

doi: 10.7690/bgzd.2016.07.018

基于二维的地面点云数据组织研究

张有亮, 王洪大, 王东华
(后勤学院模拟训练中心, 北京 100858)

摘要: 点云数据的海量性一直是制约点云数据处理方法的重要因素。为了寻找一种高效空间索引方法来管理海量地面点云数据, 笔者通过分析 LiDAR 地面点云的特点, 提出了一种基于平面直角坐标网和二叉树的二维点云数据组织方法——坐标网-二叉树联合准则法。实验结果证明: 将规则网格与二叉树的方法相结合, 较大幅度地提高了点云数据组织与索引的效率。但是, 格网间距大小的设定还没有定量的标准, 需要在以后实践中进一步进行研究总结。

关键词: 二维; 直角坐标网; 二叉树; 点云数据组织; 索引

中图分类号: TP391 **文献标志码:** A

Research on Organization of Ground Point Cloud Based on 2D

Zhang Youliang, Wang Hongda, Wang Donghua
(Simulation Training Center, Logistics Academy, Beijing 100858, China)

Abstract: The magnanimity of cloud point data is a key restricting factor for its processing. To find out an efficient method for massive ground cloud point management, the authors have analyzed the features of LiDAR cloud point data, and proposed a Cartesian coordinates – quad tree jointed method for 2D cloud point data organization, which is based on plane Cartesian coordinates and quad tree. Experiments show that the proposed method greatly improves the efficiency of data organization and data index. However, it is a future work to quantitatively determine the size of the grid interconnection.

Keywords: 2D; Cartesian coordinates; quad tree; point cloud data organization; index

0 引言

地球空间信息的快速获取和智能化处理是当前测绘领域的研究热点。激光雷达技术有效地拓宽了数据源范围, 改变了数据获取模式, 能快速获取高分辨率数字表面模型。然而, 点云数据的海量性一直是制约点云数据处理方法的重要因素, 迫切需要寻找一种高效空间索引方法来管理海量点云数据。

目前在对点云数据的组织与管理中常见的方式有规则格网、二叉树^[1]、KD 树^[2]、KDB 树^[3]、八叉树^[4]、BSP 树^[5]、R 树^[6]以及扇形网格法^[7]等, 主要是从三维空间对点云进行区域划分。空间数据索引的评价指标有索引构建效率和索引查询效率^[8]。

规则网格法的基本思想是将数据区域划分成大小相等的均匀网格, 通过网格的行列号建立索引, 用户只要判断与索引数据相交的网格, 即可得到所需点集, 其特征是构建数据组织最快。但是, 当激光点云的分布不均匀时, 则不适宜采用规则的分块索引结构。若格网取值过大, 某些单元网格内就会有过多的点, 影响查询效率; 若取值过小, 又会出现过多的空单元, 耗费大量的存储资源。

二叉树采用递归的方法, 根据网格内点云数据量将单元格一分为四, 所有点云被压入一棵二叉树

内, 存储在叶节点。此方法避免了因点云分布不均造成的空间浪费, 但是, 海量点云数据构建的二叉树往往过于庞大, 组织点云数据的时间过长或生成的二叉树过深, 影响索引效率。

笔者通过分析 LiDAR 地面点云的特点, 提出了一种基于平面直角坐标网和二叉树的二维点云数据组织方法——坐标网-二叉树联合准则法。

1 坐标网-二叉树联合准则法

1.1 LiDAR 点云数据特点

1) 地面激光点与水平面投影点一一映射。现实地形(洞穴除外)中, 水平面上一个投影点, 只对应一个地面点, 若存在 2 个或 2 个以上高程不同的点, 则此处存在建筑物、树木或其他地物。LiDAR 点云之间的比较, 只需在投影距离自身较近的区域内进行。基于此, 笔者采用二维的点云数据组织结构。

2) 不均匀性。由于云层的遮挡, 或者水面对激光波束的吸收等原因, 容易造成激光点云数据的空间分布极其不均匀。

1.2 坐标网-二叉树的数据组织

1.2.1 坐标网对点云区域的划分

平面直角坐标网是标志地形图的一种方式, 即

收稿日期: 2016-03-07; 修回日期: 2016-04-27

作者简介: 张有亮(1986—), 男, 河北人, 博士, 从事指挥信息系统、军事运筹学研究。

以公里为单位，按相等的间距作平行于纵、横轴的若干直线，便构成了图平面上的平面直角坐标网，又称方里网。笔者采用这种方法，将点云区域等间距划分，间距的大小取决于点云外接矩形的面积和点云的不均匀程度，拓扑结构及编码方式如图 1。

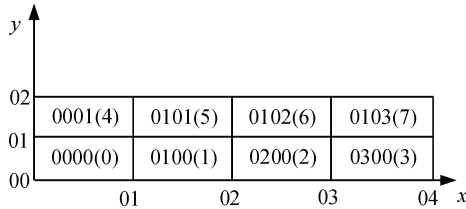


图 1 网格划分及其编码

文中方法对网格采用哈希存储的方式，将关键码值(key value)转换成一维数组 $a[num]$ 的下标 num ， $a[num]$ 为网格中四叉树的根节点，从而直接对其进行访问。转换公式为：

$$i = (\text{int})(\frac{x_i - x_{\min}}{r}); \quad j = (\text{int})(\frac{y_i - y_{\min}}{r});$$

$$y_{\text{sum}} = (\text{int})(\frac{y_{\max} - y_{\min}}{r}) + 1;$$

$$\text{num} = i * y_{\text{sum}} + j. \quad (1)$$

式中： i 为列号； j 为行号； y_{sum} 为一行内单元格数； num 为数组下标； x_{\min} 为点云在 x 轴上的最小值； y_{\min} 、 y_{\max} 分别为点云在 y 轴方向上的最大值与最小值。例如图 1 中将关键码值“0102”转换为数组下标“6”。

1.2.2 网格内创建四叉树

用平面直角坐标网对点云区域进行粗略划分后，再采用递归的方法，在单元网格内建立四叉树，如图 2。

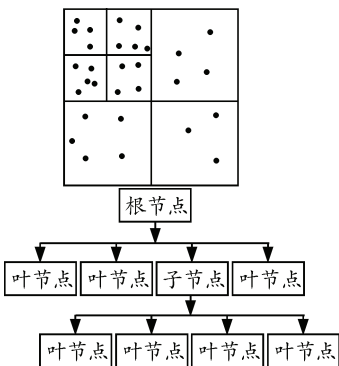


图 2 四叉树划分及拓扑结构

图中，四叉树中每个节点对应一块矩形区域，叶节点中采用链表的形式存储点云，子节点不存储，只作为数据索引的路径。单元网格内创建四叉树与

整个点云文件构建四叉树的过程相同，目前存在几种构建四叉树的方法，如文献[7,9]，原理大同小异。

四叉树构建过程中，叶节点中点数阈值(一般不大于 $\log_2 N$) 的确定也会影响数据组织与空间索引的效率。三维场景显示时，阈值可以适当偏大；而构建格网或曲面拟合需要比较 k 近邻时，阈值可相对偏小。

1.2.3 数据组织的建立

点云数据组织构建如图 3。基本流程如下：

- 1) 得出激光点坐标的最大值与最小值 $x_{\min}, x_{\max}, y_{\min}, y_{\max}$ (对于 LAS 格式的点云数据，可从其文件头直接获取)，确定点云最小外接矩形；
- 2) 输入网格间距 r 以及四叉树叶节点内点云个数阈值 n ；
- 3) 从文件读取一个激光点坐标，如果读取结束，转 7)；
- 4) 根据式 (1) 直接访问激光点所隶属四叉树的根节点；
- 5) 查询激光点隶属的叶节点，将其插入叶节点的点云链表中，判断叶节点中点云数目是否超过给定阈值 n ，超过则转为 6)；
- 6) 节点四均分，将父节点的点撒播到所属子区域，并将父节点链表赋空；
- 7) 文件读取结束，点云数据组织建立完毕。

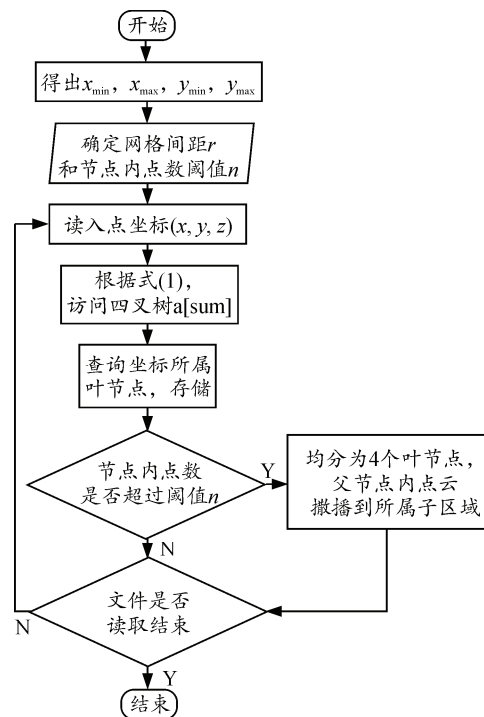


图 3 点云数据组织构建流程

2 k 近邻搜索

散乱点集的曲面重建一般需要点云滤波^[10]、网格简化^[11]、点云特征判断^[12]等处理，而面临的一个首要问题就是搜索距离一点最近的 k 个点，即 k 近邻搜索。

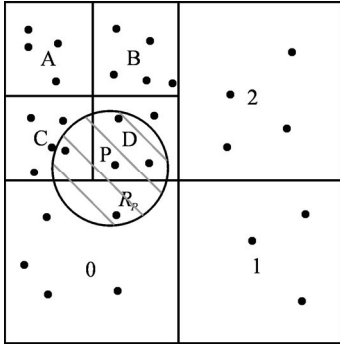


图 4 P 点的 k 近邻搜索

第 1 步，以目标点为圆心，按某一半径画圆，搜索与圆形相交或包含的叶节点，将这些叶节点中的点云作为 k 近邻候选点。搜索过程中，为了提高效率，只需要对部分节点进行判断。

第 2 步，计算 U 中点集与目标点的欧氏距离，选取距离较小的前 k 个点，作为目标点的 k 近邻。

以图 4 点 P 为例，具体说明 k 近邻叶节点搜索

的方法，为了表述方便，将叶节点表达区域记为 0、1、2、A、B、C、D，圆形区域记为 R_p ，将候选节点集合记为 U。

- 1) 将 P 点所在叶节点加入集合 U；
- 2) 判断圆形 R_p 与叶节点 D 的关系，如果 R_p 包含于 D，转 5)；否则判断圆形超出 D 的哪个边界，示例中 R_p 超出下边界和左边界，将节点 C, R 加入集合 U；
- 3) 判断圆形与 C 的关系，超出 C 的下、右边界，R, D 已经属于集合 U；
- 4) 判断圆形与 R 的关系，超出 R 的上边界，C, D 已经属于集合 U；
- 5) 搜索停止。

3 实例验证

以某野外区域地形扫描点云为例，分别采用二叉树数据组织和坐标网—二叉树数据组织进行管理，两者在构建二叉树时采用的方法相同，参考文献[9]，阈值设为 50。试验的硬件配置如下：AMD Athlon(tm) 64 X2 Dual Core Processor 5 200+ 2.71 GHz, 1.00 GB 的内存, WD320GB(5 400 转)。数据结果如表 1、图 5、图 6。

表 1 测试结果对比

点云个数	数据组织时间/ms		定空间索引时间/ms		索引点数
	二叉树	坐标网-二叉树	二叉树	坐标网-二叉树	
1 212 363	1 953	1 422	380	172	606 183
3 122 838	4 656	3 703	711	414	1 561 420
6 556 325	9 344	7 578	942	583	2 185 442
11 368 541	15 598	13 360	1 101	648	2 342 129

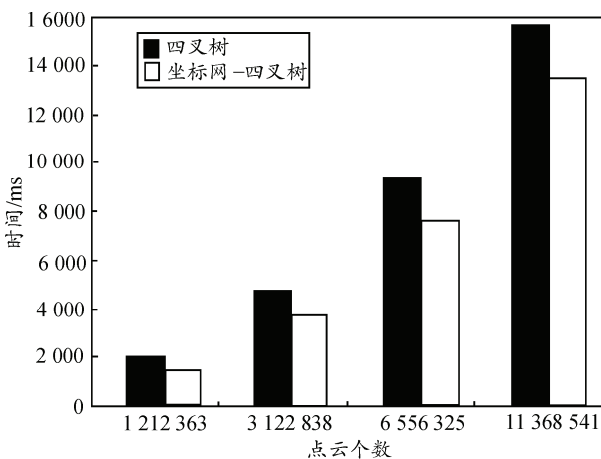


图 5 数据组织时间比较

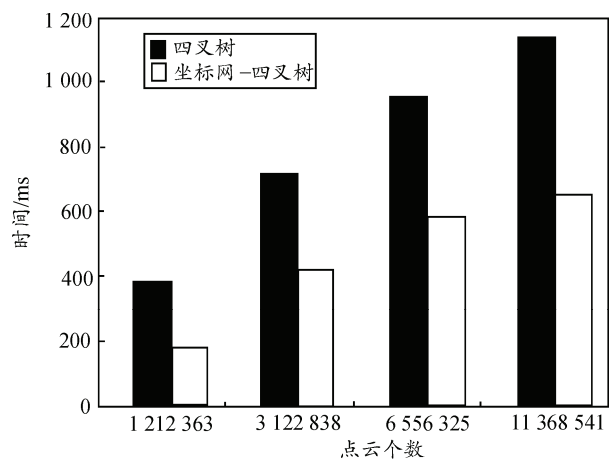


图 6 空间索引时间比较

从对比结果可以看到，文中方法在点云数据组织和空间索引效率方面都有很大程度的提高。取 4 次结果的平均值，数据组织时间提高了约 20%，数据索引时间提高了约 44%，原因在于单纯二叉树方

法将整个点云存储在一棵二叉树中，深度过深，索引路径过长；而文中方法通过点坐标运用式 (1) 迅速定位到网格，再在所属网格的二叉树内进行索引，索引路径短，大大提高了效率。