

# 基于 RANSAC 分割的点云数据 K-近邻去噪算法研究

郭宁博<sup>1</sup> 陈向宁<sup>1</sup> 何艳华<sup>2</sup>

(1. 装备学院 北京 101416; 2. 63750 部队 渭南 714000)

**摘 要:** 针对基于 TOF 深度相机的空间目标表面重建的点云源数据容易受到仪器本身、扫描环境、外界干扰等影响,而含有大量的无效点和噪声点,增加了计算负担且影响了重建质量等问题,提出了一种基于随机采样一致性背景分割的点云 K-近邻去噪方法,以消除目标数据的异常值和无效点。首先,改进 RANSAC 算法,通过设置不同的阈值对原始点云进行背景分割,以确保准确提取待重建目标的主要特征。然后,通过 K-近邻点云平均算法和双边滤波算法移除离群点,最后使用体素化网格方法实现点云大数据的下采样,简化了目标点云,保留了局部特征,加快了曲面重建速度。实验结果表明,该算法能够有效的剔除噪声点,准确率高,实时性好,满足应用的要求。

**关键词:** 点云数据;去噪;分割;随机采样一致性

**中图分类号:** TP391.9      **文献标识码:** A      **国家标准学科分类代码:** 520.1040

## K-nearest neighbor denoising algorithm for point cloud data based on RANSAC algorithm

Guo Ningbo<sup>1</sup> Chen Xiangning<sup>1</sup> He Yanhua<sup>2</sup>

(1. Equipment Academy of PLA, Beijing 101416, China; 2. Unit 63750 of PLA, 714000 Weinan, China)

**Abstract:** According to the data of space target surface reconstruction of TOF depth camera point cloud based on source vulnerable to the instrument itself, scanning the environment, the effect of external disturbances, invalid points and noise points and large amount, increases the computational burden and the impact of the reconstruction quality problems, this paper proposes a denoising method of random sample consensus background segmentation of point cloud based on k-nearest neighbor, to eliminate outliers and invalid target data. Firstly, the improved RANSAC (random sample consensus, RANSAC) algorithm, by setting different threshold of the original point cloud background segmentation, to Ensure the accuracy of extracting the main features of the reconstructed object. Then, through the K-nearest neighbor point cloud average algorithm and bilateral filtering algorithm to remove outliers, finally using voxel grid method to achieve point cloud data sampling, simplified target point cloud, retains the local characteristics, to speed up the reconstruction speed. The experimental results show that the the algorithm can effectively eliminate noise, high accuracy, good real-time performance, meet the application requirements.

**Keywords:** point cloud data; denoising; segmentation; random sample consensus

## 0 引 言

随着激光扫描技术的不断成熟与计算机科学的迅猛发展,基于点云数据的空间目标三维实体造型技术被广泛地应用于机器人自动导航<sup>[1]</sup>、文物保护、虚拟现实和逆向工程等领域。非接触式三维激光扫描技术以其高采样速度、高精度和高分辨率、易操作等优点而广泛应用。但是受到目标物体的材质、设备的固有误差、环境因素、以及数据采集中人为因素的干扰,初步获取的目标点云数据不可避免的存在大量的离群噪声点<sup>[2]</sup>。这些噪声严重影响了后期目标

特征提取、特征匹配、曲面重建、识别与定位等一系列工作的开展,因此剔除点云数据中的噪声点,获取完整逼真的三维模型数据一直是人们的研究重点<sup>[3]</sup>。

点云噪声的去除研究较多,针对 Kinect 传感器获取的点云数据噪声,苏本跃等人<sup>[4]</sup>提出使用 K-D 树快速分类点云,通过 K-means 聚类算法进行聚类,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。Wang 等人<sup>[5]</sup>针对采自光滑曲面的散乱点云,提出一种鲁棒特征点提取方法,但是该方法具有局限性,难以适应复杂的含有噪声的点云数据。Cao 等人<sup>[6]</sup>对点云数据噪声进行了分

类,然后使用改进的双边滤波算法去噪,但是依然存在过度的光滑现象。

为了更快速的解决无组织散乱点云的分割与简化去噪,本文提出了基于 RANSAC 分割的 K-近邻点云数据去噪算法,首先设置阈值改进 RANSAC 算法对获取的目标点云数据进行无效背景噪声的去除,然后在点云数据中建立小立方体格栅,利用平均近邻搜索实现特征点云的简化滤波,最后采用 K-近邻点云优化策略剔除目标的离群点,最终得到一个满足应用要求的去噪结果。

## 1 算法设计

### 1.1 改进 RANSAC 算法分割

#### 1) RANSAC 算法原理

本算法通过随机取样剔除局外点,构建一个仅由局内点数据组成的基本子集。当从采集到的点云数据中提取目标区域时,首先针对性设计出目标判读准则模型,然后利用该判断准则迭代地剔除目标区域以外的点云数据。该算法要求在一定的置信概率下,基本子集最小抽样数  $M$  与至少取得一个良性取样子集的概率  $P(P > \epsilon)$  满足如下关系<sup>[7]</sup>:

$$P = 1 - (1 - (1 - \epsilon)^m)^M \quad (1)$$

式中:  $\epsilon$  为数据错误率,  $m$  为计算模型参数需要的最小数据量。

#### 2) 改进 RANSAC 分割算法

本文用 TOF 激光扫描设备获取的圣诞老人点云数据  $P_k(x_k, y_k, z_k)$ , 具有明显的几何特征, 可以近似看做一个椭球体用以下方程表示:

$$\frac{x_k^2}{a^2} + \frac{y_k^2}{b^2} + \frac{z_k^2}{c^2} = d \quad (2)$$

式中:  $(a, b, c)$  为平面法向量, 且  $a^2 + b^2 + c^2 = 1$ ,  $d$  为坐标原点到曲面的距离。

假设要提取的目标点云为  $R = \{x_k, y_k, z_k\} (k = 1, 2, \dots, n)$ ,  $n$  为点云总数, 基本矩阵  $M$  满足公式:

$$\begin{cases} [x_k & y_k & z_k & -1]M = 0 \\ M = [a & b & c & d]^T \end{cases} \quad (3)$$

由式(2)、(3)可以得出, 基本矩阵具有 3 个自由度, 因此需要至少 3 个数据点才能计算。所以, 为了较大提高点云数据计算效率, 最大限度提高分割准确性, 本文直接利用随机选出的 3 个原始数据点作为内点得到参数初始值, 然后寻找点云集合中其他的内点。本文改进了文献中的判断准则选取办法, 通过计算圣诞老人点云数据  $P_k(x_k, y_k, z_k)$  到曲面  $Q(a, b, c, d)$  的欧氏距离和引入  $r$  半径密度来设置判断内外点的条件。

$$d(P_k, Q) = |ax_k + by_k + cz_k - d| \quad (4)$$

本算法局内点与平面之间的距离不是理论上的零值, 通过设定合适的容忍值阈值  $\delta$  近似拟合目标曲面。

#### 3) 分割算法实现

第 1 步, 从采集到的目标数据中随机选择 3 个点, 接着

计算初始设定的模型参数并设置目标点云数据中所有点距离阈值, 如果模型与点之间的距离小于该阈值, 那么此点即是局内点, 不然为局外点。第 2 步, 计算局内点的具体个数, 判断该值是否大于设定的阈值, 如果大于设定值, 则用内点重新估计模型, 作为模型输出, 并且保存所有内点作为分割的结果, 如果小于, 则与当前最大局内点个数进行对比, 大于的话则取代当前最大局内点个数, 存储当前的模型系数。第 3 步, 进行迭代计算, 直到分割出有效目标点集。

### 1.2 K-近邻点云移除离群点

通过建立目标点云间的拓扑关系后, 对其 K-近邻点云进行搜索, 利用邻域平均法进行离群点去除, 具体算法描述如下。

1) 读取进过背景分割后的三维点云;

2) 根据文献[4]提出的方法, 建立目标点云拓扑关系;

3) 搜索目标点云中每一点  $q_k \in P_k$  的 K-近邻  $Q_k(q)$ ;

4) 描述目标点  $q_k$  与其 K-近邻范围内各点的相对距离, 求取平均值,  $D_{ave}(q_k)$ , 即:

$$D_{ave}(q_k) = \frac{1}{k} \sum_{q_i \in h_{q_k}} \|q_i - q_k\| \quad (5)$$

5) 比较判断式(4)与设定阈值  $D_{reg}$  的大小, 若大于  $D_{reg}$ , 则判断该点为离群噪声点删除;

6) 重复上述过程, 并结合双边滤波算法对目标点云进行全部处理。

### 1.3 体素化网格下采样

#### 1.3.1 算法原理

根据上述算法, 得到了去除掉背景等噪声的圣诞老人点云数据, 接着采用体素化网格下采样对该点云进行简化。本步算法首先创建较小的体素栅格, 然后利用划分的三维栅格求解边长  $a$ , 根据  $a$  的值将立方栅格分成  $x \times y \times z$  个小栅格。此时, 点云数据已经分布到相应的小栅格中, 删除不包围目标点的栅格, 保留与邻近栅格重心的数据点, 删除较远的点。

#### 1.3.2 算法实现

1) 确定栅格边长  $a$

$a$  值大小与邻近点个数  $k$  正相关, 与其邻近点云的平均密度  $\rho_{q_i}$  负相关。  $a$  值过大, 搜索效率低,  $a$  值过小, 栅格有效率低, 因此将边长增加  $\Delta d$ , 即:

$$\begin{cases} a_x = (x_{\max} - x_{\min}) + \Delta d \\ a_y = (y_{\max} - y_{\min}) + \Delta d \\ a_z = (z_{\max} - z_{\min}) + \Delta d \end{cases} \quad (6)$$

得到的小立方体栅格边长为:

$$a = \lambda \sqrt[3]{sa_x a_y a_z / N} \quad (7)$$

式中:  $\lambda$  是比例因子,  $s$  是比例系数,  $N$  为目标点的总数。

2) 划分目标点云

将目标数据划分为  $x \times y \times z$  个立方体栅格, 其中  $x = \text{ceil}(a_x/a)$ ,  $y = \text{ceil}(a_y/a)$ ,  $\text{ceil}(x)$  为取整函数,  $z = \text{ceil}(a_z/a)$ 。获取点  $q_k$  被包围的栅格编码  $(x_{q_k}, y_{q_k}, z_{q_k})$ 。

3)求解数据点  $q_k$  的邻近点

根据栅格编码,计算数据点到点  $q_k$  的欧式距离,将与点  $q_k$  最近的前  $K$  个数据点存入到哈希链表中作为数据点  $q_k$  的  $K$  邻域点。

4)估算法向量

获取的目标点云表面通过采用最小二乘法实现曲面的拟合。

$$H(n,d) = \operatorname{argmin}_{(n,d)} \sum_{i=1}^k \theta(\|q_i - q\|)(n \cdot q_i - d)^2 \quad (8)$$

式中:  $\theta$  表示高斯权重,  $n$  表示平面  $H$  的法向量,满足  $\|n\|^2 = 1$ ,  $d$  表示平面  $H$  与坐标原点之间的距离。

5)目标点云下采样

求出每个数据点  $q_k$  的法向量与其邻近点法向量之间的夹角,设置阈值,选取小于该阈值的数据特征点,剔除最远离重心  $(X_c, Y_c, Z_c)$  的数据点,实现点云下采样。

2 实验结果与分析

2.1 实验环境与数据采集

实验的硬件环境为 Intel i5 双核 2.9 GHz CPU、8 GB 内存、1 GB 显存华硕 GTX570 显卡的 PC;软件环境为 Windows 7 操作系统、Microsoft Visual Studio 2010 开发环

境、Point Cloud Library 开源编程库、以及 MATLAB 2013a 等。

本课题应用的是 TOF 深度传感器,该设备采用时间飞行法(time-of-flight, TOP)技术获取目标深度数据,它有  $640 \times 480$  的像素分辨率、30 帧/s 的空间目标三维点云成像优势。以色彩分布均匀的圣诞老人为研究对象,在白天将圣诞老人放置在室内实验台上,用 TOF 成像系统采集数据。考虑到 TOF 深度传感器的探测距离为  $0 \sim 13$  m,为保证获取点云数据的精度及完整性,将 TOF 三维点云数据采集点设置为距圣诞老人 50、100、150、200 mm,每个置点采集 10 帧三维点云数据,并存储为 PCD 文档,用于后续处理。

2.2 本文算法实例与分析

如图 1(b)所示,该点云数据包括门窗、箱子等无效背景。实验中取  $\delta=0.075$  m,  $r=1$  m。根据文献[8]可知,公式中取  $\epsilon=0.25$ ,当所有取样点集中至少出现一个点云是局内点的几率大于 99%时,进行搜索迭代的次数应为 1 000 次。如图 2 所示,主要的特征已经被分割出来,但在圣诞老人上存在过度分割的现象和未识别的噪点。本文方法整个处理过程耗时 15.23 s,相比于对本数据运用欧氏聚类分割法[9]进行分割时,耗时为 20.46 s,本方法分割效率占有明显优势。

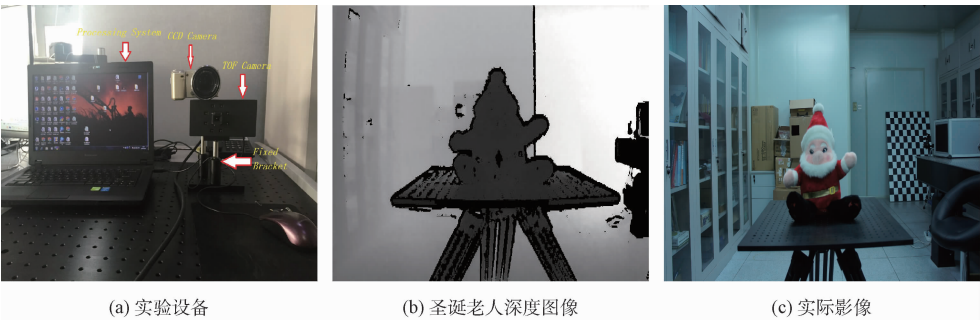


图 1 实验环境与数据

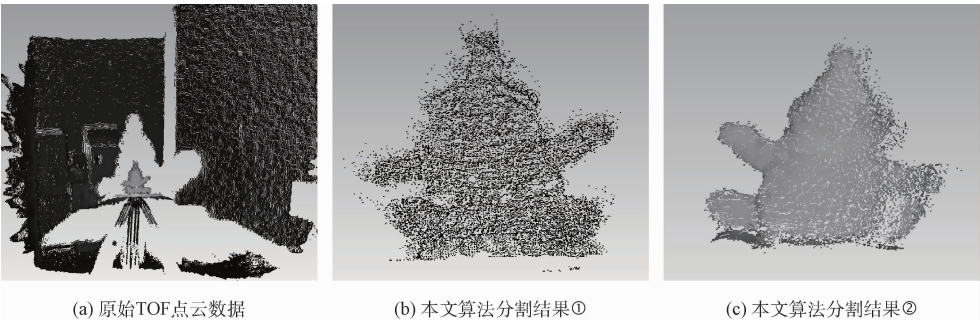


图 2 改进 RANSAC 算法分割结果

如图 3(a)所示,对 TOF 深度传感器获取的圣诞老人点云数据进行分割后的结果,选取不同的  $K$  值进行邻域平均法和双边滤波[10]去噪。去噪效果分别如图 3(b)和(c)所示,

图 3(a)中椭圆区域均为分割后未被识别无效噪声点。获取的点云进行 RANSAC 背景分割,然后基于 K-近邻和双边滤波移除离群点,最后通过 ICP 算法进行多视角

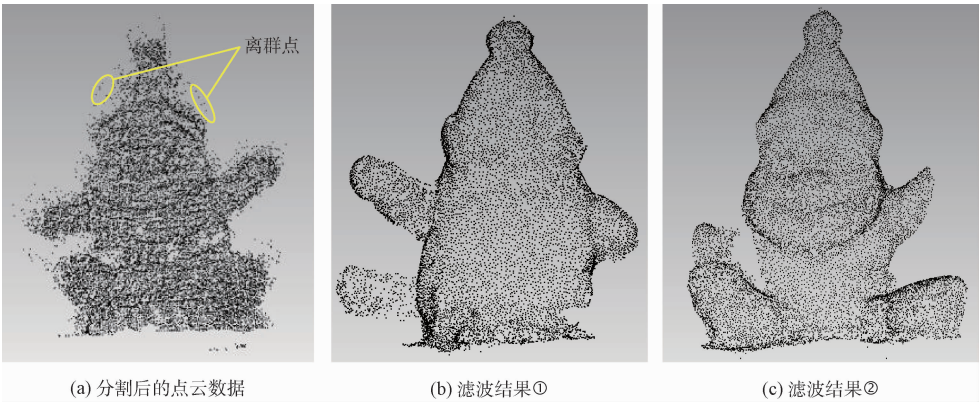


图 3 K-近邻与双边滤波结果

点云的配准,通过对重建后的圣诞老人三维点云数据模型来进行体素化网格下采样。图 4(a)所示为分割配准后的三维初始点云,对应的调用 Delaunay 算法<sup>[11]</sup>重建的效果见图 4(b),图 5(a)~图 7(a)是通过本文算法对目标初始配

准结果进行单位统一化为毫米后的栅格边长分别为  $a=5$ 、15、25 mm 3 种栅格下采样后三维点云图,对应调用 Delaunay 算法重建的效果图是图 5(b)~图 7(b)。

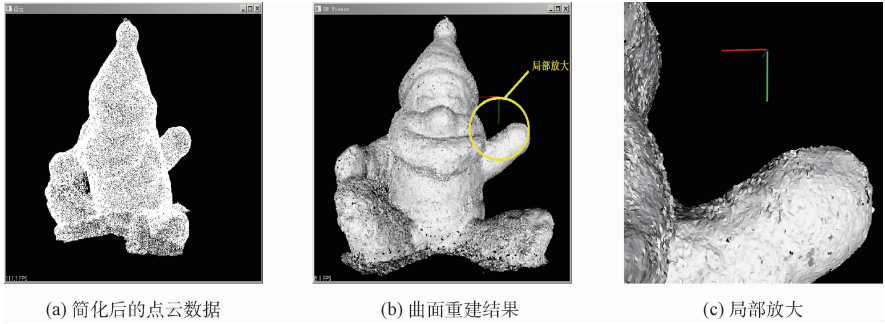


图 4 圣诞老人初始点云

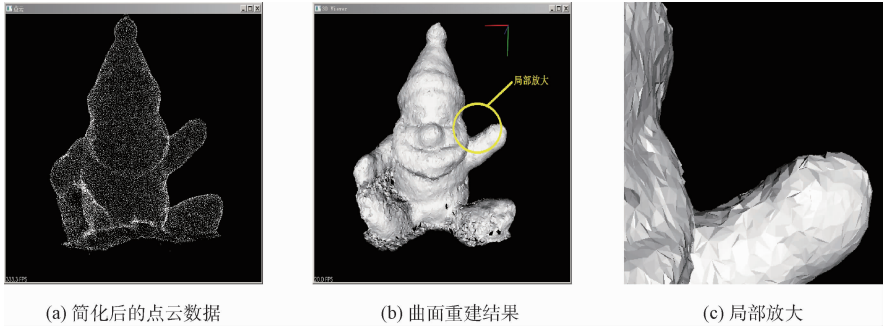
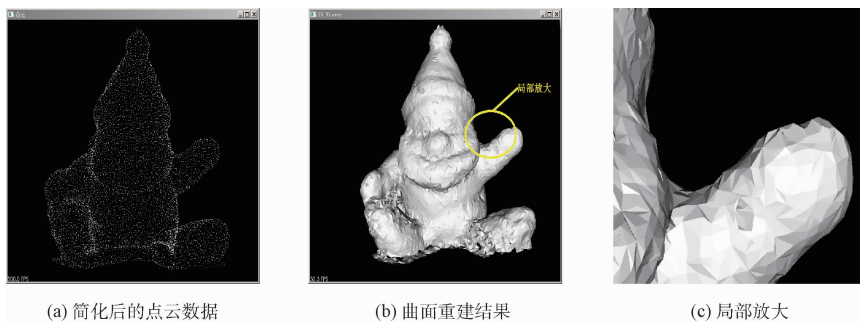
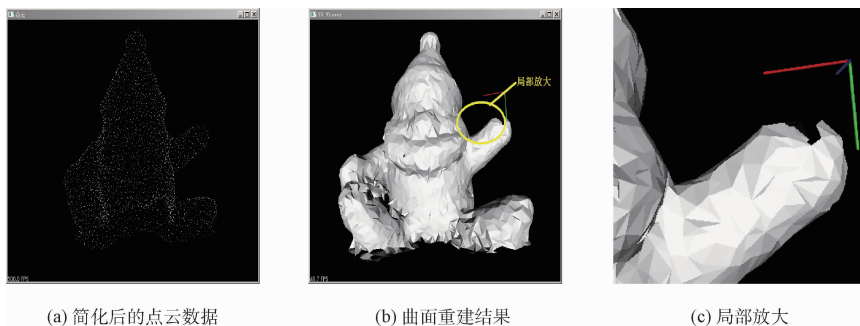


图 5  $a=5$  mm 效果

由实验结果可以看出,通过本文算法后,目标物体的采样点集密度与各局部特征有关,如三维模型表面的曲率、夹角等。从圣诞老人点云的曲面重建图中可以看出,配准后的圣诞老人经过本文算法下采样后,各曲面夹角处点云依然密集的,该部位特征点较多,点云曲率变化明显;而在特征点比较少的腹部与脸部等部位,与原始点云相比非常稀疏。从本算法的去噪效果来看,点云体素化网格下采样不仅可以去除大量的无效点,还可以较好的保留初始

三维点云模型的局部特征,减少目标特征信息的噪点。另外,栅格调节系数  $s$  影响了本算法的执行效率,当  $s$  的值大时,对应栅格的边长就越大,栅格内的目标点数变大,估算时间每个点的法向量时间就越长,导致本算法简化目标模型的时间变大。从重建效果看,不同的调节系数下都能较好的保持原始点云的曲面特征和拓扑结构不发生较大变化。而图 7 体素栅格边长选得过大,圣诞老人手部点云存在空洞,部分细节缺失,剩余点云不能完整体现



图6  $a = 15\text{ mm}$  效果图7  $a = 25\text{ mm}$  效果

原始点云的特征。

图5与6的对应的简化率分别为56.7%和43.2%，而图7的简化率为21.3%，且图7的简化时间最小为0.236 1 s。因此，应根据不同的模型需求，综合模型效果、简化时间、重建时间等因素综合考虑栅格边长的取值范围。

### 3 结 论

本文通过分析目前点云去噪现状，提出一种基于RANSAC分割的点云数据K-近邻去噪算法。通过TOF深度相机对特征复杂的圣诞老人模型进行数据采集，然后利用本文算法进行无效背景分割、离群点去除、最后利用体素化网格法对经ICP算法配准的点云进行下采样简化，滤除了无效的噪点。结果表明，本文算法可以快速识别去除离群噪声数据，能较好的保留点云数据的特征信息，具有较强的计算效率和鲁棒性，可对逆向工程、虚拟现实等邻域关键技术起到辅助作用。

### 参考文献

- [1] 邓文君, 叶景杨, 张铁. 面向机器人磨抛的激光点云获取及去噪算法[J]. 光学学报, 2016(8):172-180.
- [2] 李彩林, 郭宝云, 季铮. 多视角三维激光点云全局优化整体配准算法[J]. 测绘学报, 2015, 44(2): 183-189.
- [3] 崔微, 朱民耀, 颜柯, 等. 基于Kinect的三维重建技术在三维显示中的应用[J]. 电子测量技术, 2017, 40(3):113-116.
- [4] 苏本跃, 马金字, 彭玉升, 等. 基于K-means聚类的

RGBD点云去噪和精简算法[J]. 系统仿真学报, 2016, 28(10):2329-2334.

- [5] 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.
- [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(s2): 351-354.
- [7] 余鹏磊, 于海洋, 谢秋平, 等. 基于RANSAC算法的地铁隧道点云数据滤波[J]. 测绘工程, 2014, 23(8):28-31.
- [8] XU B, JIANG W, SHAN J, et al. Investigation on the weighted RANSAC approaches for building roof plane segmentation from LiDAR point clouds [J]. Remote Sensing, 2015, 8(1):5-5.
- [9] 吴燕雄, 李峰, 刘芳, 等. 平滑度欧式聚类算法分割点云数据[J]. 测控技术, 2016, 35(3):36-38.
- [10] 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.
- [11] 张晓蒙, 陆忠华, 张鉴. 基于OpenMP的三维并行Delaunay网格生成算法及实现[J]. 计算机应用研究, 2016, 33(12):3658-3662.

### 作者简介

郭宁博, 1992年出生, 硕士研究生, 主要研究方向为数字图像处理技术等。

E-mail: 3202569094@qq.com

何艳华, 1988年出生, 助理工程师, 主要研究方向为通信与测控技术、数字信号处理。