

基于紧支撑径向基函数与共轭梯度法的大规模散乱数据快速曲面插值

于秋则 曹 矩 柳 健 田金文

(华中科技大学图像识别与人工智能研究所 图像信息处理与智能控制教育部重点实验室 武汉 430074)

摘 要: 该文提出一种快速大规模散乱数据的曲面插值算法。在此算法中, 首先采用紧支撑径向基函数(CSRBF)作为插值基函数, 采用 CSRBF 的优点是保证构成的系数方程组是对称正定而且系数是稀疏的。这样可保证系数方程组一定可解而且可以减少内存的开销。其次采用共轭梯度法求解大规模系数方程组。该算法在系数方程组的系数矩阵 $A: N \times N$ 是对称正定的情况下, 最多迭代 N 步就可以求得方程组的解, 实验结果表明该算法的快速性, 特别适合大规模散乱数据的曲面的插值。

关键词: 散乱数据曲面插值, 紧支撑径向基函数, 约束点, 系数方程组, 共轭梯度法

中图分类号: TN911.73, TP391.7

文献标识码: A

文章编号: 1009-5896(2005)02-0298-04

Rapid Surface Interpolation from Massive Scattered Data Using Compactly Supported Radial Basis Functions and Conjugate Gradient Method

Yu Qiu-ze Cao Ju Liu Jian Tian Jin-wen

(Institute for Pattern Recognition and Artificial Intelligence, Key Laboratory of Education Ministry for Image processing & Intelligence Control, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract A novel algorithm for rapid surface interpolation from massive scattered data using Compactly Supported Radial Basis Functions (CSRBF) and conjugate gradient method is presented in this paper. CSRBF is used because it can make the coefficient equations symmetric positive definite (spd), and very sparse. So there must be a solver and a small storage memory are needed. In solving the system equations, iterative method is used. The conjugate gradient method is used to solve the system equations, because the method converges in at most N steps for a symmetric positive definite N by N matrix. Experimental results using massive scattered points demonstrate the algorithm is fast. The proposed algorithm is very appropriate for surface interpolation from massive scattered points.

Key words Surface interpolation from scattered data, Compactly Supported Radial Basis Functions (CSRBF), Constraint points, Coefficient equations, Conjugate gradient method

1 引言

散乱点曲面插值(曲面重构)是指通过散乱(不规则分布)的数据采样点构造一个光滑的曲面, 其实际是由一些已知的数据来求解一些未知的数据。散乱点曲面插值在科学和工程领域有很重要的应用: 医学图像中, 散乱点曲面插值用于由 CT 或 MRI 图像构造人体器官的表面; 在计算机视觉中, 通过对散乱的测量点(立体视觉中的图像匹配点等)进行散乱点的插值构造可视曲面达到散乱数据点可视化的目的; 在图像变形中, 散乱点曲面插值可用于获得两图像之间的光滑映射函数^[1]。尽管很多学者在散乱点曲面插值研究中做了大量的工作^[2,3], 但是散乱点曲面插值依然是一个难度大且计算复杂度很大的工作。尽管在这个领域存在大量的算法, 很多方

法在光滑性, 计算速度方面存在缺陷, 或者对数据点的分布有所限制。Shepard 方法首先由气象学家与地质工作者提出, 后来由于 Shepard 的工作被称为 Shepard 方法^[4], 其基本思想是将插值函数定义为各数据点函数值的距离倒数加权和, 缺点是 C^0 连续, 而且当增加, 删除, 修改一个点时都必须重新计算插值, 因此是一种全局插值方法。有限元方法^[5]是首先对数据点集进行最优 Delaunay 三角划分, 然后在每一个三角形中分别采用最近邻或双线性或双立方插值得到整个表面的插值, 这种方法的优点是速度快, 缺点是不够光滑, 而且这种方法对数据分布比较敏感, 如: 出现狭长的三角形。而且落在三角形以外的点无法通过插值得到。后来有学者提出基于节点基函数方法; 这种方法把要重构的曲面看作基函数的线性组合, 未知的系数通过求解一个线性方程组得到。采用节

点基函数法的曲面插值主要有 2 个问题。(1) 节点基函数的选择。用的比较多的是 Gaussian multiquadratics^[6] 与薄板样条^[7], 薄板样条是通过求插值函数在曲面积分最小时得到, 薄板样条由于其良好的视觉效果和大数据的稳定性, 被广泛应用, 但是由于薄板样条基函数是全局的, 所以存在如下不足: 当采样数据点数很多时, 构成的系数矩阵很大而且几乎所有的系数都是非零, 因而矩阵的条件数很高, 求解起来非常困难。另外任何一个采样数据点都影响所有的插值点, 所以任意采样数据点的变化都必须重新求解方程组。(2) 构造并求解一个 $N \times N$ 线性方程组, 随着散乱数据点 (N) 的增加, 求解起来非常困难且运算量非常大, 甚至有时不可解。

本文提出一种快速大规模散乱数据的曲面插值的算法, 首先采用紧支撑径向基函数作为插值基函数, 采用紧支撑径向基函数的优点是保证构成的系数方程组是对称正定的^[8], 这样保证系数方程组可以求解且系数是稀疏的。其次采用共轭梯度法求解大规模方程组, 该算法在系数方程组的系数矩阵 $A: N \times N$ 是对称正定的情况下, 最多迭代 N 步就可以求得方程组的解, 实验结果表明该算法的快速性。

2 基于紧支撑径向基函数(CSRBF)的散乱数据插值

2.1 散乱数据曲面插值的数学定义

设已知一些散乱点(约束点) $(x_i, y_i), i=1, 2, \dots, N$, 散乱数据点曲面插值的目的是寻找一个光滑的插值函数 $f: R^d \rightarrow R$, 使得: $f(x_i) = y_i, i=1, 2, \dots, N$ 。基于节点基函数的方法是选取一个 n 维的线性空间 A , 使 f 可以由 A 空间的基函数 $g_i(x), i=1, 2, \dots, N$ 线性张成, 即

$$f(x) = \sum_{i=1}^N c_i g_i(x) \tag{1}$$

其中 $c_i, i=1, 2, \dots, N$ 是加权系数。

2.2 紧支撑径向基函数与基于紧支撑径向基函数的曲面插值

径向基函数被定义为以某点为中心, 以某个数值为半径的圆对称函数, Wendland^[9]为了解决最少度的径向基函数的对称正定问题, 定义紧支撑径向基函数如下:

$$\Phi(R) = \begin{cases} (1-R)^p P(R), & R < 1 \\ 0, & R \geq 1 \end{cases} \tag{2}$$

而且针对不同的连续性 C^k 及插值变量的维数 D 得到如下一些紧支撑径向基函数(见表 1)。特别的是: $D=3$ (插值数据是三维向量)时 $R = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2} / r_0$, r_0 是紧支撑径向基函数的实际半径 (即数据点离中心点的距离超过 r_0 时, 径向基函数的值为 0)。图 1 是几个中心都在 (10,

10), 不同半径的紧支撑径向基函数图。

表 1 不同的径向基函数

$D=1$	$(1-R)_+$	C^0
	$(1-R)_+^2(3R+1)$	C^2
	$(1-R)_+^5(8R^2+5R+1)$	C^4
$D=3$	$(1-R)_+^2$	C^0
	$(1-R)_+^4(4R+1)$	C^2
	$(1-R)_+^6(35R^2+18R+3)$	C^4
	$(1-R)_+^8(32R^3+25R^2+8R+1)$	C^6
$D=5$	$(1-R)_+^3$	C^0
	$(1-R)_+^5(R+1)$	C^2
	$(1-R)_+^7(16R^2+7R+1)$	C^4

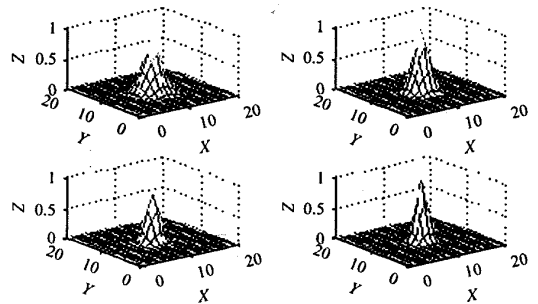


图 1 不同半径的紧支撑径向基函数图

基于紧支撑径向基函数的曲面插值是用紧支撑径向基函数作插值基函数^[9,10], 由式 (1), 可知

$$f(x) = \sum_{i=1}^N c_i \Phi(\|x-x_i\|) \tag{3}$$

为了求得加权系数, 必须满足约束点的值为已知点的函数值, 即

$$f(x_i) = \sum_{j=1}^N c_j \Phi(\|x_i-x_j\|) = y_i, \quad i=1, 2, \dots, N \tag{4}$$

那么可以构成如下的系数方程组:

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} & \dots & \Phi_{1N} \\ \Phi_{21} & \Phi_{22} & \dots & \Phi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{N1} & \Phi_{N2} & \dots & \Phi_{NN} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \tag{5}$$

其中 $\Phi_j = \Phi\|x_i-x_j\|$ 。

令 $A = [\Phi_{ij}], i, j=1, 2, \dots, N, X = [c_1, c_2, \dots, c_N]^T, b = [y_1, y_2, \dots, y_N]^T$, 那么式 (5) 可以简化为 $AX = b$ 。

由于采用了紧支撑径向基函数, 因此系数方程组是一个稀疏的方程组, 在方程组的建立上(系数的求解), 本文采用了文献[9]中的 K-d(K-dimension)算法。在求解方程组时, 通常采用系数消元方法, 针对稀疏矩阵可以用文献[9]中的 LU 分

速增大,其运算复杂度是 $O(N^{3/2})$ 。因此必须研究一种快速算法求解大规模方程组。

3 基于共轭梯度法的系数方程组迭代求解

由于采用紧支撑径向基函数作为插值基函数,系数方程组是一个稀疏的方程组。同时,该系数方程组又是对称正定的,这就保证了方程是有唯一解,因而可以采用共轭梯度方法来求解这个系数方程组,共轭梯度法是介于最速下降法与牛顿法之间的一个方法,它只需要一阶导数信息,但克服了最速下降法收敛慢的缺点^[11,12]。

设方程组为

$$AX = b \quad (6)$$

其中 A 为对称正定矩阵, X 是解向量。定义如下目标函数:

$$\varphi(X) = \frac{1}{2} X^T AX - X^T b + \frac{1}{2} b^T A^{-1} b = \frac{1}{2} (AX - b)^T A^{-1} (AX - b) \quad (7)$$

根据文献[12]:如果 A 为对称正定矩阵,则 φ 一定有唯一的最小值,且最小值的位置在

$$0 = \text{grad}(\varphi(X)) = AX - b = -r(X) \quad (8)$$

其中 $r(X)$ 是余量函数。因此方程组的求解转化为目标函数的求优,优化的迭代方程如下:

$$X_k = X_{k-1} + \alpha_k P_k \quad (9)$$

其中 X_{k-1} 是当前对解的一个逼近, P_k 是搜索方向, α_k 是搜索步长。把迭代方程式(9)代入目标函数式(7),得到

$$\begin{aligned} \varphi(X_k) = & \frac{1}{2} X_{k-1}^T AX_{k-1} + \alpha_k P_k^T AX_{k-1} \\ & + \alpha_k^2 P_k^T AP_k - X_{k-1}^T b - \alpha_k P_k^T b + \frac{1}{2} b^T A^{-1} b \end{aligned} \quad (10)$$

上式是关于 α_k 的正定二次多项式,它具有唯一的最小值。

对上式求导数可得到下式:

$$\begin{aligned} \partial\varphi(X_k)/\partial\alpha_k = & P_k^T (AX_{k-1} - b) + \alpha_k P_k^T AP_k \\ = & -P_k^T r_{k-1} + \alpha_k P_k^T AP_k = 0 \end{aligned} \quad (11)$$

那么,可以得到,搜索步长的迭代公式为

$$\alpha_k = P_k^T r_{k-1} / P_k^T AP_k \quad (12)$$

而且可以得到计算误差 $r(X)$ 迭代公式,如下:

$$r_k = b - AX_k = b - AX_{k-1} - \alpha_k AP_k = r_{k-1} - \alpha_k AP_k \quad (13)$$

由文献[9],我们可以知道,优化方向的迭代计算公式如下:

$$P_k = r_{k-1} + \beta_{k-1} P_{k-1} \quad (14)$$

其中

$$\beta_{k-1} = r_{k-1}^T r_{k-1} / r_{k-2}^T r_{k-2} \quad (15)$$

因此可以得到如下的共轭梯度算法:

PROGRAM CG ()

Begin

Step 1 设置初始解 X_0 ,并设置初始的方向 $P_1 = r_0 = b - AX_0$ 并设置收敛的阈值为 ε

Step 2 If ($\|r_0\|^2 < \varepsilon$), Exit;

Step 3 for $k=1, 2, \dots, D_0$

■ $\alpha_k = P_k^T r_{k-1} / P_k^T AP_k$

■ $X_k = X_{k-1} + \alpha_k P_k$

■ $r_k = r_{k-1} - \alpha_k AP_k$

■ If ($\|r_k\|^2 < \varepsilon$), Exit;

■ $\beta_k = r_k^T r_k / r_{k-1}^T r_{k-1}, P_{k+1} = r_k + \beta_k P_k$

End for

End

由文献[9],可以证明对于对称正定矩阵 $A: N \times N$,共轭梯度法最多迭代 N 次一定可以找到最优解,即运算复杂性为 $O(N)$,因此该方法是很快的。表2是用紧支撑径向基函数与薄板样条的计算复杂性及内存开销的比较表。

表2 采用紧支撑径向基函数与典型的薄板样条作为插值基函数的性能比较表

	薄板样条(Thin-plate)	紧支撑径向基函数(CRBF)
建立方程的运算量	$O(N^2)$	$O(N \times \lg N)$
解方程所需的运算量	$O(N^2)$	$O(N)$
所需要的内存	$O(N^2)$	$O(N)$ (稀疏矩阵存储方式)
单点误差的影响	全局	局部

4 实验结果

基于紧支撑径向基函数与共轭梯度法的大规模散乱数据曲面插值算法的快速性在很多实际散乱数据的曲面插值中得到验证。限于篇幅,这里只列举两个实验说明。实验1是激光测距仪获得的距离数据的曲面插值(重建),该距离数据来源于STANFORD大学的机器人研究所,共6012个数据点。实验中采用的紧支撑径向基函数为 $\Phi(R) = (1-R)_+$,共轭梯度的收敛阈值 $\varepsilon = 0.01$,实验结果见图2。图2(a)是激光扫描的距离数据(图中为了显示清楚进行了抽样使数据稀疏);图2(b)是采用径向基函数的插值结果;图2(c)是插值结果的光照显示。实验中用的PC机配置为P4:1.4G,256M内存,W2000操作系统,编程语言为VC。整个插值运算时间为13560ms(约14s)。而采用薄板样条作为插值基函数的插值时间为912024ms(约15min)。实验2中采用机载激光成像雷达扫描的地形数据(地形的一部分)作为散乱数据点,总共为

达扫描的地形数据(地形的一部分)作为散乱数据点, 总共为 10000 数据点, 实验采用的基函数同样为: $\Phi(R) = (1 - R)_+^2$, 共轭梯度中收敛阈值与实验 1 相同 ($\epsilon = 0.01$), 插值实验的结果见图 3。图 3(a) 是散乱数据点的分布图 (为了显示清楚, 同样进行了抽样使数据稀疏); 图 3(b) 是插值的结果。实验的 PC 机配置与实验 1 相同。插值所用时间为 25124ms (约 26s)。而采用薄板样条作为插值基函数的插值时间为 3.23×10^6 ms (约 53min)。为了研究本插值算法所用的时间与数据点数规模的关系, 做了不同数目数据点情况下的插值实验, 实验结果曲线见图 4。从时间曲线可以看出, 基于紧支撑的径向基函数的插值算法, 运算速度是很快的。

参 考 文 献

[1] Seungyong Lee, George wolberg, Sung Yong Shin. Scattered data Interpolation with multilevel B-splines. *IEEE Trans. on Visualization and Computer Graphics*, 1998, 3(3): 228 - 244.

[2] Franke R, Nielson G M. Scattered data interpolation and applications: A Tutorial and Survey, Geometric Modeling: Methods and Their Application, Hagen H and Roller D, eds., Berlin: Springer-Verlag, 1991: 131 - 160.

[3] Shapiro V. Real functions for representation of rigid solids. *Computer Aided Geometric Design*, 1994, 11(2): 153 - 175.

[4] Shepard D. A two dimensional interpolation function for irregularly spaced data. *Proc. ACM 23rd Nat'l Conf.*, New York, 1968: 517 - 524.

[5] Clough R, Tocher J. Finite element stiffness matrices for analysis of plates in bending. *Proceeding Conference Matrix Methods in Structural Mechanics*, Wright-Patterson Air Force Base, 1965: 515 - 545.

[6] Hardy R. Multiquadratic equations of topography and other irregular surfaces. *J. Geophysical Research*, 1971, 76(8): 1905 - 1915.

[7] Dyn N. Interpolation and approximation by radial and related function. Chui C, Schumaker L, Ward J, et al.. *Approximation Theory VI*, San Diego, Calif.: Academic Press, 1989: 211 - 234.

[8] Wend H. Piecewise polynomial positive definite and compactly supported radial functions of minimal degree. *AICM*, 1995, 4: 389 - 396.

[9] Morse B S. Interpolating implicit surfaces from scattered surface data using compactly supported radial basis functions, *Shape modeling Conference, Proc. SMI*, Geneva, Italy, May 2001: 89 - 98.

[10] Nikita Kojekine, Ichiro Hagiwara, Savchenko V V. Software tools using CSRBFs for processing scattered data. *Computers & Graphics*, 2003, 27(2): 311 - 319.

[11] 袁亚湘, 孙文瑜. 最优化理论与方法. 北京: 科学出版社, 1997: 183 - 199.

[12] 关治, 陈景良. 数值计算方法. 北京: 清华大学出版社, 1990: 423 - 430.

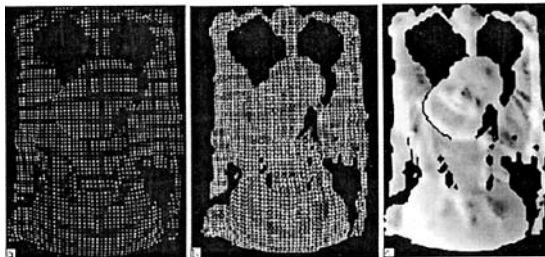


图 2 插值实验 1

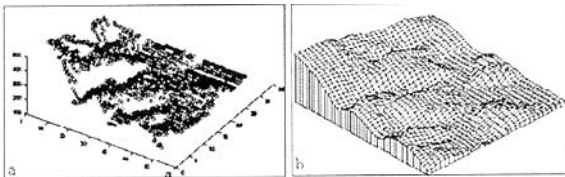


图 3 插值实验 2

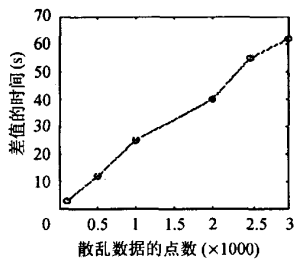


图 4 插值时间与散乱数据点数之间的关系曲线

5 结 论

本文提出一种基于紧支撑径向基函数的散乱点数据曲面插值的算法, 利用径向基函数的局部性, 使系数方程组对称正定而且稀疏, 利用共轭梯度法来快速求解大规模方程组, 实验结果表明, 算法速度快, 适合大规模散乱数据的曲面插值。

于秋则: 男, 1974 年生, 博士生, 研究方向为图像处理与模式识别、SAR 雷达图像信息处理、图像与视频压缩、地形重建与地形匹配。

曹 矩: 男, 1966 年生, 博士生, 研究方向为图像处理与模式识别。

田金文: 男, 1961 年生, 教授, 博士生导师, 研究方向为小波理