

基于无人机的道路点云数据分割提取算法^①



骆 磊, 马荣贵, 薛 昊

(长安大学 信息工程学院, 西安 710064)

通讯作者: 骆 磊, E-mail: 645471524@qq.com

摘 要: 针对机载 LiDAR 获得道路的数据信息精确度低问题, 提出基于无人机的低空扫描三维点云数据, 动态拟合提取分割道路信息的算法. 首先使用主成分分析法获得道路点数据的法向量, 之后将高程信息和法向量信息结合, 利用聚类算法获得道路的高程和法向量的范围, 提取道路点云数据; 其次利用多项式拟合对道路数据进行数学建模; 然后通过动态多项式拟合提取出所有路面数据和路面上的资产以及行人车辆数据; 最后使用区域生长算法对路面上的资产以及行人车辆数据进行分割. 实验表明算法对道路上的遮挡物有很强的抗干扰能力, 可以将路面提取出来并将路面上的数据分割进行分割, 将本文算法与区域生长算法进行对比, 本文算法对路面数据更加敏感.

关键词: 点云处理; 点云分割; 聚类算法; 多项式拟合; 区域生长; 主成分分析; 道路提取

引用格式: 骆磊, 马荣贵, 薛昊. 基于无人机的道路点云数据分割提取算法. 计算机系统应用, 2020, 29(2): 169-174. <http://www.c-s-a.org.cn/1003-3254/7277.html>

Road Point Cloud Data Segmentation and Extraction Algorithm Based on UAV

LUO Lei, MA Rong-Gui, XUE Hao

(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: Aiming at the low accuracy of road data information obtained by airborne LiDAR, an algorithm for extracting and segmenting road information by dynamic fitting based on low-altitude scanning three-dimensional point cloud data of UAV (Unmanned Aerial Vehicle) is proposed. Firstly, the principal component analysis algorithm is used to obtain the normal vector of road point data. Then, combining elevation information with normal vector information, the range of road elevation and normal vector is obtained by clustering algorithm, and the road point cloud data is extracted afterwards by range. Secondly, polynomial fitting is used to model the road data. Then the dynamic polynomial fitting is used to extract the data of the whole section of road surface, assets on the road, pedestrian and vehicle data. Finally, the region growth algorithm is used to segment the assets and pedestrian vehicle data on the road surface. The experiment shows that the proposed algorithm has a strong anti-interference ability to block objects on the road. It can extract the road surface and segment the data on the road surface. The proposed algorithm in this study is more sensitive to the road surface data by comparing with the region growth algorithm.

Key words: point cloud processing; point cloud segmentation; clustering algorithm; polynomial fitting; regional growth algorithm; principal component analysis; road extraction

LiDAR 测量技术现在广泛用于三维数字城市建设^[1-3]、道路安全设施检测^[4,5]、公路勘察设计等方面^[6,7], 车载

LiDAR 测量具有数据采集的高精度、高效率和高密度等特性^[8,9], 但是车载 LiDAR 移动速度较慢, 影响道路

① 基金项目: 陕西省技术创新引导专项 (2018XNCG-G-16)

Foundation item: Technology Innovation Guiding Project of Shaanxi Province (2018XNCG-G-16)

收稿时间: 2019-07-02; 修改时间: 2019-07-23, 2019-08-14; 采用时间: 2019-08-22; csa 在线出版时间: 2020-01-16

使用,而且当路况不好的时候,车辆无法驶入,对于特殊状况想要对道路进行检测则力不从心;机载 LiDAR 测量则是快速度、高效率,受天气、路况影响小,数据生产周期短等特性,该测量方式可分为高空扫描和低空扫描,相比于高空扫描,低空扫描点云数据间隙小,精度更高^[10],更适合用于对道路的养护。因此本文使用低空飞行的无人机对特定的道路进行扫描。

道路数据的提取方法,文献[9]中提出了使用曲线拟合道路数据的算法,但是只将路面的坡度走势进行拟合,并不能提高路面的提取精度,文献[11]提出以点云的区域分布特性为基础,通过对数据进行主成分分析(PCA),构建点云平面基元检测的新模型,虽然能将平面进行分割,但是由于道路细节近似拱形,所以效果并不理想,文献[12]使用种子区域生长算法提取道路,但是需要手动确定种子点,而且提取精度较粗,不能将路面的坑槽有效提取。

综上所述,通过对上述算法的汲取和改进,主要创新部分在于使用了聚类的算法来确定道路数据的种子点,改进了半自动确定道路的方法,并且用动态拟合的方式提取道路。

本文使用基于无人机的道路点云分割提取算法,主要内涵在于无人机低空扫描提高数据精度和动态拟合分割提高提取精度,其整体流程如图1所示,使用无人机低空扫描路面获取精度较高的点云数据,计算点云数据的法向量,丰富数据信息,为后续的聚类算法提供多维样本数据,之后使用聚类算法根据点云数据的高程信息和法向量信息确定种子数据,再对断面进行拟合,利用数据动态拟合分割提取路面数据和非路面数据,最后对非路面数据使用区域生长算法进行分割。

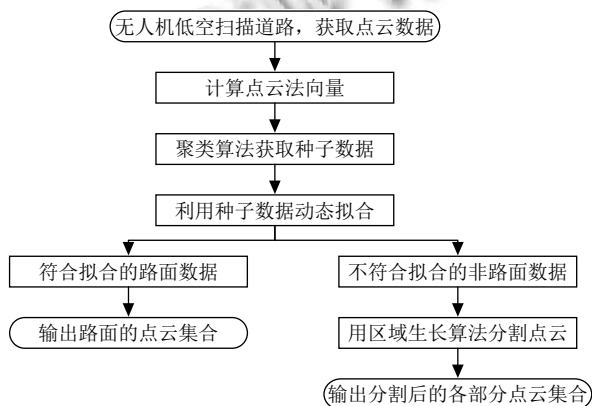


图1 算法流程图

1 丰富点云数据信息

本文采用无人机装载激光扫描设备的方法,无人机低空沿道路方向飞行,扫描指定路段,为方便后续说明,记 y 轴方向为道路方向, x 轴方向为道路横截面方向, z 轴为道路高程信息,道路三维点云数据作为集合 P 。

后续聚类算法需要使用点云的法向量信息,就要确定其法向量,本文使用主成分分析法(Principal Component Analysis, PCA)^[13-17],用来计算点云数据的法向量。对于道路的点云数据集合 P ,由于它是无序的,所以在对集合 P 进行处理之前,需要为其建立拓扑关系,本文使用 K-D 树(K-Dimensional 树的简称)^[18-20]数据结构建立点云之间的拓扑关系,之后再使用 PCA 计算法向量。PCA 计算方法如下:

设集合 P 中的点 p_i 的 k 邻域点集为 $N_k(p_i)$,构造关于点 p_i 的 3×3 协方差矩阵如式(1)所示:

$$C = \begin{bmatrix} p_1 - \bar{p} \\ \vdots \\ p_k - \bar{p} \end{bmatrix} \begin{bmatrix} p_1 - \bar{p} \\ \vdots \\ p_k - \bar{p} \end{bmatrix}^T, \quad \text{其中 } \bar{p} = \frac{1}{k} \sum_{p_j \in N_k(p_i)} p_j \quad (1)$$

矩阵 C 表示了点集 $N_k(p_i)$ 的点的分布情况,即邻域点与质心点 p 的偏离程度, p_k 为 k 邻域点集中的点, p 为数据点 p_i 及其 k 邻域所构成集合的质心。

矩阵 C 为对称半正定矩阵,有 3 个非负实数特征值,设为 $\lambda_0, \lambda_1, \lambda_2$,且 $\lambda_0 \leq \lambda_1 \leq \lambda_2$,其对应的特征向量分别为 v_0, v_1, v_2 。最小特征值 λ_0 对应的特征向量 v_0 即可近似为点 p_i 的法向量 n_i ,为方便表示,记单位法向量 $n_i = (x, y, z)$ 。

2 路面及资产分割提取

2.1 多滑动窗口的均值漂移聚类算法确定种子数据

利用聚类算法提取道路信息,存在 3 个问题,一是由于激光在扫描的时候存在噪声,所以并不是所有的点的法向量都接近聚类中心,结果只能提取道路上的部分数据;二是如果对所有断面都采用聚类算法,虽然可以提取道路,但是非常耗时;三是针对路面上高程信息不明显的的数据,同样需要提取,而聚类算法并不能提取出来。

根据上述问题,本文结合数据高程信息和法向量作为聚类算法的样本点,多滑动窗口的均值漂移聚类算法^[21-23]确定拟合道路的种子数据,之后利用动态拟

合提取所有路面。

因为高程信息与法向量的信息比例不同, 所以需要将高程数据范围限定在 0 到 1, 则使用高程数据在高程区间的相对值 h 做样本空间的高程信息, h 值的计算公式如式 (2) 所示:

$$h_i = \frac{z_i - \min(Z)}{\max(Z) - \min(Z)}, \quad h_i \in H \quad (2)$$

其中, z_i 为当前点的高程信息, Z 为所有数据点的高程信息集合。由式 (1) 得到数据点 p_i 的法向量 $n_i=(x, y, z)$, 将这两类数据投到特征空间, 则点 p_i 在特征空间的向量为 $q_i=(x, y, z, h)$ 。

得到标准化后的有 n 个样本点的特征空间 $Q=(q_1, q_2, \dots, q_n)$ 后, 使用多滑动窗口的均值漂移算法确定中心点, 在特征空间 Q 中随机选取 10 个点作为初始中心点, 其滑动窗口半径为 r , 半径 r 的取值越大则对路面数据的要求越宽泛, 越小则越严格, 本文依据数据点到距离其最近的 10 个点的距离来确定 r 值, 公式如式 (3):

$$r = \frac{1}{n} \sum_{i=1}^n \max(S_i) \quad (3)$$

其中, S_i 为数据点 q_i 的最近邻 10 个样本点的欧氏距离的集合。

由式 (3) 得到高维区域的欧氏距离半径, 由高维区域内的所有样本点可以确定各个窗口的偏移均值, 其计算公式如式 (4) 所示:

$$M(c_i) = \frac{1}{m_i} \sum_{q_j \in S_r} (c_i - q_j), \quad i = 1, 2, \dots, 10 \quad (4)$$

其中, c_i 为窗口中心, S_r 为以 c_i 为中心, 半径为 r 的高维区域, q_j 为区域内的点, m_i 为区域内 q_j 的数量, $M(c_i)$ 为以 c_i 为窗口中心的偏移均值。

根据式对 10 个滑动窗口进行中心更新, 更新公式如式 (5) 所示:

$$c_i^{t+1} = M_i^t + c_i^t \quad (5)$$

其中, M_i^t 为 t 状态下求得的偏移均值, c_i^t 为 t 状态下的中心, c_i^{t+1} 为 $t+1$ 状态下的中心。

当中心点不再变化时, 停止更新滑动窗口的中心, 则高维区域内点数最多的两个窗口内的点分别为人行道和机动车道的数据点。通常, 人行道的高度要高于机动车路面, 因此两个密度中心可以用高程信息来区分, 高程信息低的是机动车路面, 高的是人行道路面。

2.2 对路面数据动态拟合

通过对道路数据点的观察并结合实际, 机动车路面一般是中间稍高于两侧, 这样利于排除雨水。同样, 两边的人行道, 也存在一定弧度。可以将路面粗略描述为光滑曲面, 针对这种形状, 采用四次多项式拟合路面横断面曲线, 使用最小二乘法计算多项式系数。

拟合函数定义如下:

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \quad (6)$$

其中, a 为多项式系数, x 为路面横截横坐标。由式 (6) 得到均方误差, 均方误差如式 (7) 所示:

$$Q(a_0, a_1, a_2, a_3, a_4) = \sum_{i=1}^n (a_0 + a_1x_i + a_2x_i^2 + a_3x_i^3 + a_4x_i^4 - y_i)^2 \quad (7)$$

为求式 (7) 的极小值, 得到方程如式 (8) 所示:

$$A^T A \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_4 \end{pmatrix} = A^T \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \text{其中 } A = \begin{pmatrix} 1 & x_1 & \cdots & x_1^4 \\ 1 & x_2 & \cdots & x_2^4 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^4 \end{pmatrix} \quad (8)$$

解方程得到 a_0, a_1, a_2, a_3, a_4 带入式 (6) 可以得到当前断面的拟合函数 $f(x)$, 用当前断面的拟合函数 $f(x)$ 判断下一断面的点是否符合式 (9)。

$$|f(x) - y| < a \quad (9)$$

符合式 (9) 的则为道路数据点, 不符合的则为非道路数据点。

$$a = 2d + \frac{v}{n} \times 11\% \quad (10)$$

$$d = \frac{1}{n} \sum_{i=1}^n \max(D_i) \quad (11)$$

其中, a 的计算公式如式 (10) 所示, 因为需要考虑道路的坡度问题, 根据《城市道路工程设计规范》内容, 坡度最大不会超过 11%, 所以对提取范围修正。 d 由式 (11) 计算求得, D_i 为数据点 p_i 的最近邻 10 个点距离的集合, v 为无人机的移动速度, n 为激光扫描仪每秒扫描的次数, 11% 为国家规定公路最大坡度。

所谓动态拟合, 是将所有符合式的道路数据点进行如式拟合, 得到新的拟合函数 $f(x)$, 对下一断面进行重复操作, 达到提取所有道路数据点的目的。

在利用拟合曲线提取机动车道和人行道路面数据时, 为避免使用拟合曲线时将周围环境信息误提, 需要

对拟合曲线的区间进行约束。

通常情况下可以利用上一层断面的数据确定机动车路面和人行道路面的范围,但是,在有遮挡物的情况下计算的范围会受到遮挡物影响,而且不止会使当前路面的提取受到影响,还会波及后面的计算,所以,准确的计算出路面的范围是非常有必要的.针对这种情况,提出基于宽度的路面范围修正算法。

根据现实考虑,因为左右两边人行道同时边沿被遮挡的几率很小,所以可以利用左右两边人行道的宽度比修正左右两边,利用靠墙点作为基准对宽度进行修正,设计修正算法如下:

算法 1. 路面区间修正算法

- 1) 记前一断面左边人行道数据集 Q_{left_old} , 右边人行道数据集 Q_{right_old} ;
- 2) 根据集合 Q_{left_old} , Q_{right_old} 计算对应的宽度 W_{left} , W_{right} ; 左边 x 区间的最小值 $x_{leftmin}$, 右边 x 区间的最大值 $x_{rightmax}$;
- 3) 判断左右宽度的比值 ω 确定是否需要修正, 如果 $\omega > \omega'$, 则认为需要对右边进行修正, 如果 $\omega < -\omega'$, 则认为需要对左边进行修正, 利用式计算 x , 之后使用人行道的拟合函数 $f(x)$ 作为修正的 y .

其中, ω' 的值为两边人行道宽度的敏感程度, 越小越敏感, 越大越不敏感, 这里取 ω' 为 0.3, ω 的确定公式如式 (12) 所示, x 的计算公式如式 (13) 所示:

$$\omega = \frac{(W_{left} - W_{right})}{\max(W_{left}, W_{right})} \quad (12)$$

$$x = \begin{cases} x_{leftmin} + W_{right}, & w < -\omega \\ x_{rightmax} - W_{left}, & w > \omega \end{cases} \quad (13)$$

经过区间约束, 和路面拟合, 就可以利用上一断面的拟合函数和区间约束来提取当前路面。

取某一断面数据为例, 说明四次多项式可以很好的描述路面横断面的曲线. 做出如表 1 所示的以确定系数 R-square 作为对比标准的拟合效果对比表:

通过表可以看出, 无论是路面还是左右两侧的人行道, 当拟合次数达到四次时, 都可以有很好的拟合效果, 而当拟合次数达到五次时, 虽然略有提高, 但是变化不大, 在满足拟合精确的前提下, 为保证算法运算速度, 防止出现过拟合现象, 确定使用四次拟合。

表 1 拟合效果对比表

名称	一次拟合 R-square	二次拟合 R-square	三次拟合 R-square	四次拟合 R-square	五次拟合 R-square
路面	0.088 82	0.7434	0.7772	0.9468	0.9472
左侧人行道	0.8221	0.898	0.9008	0.9593	0.9595
右侧人行道	0.8379	0.9521	0.953	0.9387	0.9762

使用四次多项式拟合得到的拟合曲线, 为使图像明显, 对点云数据的横坐标比例进行了缩放, 如图 2(a)、图 2(b) 所示, 可以看到曲线在点云区间拟合效果很好。

2.3 改进区域生长算法对非路面数据分割

通过上述动态数据拟合的算法, 可以将路面数据提取出来, 达到将路面数据与非路面数据的分割, 对路面与非路面数据分割后, 再使用基于改进的区域生长三维点云分割算法^[24]将非路面的各部分分别提取. 其过程为先计算各个点的平均曲率 K_n , 计算公式如式 (14) 所示:

$$K(p_i) = \frac{1}{4A_{min}} \times \sum_{j \in N(i)} (\cot \alpha_{ij} + \cot \beta_{ij})(p_i - p_j) \times n \quad (14)$$

式中, n 为法向量, A_{min} 为 p 周围一个无限小的区域, α_{ij} 、 β_{ij} 分别为连接 p_i 和 p_j 边的对角. 选取点云数据中曲率最小的点设置为种子点。

因为路面上的行人, 车辆以及资产之间距离较大, 所以仅用距离作为阈值就可以将各部分分割, 选用式

(11) 的 d 作为阈值, 判断邻近点与种子面之间的生长半径是否小于阈值 d , 将小于阈值的点添加到当前区域, 当再无点可以添加时, 则开始分割下一区域。

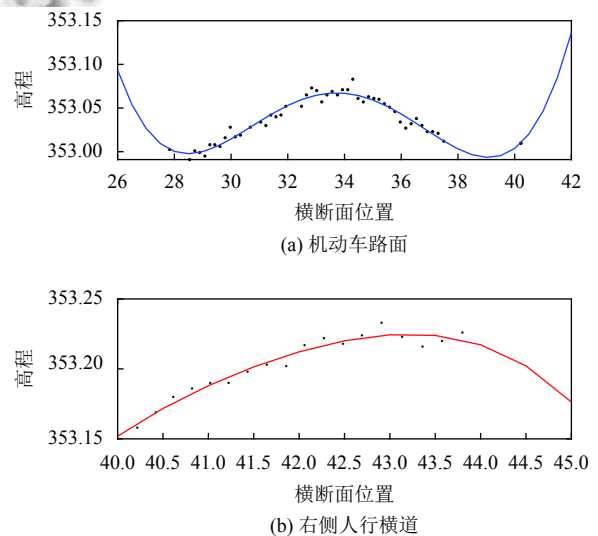


图 2 道路数据拟合图

3 实验结果分析

对西安市神州六路部分路段作为无人机数据获取的实验区,利用无人直升机 Scout B1-100 搭载 RIEGL VUX-1HA 激光扫描仪系统对实验区进行扫描,获取道路点云数据,激光扫描仪的基本参数如表 2 所示。

表 2 激光扫描仪参数表

扫描仪参数	规格
最大测距	420 m
精度/重复精度	5 mm/3 mm
激光脉冲发射频率	1000 kHz
最大扫描速度	250 线/s

将本算法与区域生长提取路面算法对比,验证算法的效果。该路段效果如图 3 所示,红色部分为机动车道,绿色部分为人行道,蓝色部分为路上资产和路面坑包,黑色部分为路边环境。

从图 3 中可以看出,两个算法很好的将道路上的

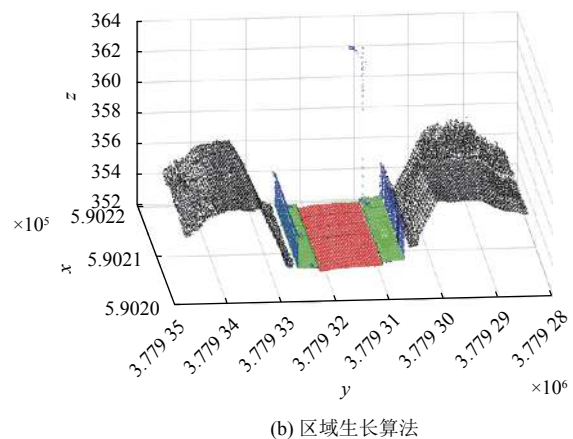
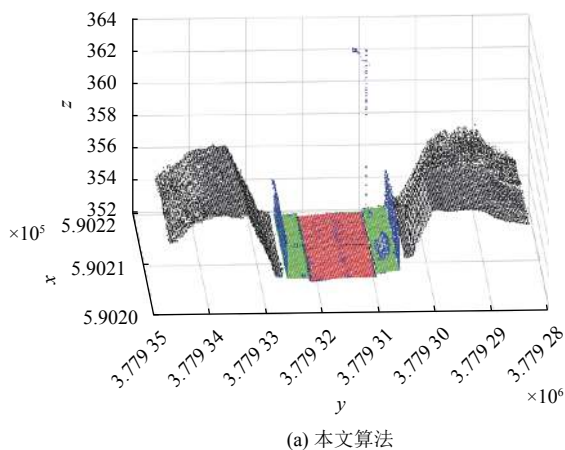


图 3 算法对比

表 3 算法效果对比表

场景	本文算法	传统区域生长提取路面算法
土包	1	0
树坑	3	3
凹槽	2	0

综上所述,本文算法可以很好的将路面数据进行识别,将路面上的资产进行提取,相比于传统的使用区域生长算法识别路面,本文算法对路面细节更加敏感。所以在路面养护和路上资产管理方面有很好的效果。

路灯,树坑,墙,蓝色部分两边为遮挡墙,蓝色竖线为路灯,绿色人行道上的几个蓝色小块为树坑。通过图 3(a)路面上有些蓝色点,人行道上有一片土包可以看出,本文算法可以将路面上的一些细节提取出来,而且对于曲率变化较小的土包也可以提取出来,而图 3(b)中的区域生长算法,对路面上的细节却不敏感,而且对于曲率变化较小的土包也无法提取。用识别路面上的树坑,土包,凹槽数量作为对比说明两种算法的效果,列出表 3。

为验证该算法的适用性,对西安市神州六路多处路段进行算法验证,结果如图 4 所示,同样,红色部分为机动车道,绿色部分为人行道,蓝色部分为路上资产和路面坑包,黑色部分为路边环境,对于多处路段,都可以很好的将路面进行提取,并且受周围环境影响较小。当路面上有车辆时,如图 4(b)、图 4(d)无论车辆是在提取路面的初始断面,还是在后续断面,都不会影响到路面的提取。

4 结语

本文在处理道路信息时,比较依赖于扫描的起始位置,虽然对遮挡物的抗干扰能力强,但是,对于遮挡物太多,道路信息过少的情况下,并不能很好的将道路数据的中心通过聚类算法聚类出来;另外,在对数据进行分割的时候,对于杂糅在一起的数据也无法进行分割,后续可以利用聚类算法,或者神经网络对数据进行更好的分割;最后对于和路面的高程数据差过小的数据也无法提取出来,这种淹没在噪声中的有用数据提取难度很大,需要后续仔细研究。

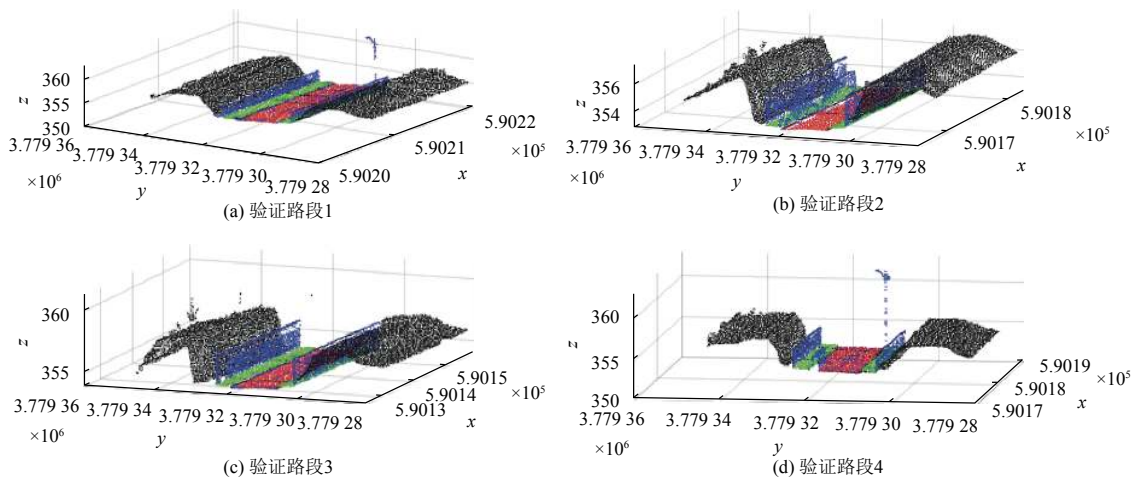


图4 实验结果

参考文献

- 李杨. 无人机摄影测量及在城市规划中的应用. 冶金与材料, 2018, 38(6): 181–182. [doi: 10.3969/j.issn.1674-5183.2018.06.109]
- 胡世敬. 浅谈无人机遥感技术在智慧城市中的应用. 中国新通信, 2019, 21(7): 99. [doi: 10.3969/j.issn.1673-4866.2019.07.081]
- 王浩天. 基于无人机航测技术的城市用地规划模型研究. 北京测绘, 2018, 32(10): 1174–1177.
- 彭检贵, 马洪超, 高广, 等. 利用机载 LiDAR 点云数据提取城区道路. 测绘通报, 2012, (9): 16–19.
- 徐景中, 万幼川, 赖祖龙. 机载激光雷达数据中道路中线的多尺度提取方法. 红外与激光工程, 2009, 38(6): 1099–1103. [doi: 10.3969/j.issn.1007-2276.2009.06.032]
- 尹恒, 裴尼松, 余梨. 无人机技术在复杂公路中的应用研究. 中外公路, 2018, 38(2): 1–5.
- 孙权, 程俊毅, 田绍鸿, 等. 基于无人机 LIDAR 数据多尺度特征的沥青路面病害提取方法. 石河子大学学报 (自然科学版), 2019, 37(1): 1–11.
- 毛寅睿, 陈雨人, 刘洋东. 基于车载激光点云的道路标线提取方法研究. 公路交通科技, 2019, 36(7): 127–135.
- 张志伟, 刘志刚, 黄晓明, 等. 基于 LIDAR 数据的道路平面线形拟合方法研究. 公路交通科技, 2009, 26(12): 17–22. [doi: 10.3969/j.issn.1002-0268.2009.12.004]
- 余飞, 余绍淮, 陈楚江. 机载激光雷达测量技术在高速公路勘察设计中的应用研究. 中外公路, 2016, 36(2): 335–338.
- 李宝顺, 岑红燕, 包亚萍, 等. 基于平面提取的点云数据分割算法. 计算机应用与软件, 2014, 31(7): 145–148, 176. [doi: 10.3969/j.issn.1000-386x.2014.07.037]
- 宋袁龙, 胡洋. 一种机载激光雷达道路点云的半自动提取方法. 测绘科学, 2015, 40(5): 92–94, 121.
- Jafarzadegan M, Safi-Esfahani F, Beheshti Z. Combining hierarchical clustering approaches using the PCA method. Expert Systems with Applications, 2019, 137: 1–10. [doi: 10.1016/j.eswa.2019.06.064]
- 王磊, 何辰, 谢江宁. 基于加权 PCA 分析的三维点云模型对称性检测算法. 山东大学学报 (理学版), 2014, 49(9): 166–170.
- 刘艳菊, 张永德, 姜金刚. 三维点云模糊分类的法向量估值算法. 华中科技大学学报 (自然科学版), 2013, 41(8): 50–54.
- 曹志民, 谷延锋, 吴云. 机载 LiDAR 点云量化局部结构信息分析. 地理空间信息, 2016, 14(2): 10–12. [doi: 10.3969/j.issn.1672-4623.2016.02.003]
- 朱东方. 基于复杂拓扑结构点云的曲线拟合研究与应用 [硕士学位论文]. 济南: 山东大学, 2015.
- Zhang J, Xiu XJ. K-d tree based approach for point location problem in explicit model predictive control. Journal of the Franklin Institute, 2018, 355(13): 5431–5451. [doi: 10.1016/j.jfranklin.2018.05.040]
- Lacerda ASM, Batista LS. KDT-MOEA: A multiobjective optimization framework based on K-D trees. Information Sciences, 2019, 503: 200–218. [doi: 10.1016/j.ins.2019.07.011]
- 梁周雁, 邵为真, 孙文潇, 等. 基于 PCL 的点云数据空间管理及近邻搜索. 北京测绘, 2018, 32(1): 52–57.
- 汪仁红. 基于聚类分析的数据流处理算法 [硕士学位论文]. 重庆: 重庆交通大学, 2013.
- 张睿, 熊金虎, 汪东兴, 等. 基于体素邻域信息的均值漂移聚类算法检测 fMRI 激活区. 江苏大学学报 (自然科学版), 2016, 37(5): 556–561.
- 刘凤梅, 葛洪伟, 杨金龙, 等. 基于均值漂移聚类的扩展目标量测划分算法. 计算机工程, 2014, 40(12): 182–187, 194. [doi: 10.3969/j.issn.1000-3428.2014.12.034]
- 李仁忠, 刘阳阳, 杨曼, 等. 基于改进的区域生长三维点云分割. 激光与光电子学进展, 2018, 55(5): 051502.