

文章编号: 1671-8860(2004)11-1017-04

文献标识码: A

基于动态 Voronoi 图的距离倒数 加权法的改进研究

颜辉武^{1,2} 祝国瑞² 徐智勇²

(1 清华大学深圳研究生院环境工程与管理研究中心, 深圳市大学城, 518055)

(2 武汉大学资源与环境科学学院, 武汉市珞喻路 129 号, 430079)

摘要: 针对距离倒数加权法的不足, 采用动态 Voronoi 图选取待插值点的邻近点, 对每个已知数据点增加反映该点梯度的增量, 并建立已知数据点的 ϵ 邻域避免计算误差, 以改进距离倒数加权法, 为距离倒数加权法的广泛应用创造条件。通过交叉检验与实际应用, 证实利用改进后的距离倒数加权法进行空间数据插值精度更高, 效果更为理想。

关键词: 动态 Voronoi 图; 距离倒数加权法; 空间数据插值

中图分类号: P283.1

散乱点数据插值在地理、地质、医学和气象等领域有着广泛的应用, 插值曲面的构造通常有三角插值和距离加权插值两种途径^[1]。距离倒数加权法的基本思想是: 假设在 $x-y$ 平面上有一离散的空间数据点集 $\{(x_i, y_i, z_i) | i = 1, 2, \dots, N\}$, 离散点 (x_i, y_i, z_i) 在 $x-y$ 平面上的投影记为 D_i , 则平面上任何一点 $P(x, y)$ 上之所以有一高程值 z , 是由于已知离散点 D_i 的高程值 z_i 影响的结果。每个 D_i 上的 z 值对点 $P(x, y)$ 上 z 值的影响是不同的, 这种影响的大小与它们的距离的大小正好相反, 即离 P 点近的点, 其 z 值对 P 点 z 值的影响大些; 离 P 点远的点, 其 z 值影响就小些。把各个点对于 P 点 z 值的影响值加起来就是 P 点的 z 值, 即数据点的 z 值按与 P 点的距离加权平均, 距离小的, 反映在权上就大; 距离大的, 反映在权上就小。由于要求曲面满足光滑性的要求, 权函数应是距离倒数的某次幂。

令 z_i 是 D_i 上的 z 值, $d[P, D_i]$ 为 P 与 D_i 之间的距离, 当 P 点确定时, 可以简写为 d_i 。根据上述思想, P 点上的 z 值为:

$$f(P) = \begin{cases} \frac{\sum_{i=1}^N (d_i)^{-\mu} z_i}{\sum_{i=1}^N (d_i)^{-\mu}}, & \text{对所有 } D_i, d_i \neq 0 \\ z_i, & \text{对某个 } D_i, d_i = 0 \end{cases} \quad (1)$$

由于 $\mu=2$ 时, 不仅给出了外观上满意的实验结果, 而且具有容易计算的优点(不需要进行开方运算), 因此, 在实际计算中幂指数 μ 常取为 2。

1 距离倒数加权法的不足

传统的距离倒数加权法方法无疑十分简单, 但也存在着明显的缺陷。

1) 当数据点的数目非常庞大时, $z=f(P)$ 的计算将变得十分复杂, 巨大的计算量甚至可能使方法无法实现。

2) 方法只考虑了从 D_i 到 P 的距离, 而没有考虑其方向。而事实上, 只考虑距离的大小是不足的, 有的已知离散点虽然距待插值点较近, 但是它对插值点的影响可能被其他的点所屏蔽掉, 如图 1, 可以看出, 点 2 虽然同点 3、4、5 距待插值点的距离相等, 但相比而言, 点 2 对插值点的影响就很小, 因为它受到了点 1 的屏蔽。

3) 在每个数据点上, 对曲面没有附加任何导数约束。

4) 在 D_i 的邻域中, 由于 $d_i \approx 0$, 计算的误差将变得非常敏感, 尤其是当有两个项形成占优而又符号相反时, 计算的误差更是如此。

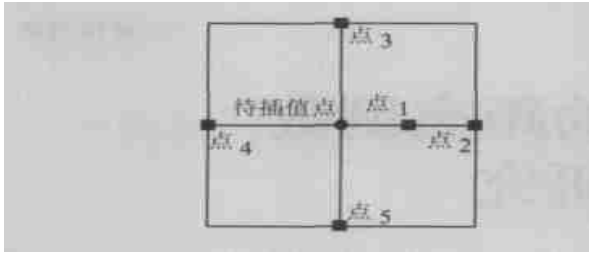


图 1 方向对权值影响

Fig. 1 Influence of Direction to the Weight

2 距离倒数加权法的改进

2.1 选择邻近点

虽然有学者针对此缺点进行过改进^[2],也取得了一定的效果,但是仍然存在着一个很重要却经常被忽视的问题,即参考点坐标或参考点所在坐标系统的微小变化会使选点结果差别很大,结果可能造成插值曲面的不连续^[3]。而造成这个问题的原因恰恰在于仅以距离为基础进行选点,事实上距离难以很好地描述空间相邻性。显然,对于离散的数据点之间空间相邻性的描述,需要给出一种较好的数学表达,Voronoi 图就是一种很好的工具^[4]。因此,笔者提出利用空间邻近关系分析的有效工具 Voronoi 图对距离倒数加权法进行邻近点选取,并对邻近点进行分级,该方法还可避免上述算法中所出现的由于方向原因而产生的对插值点的计算的影响。

2.1.1 利用 Voronoi 图求取待插值点的邻近点

首先,根据已知数据点构建 Voronoi 图,然后加入待插值点动态构建 Voronoi 图。在图中,与待插值点相邻的发生点为其 1 级邻近点,与 1 级邻近点相邻的发生点则为 2 级邻近点,依此类推,可以得到待插值点的 n 级邻近点。图 2(a)为研究区域已知数据形成的 Voronoi 图,小方块表示待插值点;图 2(b)为加入待插值点后动态生成的 Voronoi 图,其中颜色较深的区域内的发生点表示待插值点的 1 级邻近点,颜色较浅的区域内的发生点则表示待插值点的 2 级邻近点。

2.1.2 两种模型

显然,1 级邻近点比 2 级邻近点对待插值点的影响大, $n-1$ 级邻近点比 n 级邻近点对待插值点的影响大,其影响程度(w)可以采用指数模型或线性模型。

① 指数模型 $w=e^{-x}$ (图 3(a))

采用指数模型时,影响程度随邻近级数的增大而呈指数衰减,当邻近级数趋于无穷大时,其影

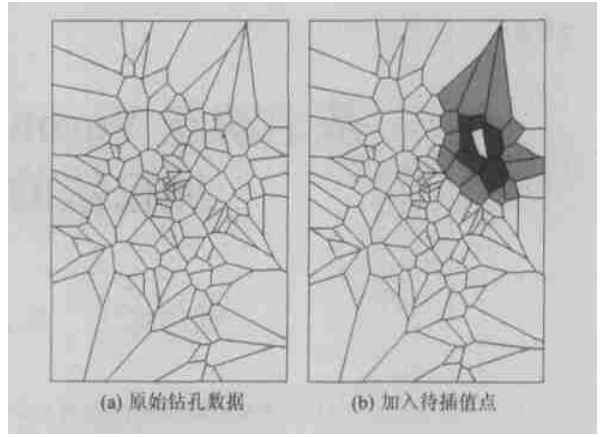


图 2 Voronoi 图

Fig. 2 Map of Voronoi

响程度趋于 0。

② 线性模型 (图 3(b))

采用线性模型时,在邻近级数小于等于某一常数时,影响程度呈线性变小,在大于该常数后,等于 0,即

$$w = \begin{cases} (a-x)/a, & x \leq a \\ 0 & x > a \end{cases} \quad (a \text{ 为考虑的最大邻近级数}) \quad (2)$$

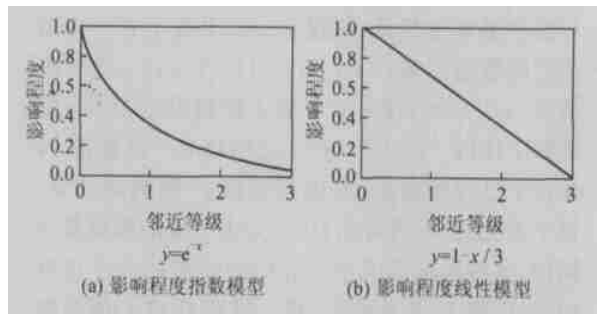


图 3 两种影响程度模型

Fig. 3 Two Types of Models for Influence Degree

由于 2 级邻近点过渡到 1 级邻近点,然后由 1 级邻近点过渡到待插值点,以考虑到数据变化的趋势,因此,邻近级数一般取 2。实验表明,指数模型与线性模型均可采用,两者结果相差不大。

2.1.3 改进后的公式

记待插值数据点 P 的邻近点为 C'_i ,简称为 C' ,考虑邻近级数对其的影响,引入新的权函数为: $W_i=(d_i)^{-2} w_i$,则 P 点的 z 值为:

$$f(P) = \begin{cases} \frac{\sum_{D_i \in C'} W_i z_i}{\sum_{D_i \in C'} W_i}, & \text{对所有 } D_i, d_i \neq 0 \\ z_i, & \text{若有某个 } D_i, d_i = 0 \end{cases} \quad (3)$$

2.2 偏导数的估计

距离倒数加权法的第三个不足之处是: 在每个数据点上都有一个零梯度。对于这一不足的改进, 笔者采用了参考文献 [2] 中提出的方法: 在每个数据点附近给函数添加上一个增量, 以便使形成的曲面在每一个 D_i 上具有一个比较合理的偏导数。

首先对偏导数做一估计, 即采用一阶差商的加权平均作为 D_i 上 X 方向与 Y 方向偏导数 A_i 与 B_i 的近似值, 令 C_i 为 D_i 的邻近点集合, 则

$$A_i = \frac{\sum_{D_j \in C_i} W_j \frac{(z_j - z_i)(x_j - x_i)}{(d[D_j, D_i])^2}}{\sum_{D_j \in C_i} W_j}$$

$$B_i = \frac{\sum_{D_j \in C_i} W_j \frac{(z_j - z_i)(y_j - y_i)}{(d[D_j, D_i])^2}}{\sum_{D_j \in C_i} W_j} \quad (4)$$

为了使所构造的曲面具有所估计的偏导数, 对于每一个属于 C' 中的 D^i , 必须计算一个关于 P 的函数的增量 Δz_i , 将此增量加到曲面上时, 曲面在 D_i 点上将拥有所希望的偏导数 (A_i, B_i) 。其中 Δz_i 可取为:

$$\Delta z_i = [A_i(x - x_i) + B_i(y - y_i)] \frac{v}{v + d_i} \quad (5)$$

插入因子 $v/(v + d_i)$ 可以保证当 d_i 从 0 变到无穷大时, 它单调地从 1 变为 0, 当 d_i 很大时, 它的性质很像 $1/d_i$, 从而控制远距离点对偏导数的影响程度。虽然如此, 变化的幅度可能仍然很大, 对此必须选取适当的 v , 使得增量的变化幅度得到限制(至少不应掩盖 z 值本身), 例如,

$$v = \frac{0.1[\max\{z_i\} - \min\{z_i\}]}{[\max\{A_i^2 + B_i^2\}]^{\frac{1}{2}}} \quad (6)$$

此时可以证明, 对于任意 d_i 均有:

$$|\Delta z_i| \leq 0.1[\max\{z_i\} - \min\{z_i\}] \quad (7)$$

其中数 0.1 只是任意地选取, 对于某些曲面, 可以变动这个常数。此时得到:

$$f(P) = \begin{cases} \frac{\sum_{D_i \in C'} W_i (z_i + \Delta z_i)}{\sum_{D_i \in C'} W_i}, & \text{若所有 } D_i, d_i \neq 0 \\ z_i, & \text{若某个 } D_i, d_i = 0 \end{cases} \quad (8)$$

显然, 曲面在 D_i 有偏导数 A_i, B_i 。

2.3 减少计算误差

当待插值点非常接近某一已知离散点时, 舍入和截断将引起很大的误差; 另外, 计算机进行计算时也会有很大的困难(如避免溢出等)。这些问

题可通过在 D_i 上建立一个 ϵ 邻域加以避免。因为

$$\lim_{P \rightarrow D_i} f(P) = z_i$$

所以建立 D_i 的 ϵ 邻域, 一旦 P 点落进这个邻域, 则令 $f(P) = z_i$, 如果 P 点同时落入几个 D_i 的 ϵ 邻域, 则它们的 z 值被平均后作为 P 点的值。这几个 D_i 组成的集合定义为 $N(P)$ 。 ϵ 的选取是依据机器的具体情况而定的, 最后得到的距离倒数加权公式为:

$$f(P) = \begin{cases} \frac{\sum_{D_i \in C'} W_i (z_i + \Delta z_i)}{\sum_{D_i \in C'} W_i}, & \text{若所有 } D_i, d_i > \epsilon \\ \frac{\sum_{D_i \in N(P)} z_i / \sum_{D_i \in N(P)} 1}{\sum_{D_i \in N(P)} 1}, & \text{若某些 } D_i, d_i \leq \epsilon \end{cases} \quad (9)$$

3 改进后的距离倒数加权法插值结果验证

距离倒数加权法插值结果的检验是由交叉检验程序来完成的。交叉检验的基本原理是: 把各样点的实测值作为真值, 而将采用已知样点相互估计出的值作为估计值, 通过比较估计值与真值之间的偏差程度来进行检验。

设采样点观测值为 Z_i , 估计值为 Z_i^* , N 为交叉检验点数。常用的检验标准^[5]是: ① 偏差均值 R_1 ; ② 偏差平方和 R_2 ; ③ 绝对偏差均值 R_3 ; ④ 偏差方差 R_4 ; ⑤ 均值误差 R_5 。经过交叉检验计算, 各项计算结果如下: $R_1 = -0.051$; $R_2 = 0.0233$; $R_3 = 0.1204$; $R_4 = 0.020699$; $R_5 = -0.0025$, 由此可知, 改进后的计算模型拟合效果较好, 可用于体数据生成的空间格网插值计算。

在南京师范大学江苏省地理信息科学重点实验室资助的“体图形学技术及其在地下水开采与地面沉降分析研究中的应用”项目中, 笔者应用改进后的距离倒数加权法进行曲面插值, 图4即是

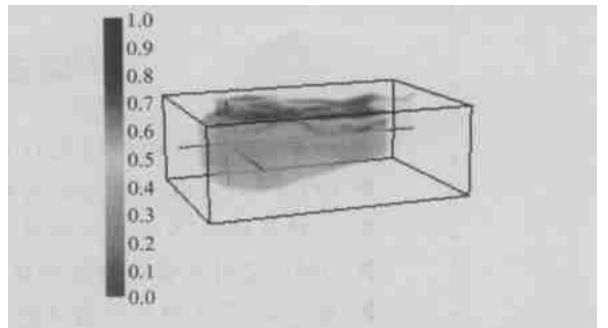


图 4 孔隙度体绘制图像及三维量测

Fig. 4 Volume Rendering and 3D Measuring of Porosity

利用改进后的距离倒数加权法插值生成常州-武

进地区地下水水文地质层孔隙度三维体数据后的体视化图像。

参 考 文 献

- 1 王靖波. 基于 Kriging 方法的空间散乱点插值. 计算机辅助设计与图形学学报, 1999, 11(6): 525~529
- 2 杨学平. 计算机绘图. 北京: 电子工业出版社, 1980. 75~80
- 3 朱怀球, 吴江航. 基于 C^∞ 基函数的自然邻近点插值 (NND) 方法在科学计算可视化上的应用. 计算机工程与应用, 2001(1): 1~3

- 4 李志林, 朱 庆. 数字高程模型. 武汉: 武汉测绘科技大学出版社, 2000
- 5 张瑞新, 于汝绶. 距离加权法的完善及其在层状矿床储量计算中的应用前景. 中国矿业大学学报, 1991, 20(4)
- 6 闫治文, 郭仁忠. 用 Voronoi 图描述空间方向关系的理论依据. 武汉大学学报·信息科学版, 2002, 27(3): 306~310

第一作者简介: 颜辉武, 副教授, 博士后. 主要研究方向: 环境地理信息系统, 地理信息三维可视化及生态城市规划。

E-mail: yhwmcy@263.net

Improvement Research on Inverse Distance Weighting Method Based on Dynamic Voronoi Diagram

YAN Huiwu^{1,2} ZHU Guorui² XU Zhiyong²

(1 Research Center for Environmental Engineering and Management, Shenzhen Graduate College,

Tsinghua University, College City of Shenzhen, Shenzhen 518055, China)

(2 School of Resource and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: In the light of the shortage of traditional inverse distance weighting method, this paper proposes to use the dynamic Voronoi method to select neighboring points to be interpolated point, to add increment reflecting the gradient to every known data points, and to create neighborhood of the known data points to avoid calculation error. The result of cross inspection shows that improved inverse distance weighting method is better than the traditional one.

Key words: Voronoi; inverse distance weighting method; spatial interpolation

About the first author: YAN Huiwu, associate professor, post-doctoral fellow. His major research orientations include environmental GIS, the 3D visualization of geo-information and the ecological urban planning.

E-mail: yhwmcy@263.net

(责任编辑: 洪远)

下期主要内容预告

- ▶ 基于设计模式的组件化 GIS 软件开发的方法
- ▶ PPGIS 及其在加拿大 安省核废料处理选址项目中的应用
- ▶ 一种新的 GIS 智能分析框架——GIS 智能体
- ▶ GML 空间数据存储映射模型研究
- ▶ 基于边-节点和原子属性的多比例尺 GIS 数据模型
- ▶ 基于 Ontology 的网络地理服务描述与发现