

DEM 内插邻域自适应确定方法

朱 莹¹,刘学军²,赵 静¹

(1. 南京晓庄学院生物化工与环境工程学院,江苏 南京 211171)
(2. 南京师范大学虚拟地理环境教育部重点实验室,江苏 南京 210023)

[摘要] 用内插方法构建 DEM 是地表建模和分析的基本思路,内插邻域的大小影响计算结果的精度,也进一步影响着地学分析的结果. 本文分析了内插邻域的特点,对于具体的内插方法,内插邻域存在基本的取值范围,基于这个范围对待插值点的邻近采样点进行计算,根据计算结果对待插值点进行自适应内插. 实验选择模拟数据和实测数据,对根据点数确定邻域的方法和自适应方法的计算结果进行比较,发现采用自适应方法计算得到的结果精度较高.

[关键词] 内插邻域,自适应,移动曲面法

[中图分类号] P28 [文献标志码] A [文章编号] 1001-4616(2013)01-0122-05

An Adaptive Method of Determining DEM Interpolation Neighbors

Zhu Ying¹, Liu Xuejun², Zhao Jing¹

(1. School of Biochemical and Environmental Engineering, Nanjing Xiaozhuang University, Nanjing 211171, China)
(2. Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Nanjing 210023, China)

Abstract: DEM interpolation is an essential tool for terrain surface modeling and analysis. The size of interpolation neighborhood affects the accuracy of calculation results, and it affects geographic analysis results further. This paper analyzes the characteristics of interpolation neighborhood. There are minimum and maximum values of interpolation neighborhood of specific interpolation methods. The elevation of sampling points adjacent to interpolation point is calculated based on the range of interpolation neighborhood. The elevation of interpolation points is calculated by adaptive methods according to the results. The simulated and measured data are selected as the experimental data. The neighborhood of essential method is determined by the number of points. The calculation results of essential method and adaptive method are compared, and the result of adaptive method has higher precision.

Key words: interpolation neighborhood, adaptive, moving surface

DEM 是对地表的数字化表达,也是地学分析的基础数据. 高度逼真的地形建模技术和高精度的内插算法是 DEM 研究领域长期的追求目标^[1,2]. DEM 内插以地表的空间自相关性为前提,在特定的插值模型下,内插邻域的大小影响着插值结果的精度,进一步影响地学空间分析结果的精度.

DEM 内插邻域的确定方法主要有两类,一类对每个待插值点采用同样大小的邻域进行内插计算,另一类对每个待插值点采用不同大小的邻域进行内插计算^[3-6]. 第一类方法较简单,一般适用于规则分布的采样点,当采样点不规则分布时,若待插值点所在的局部区域采样点过于稀疏,则在确定的邻域内不能获得足够多的采样点参与计算. 第二类方法能适应不规则分布的采样点,不存在由于局部采样点过于稀疏而选不到足够点的情况,但是该方法一般不能根据不同地形和不同分布的采样数据而确定合适的邻域大小. 本文以 DEM 内插中的移动曲面法为例,分析邻域对 DEM 内插结果的影响,以此为基础提出 DEM 内插时邻域的自适应确定方法,实现了根据不同地形和不同分布的采样数据确定适宜的内插邻域.

收稿日期:2012-08-25.
基金项目:国家自然科学基金(40971230、41201151)、江苏省教育厅自然科学基金(11KJD420001)、南京晓庄学院青年科研项目(2010KYQN26).
通讯联系人:朱莹,讲师,研究方向:数字地形分析、GIS 应用系统架构与设计. E-mail:luckyzhuying@yahoo.com.cn.

1 内插邻域自适应确定方法

1.1 内插邻域剖析

空间事物具有自相关性^[7],该特性与移动曲面法中核函数的局部性相吻合. 内插邻域的确定原则为尽量保证形函数的局部性,因此点的内插邻域应该取得足够小. 但是为了保证参数求解和矩阵求逆的需要,点的内插邻域又要取得足够大. 因此内插邻域的确定原则应为尽可能小又足够大. 根据 DEM 内插方法,内插邻域的大小存在极小值,在本研究中该值与移动曲面法中多项式的项数有关. 由于空间相近事物的影响通常大于较远事物的影响,在 DEM 内插计算时为了提高计算效率,降低计算的复杂度,可以忽略较远事物的影响. 因此内插邻域的大小存在极大值. 当内插邻域太大时,也不能保证核函数的局部性. 内插邻域的大小还与地形和采样点的疏密紧密相关.

1.2 内插邻域自适应确定方法

比较现有的几种内插邻域确定方法,在不采用 Voronoi 图的情况下,根据邻域内包含的点数确定内插邻域半径的方法具有优势. 但是该方法也存在如何自适应地确定每个待插值点邻域半径的问题. 由于每个待插值点所在区域的采样点分布不同,其地表形态变化也不同,应对每个待插值点确定邻域大小以提高内插精度. 因此本文对待插值点的邻近采样点进行分析,确定待插值点的邻域大小.

由于各待插值点的邻近采样点分布不同,因此其邻域半径变化所引起的内插结果变化规律并不相同. 对所有待插值点进行分析,发现当邻域由小到大逐渐变化时,其内插结果中误差先逐渐降低再逐渐升高. 但单个点计算值的误差变化不完全符合这个规律. 尽管如此,当内插邻域从小到大逐渐变化时,单个点计算值的误差有最小值,因此使各待插值点邻近点的邻域半径从小到大变化,计算取不同邻域半径的邻近点高程值,并记录邻近点计算值误差最小时内插邻域内包含的点数. 根据该点数确定待插值点邻域半径,进行内插计算,得到多个插值结果,取均值作为待插值点的最终高程值.

综上所述,本文提出的内插邻域自适应确定方法的主要步骤为:

- (1)对待插值点邻近的 $n(1 \leq n \leq 6)$ 个点进行分析,默认 $n=6$;
- (2)根据移动曲面法中多项式的项数确定邻近点内插邻域内点数的变化范围;
- (3)建立索引,搜索待插值点最近的 n 个点;
- (4)使 n 个邻近点内插邻域内包含的点数在设置的范围内变化,计算相应的邻域半径;
- (5)采用移动曲面法计算邻近点的高程值,根据邻近点的实测高程值,计算出误差;
- (6)记录邻近点计算值误差最小情况下相应的邻域内点数;
- (7)根据(6)中记录的邻域内点数,计算相应的待插值点邻域半径;
- (8)计算取(7)中相应邻域半径时,待插值点的高程值,求均值作为待插值点的高程.

2 实验分析

2.1 实验数据

(1)实测地形采样点数据

实验采用两种数据,一种为实测采样点,选择云南省的散点数据,等高线数据如图 1,总采样点数为 15 012,选择其中均匀覆盖研究区域的 30 个点作为待插值点. 实验数据概况如表 1 所示.

表 1 实验数据基本信息
Table 1 The information of experimental data

地貌类型	相对高差/m	坐标范围		平均坡度/(°)	平均高程/m
		x/m	y/m		
山地	154	620 000 ~ 620 421. 13	98 528. 56 ~ 99 000	26. 1	796. 6

(2)数学曲面模拟数据

由于实测采样点不可避免地带有各种误差,为了能更好地检测方法的有效性和精度,本文选用了— 123 —

面^[8,9],表达式为:

$$z=A[1-(x/m)^2]e^{-(x/m)^2-(y/n+1)^2}-B[0.2(x/m)-(x/m)^3-(y/n)^5]e^{-(x/m)^2-(y/n)^2}-Ce^{-[(x/m)+1]^2-(y/n)^2}, \quad (1)$$

式中 A 、 B 、 C 是地势起伏参数, m 、 n 为区域范围控制参数. 参数 $A=150$, $B=500$, $C=24$, 本文取 $m=n=400$, 即 $-200 \leq x, y \leq 200$. 基于该曲面离散, 随机生成已知点 384 个, 选择均匀覆盖研究区域间距为 20 的点作为待插值点, 共有 361 个点, 如图 2.

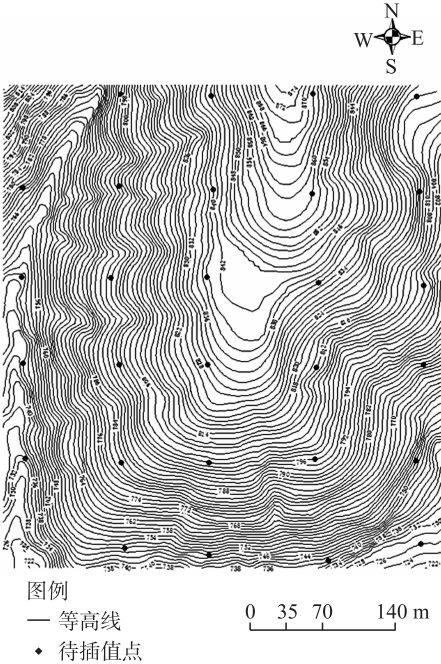


图 1 研究区域等高线数据
Fig. 1 The contour of study area

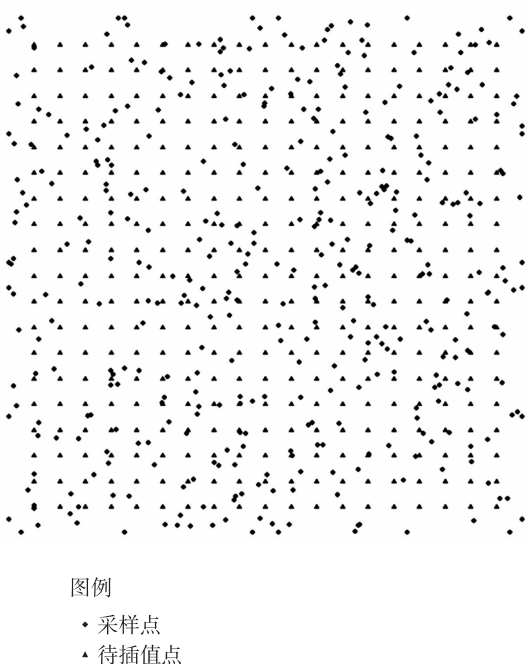


图 2 曲面离散数据
Fig. 2 The discrete points on Gaussian surface

2.2 研究方法

实验首先分析内插邻域从小到大变化时,待插值点计算值误差的变化规律,给出其中单个点误差的变化规律以及所有点误差统计值的变化规律. 在此基础上,采用本文提出的自适应方法计算待插值点高程值,并给出中误差,其计算公式为

$$\sigma = \sqrt{\frac{\sum_{l=1}^N (A_l - a_l)^2}{N}}, \quad (2)$$

式中, σ 为中误差, A_l 为待插值点的真值或测量值, a_l 为待插值点的计算值, N 为待插值点数. 将结果与采用点数确定内插邻域方法的计算结果进行比较.

实验中移动曲面法拟合的曲面选用如下形式:

$$z=A+Bx+Cy, \quad (3)$$

式中, x 、 y 、 z 为点的坐标值, A 、 B 、 C 为待定参数. 权函数采用高斯权函数, 公式为

$$w(r) = \begin{cases} \frac{e^{-r^2\beta^2}-e^{-\beta^2}}{1-e^{-\beta^2}}, & r \leq 1, \\ 0, & r > 1, \end{cases} \quad (4)$$

式中 r 为已知点与待插值点之间的距离, β 为参数, 取值为 3.0.

2.3 实验结果与分析

对两种数据分别抽取任意 4 个待插值点, 计算邻域内点数逐渐增大时, 待插值点高程的绝对误差, 结果如图 3、图 4. 单个点的误差随内插邻域增大的变化趋势并不相同. 对两种数据所有待插值点计算结果进行统计, 得到中误差的变化趋势, 如图 5、图 6. 两种数据的待插值点在邻域半径逐渐增大时, 计算结果中误差的总体趋势几乎相同, 即先逐渐减小再逐渐增大.

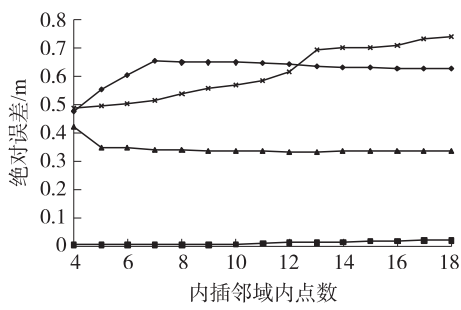


图3 云南数据中4个点误差变化折线图

Fig.3 The error trends of four points in Yunnan data

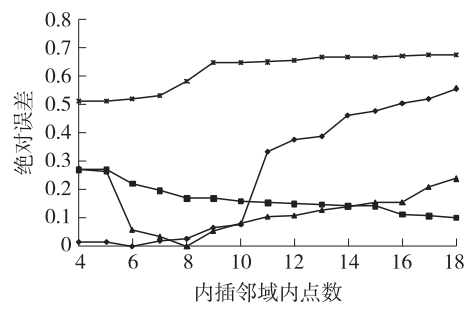


图4 数学曲面数据中4个点误差变化折线图

Fig.4 The error trends of four points on Gaussian surface

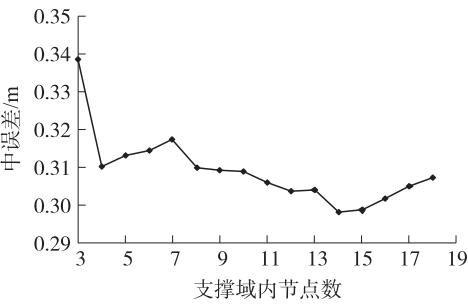


图5 云南数据待插值点中误差变化折线图

Fig.5 The mean square error trend of interpolated points in Yunnan data

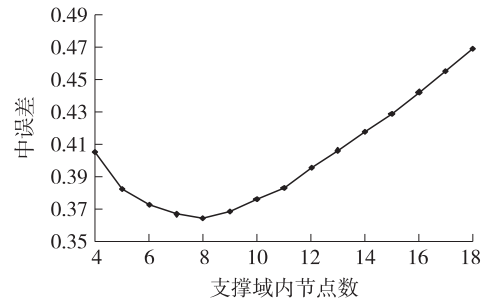


图6 数学曲面数据待插值点中误差变化折线图

Fig.6 The mean square error trend of interpolated points on Gaussian surface

基于以上结果,采用本文的自适应方法确定内插邻域,选择待插值点的邻近点,其内插邻域内点数的变化范围为4~15,云南数据待插值点计算结果的中误差为0.29 m.采用数学曲面数据计算时,邻近点内插邻域内点数的变化范围为4~11,中误差为0.35.由图5、图6可见,在根据经验估计邻域内点数,并以此确定邻域半径进行内插时,两种数据插值结果中误差的最小值均大于自适应方法所得结果.

图7、图8为云南数据一般方法与自适应法确定邻域后,计算结果的误差统计直方图.在采用自适应法确定邻域时,计算结果误差的最大值和最小值均小于一般方法的误差.数学曲面数据采用两种方法得到结果的误差统计直方图为图9和图10,在采用自适应方法确定邻域时,计算结果精度明显提高,且绝对误差在0.8以内的点较多.

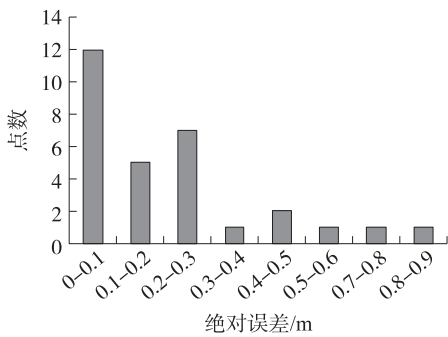


图7 云南数据待插值点邻域内包含14个点计算结果误差统计直方图

Fig.7 The error of ordinary method of Yunnan interpolated points

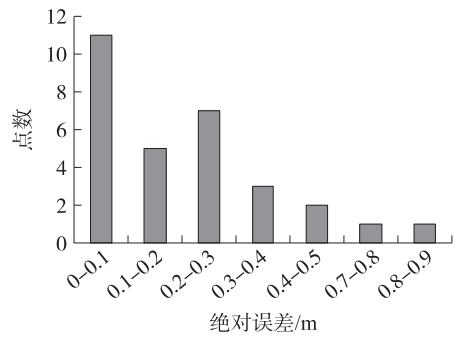


图8 云南数据采用自适应法计算结果误差统计直方图

Fig.8 The error of adaptive method of Yunnan interpolated points

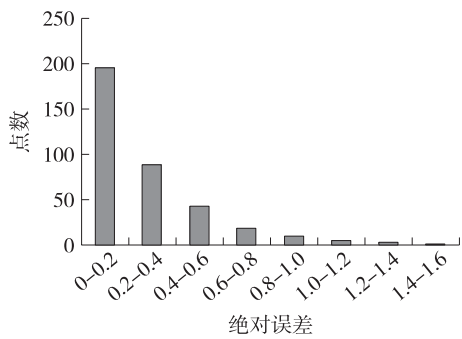


图 9 数学曲面数据待插值点邻域内包含 9 个点计算结果误差统计直方图

Fig. 9 The error of ordinary method of interpolated points on Gaussian surface

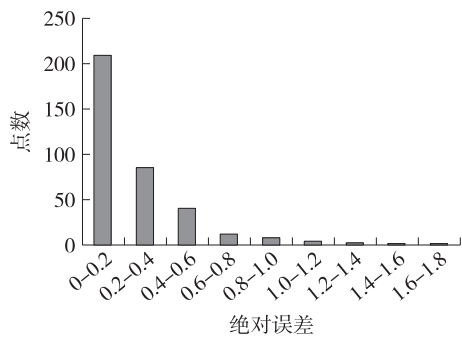


图 10 数学曲面数据采用自适应法计算结果误差统计直方图

Fig. 10 The error of adaptive method of interpolated points on Gaussian surface

3 结论与展望

本文选择根据邻域内包含的点数确定邻域半径的方法作为基本方法,分析了待插值点邻域逐渐变化时,计算结果的误差变化趋势. 以此为基础提出了自适应确定内插邻域的方法. 通过实验得出,待插值点邻域从小到大变化时,各点误差的变化趋势基本不同,而所有的点的中误差变化趋势几乎一致,先逐渐减小再逐渐增大. 采用自适应确定内插邻域的方法,计算各个待插值点的高程值,其误差小于仅根据点数确定内插邻域的方法,在精度上具有一定的优势. 但该方法还存在一些不足,如参数如何设置,效率如何提高,在下一步的研究工作中仍需继续改进.

[参考文献]

[1] 王耀革,朱长青. 基于 Coons 曲面的规则格网 DEM 表面模型[J]. 测绘学报,2008,37(2) :217-222.

[2] 涂震飏,刘永才,苏康,等. 一种面向大场景 3 维显示的地形建模方法[J]. 测绘学报,2004,33(1) :71-76.

[3] 周瑞忠,周小平,缪圆冰. 具有自适应影响半径的无单元法[J]. 工程力学,2001,18(6) :94-99.

[4] 汤国安,刘学军,闫国年. 数字高程模型及地学分析的原理与方法[M]. 北京:科学出版社,2005.

[5] Du Chongjiang. An interpolation method of grid-based terrain modelling[J]. The Computer Journal,1996,39(10) :837-843.

[6] 蔡永昌,朱合华. 基于局部搜索算法的自然邻接点方法[J]. 力学学报,2004,36(5) :623-628.

[7] Tobler W R. A computer movie simulating urban growth in the detroit region[J]. Economic Geography,1970,46:234-240.

[8] 刘学军. 基于规则格网数字高程模型解译算法误差分析与评价[D]. 武汉:武汉大学测绘遥感信息工程国家重点实验室,2002.

[9] 刘学军,龚建雅,周启鸣,等. DEM 结构特征对坡度坡向的影响分析[J]. 地理与地理信息科学,2004,20(6) :1-5.

[责任编辑:丁 蓉]