

8-13-2020

Algorithm for RGBD Point Cloud Denoising and Simplification Based on K-means Clustering

Benyue Su

1. School of Computer and Information, Anqing Normal University, Anqing 246133, China;;2. The University Key Laboratory of Intelligent Perception and Computing of Anhui Province, Anqing 246133, China;;

Jinyu Ma

1. School of Computer and Information, Anqing Normal University, Anqing 246133, China;;2. The University Key Laboratory of Intelligent Perception and Computing of Anhui Province, Anqing 246133, China;;

Yusheng Peng

2. The University Key Laboratory of Intelligent Perception and Computing of Anhui Province, Anqing 246133, China;;3. School of Mathematics and Computational Science, Anqing Normal University, Anqing 246133, China;

Sheng Min

2. The University Key Laboratory of Intelligent Perception and Computing of Anhui Province, Anqing 246133, China;;3. School of Mathematics and Computational Science, Anqing Normal University, Anqing 246133, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Algorithm for RGBD Point Cloud Denoising and Simplification Based on K-means Clustering

Abstract

*Abstract: Aiming at the problem that the point cloud data scanned by Kinect or other depth camera have a huge number and more noise, a feature preserving method for point cloud denoising and simplification was proposed. This algorithm classified the point cloud rapidly by K-D tree; find The corresponding surface curvature values were obtained using curvature estimation algorithm; The K-means clustering algorithm for point cloud clustering was used. For each point in the cluster, the Euclidean distance was depended on between the point and center of the cluster and the change of the near points curvature to determine whether the noise points. The point cloud data was simplified by the feature preserving method. The experimental results show that *the denoising and feature preserving point cloud simplification method is quickly and efficiently, for the removal of a large number of external noise has a positive effect, and the streamline point cloud data have the retention of original point cloud features.**

Keywords

K-means clustering, point cloud denoising, point cloud simplification, RGBD data

Recommended Citation

Su Benyue, Ma Jinyu, Peng Yusheng, Sheng Min. Algorithm for RGBD Point Cloud Denoising and Simplification Based on K-means Clustering[J]. Journal of System Simulation, 2016, 28(10): 2329-2335.

基于 K-means 聚类的 RGBD 点云去噪和精简算法

苏本跃^{1,2}, 马金字^{1,2}, 彭玉升^{2,3}, 盛敏^{2,3}

(1. 安庆师范大学计算机与信息学院, 安徽 安庆 246133; 2. 安徽省智能感知与计算重点实验室, 安徽 安庆 246133;
3. 安庆师范大学数学与计算科学学院, 安徽 安庆 246133)

摘要: 针对 Kinect 等深度相机扫描获取的点云数据数量庞大、噪声较多的问题, 提出一种特征保持的点云去噪和精简算法。使用 K-D 树快速分类点云; 通过曲率估计算法得到局部曲面的曲率值; 使用 K-means 聚类算法对点云进行聚类, 对每个类中的点, 根据点到聚类中心的欧式距离和邻近点曲率变化判断是否为噪声点; 通过保持特征点云精简算法实现对点云数据的简化。实验结果显示, 算法快速有效, 对于去除大量外部噪声有良好效果, 且精简后的点云数据保持了原始点云特征。

关键词: K-means 聚类; 点云去噪; 点云精简; RGBD 数据

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-731X (2016) 10-2329-07

Algorithm for RGBD Point Cloud Denoising and Simplification Based on K-means Clustering

Su Benyue^{1,2}, Ma Jinyu^{1,2}, Peng Yusheng^{2,3}, Sheng Min^{2,3}

(1. School of Computer and Information, Anqing Normal University, Anqing 246133, China;
2. The University Key Laboratory of Intelligent Perception and Computing of Anhui Province, Anqing 246133, China;
3. School of Mathematics and Computational Science, Anqing Normal University, Anqing 246133, China)

Abstract: Aiming at the problem that the point cloud data scanned by Kinect or other depth camera have a huge number and more noise, a feature preserving method for point cloud denoising and simplification was proposed. This algorithm classified the point cloud rapidly by K-D tree; find The corresponding surface curvature values were obtained using curvature estimation algorithm; The K-means clustering algorithm for point cloud clustering was used. For each point in the cluster, the Euclidean distance was depended on between the point and center of the cluster and the change of the near points curvature to determine whether the noise points. The point cloud data was simplified by the feature preserving method. The experimental results show that the denoising and feature preserving point cloud simplification method is quickly and efficiently, for the removal of a large number of external noise has a positive effect, and the streamline point cloud data have the retention of original point cloud features.

Keywords: K-means clustering; point cloud denoising; point cloud simplification; RGBD data

引言

三维重建一直是计算机图形学、计算机视觉等领域的研究热点。随着 Kinect 等深度相机设备



收稿日期: 2016-06-10 修回日期: 2016-07-14;
基金项目: 国家自然科学基金(11471093); 安徽省教育厅自然科学基金项目(KJ2014A142); 安徽省重点实验室开放课题(ACAIMI60102);
作者简介: 苏本跃(1971-), 男, 安徽芜湖, 博士, 教授, 研究方向为图形图像处理、虚拟现实、行为识别。

的推出, 获取点云数据变得越来越便捷, 而此类数据普遍存在噪声多数据量大的问题。Sun 等^[1]基于 K-means 聚类方法对三维点云数据进行分割处理, 但对点云模型形状要求较高。Wang 等^[2]针对光滑曲面的散乱点云, 提出了一种基于局部重建的特征点提取方法, 但方法具有局限性不适用于含噪声的复杂点云。文献[3]使用改进的主成分分析法将点云去噪问题转化为约束非线性最小二乘

问题, 从而实现点对云的去噪。Dutta^[4]等在去噪算法中通过估计高斯加权曲率提取点云特征点进行特征保持, 取得了较好的效果。文献[6-7]对噪声点进行分类后使用改进的双边滤波算法进行去噪处理, 但仍会有过度光滑现象。He 等^[8]提出一种对散乱点云数据计算曲面曲率的局部微分性质, 通过曲率算法对点云进行简化。Ma 等^[9]通过搜索曲率极值点的方法快速提取散乱点云的特征点, 提高了搜索效率。Wang 等^[10]针对建筑物点云, 提出基于平面特征分层聚类的点云数据精简算法, 得到了较高的精简比率。Yuan 等^[11]通过 K 均值聚类对点云数据进行聚类, 使用主成分分析法估计点云法矢和初始特征点, 结合自适应均值漂移法对数据进行分类, 实现了特征保持的点云精简效果。本文针对 Kinect 等深度相机扫描获取的点云数据数量庞大、噪声较多的问题, 提出一种特征保持的点云去噪和精简算法, 并分别使用数据库点云和 Kinect 扫描点云进行了算法实验。实验结果表面, 本文算法稳定, 对庞大点云数据有很好的去噪和精简效果, 且精简后的点云数据保留了原始特征和轮廓信息。

1 曲率估计算法

在二维空间中, 曲率用来反应曲线的弯曲程度, 而在三维空间中, 面的弯曲程度同样可以用曲率来表示, 因此根据曲率的大小可以得到几何体上的特征区域。在三维欧式空间中, 一般使用高斯曲率来反映某点上的弯曲程度, 要想获取高斯曲率首先要得到曲面的两个主曲率。对于散乱点云数据, 要想获取某点在点云中的曲率特征, 一般是将此点和领域内的点拟合曲面, 再求取曲面的主曲率和高斯曲率。

由于三维扫描获取的数据量往往很大, 因此对每个点拟合曲面将会耗费大量的内存和时间, 根据点云间的三角拓扑关系, 提出一种散乱点云的曲率估计算法。如图 1 所示, p 点为散乱点云中的其中一点, 由 K-D 树算法寻找得到点云中距

离 p 点最近的其他 8 个点。以 p 点为中心, 根据坐标大小将 9 个点排列成如图 1 的大致形状, 这 9 个点共同组成了曲面 S 。由于在三维重建中, 散乱点云通过点与点的连线形成拓扑关系, 以三角面片的形式表现模型表面的几何形状, 因此可以使用点与其领域点的三角关系估计此点在模型表面的凹凸情况, 即通过点与领域内点的三角关系估计此点的“曲率”。

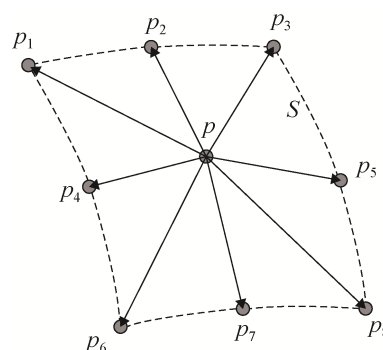


图 1 曲面构建示意图

对图 1 中的 p 点, 曲率估计算法步骤如下:

1) 分别连接 p 点和相邻的其他 8 个点, 得到 8 个向量 $\overrightarrow{pp_1}$ 、 $\overrightarrow{pp_2}$ 、 $\overrightarrow{pp_3}$ 、 $\overrightarrow{pp_4}$ 、 $\overrightarrow{pp_5}$ 、 $\overrightarrow{pp_6}$ 、 $\overrightarrow{pp_7}$ 、 $\overrightarrow{pp_8}$ 。

2) 向量 $\overrightarrow{pp_1}$ 和 $\overrightarrow{pp_8}$ 的夹角为 θ_{1-8} , 其余弦值为

$$\cos \theta_{1-8} = \frac{\overrightarrow{pp_1} \times \overrightarrow{pp_8}}{|\overrightarrow{pp_1}| \times |\overrightarrow{pp_8}|} \quad (1)$$

余弦值的大小反应了两个向量的夹角大小, 当 $\cos \theta_{1-8}$ 越大时夹角越小, 当 $\cos \theta_{1-8}$ 越小时夹角越大, 反应在三维空间中就是当 $\cos \theta_{1-8}$ 越大时点越凸出或凹进, 当 $\cos \theta_{1-8}$ 越小时点越平滑。由式(1)同理分别计算 $\overrightarrow{pp_2}$ 和 $\overrightarrow{pp_7}$ 、 $\overrightarrow{pp_3}$ 和 $\overrightarrow{pp_6}$ 、 $\overrightarrow{pp_4}$ 和 $\overrightarrow{pp_5}$ 两个向量夹角的余弦值得到 $\cos \theta_{2-7}$ 、 $\cos \theta_{3-6}$ 、 $\cos \theta_{4-5}$ 。则估计曲率

$$H = \cos \theta_{1-8} + \cos \theta_{2-7} + \cos \theta_{3-6} + \cos \theta_{4-5} \quad (2)$$

估计曲率 H 反应了此点在点与领域内点构成的几何形状的凹凸程度, 当 H 越大时形状越突出, 当 H 越小时形状越平滑。

2 基于 K-means 聚类的去噪和精简算法

K-means 算法是基于距离的聚类算法, 采用距离作为相似性的评价指标, 即认为两个对象的距离越近, 其相似度就越大。在散乱点云处理中, 使用 K-means 算法可以快速将点云数据进行分类, 得到紧凑且独立的簇。通过比较簇中各个点到簇中心的距离、分析点的估计曲率关系即可找到噪声点并去除。如果一个簇中的所有点的估计曲率都较小, 则可以使用簇中心代替簇中所有的点, 从而达到精简点云的目的。

2.1 K-means 聚类算法

假设原始数据集为 (x_1, x_2, \dots, x_N) , 并且每个 x_i 为一个 d 维向量, K-means 聚类的目地就是, 在给定分类组数 $K(K \leq N)$ 值的条件下, 将原始数据

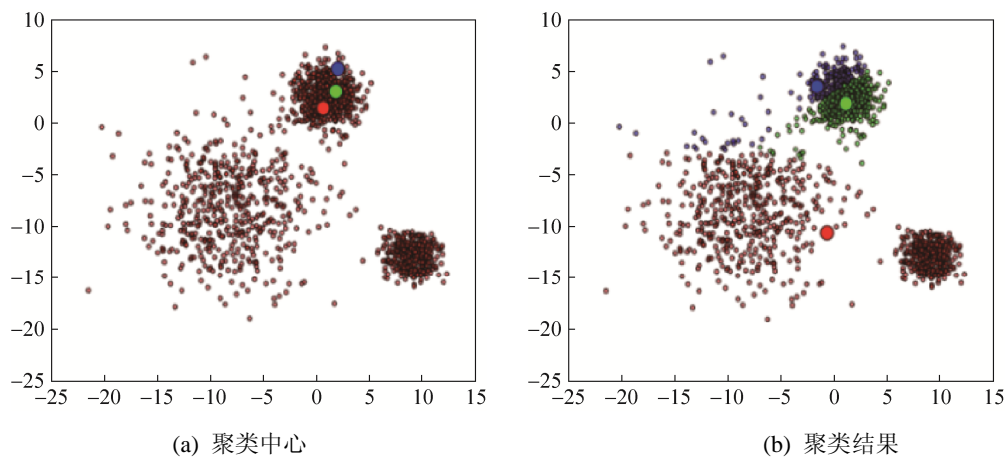


图 2 错误的初始聚类中心选择

由于 K-D 树划分方法的特殊性, 散乱点云上的点在 K-D 树上均匀的分层分散在各个结点上, 因此可以通过 K-D 树决定初始聚类中心, 使得初始聚类中心能均匀分散在点云的各个部分。对散乱点云 K-means 聚类的算法如下:

(1) 根据聚类 K 值的大小, 在 K-D 树中寻找拥有与 K 值最接近的结点个数的那一层, 以这一层中所有结点对应的点坐标作为初始聚类中心, 得到 K 个簇中心。

(2) 分别计算点云中所有点到簇中心的欧式

距离, 将这些点分别划归到距离最近的簇中, 得到 K 个簇。

$$\arg \min \sum_{i=1}^K \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (3)$$

其中: μ_i 表示分类 S_i 的平均值。

K-means 算法的两个重要问题就是 K 值的大小和初始聚类中心的选择。在本文的点云处理算法中, K 值的大小取决于想要精简的程度, K 的取值越大则精简程度越低, 越小则精简程度越大。

初始聚类中心的选择同样对聚类结果有很大的影响, 过近的初始聚类中心可能导致得到错误的聚类或者聚类不平均的问题。如图 2(a)所示为选择的三个初始聚类中心, 由于中心距离过近因此得到如图 2(b)所示的不好的聚类结果, 因此聚类中心的选择需要做到在点中平均分散并且相互距离不能太近。

距离, 将这些点分别划归到距离最近的簇中, 得到 K 个簇。

(3) 根据聚类结果, 重新计算 K 个簇各自的中心, 计算方法是取簇中所有点的重心, 得到 K 个新的簇中心。

(4) 对点云中所有点按照新的簇中心重新聚类。

(5) 重复(3)(4)步直到簇的中心不再变化, 最终得到的 K 个簇即为聚类结果。

2.2 去噪算法

在三维扫描中的噪声点一般为扫描时在模型表面外部或内部的漂移点，他们的共同特点是距离表面点云较远、与邻近点云连接形成的几何形状凹凸明显，所以可以通过欧式距离和估计曲率衡量某一点是否为噪声点。

设聚类后得到的一个簇为 $S = \{p_i, i=1, 2, \dots, n\}$ ，对该簇中点的去噪算法如下：

1) 计算簇中每个点 p_i 到簇中心 μ 的欧式距离 d_i 。

2) 由 3.3 节的曲率估计算法计算每个点的估计曲率 k_i ，并得到所有点曲率的平均值

$$\bar{k} = \frac{1}{n} \sum_{i=1}^n k_i \quad (4)$$

3) 给定权值 $\omega (\omega > 1)$ ，用来定义噪声的去除程度， ω 越小则去噪程度越强(有可能去除非噪声点云)， ω 越大则去噪程度越弱。

4) 对簇中的每个点 p_i 进行判断，若 $d_i > \frac{\omega}{n} \sum_{i=1}^n d_i$ 且 $\begin{cases} k_i > \frac{\bar{k}}{\omega}, \bar{k} < 0; \\ k_i > \omega \bar{k}, \bar{k} > 0; \end{cases}$ ，则 p_i 点为噪声点，去除该点。

对聚类后的每个簇分别使用去噪算法进行处理，即可得到去噪后的点云数据。

2.3 特征保持的点云精简算法

设聚类去噪后得到的一个簇为 $S = \{p_i, i=1, 2, \dots, n\}$ ，给定权值 $\lambda (0 < \lambda < 2)$ 用来定义点云的特征保持程度，若 λ 越小则特征保持效果越好， λ 越大则特征保持越少但精简程度更高。算法步骤如下：

(1) 计算簇中每个点的估计曲率分别为 k_i ，且其中最大曲率为 k_{\max} ，最小曲率为 k_{\min} 。

(2) 若簇中 $k_{\max} - k_{\min} < \lambda$ ，则认为簇中点的曲率变换不大，使用簇的中心代替簇中所有点。

(3) 若簇中 $k_{\max} - k_{\min} > \lambda$ ，则认为簇中含有特征部分，分别以曲率最大和最小的两个点作为初始聚类中心对簇进行基于欧式距离的 K-means 聚

类，在簇中得到两个新簇，用这两个新簇代替簇。

(4) 分别对两个新簇重复第(1)(2)(3)步，直到簇中只有一个点为止。

通过以上算法即可实现保持特征的点云数据精简，平滑区域的点云由于曲率变化不大使用中心点替代。特征区域的点云由于最大最小曲率差值大，会不断的分裂变成多个新的簇，曲率变换大的区域没有被替代掉，因此保留了原始点云的几何特征。

3 实验结果及分析

本文使用 Matlab 编程实现算法，计算机配置为 Windows7 系统、Intel Core i3 CPU、2GB 内存和 AMD HD545V 显卡。使用的 Bunny 点云和 Dragon 点云来自于斯坦福大学建立的 3D 点云数据库^[12]，格式为 PLY。

3.1 去噪算法实验及分析

如图 3(a) 为原始 Bunny 点云数据，共有 35 947 个点，图 3(b) 为在点云表面随机添加了 2 000 个噪声点后的点云数据。图 4 为对含噪声点云数据的聚类结果，不同的类使用不同灰度值表示。其中初始聚类数 K 取 512、权值 $\omega = 1.5$ ，聚类算法的运行时间为 6.16 s。

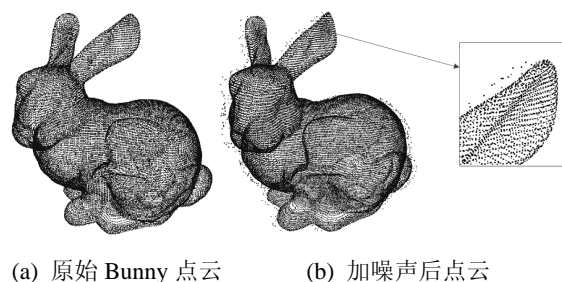


图 3 bunny 点云数据

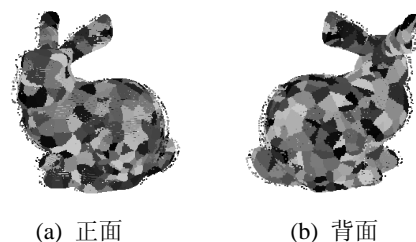


图 4 聚类结果

图 5 所示分别为使用本文曲率估计算法对点云提取 1 000 个和 4 000 个特征点的结果。由图可以看出, 点云的大部分特征都有了很好的保留。但由于噪声点为点云表面或内部的漂移点, 在本文的曲率估计算法中也会被视为特征点, 因此需要使用去噪算法做进一步处理。

图 6 所示为本文去噪算法的处理结果。从局部放大图可以看出, 在弧度较大的特征区域也有很好的去噪效果。

图 7 所示为使用文献[5]算法对 Bunny 点云的去噪处理结果。

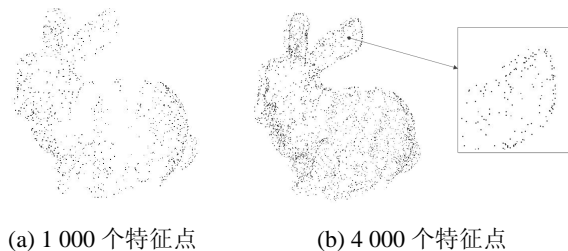


图 5 Bunny 点云特征点提取结果

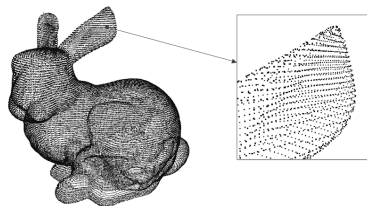


图 6 本文算法去噪结果

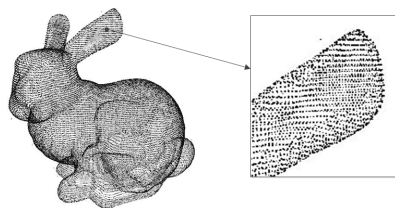


图 7 文献[5]方法去噪结果

表 1 所示为本文去噪算法与文献[5]算法对比结果。由表可知, 本文算法去除噪声点较多, 峰值信噪比更高, 且算法运行速度提高明显。

表 1 两种去噪算法对比

方法	含噪声点云个数	去噪后点云个数	PSNR	运行时间/s
本文	37 947	35 968	50.9	17.6
文献[5]	37 947	36 047	43.9	48.5

3.2 精简算法实验及分析

图 8 所示为本文算法对 Bunny 点云精简后的结果, 左图中权值 $\lambda = 0.3$, 精简后的点云个数为 13 486 个, 右图中权值 $\lambda = 0.6$, 精简后的点云个数为 4 071 个。从局部图可以看出, 本文精简算法对点云的边缘和特征信息都有很好的保持效果。

图 9 所示为使用本文算法对 Dragon 点云进行精简后的结果, 由图可以看出点云的边缘和特征信息都得到了很好的保留。

图 10 所示为使用文献[11]的算法对 Bunny 点云和 Dragon 点云进行精简后的实验结果。由图可以看出, 该算法在点云的特征区域没有很好的保留, 且点云的轮廓信息有所丢失。

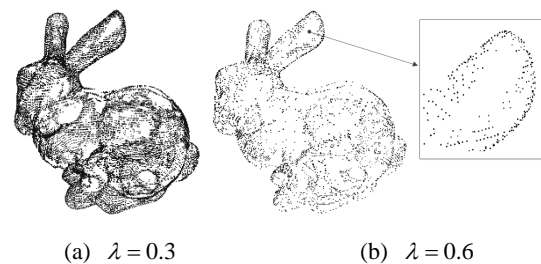


图 8 本文算法 Bunny 点云精简结果

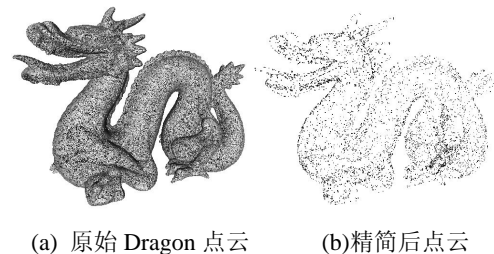


图 9 本文算法 Dragon 点云精简结果

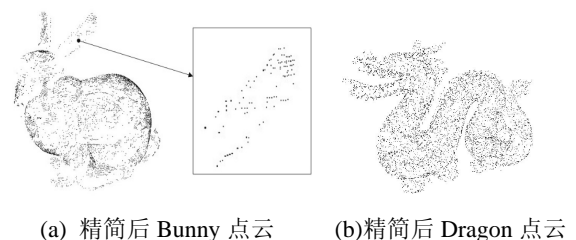


图 10 文献[11]精简结果

为了评估精简后点云的质量, 分别计算精简点云与原始点云的最大误差和平均误差^[13]。

最大误差:

$$\Delta_{\max}(S, S') = \max_{q \in S} |d(q, S')| \quad (5)$$

平均误差:

$$\Delta_{\text{avg}}(S, S') = \frac{1}{\|S\|} \sum_{q \in S} |d(q, S')| \quad (6)$$

式中: $d(q, S')$ 表示原始曲面 S 上采样点 q 到精简

点云曲面 S' 上投影点 q' 的欧式距离。

表 2 所示为本文方法与文献[11]方法对比结果。由表可知, 本文算法在精简程度更高的情况下运行时间仍有明显提高, 且最大误差和平均误差较小, 算法的精简效果更好。

表 2 两种精简算法对比

点云	方法	原始点云个数	精简后点云个数	运行时间/s	最大误差	平均误差
Bunny	本文算法	35 947	4 071	12.8	0.015	2.1×10^{-3}
	文献[11]算法		5 384	34.1	0.027	6.7×10^{-3}
Dragon	本文算法	437 645	6 871	40.8	0.043	5.2×10^{-4}
	文献[11]算法		9 748	374.3	0.277	1.6×10^{-3}

3.3 RGBD 点云实验及分析

本文同样使用 Kinect 扫描真实物体获取的 RGBD 点云数据进行了实验, 如图 11(a)所示为 Kinect 扫描获取的点云数据, 点云个数为 76 826 个, 在模型表面含有部分噪声点。图 11(b)所示为精简后的点云, 点云个数为 13 582 个, 由侧面图可以看出, 本文算法去除了大部分模型漂移噪声, 精简后保留了原始模型的特征部分, 且轮廓边缘信息保留较好。

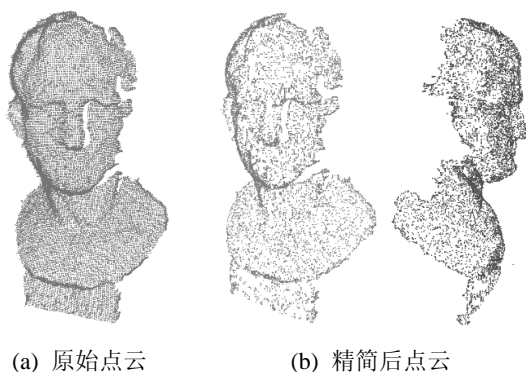


图 11 点云精简结果

4 结论

针对 Kinect 等深度相机扫描获取的点云数据数量庞大、噪声较多的问题, 提出了一种特征保持的点云去噪和精简算法。实验结果显示, 本文算法快速有效, 对于去除点云外部和内部漂移噪

声有良好的效果, 且精简后的点云数据保持了原始点云特征, 可以直接用于三维重建中, 提高了三维重建的效率。

参考文献:

- [1] 孙红岩, 孙晓鹏, 李华. 基于 K-means 聚类方法的三维点云模型分割 [J]. 计算机工程与应用, 2006, 42(10): 42-45.
- [2] 王小超, 刘秀平, 李宝军, 等. 基于局部重建的点云特征点提取 [J]. 计算机辅助设计与图形学学报, 2013, 25(5): 659-665.
- [3] Castillo E, Liang J, Zhao H. Point Cloud Segmentation and Denoising via Constrained Nonlinear Least Squares Normal Estimates [M]// Innovations for Shape Analysis. Germany: Springer Berlin Heidelberg, 2013: 283-299.
- [4] Dutta S, Banerjee S, Biswas P K, et al. Mesh Denoising Using Multi-scale Curvature-Based Saliency [M]// Computer Vision-ACCV 2014 Workshops. Germany: Springer International Publishing, 2014: 507-516.
- [5] 王勇, 唐靖, 饶勤菲, 等. 一种新的散乱点云快速去噪算法 [J]. 计算机应用与软件, 2015, 32(7): 74-78.
- [6] 曹爽, 岳建平, 马文. 基于特征选择的双边滤波点云去噪算法 [J]. 东南大学学报(自然科学版), 2013, 43(增 2): 351-354.
- [7] 袁华, 庞建铿, 莫建文. 基于噪声分类的双边滤波点云去噪算法 [J]. 计算机应用, 2015, 35(8): 2305-2310.
- [8] 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.

(下转第 2341 页)