

# 空间分割与曲率相融合的点云精简算法研究\*

葛源坤, 黎荣, 李海伦

(西南交通大学机械工程学院先进设计制造技术研究所, 成都 610031)

**摘要:** 非接触式扫描方法获得点云数据存在大量的冗余数据。为便于模型重构, 针对点云数据精简是必不可少的数据预处理手段, 提出了一种基于空间分割和曲率特征信息的点云数据精简算法。通过 K-邻域计算、二次曲面拟合、曲率估算和曲率阈值可调的数据分区等关键精简技术, 实现了对同一数据不同区域应用不同精简算法, 进行不同比例的数据精简。实验验证表明, 该算法能适应各种类型曲面数据的精简要求, 保证精简效率的同时, 很好地保留点云的特征信息。

**关键词:** K-邻域; 曲率; 空间分割; 最小距离; 包围盒

**中图分类号:** TP399      **文献标志码:** A      **文章编号:** 1001-3695(2012)05-1997-04

**doi:**10.3969/j.issn.1001-3695.2012.05.106

## Study of point cloud data reduction algorithm integrating space partition and curvature

GE Yuan-kun, LI Rong, LI Hai-lun

(Institute of Advanced Design & Manufacturing, School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

**Abstract:** There are huge amounts of redundant data in point cloud data obtained by non-contact scanning. In order to realize model reconstruction effectively, point cloud data reduction is an indispensable means of pre-processing means. This paper presented an approach of point cloud data reduction based on space partition and curvature. Through some key technologies, such as K-neighborhood search, second surface fitting, curvature estimation, and data partition by of controllable curvature threshold, it applied the different reduction algorithms in different regions of the same point cloud data, meanwhile, achieved realizable reduction proportions. So, the algorithm can ensure reduction efficiency and retain characteristic information of the point cloud data simultaneously.

**Key words:** K-neighbors; curvature; space partition; minimum distance; bounding box

逆向工程被广泛应用于汽车、航空、模具等制造领域, 是产品创新设计的重要手段之一。逆向工程中, 获取实物表面几何信息的方法分为接触式和非接触式两类。由于非接触式测量数据获取速度快、灵活性高, 单次扫描能获上万个数据点, 随着其测量精度的日益提高, 非接触式测量的应用愈加广泛。由于高精度的非接触式扫描所获取点云数据量非常庞大, 且存在数据冗余, 如一般汽车覆盖件的数据量达上百万个点, 若不进行必要的精简预处理, 则曲面重构的过程不但占用大量的计算处理时间, 也会影响所构曲面的质量, 甚至会由于存在噪声点而无法实现模型重建。因而, 在保证精度的前提下有效地对点云数据进行精简是逆向工程中的关键技术之一。

传统的数据精简方法分为基于空间分割和基于曲率两大类。例如, 基于空间分割方法有包围盒法<sup>[1]</sup>、均匀方格法<sup>[2]</sup>、基于三角网格的网格精简法<sup>[3]</sup>; 根据曲率方法有最小距离法<sup>[4]</sup>、角度偏差法<sup>[5]</sup>等。两类算法各有局限, 基于空间分割的精简算法不适应具有复杂特征和多样曲率的高密度散乱点云的数据精简<sup>[6]</sup>, 基于曲率的精简算法虽能很好地保留几何特征, 但精简效率低<sup>[7]</sup>。因此, 在适应各类曲面数据精简要求, 保留必要点云特征信息的同时, 获得尽可能高的精简效率是数据精简方法有效性最重要的衡量条件。

### 1 算法总体思路

本文提出基于空间分割精简算法与基于曲率精简算法有效融合的数据精简方法, 该方法根据曲率阈值对不同类型曲面的点云应用相适应数据进行精简。算法总体思路为: 通过 K-邻域计算、曲率估算, 获得点云曲率, 信息; 设定可调的曲率阈值, 依据阈值将同一数据源的数据点区分为平坦和突变区域, 在平坦区域应用基于空间分割的包围盒法完成数据精简, 保证整个方法的精简效率; 在突变区域应用最小距离法完成数据精简, 保证点云的几何特征信息不丢失, 从而实现针对不同类型曲面的要求, 使数据精简保留必要特征信息的同时, 具有较高的运算效率。算法的具体流程如图 1 所示。

算法主要分为五步:

- 点云数据读取。读入点云数据, 把点云数据空间分成许多大小相同的立方体子空间, 把各数据点归入到相应的子空间内, 保存子空间数组。
- K-邻域计算。计算子空间内点之间的相对距离, 对每个候选点进行邻个最近邻域的搜索, 计算出所有点的 K-邻域, 保存于点结构的 K-邻域数组中。
- 曲率估算。基于 K-邻域内的数据点采用线性最小二乘

**收稿日期:** 2011-09-26; **修回日期:** 2011-10-29      **基金项目:** 中央高校基本科研业务费专项资金资助(SWJTU09ZT06)

**作者简介:** 葛源坤(1985-)男, 山东人, 硕士研究生, 主要研究方向为数字化设计制造(geyunkun@126.com); 黎荣(1974-), 女, 重庆人, 副所长, 副教授, 主要研究方向为数控技术、虚拟样机、快速制造技术; 李海伦(1986-), 女, 辽宁人, 硕士研究生, 主要研究方向为数字化设计制造。

法拟合一张微切平面,通过计算该平面的法矢将其估算为候选点的法矢;然后,以候选点为坐标原点建立局部坐标系,将 K-邻域内的数据点集用最小二乘法拟合为一个抛物面,根据拟合的抛物面完成曲率估算。

d) 确定曲率阈值,完成数据精简。取估算曲率的平均值作为点云数据分区的阈值判据,将点云数据分为曲率较高区域——突变区域和曲率较低区域——平坦区域,在突变区域采用最小距离法精简数据,在平坦区域采用包围盒法精简数据。

e) 精简结果验证。判断点云数据精简结果,如果曲率平坦区域精简效果不好则调整包围盒的大小;如果曲率突变区域精简效果不好则修改最小距离值,完成数据精简。

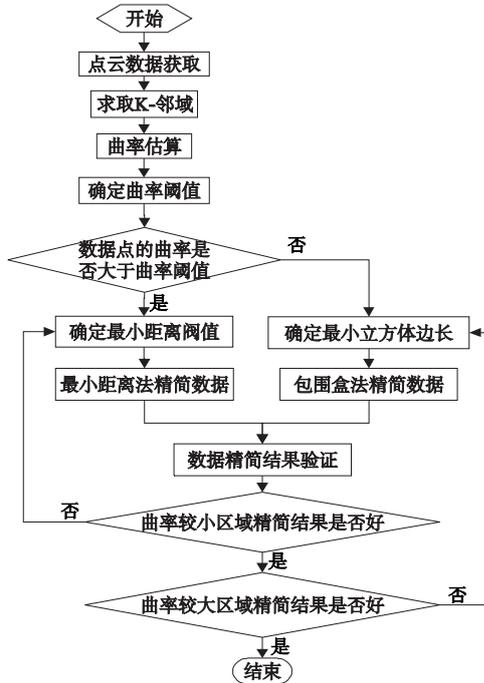


图1 算法流程

## 2 算法关键技术

### 2.1 K-邻域计算

K-邻域的计算步骤包括点云子空间划分、K-邻域计算两个方面。

#### 2.1.1 点云子空间的划分

散乱点云数据中,候选点  $p_i$  的 K-邻域是指点集中与  $p_i$  直线距离最短的  $k$  个数据点  $\{(x_i, y_i, z_i)\}_{i=1}^k$ , 称为点  $p_i$  的 K-邻域, 记为  $N(p)$ 。当  $k = 24 - 32$  时, 拟合抛物面会获取较高精度, 如果数值过小则不能充分反映点云数据局部范围内的几何特性, 从而降低法矢和曲率估算精度; 过多的邻域数据量会加大曲率估算的计算量。本文算法中  $k$  的取值为 32。

K-邻域计算步骤如下:

a) 计算包含所有数据的最小长方体空间;

b) 估算立方体子空间的边长  $L$ ; 最后把数据点集的最小长方体空间分成  $n_{\text{cube}}$  个立方体, 则

$$n_{\text{cube}} = \lceil (x_{\max} - x_{\min}) / L \rceil \lceil (y_{\max} - y_{\min}) / L \rceil \lceil (z_{\max} - z_{\min}) / L \rceil \quad (1)$$

其中,  $\lceil \cdot \rceil$  表示向上取整。假设每个立方体子空间中平均数据点的个数是  $K$  的函数, 则有  $n/n_{\text{cube}} = ak$ , 其中  $a$  是一个标量。为便于计算, 把式(1)代入式(2), 整理后可得到边长  $L$  的计算式<sup>[8]</sup>:

$$L = \beta \cdot \sqrt[3]{\frac{k}{n} (x_{\max} - x_{\min}) (y_{\max} - y_{\min}) (z_{\max} - z_{\min})} \quad (2)$$

其中:  $\beta = \sqrt[3]{a}$ , 用于调节立方体边长的大小,  $\beta$  的最优值为  $0.8 - 1.2$ <sup>[8]</sup>;  $n$  为点云数据总数。立方体边长  $L$  确定后, 最小长方体空间在  $x, y, z$  方向上的分辨率也就确定了。

c) 完成子立方体划分, 把点云数据的序号归入到相应的立方体子空间中。

#### 2.1.2 K-邻域的搜索

利用上面得到的立方体子空间结构, 对每个候选点  $p_i$  搜索并计算 K-邻域的步骤, 如图 2 所示。

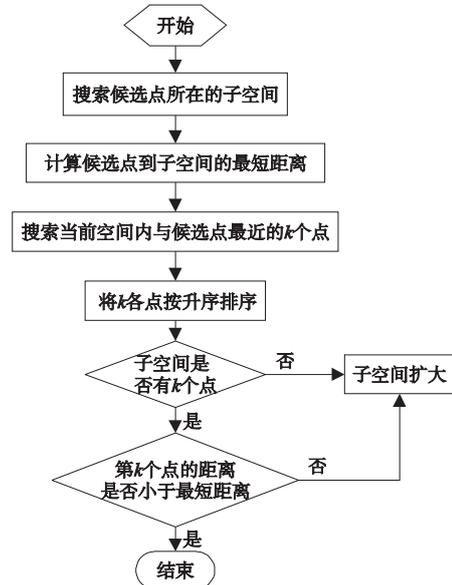


图2 K-邻域搜索流程

a) 根据候选点  $p_i$  的三维坐标值, 找到其所在的立方体子空间。

b) 计算候选点到当前子空间环六壁的最短距离  $d_{\text{short}}$ 。

c) 在当前子空间环内搜索与候选点最近的  $k$  个点, 按距离升序的方式, 记录  $k$  个邻近点与候选点间的距离及其序号, 同时对访问过的立方体子空间进行标记。

d) 在当前子空间环内, 如果候选点的  $k$  个最近邻域已找到, 并且候选点到  $k$  个最近点的距离小于  $d_{\text{short}}$ , 则候选点的  $k$  个最近邻域搜索结束, 记录它们的序号, 并对已标记的立方体子空间进行复位, 以便于下一个候选点的  $k$  个最近邻域搜索; 否则, 子空间环向与候选点距离最近的壁方向扩展  $L$ , 进行判断。若不满足条件则子空间环向次最近的壁扩展  $L$  继续搜索, 直至满足条件。最后遍历所有的点, 完成 K-邻域搜索。

### 2.2 曲率估算

算法根据曲率信息作为阈值实现点云分区, 因此需对曲率进行估算。曲率估算算法主要包括法矢估算、抛物面拟合和曲率估算三步。

#### 2.2.1 法矢估算

应用线性最小二乘法来拟合一张平面(微切平面), 候选点的法矢由该平面的法矢估算来确定。拟合最小二乘法的方法为: 首先由散乱点集  $p_i(0, 1, \dots, n)$  求一张平面  $S(u_i, v_i)$ , 使数据点集  $p_i(i=0, l, \dots, n)$  中的所有点到该平面距离的平方和最小:

$$\sum_{i=0}^n \|S(u_i, v_i) - p_i\|_2 \rightarrow \min \quad (3)$$

平面方程的一般形式为

$$ax + by + cz + d = 0 \quad (4)$$

则平面法矢可表示为  $\vec{N}(a, b, c)$ , 由平面方程可知参数  $a, b, c, d$  为线性表示。当求得平面参数  $(a, b, c, d)$  后, 由数据点集  $\{p_i(x_i, y_i, z_i), i = 1, \dots, n\}$  拟合的最小二乘平面就已完成了。由于参数  $a, b, c, d$  不是相对独立的, 大小不是绝对的, 所以在求解最小二乘平面法矢时必须对矢量进行单位化处理。至此, 初步估算的法矢已经确定。法矢确定后, 应用最小生成树法对法矢进行修正。若法矢不进行修正, 法矢方向错乱, 影响曲率估算的正确性。

### 2.2.2 抛物面拟合

建立局部坐标系  $(p_i, u, v, h)$ : 设候选点  $p_i$  为局部坐标系原点, 把  $p_i$  点的法矢设为坐标系的  $h$  轴,  $u$  和  $v$  坐标轴可以由过原点垂直于法矢的平面  $T$  中任意设定。由于在第一步中法矢已经求得, 所以局部坐标系可以建立。将候选点  $p_i$  的  $K$ -邻域转换到局部坐标系下得数据点集  $\{(x_i, y_i, z_i)\}_{i=1}^k$ , 并对坐标转换后的点集数据应用  $S(u, v) = au^2 + buv + cv^2$  拟合抛物面。拟合抛物面的目标函数为

$$Q^2 = \min \sum_{j=1}^k [h_j - (au_j^2 + bu_jv_j + cv_j^2)]^2 \quad (5)$$

应用奇异值分解法可获得拟合抛物面的最小二乘解:

$$\begin{cases} \frac{\partial Q^2}{\partial a} = \sum_j 2u_j^2 (au_j^2 + bu_jv_j + cv_j^2 - h_j) = 0 \\ \frac{\partial Q^2}{\partial b} = \sum_j 2u_jv_j (au_j^2 + bu_jv_j + cv_j^2 - h_j) = 0 \\ \frac{\partial Q^2}{\partial c} = \sum_j 2v_j^2 (au_j^2 + bu_jv_j + cv_j^2 - h_j) = 0 \end{cases} \quad (6)$$

### 2.2.3 曲率估算

计算抛物面  $S$  在原点处的高斯曲率。拟合抛物面  $S(u, v) = au^2 + buv + cv^2$  在局部坐标系原点的高斯曲率为

$$K = (LN - M^2) / (EG - F^2) = 4ac - b^2 \quad (7)$$

### 2.3 点云数据精简

点云数据精简算法的基本思路是在突变区域采用基于曲率的最小距离法精简, 很好地保留了数据的曲率等特征信息; 在平坦区域采用基于空间分割的包围盒法精简数据, 以确保算法有较快的运算速度和较高的精简效率。

#### 2.3.1 数据分区

在曲率估算完成后, 求取曲率绝对值的平均值:

$$\bar{k} = \frac{1}{n} \sum_{i=1}^n |k_i|$$

其中:  $n$  为点云数据总数,  $k_i$  为每点的高斯曲率。初步确定曲率阈值为  $\bar{k}$ , 把大于  $\bar{k}$  的区域作为突变区域, 而小于  $\bar{k}$  的区域作为平坦区域。由于不同模型点云数据中平坦区域和突变区域所占的比例不同, 所以要对初步的法矢进行修正。修正后的曲率阈值为  $k' = \alpha \bar{k}$ ,  $\alpha$  根据数据模型表面特征取值, 取值为 0.6 ~ 1, 用修正后的曲率阈值对点云数据重新分区。

#### 2.3.2 数据精简

对平坦区域的点数据应用包围盒法<sup>[9]</sup>精简数据。首先利用  $K$ -邻域的计算方法将平坦区域的点数据归入到边长  $L = 1$  mm 的子立方体中。计算子立方体空间的点  $p_i$  到其中心  $O$  的距离  $d_i$ 。设  $O(x, y, z)$  为子空间的中心,  $X_1(x_1, y_1, z_1), X_2(x_2, y_2, z_2)$  为子空间的体对角线上的两顶点, 则子空间的中心坐标可表示为

$$x = \frac{x_1 + x_2}{2}, y = \frac{y_1 + y_2}{2}, z = \frac{z_1 + z_2}{2} \quad (8)$$

即  $O(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2})$ , 子空间内任意一点  $p_i(x_i, y_i, z_i)$  到中心的距离  $d_i$  为

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} \quad (9)$$

将式(8)代入(9)即可求出  $d_i$ 。对求得的  $d_i$  进行比较, 找到最小距离  $d_{\min}$ , 保留  $d_{\min}$  对应的点, 子立方体内其他点删除。依次遍历所有平坦区域内的子立方体, 完成所有包围盒的数据精简。

对突变区域的点数据应用最小距离法精简。最小距离精简原理是: 首先设定一个两点间的最小距离  $d_{\min}$ ; 后在  $K$ -邻域内所有点之间的距离  $d_i$  与  $d_{\min}$  进行比较, 若  $d_i < d_{\min}$ , 则两点中的一个点将被删除, 否则保留两个点; 依次判断所有突变区域的数据点, 完成数据精简。

### 3 实例应用

为了评估算法的实用效果, 应用 Stanford 大学 Bunny 测试模型数据集和一汽车覆盖件两个点云数据集来验证本文的算法。Bunny 模型曲率变化复杂, 原始点云数据有 294 408 个点, 如图 3 所示。汽车覆盖件平面区域所占比例较大, 但需保证钣金弯曲的倒圆角和冲孔圆柱面等特征信息, 点集原始数据共有 134 913 个点。

图 4 为 Bunny 数据集单纯采用包围盒法精简后的数据, 图 5 为该数据集单纯采用最小距离法精简后的数据; 图 6 为本算法精简后的数据。其中包围盒子空间参数与包围盒法所用参数相同, 最小距离控制值与最小距离法完全一致。三种方法精简率比较如表 1 所示。

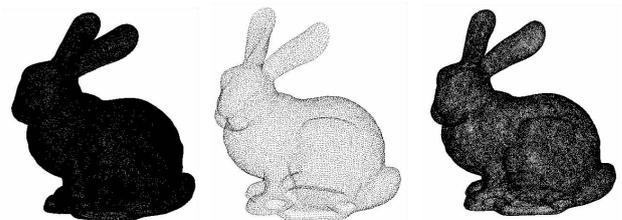


图3 精简前的数据

图4 包围盒法精简后的数据

图5 最小距离法精简后的数据

表 1 Bunny 数据集精简率比较

模型	点数	精简率/%
原始模型	294 408	-
包围盒法	18 904	93.6
最小距离法	283 515	3.7
本算法	48 087	83.7

比较分析图 4~6 可以得出: 包围盒精简的数据量比较大, 但是数据集的表面特征丢失比较严重, 除基本轮廓外, 在眼睛、脚趾和毛发等曲率变化较大的区域, 特征完全丢失; 最小距离法虽然能很好地保留数据的表面特征, 但是精简的数据量少、精简效率低; 采用本算法在 Bunny 背部、腹部等较平坦区域精简了大量数据, 又保留了眼睛、脚趾和毛发等曲率变化大的区域的特征信息。

图 8 为汽车覆盖件单纯采用包围盒法精简后的数据; 图 9 为该数据集单纯采用最小距离法精简后的数据; 图 10 为本算法精简后的数据, 精简率比较如表 2 所示。

汽车覆盖件精简结果与 Bunny 数据集结果一致。由于本算法中曲率阈值、包围盒大小、最小距离值的大小均可调整, 因

此可以灵活方便地控制算法的精简精度和速度,因而算法对表面特征复杂的点云数据和表面曲率变化较小的曲面都有很好的适应性。

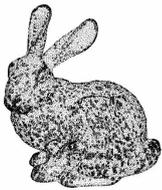


图6 本算法精简后的结果

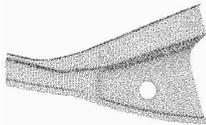


图7 精简前的数据

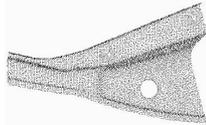


图8 包围盒法精简后的数据

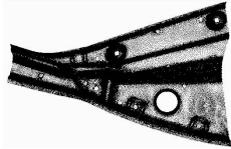


图9 最小距离法精简后的数据

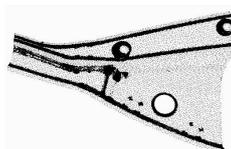


图10 本算法精简后的数据

表2 汽车覆盖件数据集精简率比较

模型	点数	精简率/%
原始模型	134 913	-
包围盒法	9 634	92.9
最小距离法	133 522	1.0
本算法	63 932	52.6

## 4 结束语

散乱点云数据精简算法中基于空间分割的精简算法具有较高的精简效率,但不适应具有复杂特征和多样曲率的高密度散乱点云的数据,基于曲率的精简算法虽能很好地保留几何特

征,但精简效率较低。本算法将两者有效融合,通过可调整的曲率阈值灵活控制平坦和突变区域数据点的数量,从而使算法具有较好的精简率、计算效率,而且能很好地保留零件的表面特征。实例验证表明,本文算法对表面特征复杂和表面曲率变化较简单的散乱点云数据都有很好的适应性。

## 参考文献:

- [1] SUN W, BRADLEY C, Zhang Y F, *et al.* Cloud data modeling employing a unified non-redundant triangular mesh [J]. *Computer Aided Design*, 2001, 33(2): 183-193.
- [2] MARTIN R R, STROUD I A, MARSHAL A D. Data reduction for reverse engineering, No. 1068 [R]. [S. l.]: Computer and Automation Institute of Hungarian Academy of Science, 1996: 63-69.
- [3] CHEN Y H, NG C T, WANG Y Z. Data reduction in integrated reverse engineering and rapid prototyping [J]. *International Journal of Computer Integrated Manufacturing*, 1999, 12(2): 97-103.
- [4] 刘德平, 陈建军. 逆向工程中数据精简技术的研究 [J]. *西安电子科技大学学报*, 2008, 35(2): 334-339.
- [5] LEE K H, WOO H, SUK T. Point data reduction using 3D grids [J]. *The International Journal of Advanced Manufacturing Technology*, 2001, 18(3): 201-210.
- [6] 周煜, 张万兵, 杜发荣, 等. 散乱点云数据的曲率精简算法 [J]. *北京理工大学学报*, 2010, 30(7): 785-789.
- [7] 孙殿柱, 朱昌志, 范志先, 等. 基于型面特征的三维散乱点云精简算法 [J]. *中国机械工程*, 2009, 20(23): 2840-2843.
- [8] 熊邦书, 何明一, 俞华琛. 三维散乱数据的  $k$  个最近邻域快速搜索算法 [J]. *计算机辅助设计与图形学学报*, 2004, 16(7): 909-917.
- [9] FILIP D, MAGEDSON R, MARKOT R. Surface algorithms using bounds on derivatives [J]. *Computer Aided Geometric Design*, 1987, 3(4): 295-311.