

矢量 GIS 平面一般曲线等概率密度误差模型的几何特征

汤仲安^{1,2}

(1. 中南大学 信息物理工程学院,湖南 长沙 410083; 2. 湖南省国土资源信息中心,湖南 长沙 410011)

Geometrical Characteristic for The Equivalent Probability Density Error Model of 2D Generic Curve in Vector GIS

TANG Zhong-an^{1,2}

(1. School of Infophys and Geomatics Engineering, Central South University, Changsha 410083, China; 2. Info Center for National Land & Resources of Hunan Province, Changsha 410011, China)

Abstract:Based on the modeling theory of the equivalent probability density error model (marked as "Tepdem") and the numerical algorithm, this paper provides an approach on how to ascertain a point location with the highest coordinates precision between both of the adjacent characteristic points on a generic curve accurately by use of the theory of function extremum as well as the iterative algorithm, and it also provides the method on how to obtain the least width of "Tepdem" exactly, from which, the explanation of geometrical characteristic on "Tepdem" has been brought forward theoretically. Accordingly, the positional uncertainty theory for vector GIS will be more consummate. Test validates the correctness on theoretical deduction by case study and visualization analysis.

Key words: GIS; generic curve; error model; geometrical characteristic

摘 要:基于等概率密度误差模型建模原理和数值算法,运用函数极值理论和迭代方法来求解平面一般曲线上两相邻特征点间位置精度最高的点,以精确确定误差模型的最小带宽,从理论上给出等概率密度误差模型的几何特征,从而进一步完善矢量 GIS 的位置不确定性理论。通过实例计算与可视化分析,验证了理论推导的正确性。

关键词:地理信息系统;一般曲线;误差模型;几何特征

汤仲安^[1,2]定义了“等概率密度误差模型”(The Equivalent Probability Density Error Model,记为“Tepdem”)、“形状因子”(Shape factor)、“尺度因子”(Scale factor)等概念,提出了矢量 GIS 线状实体位置不确定性误差模型的建模原理;汤仲安^[3]研究了矢量 GIS 中平面一般曲线误差模型包络线;汤仲安、朱建军^[4]基于矢量 GIS 平面线状实体的法平面提出了误差模型的概率算法。基于上述研究成果汤仲安、朱建军等^[5]研究了矢量 GIS 平面随机线元误差模型的几何特征。

本文拟研究平面一般曲线“Tepdem”的几何特征,精确确定各分段曲线上坐标精度最高点的位置,以期提供曲线“Tepdem”几何特征的理论解释。因为平面规则曲线(如圆曲线、缓和曲线等)可以采

用解析函数来表达,无需拟合过程,因而可以看作平面不规则曲线误差模型几何特征研究的特例。

1 相关概念及其公式推导

给定 $n + 1$ 个特征点 $P(x_j, y_j)$, $j = 0, 1, 2, \dots, n$, 设其对应的方差-协方差矩阵为

$$xy = \begin{bmatrix} 2 & & & & \\ x_0 & x_0 y_0 & \cdots & x_0 x_n & x_0 y_n \\ & 2 & & & \\ y_0 x_0 & y_0 & \cdots & y_0 x_n & y_0 y_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & 2 & \\ x_n x_0 & x_n y_0 & \cdots & x_n & x_n y_n \\ & & & & 2 \\ y_n x_0 & y_n y_0 & \cdots & y_n x_n & y_n \end{bmatrix} \quad (1)$$

虽然不规则曲线拟合方法众多,但是三次样条插值方法(Spline Method)平滑性最好^[6]。由于

收稿日期:2006-04-24; 修回日期:2006-09-08

基金资助:中南大学博士后科研基金项目(2006);国土资源部百名优秀青年科技人才计划项目(2005)

作者简介:汤仲安(1971-),男,湖南浏阳人,高级工程师,从事 GIS 与遥感信息的不确定性研究、现代测量数据处理理论与应用研究。E-mail:whtia@163.com

很多情况下三次样条插值方法的插值效果最好, 为便于原理描述, 本文选用三次样条插值函数确定各相邻特征点间拟合曲线分段函数解析式 $y = S_i(x)$, ($i = 1, 2, \dots, n$)。根据易大义等^[7], 对于任意的 $i \in N, i \in n$, 相邻特征点 (x_{i-1}, y_{i-1}) , (x_i, y_i) 之间的三次样条插值函数为

$$S_i(x) = M_{i-1} \frac{(x_i - x)^3}{6 h_i} + M_i \frac{(x - x_{i-1})^3}{6 h_i} + (y_{i-1} - \frac{M_{i-1}}{6} h_i^2) \frac{x_i - x}{h_i} + (y_i - \frac{M_i}{6} h_i^2) \frac{x - x_{i-1}}{h_i} \quad (2)$$

其中, $x \in [x_{i-1}, x_i]; i = 1, 2, \dots, n, h_i = x_i - x_{i-1}$ 。拟合曲线的图形如图 1 所示。

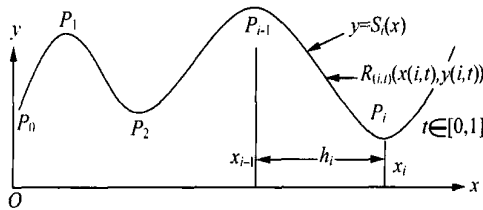


图 1 不规则曲线特征点、拟合曲线及其分段区间示意图

Fig. 1 Feature points and its corresponding anomalous fitting curve

为了讨论的方便, 此处选择自然边界条件, 即: $M_0 = S(x_0) = S(x_n) = M_n = 0$, 根据易大义等^[7], 可以唯一地确定三次样条插值函数的系数 $M_i (i = 1, \dots, n - 1)$ 。

1.1 拟合曲线上任意点方差-协方差矩阵的确定

设第 i 个特征点区间 $(x_{i-1}, x_i), i = 1, 2, \dots, n$ 对应拟合曲线上任意点为 $R(i, t), i \in N, t \in [0, 1]$, 且 $x(i, t) = tx_i + (1 - t)x_{i-1}$ 。显然, 当 $t = 0$ 或 $t = 1$ 时, $R(i, t)$ 即分别为拟合曲线的两端点。将 $x(i, t)$ 代入式 (2), 并顾及 $h_i = x_i - x_{i-1}$, 化简、微分并整理得

$$\begin{bmatrix} dx_{(i,t)} \\ dy_{(i,t)} \end{bmatrix} = \begin{bmatrix} (i,t) & 0 & (i,t) & 0 \\ (i,t) & (i,t) & \vartheta(i,t) & (i,t) \end{bmatrix} \begin{bmatrix} dx_{i-1} \\ dy_{i-1} \\ dx_i \\ dy_i \end{bmatrix}$$

其中 $(i, t) = -\frac{h_i}{3} [M_{i-1}(1-t)^3 - M_i(1-t) +$

$M_i t^3 - M_i t]$; $(i, t) = 1 - t; \vartheta(i, t) = - (i, t) t; (i, t) = t$ 。

由方差-协方差传播率得 $R(i, t)$ 处方差-协方差公式:

$$x(i,t)y(i,t) = \begin{bmatrix} x_{(i,t)}^2 & x_{(i,t)}y_{(i,t)} \\ y_{(i,t)}x_{(i,t)} & y_{(i,t)}^2 \end{bmatrix} \quad (3)$$

为便于书写, 将 (i, t) 记为 \cdot , 其余类似。将式 (3) 展开, 得

$$\begin{aligned} x_{(i,t)}^2 &= x_{i-1}^2 + x_{i-1}x_i + x_i x_{i-1} + x_i^2 \\ x(i,t)y(i,t) &= x_{i-1}^2 y_{i-1} + x_{i-1}^2 y_i + \vartheta x_{i-1} x_i + \vartheta x_i^2 + x_{i-1} y_i + x_i y_{i-1} \\ y(i,t)x(i,t) &= x_{i-1}^2 y_{i-1} + x_{i-1}^2 y_i + \vartheta x_{i-1} x_i + \vartheta x_i^2 + y_{i-1} x_i + y_i x_{i-1} \\ y_{(i,t)}^2 &= y_{i-1}^2 + 2 y_{i-1} y_i + \vartheta^2 x_i^2 + y_i^2 + \vartheta(x_{i-1} x_i + x_i x_{i-1}) + (x_{i-1} y_{i-1} + y_{i-1} x_{i-1}) + (x_{i-1} y_i + y_i x_{i-1}) + \vartheta(x_i y_{i-1} + y_{i-1} x_i) + (y_{i-1} y_i + y_i y_{i-1}) + \vartheta(x_i y_i + y_i x_i) \end{aligned}$$

特征点 $P(x_j, y_j), j = 0, 1, 2, \dots, n$ 的方差-协方差矩阵由式 (1) 查寻。

1.2 变换坐标系中空间概率密度分布函数的简化表达式

根据 1.1, 由宗俦、鲁林成^[8], 得点 $R(i, t)$ 处标准误差椭圆的主轴方位角

$$(i, t) = 0.5 \arctan \frac{2 x_{(i,t)} y_{(i,t)}}{x_{(i,t)}^2 - y_{(i,t)}^2} \quad (4)$$

将坐标系 Oxy 的原点平移至 $R(i, t)$ 并顺时针方向旋转 (i, t) , 得坐标系 $O(i, t) uv$ 。由汤仲安^[1]知, 在 $O(i, t) uv$ 中, 点 $R(i, t)$ 对应的信息矩阵为

$$\begin{bmatrix} u_{(i,t)}^2 & u_{(i,t)}v_{(i,t)} \\ v_{(i,t)}u_{(i,t)} & v_{(i,t)}^2 \end{bmatrix}$$

式中

$$u_{(i,t)}^2 = \cos^2(i, t) x_{(i,t)}^2 + \sin^2(i, t) x_{(i,t)} y_{(i,t)} + \sin^2(i, t) y_{(i,t)}^2 \Rightarrow u_{(i,t)} \quad (5)$$

$$v_{(i,t)}^2 = \sin^2(i, t) x_{(i,t)}^2 - \sin^2(i, t) x_{(i,t)} y_{(i,t)} + \cos^2(i, t) y_{(i,t)}^2 \Rightarrow v_{(i,t)} \quad (6)$$

由汤仲安^[1,2]知: 点 $R(i, t)$ 在变换坐标系 $O(i, t) uv$ 中的空间概率密度分布函数的简化表

达式为

$$(u_{(i,t)}, v_{(i,t)}) = \frac{1}{2^{u_{(i,t)} v_{(i,t)}}} \exp \left[-\frac{1}{2} \left(\frac{u_{(i,t)}^2}{u_{(i,t)}} + \frac{v_{(i,t)}^2}{v_{(i,t)}} \right) \right] = \frac{1}{2^{u_{(i,t)} v_{(i,t)}}} \exp \left[-\frac{u_{(i,t)} + v_{(i,t)}}{2} \right]$$

其中, (i, t) R^+ 为点 $R_{(i,t)}$ 处对应的“标准误差椭圆缩放系数”^[2]。

由此可见, 矢量 GIS 平面一般曲线空间概率密度分布曲面的高度取决于该曲线上任意点 $R_{(i,t)}$ 对应的空间概率密度分布函数的振幅

$$\frac{1}{2^{u_{(i,t)} v_{(i,t)}}}$$

2 几何特征确定思路

在对“-带”的研究中, Zhang 和 Tulip, Caspary 和 Scheuring 曾利用误差传播率推算出线段中间点均方差小于其两端点均方差^[9,10]; 之后, Dutton 采用 Monte Carlo 方法证实了这一点^[11]。因此, 线元的空间概率密度分布曲面呈现出“两端低、中间高”的拱形。研究发现, 区间 (x_{i-1}, x_i) , $i = 1, 2, \dots, n$ 内平面曲线的空间概率密度分布曲面同样具有这一特征。由此可知: 平面曲线上任意点在此区间内对应的空间概率密度分布函数振幅最低的点位于曲线段的某一端点处, 且精度最高点位于除端点外的曲线上, 即必然存在 $t \in [0, 1]$, 使得点 $R_{(i,t)}$ 的点位精度最高, 或对应曲线段的空间概率密度分布曲面的振幅达到最高。因此, 有:

曲线段上点 $R_{(i,t)}$ 处点位精度最高 \Leftrightarrow 点 $R_{(i,t)}$ 处对应的空间概率密度分布函数的振幅达到最高。由此, 基于“空间概率密度分布函数的振幅达到最高”, 即可确定出点 $R_{(i,t)}$ 的精确位置。

3 几何特征确定方法

由于区间 (x_{i-1}, x_i) , $i = 1, 2, \dots, n$ 内平面曲线上精度最高点对应的空间概率密度分布函数的振幅达到最大, 即存在 $t \in [0, 1]$, 使得

$$\frac{1}{2^{u_{(i,t)} v_{(i,t)}}} = \max \Leftrightarrow u_{(i,t)} v_{(i,t)} = \min$$

因此, 参数 t 的确定依赖于方程 $(u_{(i,t)} v_{(i,t)})' = 0$, 即

$$(u_{(i,t)} v_{(i,t)})' + u_{(i,t)} (v_{(i,t)})' = 0$$

令

$$Z(i, t) = (u_{(i,t)})' v_{(i,t)} + u_{(i,t)} (v_{(i,t)})'$$

则

$$Z(i, t) = (u_{(i,t)})' v_{(i,t)} + 2(u_{(i,t)})' (v_{(i,t)})' + u_{(i,t)} (v_{(i,t)})''$$

由 1.1, 有

$$= -\frac{h_i}{3} [-3(1-t)^2 M_{i-1} + M_{i-1} + 3M_i t^2 -$$

$$M_i] \Rightarrow = -2h_i [(1-t) M_{i-1} + t M_i]$$

且

$$= -1, \quad = 0; \quad \delta = -; \quad \delta = -; \quad = 1, \quad = 0$$

于是, 根据式(3), 式(4), 可求得如下各值

$$\left(\frac{2}{x_{(i,t)}} \right), \left(\frac{2}{x_{(i,t)}} \right), \left(x_{(i,t)} y_{(i,t)} \right), \left(x_{(i,t)} y_{(i,t)} \right), \left(y_{(i,t)} \right), \left(\frac{2}{y_{(i,t)}} \right) \left((i, t) \right) \left((i, t) \right)$$

由式(5), 可得 $\left(\frac{2}{u_{(i,t)}} \right)$, 由于 $\left(\frac{2}{u_{(i,t)}} \right) = 2 u_{(i,t)} \left(u_{(i,t)} \right)$, 可得 $\left(u_{(i,t)} \right)$, 进而可得 $\left(u_{(i,t)} \right)$ 。同理, 由式(6), 可得 $\left(v_{(i,t)} \right)$, $\left(v_{(i,t)} \right)$ 。综上, 根据式(5), 式(6), 可得 $Z(i, t)$, $Z(i, t)$ 。

根据上述推导, 对于第 i 段曲线而言, 由于参数 t 是必然存在的, 因此, 利用牛顿-雷扶生 (Newton-Raphson) 迭代方法可以保证迭代方程的收敛, 于是, 有

$$t(i, n_0 + 1) = t(i, n_0) - \frac{Z(i, t(i, n_0))}{Z(i, t(i, n_0))'}$$

$$(i = 1, 2, \dots, n; n_0 = 0, 1, 2, \dots)$$

参数 t 取初值 $t(i, 0) = 0.5$ 。根据不同精度要求, 选定不同的迭代终止条件。此处, 我们选择的迭代终止条件为 $|t(i, n_0 + 1) - t(i, n_0)| < 10^{-6}$ 。

求出参数 t 值后, 即可确定曲线上分段区间内精度最高点 $R_{(i,t)}$ 的坐标。至此, 平面一般曲线“Tepdem”的几何特征点即已确定。

4 实例计算与可视化分析

设某地类界由 5 个特征点测定, 在相应的矢量数据中, 特征点坐标分别为^[11] $Z_0(1\ 085.\ 102, 2\ 187.\ 546)$, $Z_1(1\ 105.\ 004, 2\ 200.\ 456)$, $Z_2(1\ 115.\ 569, 2\ 187.\ 314)$, $Z_3(1\ 130.\ 964, 2\ 197.\ 854)$, $Z_4(1\ 140.\ 597, 2\ 187.\ 746)$; 单位为 m。

其特征点间方差-协方差矩阵为

$$xy = \begin{pmatrix} 6.171 & 0.022 & 0.099 & 0.078 & 0.035 & 0.054 & 0.049 & -0.023 & 0.017 & 0.074 \\ 0.022 & 5.090 & -0.034 & 0.067 & 0.045 & 0.083 & -0.054 & 0.061 & 0.021 & -0.044 \\ 0.099 & -0.034 & 7.919 & 0.287 & 0.038 & 0.025 & -0.051 & -0.073 & 0.037 & -0.014 \\ 0.078 & 0.067 & 0.287 & 8.971 & 0.036 & -0.027 & -0.049 & 0.073 & 0.048 & -0.043 \\ 0.035 & 0.045 & 0.038 & 0.036 & 5.049 & -0.464 & 0.071 & -0.069 & 0.052 & 0.011 \\ 0.054 & 0.083 & 0.025 & -2.070 & -0.464 & 6.131 & -0.059 & 0.032 & -0.061 & 0.013 \\ 0.049 & -0.054 & -0.051 & -0.049 & 0.071 & -0.059 & 4.029 & 0.371 & -0.047 & 0.035 \\ -0.023 & 0.061 & -0.073 & 0.073 & -0.069 & 0.032 & 0.371 & 3.071 & -0.048 & 0.076 \\ 0.071 & 0.021 & 0.037 & 0.048 & 0.052 & -0.061 & -0.047 & -0.048 & 7.189 & 0.319 \\ 0.074 & -0.044 & -0.014 & -0.043 & 0.011 & 0.013 & 0.035 & 0.076 & 0.319 & 9.231 \end{pmatrix} \text{ (cm}^2\text{)}$$

给定拟合曲线落入其相应等概率密度误差模型内的概率置信水平为 95%。

根据汤仲安、朱建军^[11], 计算知: 点位精度最低的特征点为第二个特征点 Z_1 , 且在给定的概率置信水平 (95%) 条件下, 其对应的“标准误差椭圆缩放系数”为 $t_1 = 2.05$ (此即为平面拟合曲线“Tepdem”的“尺度因子”, 见参考文献[1],[2])。

4.1 几何特征点特征值的计算

实例给定了 5 个特征点, 共有 4 段拟合曲线区间, 每个区间内曲线上几何特征点的特征值可以按照文章介绍的上述算法确定, 具体计算结果见表 1。

表 1 矢量 GIS 平面一般曲线上几何特征点特征值计算一览表

Tab. 1 The eigenvalue calculation for the geometrical characteristic points on 2D generic curve in vector GIS

曲线上相邻特征点区间	t 的计算值	横坐标 x	纵坐标 y
$Z_0 \sim Z_1$	0.183 199	1 088.748 0	2 192.983 0
$Z_1 \sim Z_2$	0.562 299	1 110.944 7	2 192.108 1
$Z_2 \sim Z_3$	0.599 115	1 124.792 4	2 193.540 6
$Z_3 \sim Z_4$	0.166 159	1 132.564 6	2 197.445 4

4.2 平面一般曲线主要特征值的可视化

特征点对应的拟合曲线、等概率密度误差模型及其边界包络线如图 2, 图 3, 图 4 所示。为便于可视化, 误差模型及其边界包络线都是在其原有规模的基础上放大了 100 倍。

4.3 可视化分析

1. 几何特征点均位于矢量 GIS 平面曲线上误差模型宽度最窄的位置处, 这与实情符合;

2. 在平面曲线的同一区间段上, 当曲线上两相邻特征点点位精度相差较大时, 几何特征点点位靠近平面曲线上点位精度较高的某一特征点

处, 如图 4 中平面曲线段 $Z_3 \sim Z_4$;

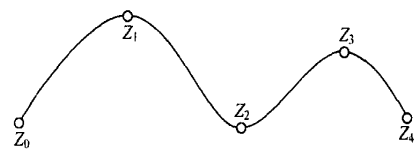


图 2 特征点及其拟合曲线

Fig. 2 The characteristic points and its corresponding fitting curve

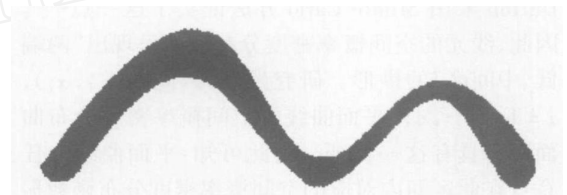


图 3 拟合曲线的等概率密度误差模型

Fig. 3 “Tepdem” to the fitting curve

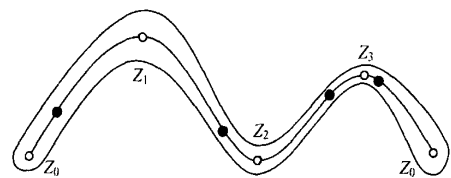


图 4 几何特征点 (图中实心点) 与等概率密度误差模型包络线之间的位置关系

Fig. 4 The positional relations between the geometrical characteristic points and the envelop of “Tepdem”

3. 在平面曲线的同一区间段上, 当曲线上两相邻特征点点位精度相差不大时, 特征值 t 在 0.5 处变化, 几何特征点点位靠近平面曲线的中点处, 如图 4 中平面曲线段 $Z_1 \sim Z_2$;

4. 由参考文献[4]知: 平面随机线元误差模型的几何特征点位置仅取决于线元端点处的坐标精

度,与线元端点处的坐标及其相应的“标准误差椭圆缩放系数”无关——平面一般曲线误差模型几何特征点位置不但取决于平面曲线特征点点位坐标,而且还取决于特征点点位精度,但与点位精度最低的特征点处“标准误差椭圆缩放系数”无关。

5 结束语

基于连续的观点和“Tepdem”建模原理,研究了平面一般曲线“Tepdem”的几何特征,提供了几何特征的具体算法,为矢量 GIS 线状实体误差模型几何特征的确定奠定了理论基础。上述研究在理论上进一步完善了矢量 GIS 平面一般曲线“等概率密度误差模型”理论,使其更为严密;在方法上为 GIS 空间数据位置不确定性研究提供了新的手段;在应用上能够更方便地表示 GIS 平面一般曲线的不确定图形,有利于 GIS 的高精度应用。

参考文献:

- [1] TANG Zhong-an. The Equivalent Probability Density Error Model to Linear Entity in Vector GIS[D]. Wuhan: Wuhan University, 2004. (汤仲安. 矢量 GIS 线状实体等概率密度误差模型[D]. 武汉: 武汉大学, 2004.)
- [2] TANG Zhong-an, WANG Xin-zhou, JI Xian-hua. Modeling Theory to Error Model of Linear Segment in Vector GIS[J]. Geomatics and Information Science of Wuhan University, 2004, 29(11): 968-972. (汤仲安, 王新洲, 纪现华. 矢量 GIS 平面随机线元误差模型建模机理[J]. 武汉大学学报(信息科学版), 2004, 29(11): 968-972.)
- [3] TANG Zhong-an, WANG Xin-zhou, CHEN Zhi-hui. Envelop to Error Model of 2D Spatial Curve in GIS[J]. Acta Geodaetica et Cartographica Sinica, 2004, 33(2): 151-155. (汤仲安, 王新洲, 陈志辉. GIS 中平面一般曲线误差模型包络线[J]. 测绘学报, 2004, 33(2): 151-155.)
- [4] TANG Zhong-an, ZHU Jian-jun. Probability Algorithm for 2-Dimension Random Line Segment Based on the Normal Plane [J]. Acta Geodaetica et Cartographica Sinica, 2006, 35(1): 90-94. (汤仲安, 朱建军. 基于法平面的 GIS 平面随机线元概率算法[J]. 测绘学报, 2006, 35(1): 90-94.)
- [5] TANG Zhong-an, ZHU Jian-jun, et al. Geometric Characteristic for the Equivalent Probability Density Error Model of 2D Random Linear Segment in Vector GIS [J]. Acta Geodaetica et Cartographica Sinica, 2006, 35(2): 149-152. (汤仲安, 朱建军, 等. 矢量 GIS 平面随机线元等概率密度误差模型的几何特征[J]. 测绘学报, 2006, 35(2): 149-152.)
- [6] LI Hai-tao, Deng Ying. MATLAB Program Design [M]. Beijing: Higher Education Publishing House, 2002. (李海涛, 邓樱. MATLAB 程序设计教程[M]. 北京: 高等教育出版社, 2002.)
- [7] YI Da-yi, SHEN Yun-bao, LI You-fa. Calculation Methods [M]. Hangzhou: Publishing Company of Zhejiang University, 1989. (易大义, 沈元宝, 李有法. 计算方法[M]. 杭州: 浙江大学出版社, 1989.)
- [8] YU Zong-chou, LU Lin-cheng. Theory Basis for Surveying Adjustment [M]. Beijing: Publishing House of Surveying and Mapping, 1982. (於宗倬, 鲁林成. 测量平差基础(增订本). 北京: 测绘出版社, 1982.)
- [9] ZHANG G, TULIP J. An Algorithm for the Avoidance of Sliver Polygons and Clusters of Points in Spatial Overlay[A]. Proceedings of 4th International Symposium on Spatial Data Handling(1) [C]. [s. l.]: [s. n.], 1990, 141-150.
- [10] CASPARY W, SCHEURING R. Error-bands as Measures of Geometrical Accuracy [A]. EGIS '92 [C]. [s. l.]: [s. n.], 1992. 226-233.
- [11] DUTTON G. Handling Positional Uncertainty in Spatial Databases [A]. Proceedings of 5th International Symposium on Spatial Data Handling [C]. South Carolina: [s. n.], 1992. 460-469.

(责任编辑: 丛树平)