

二次参数曲线拟合平面点的一种新方法^{*})

郭凤华 杨兴强

(山东大学计算机科学与技术学院 济南250061)

摘要 对平面上给定的一组数据点进行了研究,提出了构造参数曲线拟合数据点的一种新方法。所构造的拟合参数曲线是 C^1 连续的分段二次参数曲线。本文以实例对新方法与二次插值样条曲线进行了比较。

关键词 参数曲线,样条曲线,拟合

A New Method for Fitting Planar Data Points with Quadratic Parametric Curve

GUO Feng-Hua YANG Xing-Qiang

(School of Computer Science and Technology, Shandong University, Jinan 250061)

Abstract A method of constructing parametric curve to fit a set of planar points is presented. The constructed fitting curve is piecewise quadratic parametric polynomial. Experiments for comparing the efficiency of the new method with quadratic spline curve are also included.

Keywords Parametric curve, Spline curve, Fitting

1 引言

给定平面上大量有序的数据点,这些点可以是某个形状上测量得到的,也可以是设计员给出的,要求构造一条曲线拟合这些数据点,这一直是CAGD的关键问题。拟合又分为逼近与插值。逼近是构造一条曲线使之在某种意义上最为接近给定的数据点。逼近比插值更困难。在插值问题里,控制顶点的数目由选取次数和数据点数目自动地确定,由数据点的参数化直接决定节点矢量,不存在误差问题。而在逼近问题里,曲线误差界 E 与要被拟合的数据点一起给出。因此,逼近一般是一个迭代过程^[1,2],逼近比插值开销更大。

插值是构造一条曲线顺序通过这些数据点。在数据点数目较多时,将导致一般的多项式插值曲线次数很高,而次数高往往会出现扭摆的问题,甚至出现插值曲线不收敛到被插曲线^[3]。作为分段样条插值曲线可以较好地解决这个问题^[4]。但是当数据量很大时,样条插值曲线分的段数太多。本文提出了构造参数曲线拟合数据点的新方法:采用多边形逼近数据点,多边形顶点把数据点分成数据点子集,在误差允许的范围每个子集内的数据点近似在一条直线上。每一个子集的数据点在最小二乘意义上由一段二次曲线拟合,二次曲线具有二次精度,可以拟合近似在一条直线上的数据点。全部数据点由在连接处 C^1 连续的分段二次参数曲线拟合,分段数目远远少于二次样条插值曲线。

2 参数曲线拟合

给定平面上大量有序的数据点 $p_i(x_i, y_i) (i=0, 1, \dots, m)$ (假定以 X 坐标排序)。采用多边形逼近数据点,多边形顶点把数据点分成数据点子集,在误差允许的范围每个子集内的数据点近似在一条直线上。具体方法如下:对于有序数据点集 $\mathcal{P}(p_0, \dots, p_p)$

1. \mathcal{P} 初始化一个栈
2. while 栈非空 do
 - 2.1 $\mathcal{P} \leftarrow$ 出栈
 - 2.2 if \mathcal{P} 内的点是直线,输出直线 $l[p_0, p_p]$
 - 2.3 else 点集 \mathcal{P} 根据 X 坐标范围一分为二
 - 2.4 新分的左右数据点集入栈。

子数据点集 $\mathcal{P}(p_0, \dots, p_p)$ 满足下面条件就声明为直线:

距离函数 $d(p_i, l[p_0, p_p]) < \epsilon \quad i=1, \dots, p-1, l[p_0, p_p]$

是过 p_0 和 p_p 的直线, ϵ 是给定的误差。

再假定全部数据点 $p_i(x_i, y_i) (i=0, 1, \dots, m)$ 分为 k 个子集,任一个子集数据点在最小二乘意义上由一段二次曲线拟合,全部数据点 $p_i(x_i, y_i) (i=0, 1, \dots, m)$ 由在连接处 C^1 连续的分段二次参数曲线拟合。具体方法如下:

取每个数据点子集的端点作为参数的节点给定参数的一个分割 $x_0=t_0 < t_1 < \dots < t_{k-1} < x_m=t_k$ 。构造分段的二次参数曲线 $S(t)$ 拟合每个数据点子集。 $S(t)$ 满足以下条件:

1) 每个小区间 $[t_{i-1/2}, t_{i+1/2}]$ 内, $S(t)$ 是二次多项式,记为 $S_i(t)$,其中 $t_{i-1/2} = (t_i + t_{i-1})/2$,称为半节点 $i=0, 1, \dots, k$;

2) 在整个区间 $[t_0, t_k]$ 上, $S(t)$ 是一阶连续可导,即在半节点 $t_{i-1/2}, i=0, 1, \dots, k$ 处成立

$$S_{i-1}(t_{i-1/2}) = S_i(t_{i-1/2}) \quad S'_{i-1}(t_{i-1/2}) = S'_i(t_{i-1/2}) \quad (1)$$

3) 在每个小区间 $[t_{i-1/2}, t_{i+1/2}]$ 内

$$S_i(t) = a_i(t-t_i)^2 + b_i(t-t_i) + y_i \quad (2)$$

其中 a_i, b_i, y_i 待求,这里记 $t_{-1/2} \equiv t_0, t_{k+1/2} \equiv t_k$

对于区间内的每个点 $p_j \in [t_{i-1/2}, t_{i+1/2}]$,根据最小二乘法,

$$(S_i(t_j) - a_i(t_j-t_i)^2 - b_i(t_j-t_i) - y_i)^2 = 0$$

$$\text{即 } S_i(t_j) - a_i(t_j-t_i)^2 - b_i(t_j-t_i) - y_i = 0$$

又因为 $S_i(t_j) = y_j$,所以

$$y_j - a_i(t_j-t_i)^2 - b_i(t_j-t_i) - y_i = 0$$

对区间上的所有点 $p_j \in [t_{i-1/2}, t_{i+1/2}], j \in [0, q]$:

^{*})本文受到国家自然科学基金(30470905)的资助。郭凤华 讲师,博士生,研究方向为计算机辅助几何设计和计算机图形学。杨兴强 副教授,研究方向为计算机辅助几何设计和计算机图形学。

$$\sum_{j=0}^q y_j - a_i \sum_{j=0}^q (t_j - t_i)^2 - b_i \sum_{j=0}^q (t_j - t_i) - (q+1)y_i = 0$$

整理得

$$y_i = (\sum_{j=0}^q y_j - a_i \sum_{j=0}^q (t_j - t_i)^2 - b_i \sum_{j=0}^q (t_j - t_i)) / (q+1)$$

令 $u_i = \sum_{j=0}^q y_j / (q+1)$

$$v_i = \sum_{j=0}^q (t_j - t_i)^2 / (q+1)$$

$$k_i = \sum_{j=0}^q (t_j - t_i) / (q+1)$$

则 $y_i = u_i - a_i v_i - b_i k_i$ (3)

把(3)代入(2)

$$S_i(t) = a_i(t-t_i)^2 + b_i(t-t_i) + u_i - a_i v_i - b_i k_i$$
 (4)

$$S_i'(t) = 2a_i(t-t_i) + b_i$$
 (5)

于是连续性条件(1)可以写成

$$\frac{a_i h_i^2}{4} - \frac{b_i h_i}{2} + u_i - a_i v_i - b_i k_i = \frac{a_{i-1} h_i^2}{4} + \frac{b_{i-1} h_i}{2} + u_{i-1} - a_{i-1} v_{i-1} - b_{i-1} k_{i-1} - a_i h_i + b_i = a_{i-1} h_i + b_{i-1}$$
 (6)

即 $b_i = (a_i + a_{i-1})h_i + b_{i-1}$ (7)

其中 $h_i = t_i - t_{i-1}$

把(7)代入(6)解得

$$b_{i-1} = ((a_i + 3a_{i-1})h_i^2 / 4 + u_{i-1} - u_i + a_i v_i - a_{i-1} v_{i-1} + (a_i + a_{i-1})h_i k_i) / l_i$$
 (8)

$$b_i = (a_i + a_{i-1})h_i + ((a_i + 3a_{i-1})h_i^2 / 4 + (u_{i-1} - u_i) + a_i v_i - a_{i-1} v_{i-1}) / l_{i+1}$$
 (9)

同理有

$$b_i = ((a_{i+1} + 3a_i)h_{i+1}^2 / 4 + u_i - u_{i+1} + a_{i+1} v_{i+1} - a_i v_i + (a_{i+1} + a_i)h_{i+1} k_{i+1}) / l_{i+1}$$
 (10)

其中 $l_i = k_{i-1} - k_i - h_i$

根据(9)=(10)得

$$(a_i + a_{i-1})h_i + ((a_i + 3a_{i-1})h_i^2 / 4 + (u_{i-1} - u_i) + a_i v_i - a_{i-1} v_{i-1} + (a_i + a_{i-1})h_i k_i) / l_i = ((a_{i+1} + 3a_i)h_{i+1}^2 / 4 + u_i - u_{i+1} + a_{i+1} v_{i+1} - a_i v_i + (a_{i+1} + a_i)h_{i+1} k_{i+1}) / l_{i+1}$$
 (11)

整理(11)得到

$$a_{i-1}(h_i l_i l_{i+1} + (\frac{3}{4}h_i^2 - v_{i-1} + k_i h_i)l_{i+1}) - a_i((\frac{3h_{i+1}^2}{4} - v_i + h_{i+1}k_{i+1})l_i - h_i l_i l_{i+1} - (\frac{h_i^2}{4} + v_i + k_i h_i)l_{i+1}) - a_{i+1}(\frac{h_{i+1}^2}{4} + v_{i+1} + h_{i+1}k_{i+1})l_i = (u_i - u_{i+1})l_i - (u_{i-1} - u_i)l_{i+1}$$

(上接第171页)

小结 本文基于 Gap 理论思想在多尺度灰度 Gap 统计模型的基础上,考虑了多尺度的加权邻域,使得 Gap 算子在图像边缘检测方法更加符合人对图像边缘的认识规律,并详细地分析了加权 Gap 算子的模型算法.分析并比较了 Gap 算子与 Sobel 边缘检测算子之间在边缘检测方面的相互关系和不同之处.

参考文献

- 1 Rudin L, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithm. In: The eleventh annual intl. conf. of the Center for Nonlinear Studies on Experimental mathematics, Los Alamos, 1992. 259~268
- 2 Meyer Y. Oscillating patterns in image processing and nonlinear evolution equations. Boston: American Mathematical Society, 2001
- 3 Richardson W H. Bayesian-based iterative method of image

仿照二次样条插值函数, 添上相应的边界条件: $a_0 = a_q = 0$; 构成完整的连续性方程, 使用追赶法求解 a_i , 由于方程组系数矩阵对角严格占优, 解是唯一存在的. 把 a_0, a_1 代入(8)式求得 b_0 , 由(7)式迭代求得 b_1 , 然后由(3)式求得 y_i , 最后把 a_i, b_i, y_i 代入(2)式求得 $S_i(t)$, 从而求得整条二次参数曲线.

3 实例

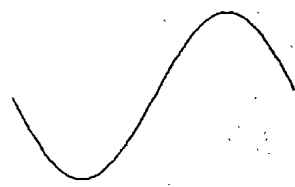


图1 一条正弦曲线

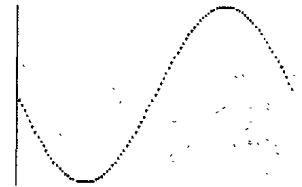


图2 将图1的曲线离散为100个数据点

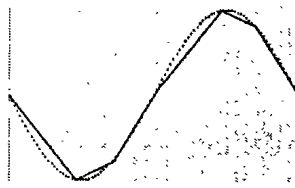


图3 在给定误差范围内用直线逼近, 将数据点分为6个子集



图4 新方法构造二次参数曲线

实例表明, 给定平面上100个数据点, 采用本文的方法划分成6个子集, 整条二次参数曲线由7段二次曲线组成, 拟合曲线逼近原始曲线. 如果采用二次样条曲线插值, 需要101段二次曲线. 本文的方法大大提高了效率, 并且拟合精度满足实际需要.

结论 大量数据点采用逼近拟合需要很大的开销; 构造样条插值曲线拟合, 分段数目太多. 采用本文的方法拟合数据点, 分段数目少, 拟合效率高.

参考文献

- 1 施法中. 计算机辅助几何设计与非均匀有理 B 样条. 高等教育出版社, 2001
- 2 Piegl L, Tiller W. The NURBS book. Second Edition. Berlin Heidelberg: Springer-Verlag, 1997
- 3 Yamaguchi F. Curves and surfaces in computer aided geometric design. Berlin: Springer-Verlag, 1988
- 4 Schoenberg I J. Contributions to the problem of application of equidistant data by analytic functions. Quart. Appl. Math., 1946
- restoration. Journal of Optical Society of America, 1972, 62(1): 55~59
- 4 Cross G, Jain A. Markov random field texture models. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1983, 5(1): 25~39
- 5 Poggio T, Voorness H, Yuille A L. A regularized solution to edge detection. MIT AI Laboratory: [Technical Report AIM833]. 1985
- 6 Canny J. A computational approach to early detection. IEEE Trans Pattern Analysis & Machine Intelligence, 1986, 8(6): 679~698
- 7 Shen J, Castan S. Fast filter transform theory and design for image processing. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Francisco, USA, 1985
- 8 Deriche R. Using Canny's criteria to derive a recursively implemented optimal edge detector. Int J Computer Vision, 1987, 6(1): 167~187
- 9 Sarkar S, Boyer K L. Optimal infinite impulse response zero crossing based edge detection. CVGIP: Image Understanding, 1991, 54(2): 224~243
- 10 Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a dataset via the Gap statistic. Palo Alto: Stanford University, SN(2000) JRSSB, 2000