

层次多结点样条曲线逼近*

蔡占川, 郑才目
(澳门科技大学资讯科技学院, 澳门)

摘要: 提出层次多结点样条曲线逼近算法, 使得逼近的过程能够根据设定阈值自动的选择最优逼近段数。同时, 引入最优逼近曲线判定方法, 实验表明该方法对于 γ 射线能谱等离散数据的平滑具有良好效果。

关键词: 多结点样条; 自由曲线; 层次逼近; 最小二乘法

中图分类号: TP391.1 **文献标志码:** A **文章编号:** 0529-6579 (2014) 01-0018-05

Curves Approximation with Hierarchical Many-knot Splines

CAI Zhanchuan, ZHENG Caimu

(Faculty of Information Technology, Macau University of Science and Technology, Macau, China)

Abstract: Based on many-knot splines, a new curves approximation algorithm is proposed in this paper. The algorithm makes use of a hierarchy of control lattices to generate a sequence of functions whose sum approaches the desired approximation function. This algorithm can automatically select the proper control points according to the setting threshold. The satisfactory result is obtained for the fitting of experimental spectroscopic data using the optimization methods of cubic many-knot splines.

Key words: many-knot splines; free curves; hierarchical approximation; least square method

实际应用特别是逆向工程中, 通过测量工件原型形状建立数学模型, 并利用曲线插补算法生成数控加工指令, 例如人工骨、假牙的数控加工及已知各种零件反求等。对于工件的测量, 一般采用扫描测量方法, 并根据采样数据建立数学模型。对离散数据进行光滑而精确的曲线曲面重构, 即数学建模, 以用于 CAD/CAM 系统。因此, 研究曲线的最优逼近具有重要意义。

B 样条曲线曲面方程是描述自由曲线曲面的最常见数字模型之一。在文献 [1-2] 的一系列论述中, 首先构造了多结点样条基函数, 这类基函数是基数型的, 使得插值过程无须求解方程组。针对常见方法的不方便之处, 多结点样条插值法的设计

适应了某些重要实际问题的需要, 并已做出了一些广泛、灵活地应用, 如飞机外形^[3]、字体设计^[4]、地震数据及漏磁数据去冗余^[5-6]、动画片的计算机制作等^[7-8]。文献 [9-10] 注意到多结点样条插值的理论价值, 给出了深入的讨论, 并将多结点样条插值在理论上推广到更一般情形。文献 [11] 研究了多结点样条的自由曲线最小误差逼近, 并得到良好的效果。本文基于多结点样条函数, 提出层次多结点样条算法^[12], 使得逼近的过程能够根据设定的阈值自动的选择最优逼近段数。

首先简要介绍多结点样条, 其次是多结点样条逼近算法, 然后提出层次多结点样条算法, 接着是实验, 最后是结论。

* 收稿日期: 2013-05-12

基金项目: 国家重点基础研究发展计划“973”资助项目 (2011CB302400); 澳门科技发展基金资助项目 (084/2012/A3, 004/2011/A1, 015/2010/A); 浙江大学 CAD & CG 国家重点实验室开放课题资助项目 (A1310); 国家自然科学基金资助项目 (61170320, 61272364)

作者简介: 蔡占川 (1974 年生), 男; 研究方向: 计算机图形学; E-mail: zccai@must.edu.mo

1 多结点样条

多结点基本样条函数是通过等距 B 样条基本函数的平移及迭加构成^[1-2]。定义多结点样条基函数

$$L_k(t) = \sum_{j=0}^{k-1} c_j \mu^{h_j} \Omega_k(t) \quad (1)$$

式中 $h_0 = 0, 0 < h_1 < h_2 < \dots < h_{k-1} = \frac{k-1}{2}, k = 1, 2, 3, \dots, c_0, c_1, \dots, c_{k-1}$ 为待定常数, μ 为平均算子, 给定任意常数 ξ , 则有

$$\mu^\xi f(x) \equiv \frac{1}{2}(f(x + \xi) + f(x - \xi))$$

$\Omega_k(t)$ 为 k 次 B 样条基函数

$$\Omega_k(t) = \bar{\Delta} \int_{-\infty}^t \Omega_{k-1}(t) dt = \frac{1}{k!} \bar{\Delta}^{k+1} t_+^k = \frac{1}{k!} \sum_{j=0}^{k+1} (-1)^j \binom{k+1}{j} (t + \frac{k+1}{2} - j)_+^k$$

式中: $(\cdot)_+ = \max\{\cdot, 0\}$;

对于公式 (1), 显然 $k = 1$ 时, 有 $L_1(t) = \Omega_1(t)$, 当 $k > 1$ 时, 令 $L_k(t)$ 满足

$$L_k(0) = 1, L_k(i) = 0, i = 1, 2, \dots, k-1$$

于是得到关于 c_0, c_1, \dots, c_{k-1} 的线性方程组, 由 $\{\mu^{h_j} \Omega_k(t)\}$ 得线性独立性质, 可知这样的解存在且唯一, 所求得的 $L_k(t)$ 即为多结点的基本样条函数。依此求得 $k = 2, 3$ 时, 多结点样条的基函数为

$$L_2(t) = (2I - \mu^{\frac{1}{2}}) \Omega_2(t);$$

$$L_3(t) = (\frac{10}{3}I - \frac{8}{3}\mu^{\frac{1}{2}} + \frac{1}{3}\mu) \Omega_3(t)$$

基函数图形如图 1。

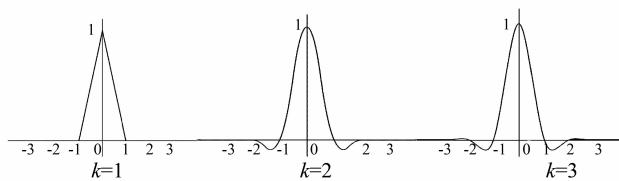


图 1 多结点样条基函数
Fig. 1 Many-knot spline functions

设给定型值点 r_0, r_1, \dots, r_n , 那么, 多结点样条插值曲线可表示为

$$f(t) = \sum_i L_k(t-i) r_i, t \in [0, n] \quad (2)$$

通常取 $k = 1, 2$ 或 3 , 本文主要以 $k = 3$ 的三次多结点样条为例, 三次多结点样条基函数为

$$L_3(t) = \begin{cases} 1 - \frac{5}{2}|t|^2 + \frac{14}{9}|t|^3, & 0 \leq |t| < \frac{1}{2}; \\ \frac{19}{18} - \frac{1}{3}|t| - \frac{11}{6}|t|^2 + \frac{10}{9}|t|^3, & \frac{1}{2} \leq |t| < 1; \\ \frac{37}{12} - \frac{77}{12}|t| + \frac{17}{4}|t|^2 - \frac{11}{12}|t|^3, & 1 \leq |t| < \frac{3}{2}; \\ \frac{5}{6} - \frac{23}{12}|t| + \frac{5}{4}|t|^2 - \frac{1}{4}|t|^3, & \frac{3}{2} \leq |t| < 2; \\ -\frac{49}{18} + \frac{41}{12}|t| - \frac{17}{12}|t|^2 + \frac{7}{36}|t|^3, & 2 \leq |t| < \frac{5}{2}; \\ \frac{3}{4} - \frac{3}{4}|t| + \frac{1}{4}|t|^2 - \frac{1}{36}|t|^3, & \frac{5}{2} \leq |t| < 3; \\ 0, & 3 \leq |t| \end{cases}$$

下面介绍多结点样条逼近算法。

2 多结点曲线逼近算法

假设多结点样条函数 f 逼近数据点集 $P(x, y) (x \in [a, b])$, 函数 f 由控制点 φ_i 唯一确定, 因此只要计算区间 $[a, b]$ 的控制点, 就可以来逼近数据点集 P 。

考虑 P 中的某个点 $p_c(x_c, y_c)$, 由多结点曲线插值函数, 易知插值函数 $f(x_c)$ 只跟周围的 6 个控制点有关, 不妨假设 $0 < x_c < 1$, 则其两边的控制点为 $\varphi_i (i = -2, -1, 0, 1, 2, 3)$, 根据多结点插值公式得:

$$y_c = \sum_{i=-2}^3 w_i \varphi_i$$

式中 $w_i = L_3(u)$, $u = x_c - \lfloor x_c \rfloor$ 。

这是以 φ_i 为变量的不定方程, 满足 φ_i 的值有很多, 选一组能使逼近函数 f 到 x 轴距离最小的解, 依最小二乘法, 即 $\sum_{i=-2}^3 \varphi_i$ 最小。系数的伪逆矩阵为该方程组的最小二乘解

$$\varphi_i = \frac{w_i y_c}{\sum_{a=-2}^3 w_a^2} \quad (3)$$

现在考虑点集中所有点, 由式 (3) 可得, 对于每个不同的点 (x_c, y_c) , 都可以得出一组不同的控制点 φ_c , 当数据点密度较大时, 就会产生控制点的重叠, 如图 2 所示。



图 2 控制点的重叠

Fig. 2 The overlap of the control point

即多个数据点都对同一个网格点有贡献。由于多结点样条的局部性, 只有在控制点 φ_i 两旁 3 个控制网格内的数据点对 φ_i 有影响, 将这些值称为 φ_i 的邻近点。设第 j 网格点的最优解为 φ_j , 则 φ_j 需使得误差 $e(\varphi_j) = \sum_c (w_c \varphi_j - w_c \varphi_c)^2$ 最小。函数 $(w_c \varphi_j - w_c \varphi_c)$ 表示第 j 个控制点对曲线 f 在 Ω 处实际权重与期望权重的差异。 $e(\varphi_j)$ 对 φ_j 求偏导数, 得 φ_j 最优解为

$$\varphi_j = \frac{\sum_c w_c^2 y_c}{\sum_c w_c^2} \quad (4)$$

这样当控制点 φ_j 具有多个邻近点时, 公式 (4) 就可以依最小二乘法取得最优值, 当控制点只有一个邻近点时, 公式 (4) 便退化成公式 (3)。当控制点没有邻近点时, 则表明控制点 φ_j 对数据集中任何点 (x_c, y_c) 均没影响, 意味着 φ_j 可以设置成任意值。这里将其设置成零, 这对后面层次多结点样条具有重要作用。当控制点数较少时, 可以生成较光滑曲线; 增多控制点, 可以增强曲线逼近精度, 但是光滑性会减少。不易找到一个适当的控制点数, 使得曲线既有较好光滑性又有较理想精度, 为此引入层次样条概念^[12], 构建层次多结点样条以使逼近能取得兼顾光滑性与精度。

3 层次多结点逼近

考虑定义域上控制点序列 $\Phi_0, \Phi_1, \dots, \Phi_h$, 给定 Φ_0 的控制点数, 其它各级控制点密度为前一级控制点数 2 倍, 即第 k 级控制点数为 $n+5$, 则第 $k+1$ 级控制点数为 $2n+5$, 算法由最稀疏控制点 Φ_0 开始, 对区间上所有数据点应用第 2 节算法, 在控制点 Φ_0 作用下生成 f_0 , 计算 f_0 到每个点 (x_c, y_c) 误差值 $\Delta^1 y_c = y_c - f_0(x_c)$ 。

将误差值 $\Delta^1 y_c$ 与对应坐标 x_c 组成新数据集 $P_1 = \{(x_c, \Delta^1 y_c)\}$ 作为下一级控制点 Φ_1 的逼近数据, 得到逼近函数 f_1 。因此, $f_0 + f_1$ 将在 x_c 处产生较小误差 $\Delta^2 y_c = y_c - f_0(x_c) - f_1(x_c)$ 。

这一过程由最稀疏控制点 Φ_0 开始, 逐次增加控制点密度, 使逼近误差逐渐减小, 至到逼近误差值小到容差范围内。最后将各级生成的逼近函数 f_k 相加, 得到逼近函数 f , 即 $f = \sum_{k=0}^n f_k$ 。这样得到的曲线即保留了光滑性, 又具有较好的精度。这里 f_0 控制了 f 的总体形状, 往后各级函数 f_k 逐渐减小逼近误差。

算法的伪码如下:

Input: 散乱数据 $V = \{(x_c, y_c)\}$

Output: 控制网格 R

初始化 $\delta_i = 0$ 及 $\bar{\omega}_i = 0$

for 点集 V 中的每个点 (x_c, y_c)

设 $i = \lfloor x_c \rfloor, u = x_c - \lfloor x_c \rfloor$,

计算 ω_k 及 $\sum_{a=-2}^3 w_a^2$

for $k = -2, -1, 0, 1, 2, 3$ do

依公式计算 r_k

将 $\omega_k^2 r_k$ 累加到 $\delta_{(i+k)}$

将 ω_k^2 累加到 $\bar{\omega}_{(i+k)}$

end

end

for all i do

if $\bar{\omega}_i \neq 0$ then

计算 $r_i = \delta_i / \bar{\omega}_i$

else let $r_i = 0$

end

若散乱数据点集的数据点数为 p , 控制网格大小为 $m+5$, 则多结点样条逼近的时间复杂度为 $O(p+m)$ 。

实际运用中也有很多情况下并不知道容差, 或是要求逼近精度不是很高, 但对光滑性要求相对较高。这种情况下本文引入另一种最佳逼近的判别方法。

依上文所述, 令 $S_i^*(x) = f = \sum_{k=0}^{i-1} f_k$, 则 $S_i^*(x)$ 为 i 层的逼近曲线。计算相邻两条逼近曲线 $S_i^*(x)$ 和 $S_{i+1}^*(x)$ 残差平方和

$$\|\varepsilon_i\|^2 = \sum_{x=1}^m [S_{i+1}^*(x) - S_i^*(x)]^2$$

这里 $\|\varepsilon_i\|^2$ 可以看作 $S_i^*(x)$ 相对于 $S_{i+1}^*(x)$ 的偏差, $\|\varepsilon_i\|^2$ 值越小, 表示两条相邻逼近曲线越接近。此外, 从光滑性方面考虑, $\|\varepsilon_i\|^2$ 可以看作曲线 $S_{i+1}^*(x)$ 相对于 $S_i^*(x)$ 的噪声。当 $\|\varepsilon_i\|^2$ 达到最小的时候, 曲线 $S_i^*(x)$ 为最优逼近曲线。

4 实验

这里从限定容差与不限容差两个方面分别举例验证层次多结点样条的曲线逼近。限定容差是为了使数据逼近符合使用者的客观要求, 侧重于逼近精度; 不限容差是为了拟合数据的变化趋势, 侧重于光滑性。

4.1 限定容差曲线逼近

下面将以激光高度计测高数据为例, 阐述在容差条件下的层次多结点样条的逼近效果。

激光高度计测高数据是激光高度计定时向地面发射激光窄脉冲, 通过计算反射回来的激光信号的延迟而得到高程数据, 图 3 中点显示了嫦娥一号激光高度计所测得的莫斯科海某一剖面的高程数据, 即第 0253 轨中的一部分。由系统误差估计得到激光高度计的系统总误差 $< 31 \text{ m}$, 参考文献 [13]。因此将容差设为 30 m , 通过层次多结点样条拟合测高数据, 得到莫斯科海的剖面图, 如图中线段所示。逼近效果显示层次多结点样条能够较好地表现出莫斯科海的剖面形状。

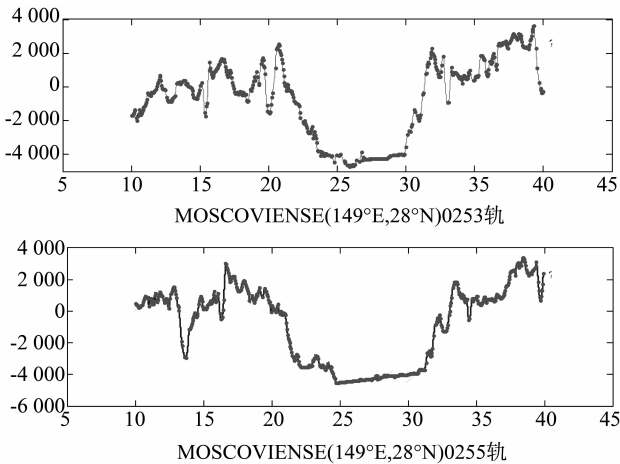


图 3 激光测高数据的重构

Fig. 3 Reconstruction with Chang'E-1 laser altimetry data

4.2 不限容差曲线逼近

下面将以 ^{60}Co 的伽玛频谱数据的平滑为例, 阐述在不限容差条件下的层次多结点样条的逼近效果。

在伽玛射线能谱分析中, 需要采用平滑方法来消除数据统计的涨落。如图 4(a) 显示了嫦娥一号 NaI 探测器对 ^{60}Co 元素积累超过 2 h 的统计数据, 为了减小统计噪声, 需要对统计数据做平滑处理以得到光滑的曲线, 这里采用层次多结点样条方法对统计数据进行逼近, 得到图 4(b) 中的线段所示, 取得良好的效果。

5 结论

本文基于多结点样条, 提出了离散数据逼近的方法, 充分利用了多结点样条的局部性和插值性, 逼近过程只考虑局部数据对控制点的影响, 将求解控制点转移到求解每个数据点对控制点的影响, 避免了求解巨大矩阵系数的线性方程组所带来的麻烦。此外, 算法还能根据设定的容差值, 自动的选

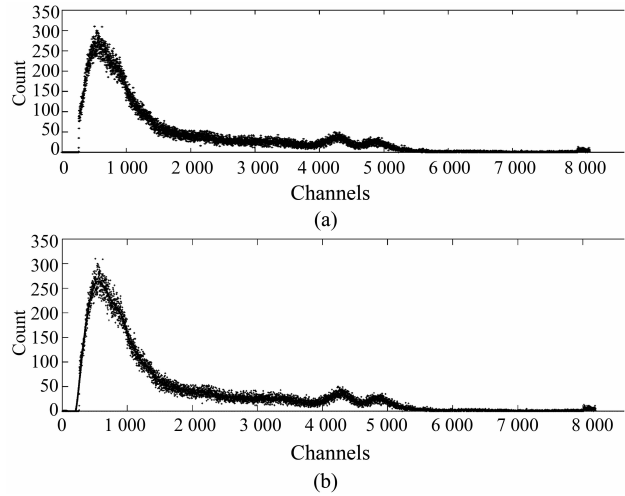


图 4 ^{60}Co 伽玛频谱数据的平滑

Fig. 4 Smoothing the gamma spectrum data for ^{60}Co

择逼近数据所需要的段数, 不需要重复试验以选取较好的分段段数。且对于不设容差的数据拟合, 引入最佳逼近曲线判定方法, 试验表明了该方法能有效的选择出最优逼近曲线。

参考文献:

- [1] 齐东旭. 关于多结点基函数型 δ -spline 插值 (I), (II), (III) [J]. 吉林大学学报: 自然科学版, 1975, 3: 70-81; 1976, 2: 36-44; 1979, 3: 1-8.
- [2] QI D X, LI H S. Many-knot spline technique for approximation of data [J]. Science in China (E), 1999, 42 (4): 383-387.
- [3] 齐东旭, 田自贤, 张玉心, 等. 样条函数在飞机理论外形设计中的应用 [J]. 飞机设计, 1982, 4: 45-51.
- [4] YAN W Q, QI D X. Many-knot spline interpolating curves and their applications in font design [J]. CADDM (Computer Aided Drafting, Design and Manufacturing), 1999, 9(1): 1-8.
- [5] 孙伟, 许君一, 齐东旭. 数据去冗余的多尺度多结点技术 [J]. 计算机辅助设计与图形学学报, 2004, 16 (5): 619-624.
- [6] 金涛, 阙沛文. 基于多结点样条理论的漏磁数据去冗余压缩算法 [J]. 上海交通大学学报, 2005, 39(4): 539-543.
- [7] 北方工业大学 CAD 研究中心与北京电视台. 计算机卡通片《咪咪钓鱼》[CP]. 北京: 北京电视台, 1991.
- [8] 北方工业大学 CAD 研究中心与北京科学教育电影制片厂. 计算机电影片《相似》[CP]. 北京: 北京科学教育电影制片厂广播电影电视部电影发行局, 1992.

(下转第 27 页)

$$\int_0^{\bar{a}} \int_{-\tau}^0 e^{\lambda\sigma} \beta_2(\sigma, y) u(y) d\sigma dy \cdot \int_0^{\bar{a}} \int_{-\tau}^0 e^{\lambda\sigma} \beta_2(\sigma, x) \int_0^x \frac{E_{1\lambda}(x) \beta_1(s)}{E_{1\lambda}(s) \gamma_1(s)} ds d\sigma dx \quad (19)$$

由于 $\int_0^{\bar{a}} \int_{-\tau}^0 e^{\lambda\sigma} \beta_2(\sigma, x) u(x) d\sigma dx \neq 0$ ，则由 (19) 式可得

$$1 = \int_0^{\bar{a}} \int_{-\tau}^0 e^{\lambda\sigma} \beta_2(\sigma, x) \int_0^x \frac{\beta_1(s)}{\gamma_1(s)} \cdot \exp\left\{-\int_s^x \frac{\lambda + \mu(s) + \rho_1(s) + \gamma'_1(s)}{\gamma_1(s)} ds\right\} ds d\sigma dx \quad (20)$$

令 $K(\lambda) = \int_0^{\bar{a}} \int_{-\tau}^0 e^{\lambda\sigma} \beta_2(\sigma, x) \int_0^x \frac{\beta_1(s)}{\gamma_1(s)} \cdot \exp\left\{-\int_s^x \frac{\lambda + \mu(s) + \rho_1(s) + \gamma'_1(s)}{\gamma_1(s)} ds\right\} ds d\sigma dx$ 函数 $K(\lambda)$ 在 \mathbb{R} 上是连续严格递增的，并且有 $\lim_{\lambda \rightarrow +\infty} K(\lambda) = 0$ 以及 $\lim_{\lambda \rightarrow -\infty} K(\lambda) = +\infty$ 。于是，方程 $K(\lambda) = 1$ 有唯一实数解 λ_0 。因此， \mathbb{A}^* 的谱不是空集。此引理得证。

由半群 $(\mathbb{T}(t))_{t \geq 0}$ 的正性和 $s(\mathbb{A}) > -\infty$ (见文 [8] 中的定理 VI. 1. 10)，以及引理 7 和引理 8，可得如下结论：

推论 1 $s(\mathbb{A}) \leq \sigma(\mathbb{A})$ 且 $\omega_{ess}(\mathbb{A}) < s(\mathbb{A})$ 。

由引理 7 和推论 1，可得存在一个 \mathfrak{E} 上的投影算子 \mathbb{P} 以及常数 $\varepsilon > 0$ 和 $M \geq 1$ 使得

$$\|e^{-s(\mathbb{A})t} \mathbb{T}(t) - \mathbb{P}\| \leq M e^{-\varepsilon t}, \text{ 对于所有的 } t \geq 0$$

其中 $\|\cdot\|$ 表示 \mathfrak{E} 上的算子范数。于是定理 2 得证。

参考文献：

[1] FARKAS J Z, HINOW P. On a size-structured two-phase population model with infinite states-at-birth[J]. Positivity, 2010, 14(3): 501–514.

[2] FARKAS J Z, GREEN D W, HINOW P. Semigroup analysis of structured parasite populations[J]. Mathematical Modelling of Natural Phenomena, 2010, 5: 94–114.

[3] AUSLANDER D M, OSTER G F, HUFFAKER C B. Dynamics of interacting populations[J]. J Franklin Inst, 1974, 297: 345–376.

[4] PIAZZERA S, TONETTO L. Asynchronous exponential growth for an age dependent population equation with delayed birth process[J]. J Evol Equ, 2005, 5: 61–77.

[5] PIAZZERA S. An age dependent population equation with delayed birth process[J]. Math Methods Appl Sci, 2004, 27: 427–439.

[6] BÁT KAI A, PIAZZERA S. Semigroups and linear partial differential equations with delay[J]. J Math Anal Appl, 2001, 264: 1–20.

[7] THIEME H R. Balanced exponential growth of operator semigroups[J]. J Math Anal Appl, 1998, 223: 30–49.

[8] ENGEL K J, NAGEL R. One-parameter semigroups for linear evolution equations [M]. New York: Springer, 2000.

[9] FARKAS J Z, HAGEN T. Stability and regularity results for a size-structured population model[J]. J Math Anal Appl, 2007, 328: 119–136.

[10] ARENDT W, GRABOSCH A, GREINER G, et al. One-parameter semigroups of positive operators [M]. Berlin: Springer-Verlag, 1986.

[11] CLÉMENT P, HEIJMANS H, ANGENENT S, et al. One-parameter semigroups [M]. Amsterdam: North-holland, 1987.

(上接第 21 页)

[9] DAHMEN W, GOODMAN T N T, MICCHELLI A. Compactly supported fundamental functions for spline interpolation[J]. Numerische Mathematik, 1988, 52(6): 639–664.

[10] RIEMENSCHNEIDER S D, SHEN Z W. General interpolation on the lattices: compactly supported fundamental solutions [J]. Numerische Mathematik, 1995, 70: 331–351.

[11] 余建德, 黄静. 基于多结点样条的自由曲线最小误差

逼近及其应用[C]//全国 15 届电脑辅助设计与图形学学术会议 CADCG2008, 2008.

[12] FORSEY D R, BARTELS R H. Surface fitting with hierarchical spline [J]. ACM Transactions on Graphics, 1995, 14(2): 134–161.

[13] 平劲松, 黄倩, 鄢建国, 等. 基于嫦娥一号卫星激光测高观测的月球地形模型 CLTM-s01 [J]. 中国科学, 2008, 53 (11): 1601–1612.