

文章编号:1001-9081(2010)02-0348-03

保留边界的点云简化方法

黄文明¹,肖朝霞¹,温佩芝¹,吴晓军²

(1. 桂林电子科技大学 计算机与控制学院,广西 桂林 541004;
2. 哈尔滨工业大学深圳研究生院 自动控制与机电工程学科部,广东 深圳 518055)
(xiaozaoxia5@163.com)

摘要:针对点云简化算法中边界点丢失的问题,提出了一种保留边界的三维散乱点云的非均匀简化算法。首先利用 kd-tree 建立散乱数据点云的空间拓扑关系,计算出每个数据点的 k 邻域;然后针对目前依据点云分布均匀性算法提取边界效率低的问题,提出一种改进的点云边界点判定算法;最后保留所有边界点,对非边界点,根据曲面变分值和 k 邻域点已保留比例,进行点云的非均匀简化。实验结果表明,该算法精度高,空间复杂度低,而且简化后点云边界保留完整。

关键词:边界点;非均匀简化;散乱点云;kd-树

中图分类号: TP391.41 **文献标志码:**A

Point cloud simplification with boundary points reservation

HUANG Wen-ming¹, XIAO Zhao-xia¹, WEN Pei-zhi¹, WU Xiao-jun²

(1. School of Computer and Control, Guilin University of Electronic Technology, Guilin Guangxi 541004, China;
2. Division of Control and Mechatronics Engineering,
Shenzhen Graduate School of Harbin Institute of Technology, Shenzhen Guangdong 518055, China)

Abstract: A non-uniform simplification approach with boundary points reservation was proposed concerning the boundary points loss caused by most point cloud simplification algorithms. First, kd-tree was used to represent the spatial topology relationship of the scattered point cloud and to calculate the k -nearest neighbors for each data point. Then an improved algorithm for boundary points detection of point cloud was presented to solve the low efficiency when the current algorithms extract boundary according to uniform distribution of point cloud. Consequently, all the resulted boundary points were reserved, while the non-boundary points were reserved for final non-uniform simplification according to the surface variation and their nearest neighbors. The experimental result shows that the proposed simplification approach has high accuracy and low space complexity, and can reserve the boundary points.

Key words: boundary point; non-uniform simplification; scattered point cloud; kd-tree

0 引言

随着近几年曲面获取技术的提高,采样点的精度达到亚毫米级,高精度的三维扫描仪包含大量的点,使用这些点进行曲面重建以及后续处理,将会消耗大量时间和空间^[1]。因此,在保证能为模型重建提供必需信息的前提下,简化测量数据点云是十分必要的。国内外许多学者对三维散乱数据点集的简化进行了研究,Wang 等人^[1]先用几何图像表示采样点,然后根据简化密度和曲面变分将点云进行简化;马磊等人^[2]和王洪涛等人^[3]根据点数和表面变化系数实现测量数据点集的快速递归分块,然后在每个分块子集中保留一个与该子集重心点最近的测量点实现数据简化。贺美芳等人^[4]先计算出所有点的法向量,然后把曲率值小于给定阈值的数据点提取出来,按照数据个数进行简化,把曲率值大于给定阈值的数据点提取出来,进行排序,再按照给定百分比提取;Pauly 等人^[5]运用递增式和分割式聚类方法实现了数据点集的快速递归划分,在每个类中保留一个点实现了简化。然而,以上点云简化算法都没有检测边界点,当点云数据包含边界时,边界

点可能会丢失,从而影响后续的曲面重建。

文献[6-7]认为如果数据点 p 的 k 邻近点的分布偏向一侧,则点 p 为边界特征点。本文根据其对边界点的定义,对其邻近点分布均匀性的判定方法进行改进。在边界提取的基础上,保留所有边界点,再利用曲面变分将其他非边界点区分为特征点和非特征点。对特征点,在其邻域内按比例保留,非特征点在邻域内保留一个点。

1 数据点邻域检索

散乱点云没有明显的几何分布特征,即没有规律性,都呈散乱无序状态,因此必须建立数据点之间的空间拓扑关系才有利于查找每个点的 k 近邻。目前搜索 k 近邻常见的方法有三种:八叉树、空间单元格和 kd-tree 法。本文首先采用 kd-tree 建立数据点之间的拓扑关系,从而找出各数据点的 k 近邻,通过邻近点的位置反映出待重建曲面在该测点处的形状信息。

kd-tree^[8]是 Friedman 等人于 1977 年提出的一种高维二叉树,在其上可实现对给定目标点的快速最近邻查找。kd-

收稿日期:2009-08-24;修回日期:2009-10-10。

基金项目:国家自然科学基金资助项目(50805031);广西科学基金资助项目(桂科自0991240)。

作者简介:黄文明(1963-),男,江苏吴县人,副教授,主要研究方向:图形图像处理、网格计算、软件工程; 肖朝霞(1983-),女,江西高安人,硕士研究生,主要研究方向:图形图像处理; 温佩芝(1963-),女,广西灵山人,副教授,博士,主要研究方向:图形图像处理、模式识别、科学可视化; 吴晓军(1975-),男,甘肃人,博士,主要研究方向:计算机图形学、CAD 技术、科学可视化。

tree 的每一层将空间分成两个。树的顶层节点按一维进行划分,下一层节点按另一维进行划分,以此类推,各个维循环往复。划分要使得在每个节点,大约一半存储左子树中,而另一半存储在右子树中。当一个节点中的点数少于给定的最大点数时,划分结束。二维 kd-tree 如图 1 所示,左图为剖分规则,其中水平线表示按 X 轴剖分,竖线表示按 Y 轴剖分。

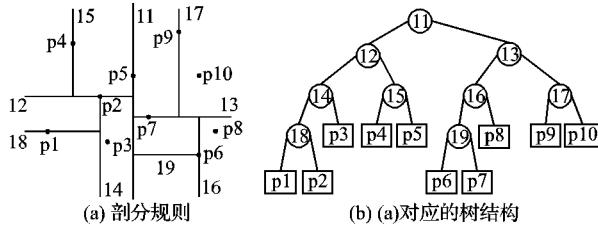


图 1 二维 kd-tree 示意图

本文选取了一个合适的分割规则,即每次分割时选择最长轴作为分割轴。散乱点云在 X, Y, Z 轴方向的最小和最大值分别为: $X_{\min}, X_{\max}, Y_{\min}, Y_{\max}, Z_{\min}, Z_{\max}$, 则分割轴长分别为: $\Delta X = X_{\max} - X_{\min}$, $\Delta Y = Y_{\max} - Y_{\min}$, $\Delta Z = Z_{\max} - Z_{\min}$ 。例如 ΔZ 最大,即选择 Z 轴进行分割。计算出位于 Z 坐标中间值的点,小于中间值的点位于节点的左子树,大于中间值的点则位于节点的右子树。

2 曲面变分

以曲率为准则的简化算法,能够保留模型的特征,其绘制效率显著提高^[4],但是对于散乱点集模型进行曲率估算都比较复杂而耗时。Pauly 等人^[5]使用曲面变分来估计点云的局部曲面变化,该方法简单且能有效地反应曲面信息,他指出曲面变分值越大,表示该处的曲面越弯曲,反之亦然。对于局部区域中某个点 p_i , p_i 周围邻近点记为 $p_m \in Nb(p_i)$, $Nb(p_i)$ 表示离 p_i 最近的 k 个点, p_i 周围邻近点均值为 $\bar{p}_i = \frac{1}{k} \sum_1^k p_i$, 点 p_i 的协方差矩阵定义为:

$$C_i = \begin{bmatrix} (p_1 - \bar{p}_i)^T & (p_1 - \bar{p}_i) \\ \vdots & \vdots \\ (p_m - \bar{p}_i)^T & (p_m - \bar{p}_i) \end{bmatrix}$$

C_i 是对称半正定矩阵,存在 3 个实数特征值,设其分别为 $\lambda_1, \lambda_2, \lambda_3$ (假设 $\lambda_1 \leq \lambda_2 \leq \lambda_3$),它们对应的特征向量分别为 v_1, v_2, v_3 , v_2 和 v_3 张成了 p_i 处的最小二乘拟合平面, v_1 即为该平面的法矢方向, λ_1 的大小表明了待重建曲面在 p_i 处沿法向方向的变化量。定义 p_i 处的曲面变分为: $\sigma_i = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3}$ 。为了确定曲面变分的阈值,本文采取随机采样法估计点云的曲面变分平均值。在点云中随机取出 N 点,计算这些点的曲面变分,然后计算平均值,即得到曲面变分的估计值, $\sigma = \sum_{i=0}^{N-1} \sigma_k$, 曲面变分的阈值为: $\eta = d * \sigma$, d 为期望保留点占总的点数的比值。

3 边界点判定

文献[6~7]利用数据点的拓扑关系以及数据点法矢等信息构造最小二乘平面,然后根据邻域内数据点在平面上的投影点之间分布均匀性提取边界,如图 2 中(a)所示。文献[6]投影点均匀性的判断标准是角度标准差。该方法首先要找到离该点最近的点作为基准点,其次要对邻近点的夹角进

行排序从而计算出夹角标准差,当数据量很大时,上述两个步骤都会耗费大量时间。文献[7]投影点均匀性的判断标准是角度间隔,该方法同样要依据角度对数据点进行排序,然后计算出相邻数据点之间的角度差,效率也很低。本文基于邻域内数据点在平面上的投影点之间分布均匀性的思想,但是不计算夹角,而是直接比较坐标值,因此可以节省很多时间。设点 p 为点云中的某一点,先利用 kd-tree 检索该点的 k 邻域,然后构建最小二乘平面,并将邻近点投影到其最小二乘平面上(如图 2(b)所示),过 p 点作平行于坐标轴平面 xoy, xoz, yoz 的平面 xpy, xpz, ypz ,以 xpy, xpz, ypz 为参考平面,当其邻近点中位于某一平面一侧的点与位于另外一侧的点数差与邻近点总数的比大于一定阈值时,该点即为边界点。以 p 点和平面 xpy 为例,位于 xpy 上方的点有 13 个,位于下方(包括平面上)的点有 2 个,所以位于两侧的点数之差与邻近点总数的比例为 0.7333。

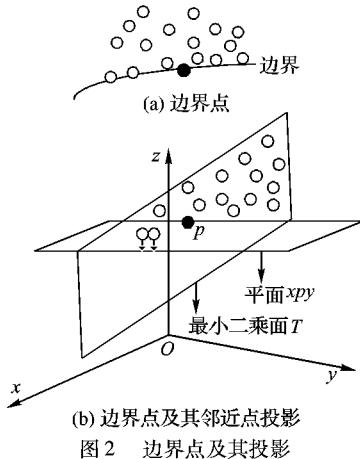


图 2 边界点及其投影

边界点的具体计算过程如下:

- 1) 由 p 点的邻近点,计算出 p 点的法向量,然后构造最小二乘平面 T ;
- 2) 将 p 点的邻近点投影到平面 T 上;
- 3) 过 p 点作平行于 xoy, xoz, yoz 的平面,并计算出位于三个平面两侧的点;
- 4) 若位于平面两侧的点数之差占邻近点总数的比例大于或等于 f ,该点即为边界点。

4 点云简化

文献[2~3]根据局部区域中曲面变分和包含的点数进行散乱点云简化。要将原始点集细分成许多子集,这样会占用大量的空间,而且当曲面存在比较大的平坦区域时,文献[3]的方法会产生空洞。保留所有边界点,对非边界点,本文根据曲面变分及邻近点保留情况进行散乱点云的简化。将曲面变分大于给定阈值点称为特征点,而局部法向量大于给定阈值点称为非特征点。如果某点的一邻近点已经确定为要保留的点,那么该邻近点即为邻近保留点。对特征点,根据其邻近保留点数的比例进行简化;对非特征点, k 邻域中只保留一个点。这样就可以在曲面变分大的地方保留更多的点,而曲面变分小的地方保留更少的点,输出点集既不产生空洞,同时也能保留足够的细节。简化步骤如下:

- 1) 输入散乱点云,建立 kd-tree,从而建立点云的拓扑关系,找到数据点的 k 邻域;
- 2) 判断该点是否为边界点,是则保留该点,否则跳至下一步;

3) 计算该点的曲面变分和邻近点保留点的个数, 将曲面变分跟曲面变分阈值 η 比较, 大于则跳至 4), 否则跳至 5);

4) 将邻近保留点数占邻近点总数的比例与特征点比例系数 e 进行比较, 小于则保留该点, 同时将该点的保留标识置为 true, 否则删除该点, 直接跳至 6);

5) 比较邻近保留点数是否等于 0, 是则保留该点, 同时将该点保留标识置为 true, 否则删除该点;

6) 重复 3) ~ 5), 直到遍历完点云数据中的所有点结束。

5 实验结果和分析

应用 VC ++ 6.0 和 OpenGL 工具, 在 P4 3.0 GHz, 内存 2.0 GB 的计算机上实现了本文上述算法, 并对几组数据进行了边界点的提取及保留边界点的简化。本文点云数据均来自 Hoppe 的个人主页^[9]。

5.1 点云边界提取

图 3 为汽车模型和 hypersheet 模型的边界提取。出于比较的目的, 本文也可实现文献[6~7]中基本角度判定的算法, 表 1 为本文算法和与角度判定法的耗费时间对比。从图中可以看出汽车模型的车窗和车轮等底部的边界非常清晰, hypersheet 的外边界圈也非常明朗, 这表明本文的边界提取算法效果良好。汽车模型的顶部突起部分之所以没有保留, 是因为本文定义的边界是指开区域的边界, 并不包含边界拐点。调整算法参数, 使得两种算法提取的边界点数目相等, 见表 1, 从表中可以本文算法的效率明显优于角度判定算法。这主要是因为本文算法无需计算各邻近点之间的角度并进行角度差的排序, 而是直接计算投影点的位置, 从而可以节省很多时间。

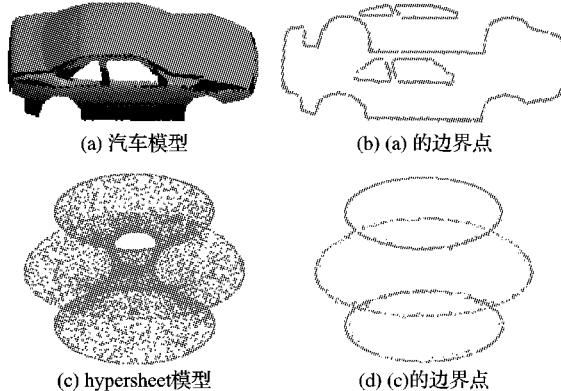


图 3 汽车模型和 hypersheet 模型边界提取

表 1 实验参数及效率比较

| 模型 | 点数 | f | 输出 | | 耗时/s |
|------------|-------|-----|-----|-------|-------|
| | | | 点数 | 本文方法 | |
| 汽车 | 20621 | 0.5 | 906 | 0.531 | 0.796 |
| hypersheet | 6753 | 0.5 | 921 | 0.125 | 0.187 |

5.2 点云简化

为了突出模型边界, 本文算法提取到的边界点用灰色显示出来, 如图 4 和图 5 中(b)所示。从图中可以看出, 文献[2]的方法由于没有考虑边界问题, 在汽车模型的边界部分丢失了一些点, 使得模型的边界很模糊甚至变形。汽车模型的车轮由于丢失了一些点使得车轮形状发生了改变, 而本文先提取边界点, 然后在简化时, 将边界点全部保存下来, 因此整个车模型的轮廓非常清楚。hypersheet 也发生了变形, 尤其是第二个圈由原来的圆变成了不规则的形状。曲面变分的大小反应了曲面的变化程度, 在曲面变化大的地方按照比例保留点云, 而在较平坦的区域邻域内只保留一点, 从而既可以保留足够的特征信息也可以达到点云简化的目的。

调节本文算法的系数和文献[2]聚类法的系数, 使得简化的点数相同, 见表 2。从表中可以看出, 本文的简化时间跟聚类法差不多, 但是占用内存少。文献[2]之所以占用的内存大于本文占用的内存, 这是因为聚类的思想都要把点云剖分成许多小的点集, 每个子集都占用一定的内存, 因此占用的内存会很大; 而本文逐个点进行简化, 不需要额外的内存开销。

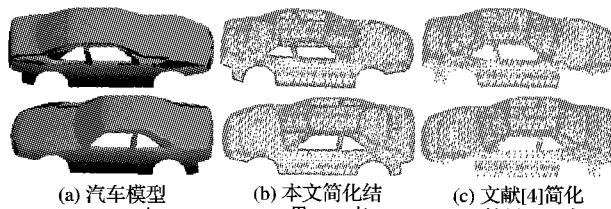


图 4 汽车模型

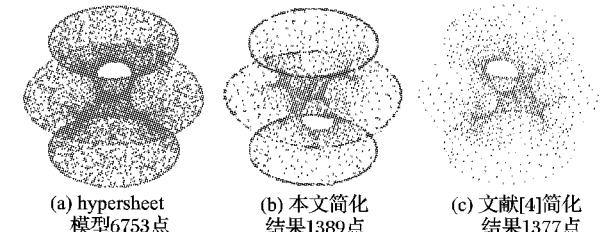


图 5 hypersheet 模型

表 2 算法效率比较

| 模型 | η | e | 输出点数 | 占用空间/MB | | 占用时间/s | |
|------------------------|--------|-----|------|---------|-----|--------|-------|
| | | | | 本文方法 | 聚类法 | 本文方法 | 聚类法 |
| 汽车 (20621 点) | 0.006 | 1/4 | 3940 | 14 | 21 | 0.490 | 0.488 |
| | 0.008 | 1/4 | 3517 | 14 | 20 | 0.490 | 0.485 |
| hypersheet (6753 点) | 0.006 | 1/2 | 1389 | 11 | 15 | 0.082 | 0.078 |
| | 0.008 | 1/2 | 1256 | 11 | 14 | 0.082 | 0.077 |

6 结语

本文首先改进了一种点云边界的提取算法, 实验证明本文的边界提取效率高, 同时效果良好。然后在边界提取的基础上, 提出了一种保留边界的点云简化算法, 使得简化后的点云边界非常清晰。目前许多点云的曲面重建算法, 都有修补

残缺数据的功能, 但是却并不能识别点云中的残缺数据是模型本身固有的还是空洞。如果本身就包含空洞或本身就是非闭合的模型, 那么在曲面重建时并不需要修补这些残缺数据。下一步的工作将主要集中在交互式选择修补空洞的曲面重建上。

(下转第 384 页)

Goose 通过启发优化来进行服务组合调整, 即通过服务调用结果来进行组合调整, 类似于基于服务 QoS 调整策略; 但 Goose 没有分析平台已有的复杂服务能力, 因此不支持复杂服务发现。

MSWC 则是通过复杂服务来分析服务之间的转移概率, 以最小熵为目标通过 N -最短路径进行组合。因为服务统计关系学习包括服务间转移概率和服务间逻辑概率统计, 均是后台计算的, 因此不影响服务生成的算法复杂度。服务生成时的算法复杂度, 主要包括一个最短路径分解和一个简单分类, 前者是线性时间^[2], 后者算法复杂度为 $O(1)$, 因此复杂服务生成的算法复杂度为 $O(n)$, 对应于 MSWC(New)。而且平台已有目标服务, 则服务起点和终点间有一条或者多条连接路径, 因此选择熵值最小的一条即为所求, 算法复杂度接近于常数, 对应于 MSWC(Existed)。

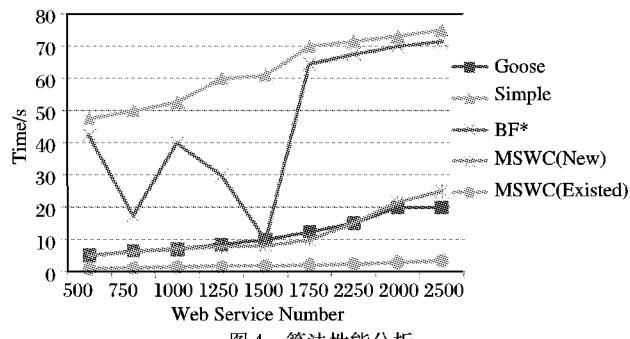


图 4 算法性能分析

4 结语

通过平谷旅游服务系统的验证, MSWC 基于服务关系统计解决了传统工作流中服务自动组合和复杂服务重用问题。

但是这种组合方式还有待完善, 因为在实际的业务例子中, 用户的服务需求可能不是平台中最常用的服务组合方式(熵值最小的服务组合)而是以某些约束或者服务质量为最优进行服务组合, 如旅游中花费最小、景点最多等。针对基于约束的服务组合有两个办法: 1) 通过用户进行服务组合进行简单修改来满足约束; 2) 通过自动验证生成的服务组合特征值来满足服务约束。如何定义服务约束, 以及如何将服务约束与统计服务组合方法进行结合是 MSWC 系统接下来的研究重点。

参考文献:

- [1] CASATI F, ILNICKI S, JIN L J, et al. Adaptive and dynamic service composition in eFlow[C]// Proceedings of the 12th International

Conference on Advanced Information Systems Engineering, LNCS 1789. London: Springer, 2000: 13–31.

- [2] ZHANG HUA-PING, YU HONG-KUI, XIONG DE-YI, et al. HHMM-based Chinese lexical analyzer ICTCLAS[C]// Proceedings of the Second SIGHAN Workshop on Chinese Language Processing. Morristown: Association for Computational Linguistics, 2003: 184–187.
- [3] BROWN P F, deSOUZA P V, MERCER R L, et al. Class-based n-gram models of natural language[J]. Computational Linguistics, 1992, 18(4): 467–479.
- [4] DODDINGTON G. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics[M]. San Francisco: Morgan Kaufmann, 2002.
- [5] LIN C-Y, HOVY E. Automatic evaluation of summaries using N-gram co-occurrence statistics[C]// NAACL'03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology. Morristown: Association for Computational Linguistics, 2003: 71–78.
- [6] BARRON A R, COHEN A, DAHMEN W, et al. Approximation and learning by greedy algorithms[J]. Annals of Statistics, 2008, 36(1): 64–94.
- [7] AHUJA R K, MEHLHORN K, ORLIN J B, et al. Faster algorithms for the shortest path problem[J]. Journal of the Association for Computing Machinery, 1990, 37(2): 213–223.
- [8] ZHAN B F, NOON C E. Shortest path algorithms: An evaluation using real road networks[J]. Transportation Science, 1998, 32(1): 65–73.
- [9] SEDGEWICK R. Algorithms in C, Part 5: Graph algorithms[M]. [S. l.]: Addison-Wesley, 2001.
- [10] 范焱, 郑诚, 王清毅, 等. 用 Naive Bayes 方法协调分类 Web 网页[J]. 软件学报, 2001, 12(9): 1386–1392.
- [11] KONONENKO I. Semi-Naive Bayesian classifier[C]// Proceedings of the European Working Session on Machine learning. New York: Springer – Verlag, 1991: 206–219.
- [12] YANG HAO, CHEN JUNLIANG, MENG XIANGWU, et al. Dynamically traveling Web service clustering based on spatial and temporal aspects[C]// Advances in Conceptual Modeling – Foundations and Applications, LNCS 4802. Berlin: Springer, 2007: 348–357.
- [13] 温嘉佳, 陈俊亮, 彭泳, 等. 基于目标距离评估的启发式 Web Services 组合算法[J]. 软件学报, 2007, 18(1): 85–93.
- [14] OH S-C, ON B-W, LARSON E J, et al. BF*: Web services discovery and composition as graph search problem[C]// Proceedings of the 2005 IEEE International Conference on E-Technology, E-Commerce and E-Service. Washington, DC: IEEE Computer Society, 2005: 784–786.

(上接第 350 页)

参考文献:

- [1] WANG RENFANG, ZHANG SANYUAN, YE XIUZI. A novel simplification algorithm for point-sampled surfaces [C]// MUE'07: International Conference on Multimedia and Ubiquitous Engineering. Soul: IEEE Computer Society, 2007: 573–578.
- [2] 马磊, 彭国华, 耿东芳. 基于八叉树的海量测量数据的非均匀简化[J]. 计算机应用, 2007, 27(8): 2027–2029.
- [3] 王宏涛, 张丽艳, 杜信, 等. 测量点集的简化及其隐式曲面重建误差分析[J]. 中国图象图形学报, 2007, 12(11): 2114–2118.
- [4] 贺美芳, 周来水, 神会存. 散乱点云数据的曲率估算及应用[J]. 南京航空航天大学学报, 2005, 37(4): 515–519.
- [5] PAULY M, CROSS M, KOBELT L P. Efficient simplification of

point-sampled surfaces [C]// Proceedings of IEEE Visualization 2002. Washington, DC: IEEE Computer Society, 2002: 163–170.

- [6] 钱锦锋, 陈志杨, 张三元, 等. 点云数据压缩中的边界特征检测[J]. 中国图象图形学报, 2005, 10(2): 164–169.
- [7] 陈飞舟, 陈志杨, 丁展, 等. 基于径向基函数的残缺点云数据修复[J]. 计算机辅助设计与图形学学报, 2006, 18(9): 4141–419.
- [8] FRIEDMAN J H, BENTLEY J L, FINKEL R A. An algorithm for finding best matches in logarithmic expected time[J]. ACM Transactions on Mathematical Software, 1977, 3(3): 209–226.
- [9] HOPPE H. Hugues Hoppe's Homepage [EB/OL]. [2006-12-13]. <http://research.microsoft.com/~hoppe/>.