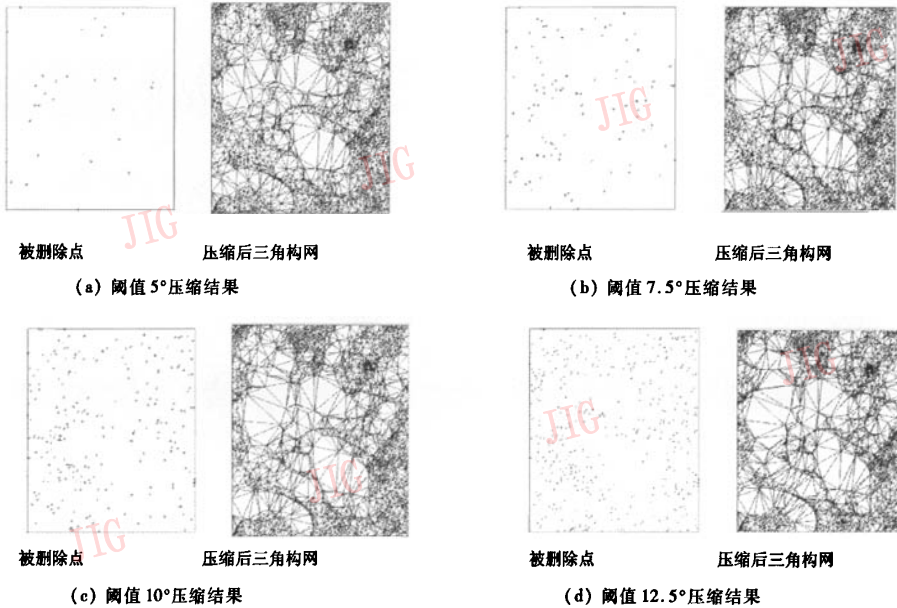


图 4 压缩结果示意图

Fig. 4 Compression results scheme

在数据范围内选取一个断面(图 3(a)的左下到右上对角线),对这一断面在各个阈值下的压缩分析结果如图 5 以及表 2 所示。

在表 2 中,由于压缩过程中,总的三角形数减

少,因此压缩后的点数(保留点数)、未变化点数等随着阈值的增大而减少。而平均高程差则因为采样点高程差有正有负,阈值的增大并不一定引起平均高程差的增大,但是会引起方差的增大。

由图 5 的断面图可以看出,压缩前后高程起伏特征基本上被保持了。此外,对地形表面的面积也进行了分析,其分析结果如表 3 所示(原始面积: 9.436 959 km<sup>2</sup>)。由表 3 可见,压缩过程没有引起面积的显著变化,也即从总体上没有破坏地形表面的起伏特征。

表 4 是对该 DEM 数据采用不同的阈值进行压缩后,根据不同的检查格网边长进行采样得到的采样点高差的方差情况。可以看出,同样的边长下,阈

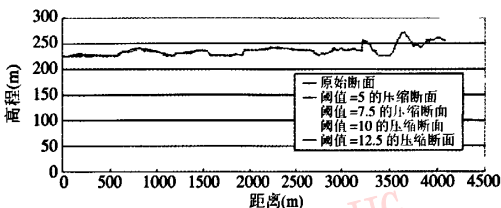


图 5 多阈值下压缩后的断面高程图

Fig. 5 Height picture of compressed profile on different thresholds

表 2 各种阈值下断面的采样点偏差情况(初始点:165)

Tab. 2 Warps of sampling point on profiles at different thresholds( original point:165)

阈值	压缩后点数 $N_1$	未变化点数 $N_2$	未变化率(%)	最大高程差(m)	平均高程差(m)	高程方差( $m^2$ )
5	162	157	96.91	0.8	0.00	0.01
7.5	157	138	87.89	3.80	0.18	0.38
10	152	127	83.55	3.85	0.17	0.47
12.5	143	108	75.52	4.19	0.26	1.02

注:表 2 的未变化率  $R$  采用如下算式:  $R = N_2/N_1 \times 100\%$

表 3 各种阈值下面积变化

Tab. 3 Difference of areas on different thresholds

阈值	面积( $km^2$ )	变化率(%)
5	9.436 883	0.000 81
7.5	9.436 571	0.004 11
10	9.435 655	0.013 82
12.5	9.434 063	0.030 69

表 4 各种阈值下不同网格宽度采样点高差的方差

Tab. 4 The square deviation of difference grid line on different thresholds

阈值	网格边长		
	1 倍平均边长下的方差	2 倍平均边长下的方差	3 倍平均边长下的方差
5	0.212	0.223	0.281
7.5	0.713	0.743	0.817
10	1.379	1.421	1.149
12.5	2.686	2.908	2.429

值越大,方差越大;同样的阈值下,不同的边长对应的方差相差并不大。其中的平均边长指的是原始三角网所有边的平均长度。

如果采用 2.3 节中的流程进行自适应压缩,并取  $W = 1m^2$  (即单位方差),  $e = 0.01$ ,  $\Delta T = 0.1^\circ$ , 当网格边长选为 1 倍平均边长,可以得到压缩阈值为  $8.9^\circ$ ;当检查网格边长选为 2 倍平均边长时,压缩阈值为  $8.7^\circ$ ;当检查网格边长选为 3 倍平均边长时,压缩阈值为  $8.9^\circ$ 。由以上分析可以得到,该 DEM 在单位方差下的压缩阈值约为  $8.8^\circ$ 。

需要指出的是,压缩效果的好坏除了和阈值选取有直接关系外,还和地面模型的复杂程度有关。

## 4 结 论

随着科学技术的发展,数字地面模型(DEM)在地理信息系统等各个领域具有普遍应用。对于大区域野外实测或者通过激光扫描技术等其他方法得到

的 DEM 原始数据量非常大,并且存在不同程度的数据冗余,因此研究 DEM 数据压缩对提高计算机处理速度,减少数据存储空间具有重要的现实意义。本文采用了对已经建立的三角网地面模型,对每个三角面的法线向量进行分析的方法进行 DEM 的数据压缩,并建立了该方法的数学模型。此外。为了说明该方法的压缩可靠性,本文又通过断面采样点分析、曲面面积变化以及 DEM 网格采样点高差的方差进行比较和论证。由比较结果可见,该方法对基于三角网的 DEM 模型压缩不仅可以压缩数据量,同时能比较好地保持原来的地面起伏情况,说明该方法是可行并且可靠的。然而在实际应用该方法的过程中,需要考虑三角面的空间分布情况。在悬崖峭壁等一些特殊的地面模型中,三角形顶点的按顺序存储会导致法线向量与实际情况相差  $180^\circ$ ,所以该方法仍然需要进一步研究。

## 参考文献(References)

- Li Zhi-lin, Zhu Qing. Digital Elevation Model [M]. Wuhan: Publication of Wuhan University of Surveying and Mapping, 2000. [李志林,朱庆. 数字高程模型[M]. 武汉:武汉测绘科技大学出版社,2000.]
- Cai Xian-hua, Zhen Tan-dong. Research of data compress and algorithm on DEM[J]. Bulletin of Surveying and Mapping, 2003, 10: 45 ~ 49. [蔡先华,郑天栋. 数字高程模型数据压缩及算法研究[J]. 北京:测绘通报,2003, 10:45 ~ 49.]
- Liu Chun, Wang Jia-lin, Liu Da-jie. Generalization of DEM grid data based on multi-band wavelet analysis [J]. Journal of Image and Graphics, 2004, 9(3): 340 ~ 344. [刘春,王家林,刘大杰. 多尺度小波分析用于 DEM 网格数据综合[J]. 中国图象图形学报, 2004, 9(3): 340 ~ 344.]
- Liao Duo-duo. Research and Application of Real Time Dynamic Virture Reality Landscape Construction [D]. Zhengzhou: Institute of Surveying and Mapping of PLA, 1996. [廖朵朵. 实时动态虚拟地景生成技术的研究与实践[D]. 郑州:解放军测绘学院,1996.]
- Xu Qing. Technology of Three Dimension Visualization of Terrain [M]. Beijing: Surveying and Mpping Press, 2000. [徐青. 三维地形可视化技术[M]. 北京:测绘出版社,2000.]