

基于注意力机制的树木点云补全网络

尤磊^{1,2,3}, 孙毅安¹, 常潇洒¹, 杜黎明⁴

¹⁾ (信阳师范大学计算机与信息技术学院 信阳 464000)

²⁾ (河南大别山森林生态系统国家野外科学观测研究站 郑州 450046)

³⁾ (信阳生态研究院 信阳 464000)

⁴⁾ (中国林业科学研究院资源信息研究所森林遥感技术与应用研究室 北京 100091)

(leiyou@xynu.edu.cn)

摘 要: 受到三维激光扫描仪分辨率限制、环境遮挡等因素的影响, 扫描获取的树木点云通常存在缺失, 尤其是单视角扫描的点云缺失更为严重. 根据树木复杂的几何结构, 提出一种树木点云补全网络. 利用交叉注意力和自注意力机制充分地学习输入树木缺失点云的潜在特征, 并通过解码该特征从稀疏到精细地预测完整树木点云; 针对树木补全数据集难以获取的问题, 采用模拟单视角扫描的方式构建有真值且包括不同树木类型的树木点云补全数据集. 在所构建的数据集上的实验结果表明, 与点云补全网络 AdaPoinTr 相比, 所提网络的平均倒角距离降低 0.62, 平均 F 分数增加 0.04; 该网络可以有效地补全不同种类的缺失树木点云.

关键词: 树木点云; 点云预处理; 点云补全; 注意力机制

中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2024-00044

Tree Point Cloud Completion Network Based on Attention Mechanism

You Lei^{1,2,3}, Sun Yian¹, Chang Xiaosa¹, and Du Liming⁴

¹⁾ (School of Computer and Information Technology, Xinyang Normal University, Xinyang 464000)

²⁾ (Henan Dabieshan National Field Observation and Research Station of Forest Ecosystem, Zhengzhou 450046)

³⁾ (Xinyang Academy of Ecological Research, Xinyang 464000)

⁴⁾ (Forest Remote Sensing Technology and Application Research Laboratory, Institute of Forest Resources Information Techniques, Chinese Academy of Forestry, Beijing 100091)

Abstract: Due to the limitations of three-dimensional laser scanners, such as resolution and environmental occlusion, the scanned point clouds of trees are usually incomplete, especially for single view scans. In light of the complex geometric structures of trees, a tree point cloud completion network is proposed in this paper, which fully learns the potential features of missing point clouds through cross-attention and self-attention mechanisms, and predicts complete tree point clouds from sparse to refined by decoding those features. To address the problem of difficult acquisition of tree completion datasets, a tree point cloud completion dataset with ground truth from different tree type using simulated scans of the single view is constructed in this paper. Experiments on this dataset show that compared to the point cloud completion network AdaPoinTr, the proposed network in this paper reduces the average Chamfer distance by 0.62 and increases the average F -score by 0.04; Thus, the proposed network can effectively complete missing point clouds of trees of different types.

Key words: tree point cloud; point cloud processing; point cloud completion; attention mechanism

收稿日期: 2024-01-21; 修回日期: 2024-08-30. 基金项目: 国家重点研发计划(2022YFF1302100); 国家自然科学基金(31872704); 河南省高等学校青年骨干教师培养计划(2020GGJS157); 河南省高等学校重点科研项目资助计划(24A520040); 信阳生态研究院开放基金(2023XYMS12); 信阳师范大学“南湖学者奖励计划”青年项目. 尤磊(1981—), 男, 博士, 副教授, 硕士生导师, CCF 会员, 主要研究方向为计算机图形学、点云分析与应用; 孙毅安(1994—), 男, 硕士研究生, 主要研究方向为点云分析与应用; 常潇洒(1998—), 男, 硕士研究生, 主要研究方向为点云分析与应用; 杜黎明(1985—), 女, 博士, 助理研究员, 主要研究方向为激光雷达数据处理与林业应用.

点云是三维坐标空间中具有无序性、无连接性的一系列坐标点的集合。采用点云表示的三维模型因具有占用空间小、表示能力强等优势,已得到广泛的应用^[1]。在林业研究领域,点云数据用于树木的形态模拟、重建及树木参数提取等^[2]。激光雷达是获取树木点云数据的主要设备之一,能够直接、快速、精确地获取单株树木或一定区域内森林的空间信息^[3]。然而,由于树木具有天然的复杂几何结构,且扫描树木过程中易受到设备与环境的因素限制,如树木自身遮挡与树木间的相互遮挡、设备分辨率限制以及微风侵扰等,导致获取的树木点云通常存在缺失。对树木进行不同视角的多次扫描并配准,可以在一定程度上减少点云缺失,但多次扫描耗时较长,易受到林地地形限制,导致依然可能存在遮挡。缺失树木点云将影响后续的研究与应用,如在单木三维表面重建中,如果忽略点云缺失区域构建模型,往往会导致分支拓扑连接错误、分支半径与真实值不一致等问题^[4]。因此,对单视角扫描的树木点云进行补全以得到完整的树木点云,不仅可以提高扫描效率,而且能为后续研究与应用提供翔实的数据。

现有的点云补全方法可以分为传统方法和基于深度学习的方法^[5]。传统的点云补全方法通过寻找输入点云的几何规律来推断缺失部分^[6],或是通过将缺失点云和大规模完整点云数据库进行匹配以补全形状^[7]。前者对于原始点云的对称性和规则性有一定要求;后者需要构建详尽的完整模型数据库,其匹配和优化过程也比较耗时。基于深度学习的方法利用神经网络强大的学习能力从有缺失的输入点云中提取特征,再通过该特征预测出完整的点云。由于点云是非结构化数据,因此难以直接被神经网络处理。早期的点云补全工作将点云映射到体素网格中,再使用三维卷积网络进行处理^[8];但此类方法存在细节特征易丢失和计算成本高昂等问题。随着 PointNet^[9]的提出,深度神经网络直接处理三维点云数据成为可能,它使用多层感知机(multi-layer perceptron, MLP)和最大池化操作提取输入点云的全局特征,以解决输入点云的排列不变性问题。在 PointNet 的基础上, PointNet++^[10]通过层层聚合多尺度特征首次实现了提取点云的局部特征。Yuan 等^[11]使用 PointNet 编码器构建补全网络 PCN,从缺失的点云中预测粗略的完整点云,然后使用基于 FoldingNet^[12]的解码器对粗略点云进行细化;TopNet^[13]用分层树状结构的增强解码器重构点云的拓扑结构;PF-Net^[14]创

新性地引入多分辨率编码器和金字塔结构解码器,使得预测的完整点云更加精细;VRCNet^[15]采用双通道网络结构和基于变分自编码器的关系增强模块进行概率建模,以提升补全效果。随着 Transformer 模型^[16]在自然语言处理和计算机视觉等领域的广泛应用,研究人员发现,其中的注意力机制具有排列不变性和较强的特征学习能力,非常适合用于点云的处理和分析^[17]。基于 Transformer 的 PCT^[18]使用偏移注意力模块和自注意力模块提取点云特征,用于场景分割和目标分类任务,并展现出较好的性能。PoinTr^[19]将缺失点云表示为一组点代理,并采用 Transformer 的编码器-解码器结构对点云缺失部分进行修复。Xiang 等^[20]提出 SnowflakeNet,其中的雪花点反卷积层可以使每个初始点像雪花一样在三维空间中分裂展开,以对缺失点云进行补全和细化,并在每层中引入 skip-Transformer 捕捉局部形状特征以及统一相邻层的分裂模式。

现有方法的补全对象多为物体,如飞机、汽车、桌椅等,这些物体形状规则、表面平整且多为对称结构。而树木结构十分复杂,每个分支的形状和方向都可能不相同,这对补全网络的性能提出了更高的要求。早期,基于深度学习的点云补全工作通过共享全连接层和最大池化层提取全局特征,但会忽略每个点与其邻域点之间的局部位置关系。在注重局部特征的方法中,大多数是通过点对点云进行下采样后,以采样点为中心进行 K 近邻搜索(k -nearest neighbor, KNN),或通过球查询等方法将点云分组,再分别对每个分组提取特征;然而,对于具有复杂结构的树木点云,若每个分组中包含的点太少,则很难反映出局部几何特征,而若每个分组中包含的点太多,则可能将不同分支上的点分为一组,同样不利于反映局部特征。另外,此类方法的感受野有限,难以从长距离的上下文信息中获取特征^[21]。因此,对于复杂树木点云补全有较大的局限性。

针对以上问题,本文受 Transformer^[16]以及其他点云补全相关工作的启发,提出基于注意力机制的树木点云补全网络。该网络采用编码器-解码器结构,使用基于注意力机制的多功能注意力特征增强(multi-functional attention feature enhancement, MAFE)模块,解决模型对树木点云的局部密度敏感,以及由感受野有限导致的难以获取长距离依赖关系的问题,获取一个充分包含输入点几何信息的形状编码;再使用包含 MAFE 模块的稀疏点云生成器将特征编码解码成稀疏而完整的树

木点云; 最后使用由雪花点反卷积(snowflake point deconvolution, SPD)模块组成的高密度点云生成器对稀疏点逐步进行细化和上采样, 输出完整而密集的树木点云. 为了训练和测试本文网络, 构建了一套基于三维建模软件 blender 的虚拟树木点云补全数据集(简称为树木补全数据集). 在树木补全数据集上的实验结果表明, 本文网络的树木点云补全性能优于对比的主流点云补全网络.

1 本文网络

本文提出的基于注意力机制的树木点云补全网络的框架如图 1 所示. 其采用编码器-解码器结

构, 使用 MAFE 模块构建特征提取器、稀疏点云生成器和高密度点云生成器. 特征提取器将输入的有缺失的点云通过多个串联的 MAFE 映射为一个形状编码; 稀疏点云生成器将形状编码进行特征增强并解码为稀疏的完整点云; 高密度点云生成器使用多个 SPD 层将稀疏点云进行分裂和位移, 以生成高密度的完整树木点云, 作为补全网络的输出. 图 1 中, SPD 层中乘号后面的数字为该层 SPD 的上采样比率.

1.1 MAFE 模块

MAFE 模块的结构如图 2 所示, 以自注意力机制和交叉注意力机制为核心进行构建, 针对不同输入可以实现不同功能.

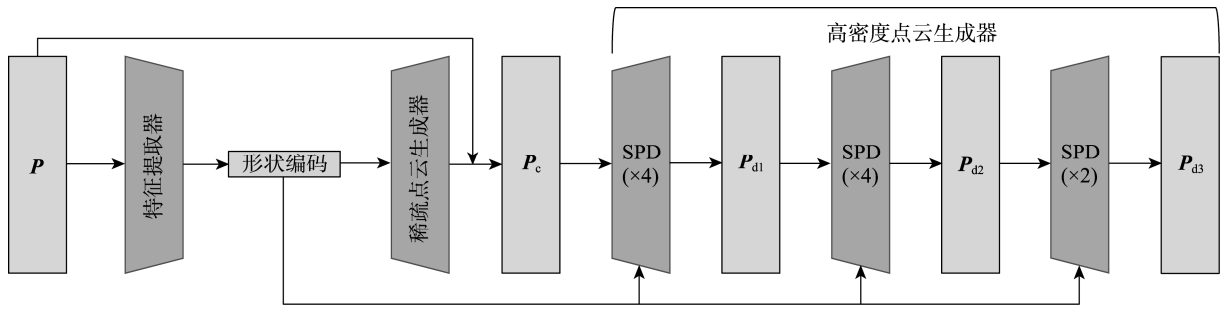


图 1 本文网络框架

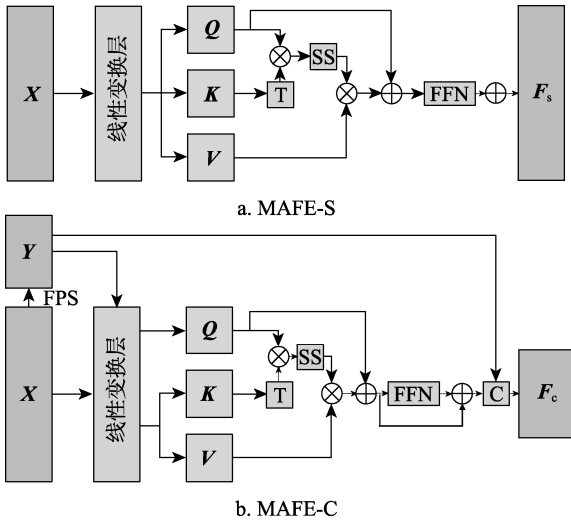


图 2 MAFE 模块结构示意图

当 MAFE 模块的输入为点云特征矩阵 $\mathbf{X}$ 时, 该模块称为 MAFE-S. 该模式下, 使用多头自注意力机制计算输入点云特征矩阵中每个点与其他点之间的关联程度, 并将每个点特征表示为点云中所有点特征的加权和, 使模型聚焦于与目标任务更相关的元素, 并减少对无关特征的依赖, 实现对于全局重要的特征的选择性提取. 如图 2a 所示,

将点云特征矩阵 $\mathbf{X}$ 输入 MAFE 模块, 乘以一组可学习的变换矩阵 $\mathbf{W}$, 计算一组查询矩阵 $\mathbf{Q}$ 、键矩阵 $\mathbf{K}$ 和值矩阵 $\mathbf{V}$. 为了减少模型参数量以简化计算, 与其他方法分别乘以 $\mathbf{W}_q$, $\mathbf{W}_k$ 和 $\mathbf{W}_v$ 这 3 个变换矩阵不同, 本文仅使用一个共享权重矩阵 $\mathbf{W}$ 来计算 $\mathbf{Q}$, $\mathbf{K}$ 和 $\mathbf{V}$ 矩阵. 通过 $\mathbf{Q}$ 和 $\mathbf{K}$ 矩阵进行点积、缩放和 Softmax 归一化(scale+Softmax, SS), 再与 $\mathbf{V}$ 矩阵相乘, 得到自注意力权重矩阵 $\mathbf{A}_s$. 这个过程表示为

$$\begin{cases} (\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathbf{X} \cdot \mathbf{W} \\ \mathbf{A}_s = \text{Softmax}\left(\frac{\mathbf{Q} \times \mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V} \end{cases}$$

其中, d_k 表示 $\mathbf{K}$ 矩阵的维度. 为了优化梯度传播、防止信息损失和增强模型的表达能力, 从而进一步提高模型性能和训练效果, 在模块末端使用 Transformer 中的残差连接和前馈神经网络(feed forward network, FFN)层, 最终得到特征矩阵 $\mathbf{F}_s$. 该模式下输入特征矩阵的形状不会被改变. 由于从点坐标得到的点特征已经包含足够的位置信息, 因此计算自注意力输出时不再采用 Transformer 中的位置嵌入^[22].

当 MAFE 模块的输入为点云特征矩阵 $\mathbf{X}$ 和其下采样特征矩阵 $\mathbf{Y}$ 时, 模块采用交叉注意力机制, 此时该模块称为 MAFE-C. 通常, 不同尺度或分辨率的点云特征中包含不同的语境或语义信息^[14]. 基于该理论, MAFE-C 模块通过使用交叉注意力机制计算不同尺度输入点云间的相关性, 如 2 个点云中的各个点之间距离上相近或形状的相似性, 实现特征自顶向下的传递, 以此来隐式地提取输入点云的局部特征. 如图 2b 所示, 首先对 $\mathbf{X}$ 使用最远点采样 (farthest point sampling, FPS), 得到 $(N/d) \times c$ 大小的下采样点云特征矩阵 $\mathbf{Y}$, 其中, N 表示 $\mathbf{X}$ 的点数, d 表示下采样比率, c 表示特征维度; 然后通过分别乘以共享权重矩阵 $\mathbf{W}$, 将 $\mathbf{X}$ 变换为 $\mathbf{K}$ 矩阵和 $\mathbf{V}$ 矩阵, 将 $\mathbf{Y}$ 变换为 $\mathbf{Q}$ 矩阵, 计算交叉注意力权重矩阵 $\mathbf{A}_c$. 这个过程表示为

$$\begin{cases} \mathbf{Q} = \mathbf{Y} \cdot \mathbf{W} \\ (\mathbf{K}, \mathbf{V}) = \mathbf{X} \cdot \mathbf{W} \\ \mathbf{A}_c = \text{Softmax}\left(\frac{\mathbf{Q} \times \mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V} \end{cases}$$

该模块同样引入残差连接和 FFN. 为了保留一部分原始特征以提供更全面的特征表示, 将 $\mathbf{Y}$ 与 $\mathbf{A}_c$ 拼接, 最终得到 $(N/d) \times 2c$ 大小的特征矩阵 $\mathbf{F}_c$.

1.2 特征提取器

特征提取器的功能是提取输入缺失点云的局部与全局特征, 并将其映射为一个形状编码. 如图 3 所示, 特征提取器由一个 MLP, 3 组相互交替串联的 MAFE-C 和 MAFE-S 层, 以及一个最大池化层组成. 首先由 MLP 将输入的点坐标映射为特征向量; 然后将 3 个下采样比率分别设置为 4, 2 和 2 的 MAFE-C 提取输入点的局部特征, 并在每个 MAFE-C 后串联一个 MAFE-S 层, 进一步增强特征提取器对目标全局几何结构的特征感知; 最后使用最大池化层聚合特征, 得到形状编码.



图 3 特征提取器结构图

1.3 稀疏点云生成器

稀疏点云生成器对特征提取器学习到的输入点云的形状编码进行解码, 得到一个稀疏但保留

目标点云完整几何结构的点云框架, 如图 4 所示. 首先通过 MLP 层和反卷积层将形状编码上采样, 通过另一个 MLP 层扩展特征维度后, 使用 3 个串联的 MAFE-S 模块增强特征矩阵的表征能力; 然后通过矩阵变形和 MLP 层将点特征逐个转换成点三维坐标; 最后保留输入部分的结构, 将生成的稀疏点云与输入点云 $\mathbf{P}$ 拼接, 为使拼接点云的密度更均匀, 对其进行一次 FPS 下采样, 得到包含 512 个点的稀疏点云 $\mathbf{P}_c$.

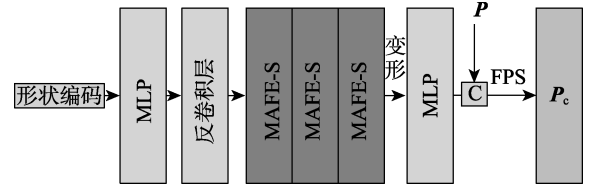


图 4 稀疏点云生成器结构图

1.4 高密度点云生成器

高密度点云生成器将稀疏点云 $\mathbf{P}_c$ 在原有结构的基础上逐步进行上采样, 以生成稠密、完整的点云, 由 SnowflakeNet^[20] 中的 SPD 模块组成. SPD 可将输入点云中的每个点分裂出多个子点, 并对它们添加一个位移, 使点云像雪花一般在三维空间中生长, 并将稀疏点云上采样为稠密点云. 为了在点分裂的过程中更好地恢复完整树木结构的局部细节, 并保证相邻 SPD 之间的分裂模式的一致性, 每个 SPD 会将本层点特征和上一个 SPD 的位移特征输入到 skip-Transformer 中; 最终根据需要生成点数的不同设置多个 SPD 层逐步生成完整点云. 第 i 层 SPD 的结构如图 5 所示, 首先将输入的来自第 $i-1$ 层的较稀疏的点云 $\mathbf{P}_{i-1}$ 通过 PointNet 编码器提取特征, 并将每个点特征与特征提取器生成的形状编码进行拼接, 得到特征矩阵 $\mathbf{Q}_i$, 将其和来自上层的位移特征 $\mathbf{K}_{i-1}$ 作为 skip-Transformer 的输入, 得到形状上下文特征 $\mathbf{H}_i$; 对于第 1 层 SPD, 其 $\mathbf{K}_{i-1}$ 的值取 $\mathbf{Q}_i$, 再将 $\mathbf{H}_i$ 经过逐点分割操作 (point-wise splitting, PS) 为每个点生成多个子点特

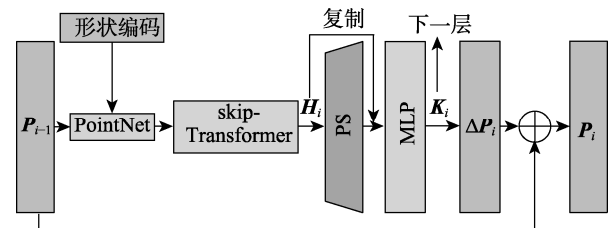


图 5 SPD 模块结构示意图

征, 并将子点特征与复制后的 H_i 拼接, 馈送到 MLP 层得到位移特征 K_i , K_i 再通过一个 MLP 和 tanh 激活函数转换为点位移 ΔP_i , 同时也被送入下一个 SPD 中的 skip-Transformer 模块中; 最后将点位移 ΔP_i 加到 K_{i-1} 的复制点上, 得到第 i 层的上采样输出点云 P_i , 如需要进一步上采样, 便将 P_i 输

入到下一个 SPD 块中. 本文中的高密度点云生成器由 2 个上采样比率分别为 4, 4 和 2 的 SPD 组成, 可将来自稀疏点云生成器的包含 512 个点的稀疏点 P_c 逐步上采样和细化, 得到 2 048 点的 P_{d1} , 8 192 点的 P_{d2} 和 16 384 个点的 P_{d3} . 补全流程的可视化效果如图 6 所示.

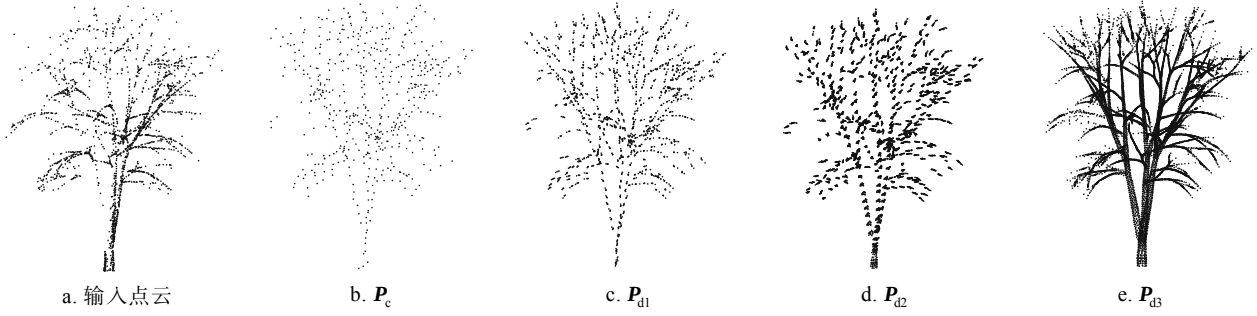


图 6 完整树木点云生成流程

1.5 损失函数

本文使用倒角距离(chamfer distance, CD)作为损失函数, 通过计算一个点云中每个点到另一个点云中最近点的平均距离之和来度量 2 个点云的相似度. 假设 S_1 为通过网络补全的点云, S_2 为对应真值点云, x 和 y 分别为 S_1 和 S_2 中的任意一点, 则 CD 的计算公式为

$$CD(S_1, S_2) = \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2 + \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} \|y - x\|_2.$$

CD 越小, 则网络预测的完整点云和真值越接近. 设稀疏点云生成器的输出点云为 P_c , 高密度点云生成器的 3 个 SPD 模块的输出分别为 P_{d1} , P_{d2} 和 P_{d3} . 将真值点云分别用 FPS 下采样到与 P_c , P_{d1} , P_{d2} 和 P_{d3} 相同的点数, 得到 Q_c , Q_{d1} , Q_{d2} 和 Q_{d3} ; 再分别计算 CD 并求和, 从而在生成完整点云的每个步骤中约束模型参数. 本文网络总损失函数定义为

$$L = CD(P_c, Q_c) + CD(P_{d1}, Q_{d1}) + CD(P_{d2}, Q_{d2}) + CD(P_{d3}, Q_{d3}).$$

2 实验及结果分析

2.1 数据集

考虑现实中难以使用激光扫描仪获取大量、无缺失的树木点云, 本文借鉴 PCN^[11]在虚拟三维模型数据集 ShapeNet^[23]的基础上构建点云补全数据集的做法, 构建一套树木补全数据集. 构建流程如图 7 所示:

Step1. 利用三维建模软件 blender 的 Sapling Tree Gen 插件, 在坐标原点处批量随机生成树木三维网格模型, 并将其缩放至以坐标原点为中心的单位球范围内, 如图 7a 所示.

Step2. 对树木网格模型进行均匀采样 16 384 个点, 得到完整的树木点云作为训练真值, 真值点云如图 7b 所示.

Step3. 将 blender 的相机设置在以坐标原点为中心的单位球面的随机位置上, 并使其始终朝向坐标原点; 然后通过渲染得到 EXR 格式的 depth 信息, 如图 7c 所示, 其中, 红色点为相机、黑色点为坐标原点, 深度信息记录当前相机视角下可视的树木表面到相机的距离信息, 而被遮挡的部分信息会丢失.

Step4. 使用 Python 的 OpenEXR 库读取 EXR 文件并转换成指定高度和宽度的深度图像, 如图 7d 所示.

Step5. 将深度图像转换成点云, 以模拟现实中单视角扫描得到的有缺失的点云, 如图 7e 所示. 为了更充分地利用训练数据, 每棵树木都会生成 8 个随机视角下获取的深度图像及对应的点云.

Step6. 为了统一输入点数、加快训练速度并提高训练难度, 将缺失点云下采样至 2 048 点, 如图 7f 所示. 生成的所有点云均归一化至同一坐标范围内, 即通过缩放和平移使点云的中心点位于坐标原点, 且每个点在各轴的坐标值均在 -1~1.

通过以上方法, 本文生成共 4 类(包括 Japanese Maple, White Birch, Callistemon 和 Small Maple)树木各 3 000 棵, 其中, 将每类中的 2 250 棵作为训练集、300 棵作为验证集、450 棵作为测试集. 为了测试各种方法对未训练树木种类的补全效果, 另外生成类 2 类(Willow 和 Weeping Willow)树木各 450 棵, 并添加到测试集中.

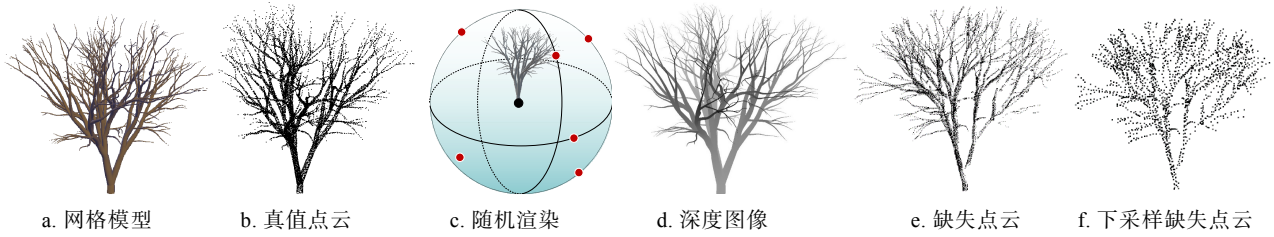


图 7 树木补全数据集生成流程

2.2 实验设置

本文实验的硬件环境为 Intel(R) Xeon(R) E5-2650 v4 CPU, 128 GB 内存, 4 块 12 GB 显存的 NVIDIA Titan Xp GPU; 操作系统为 Ubuntu22.04; 实验代码由 Python 语言和 Pytorch 深度学习框架实现. 使用 Adam 优化器训练模型共 400 轮; 初始学习率为 0.001, 并设置每 40 轮衰减至 70%; 批量大小设置为 32. 实验中, 高密度点云生成器采用 3 个上采样比率分别为 4, 4 和 2 的 SPD 层.

2.3 评价指标

参考现有点云补全方法在公共数据集上采用的评价指标, 本文选用 L_1 倒角距离($CD-I_1$)^[19]和 F 分数^[24]作为点云补全效果的评价指标. 计算预测完整点云和真值点云的 $CD-I_1$, 结果越小, 表示预测值越接近真值, 则网络补全性能越好. F 分数由

重建点云中正确重建的点所占的比例决定, 分值越高, 补全效果越好.

2.4 对比实验

本文选用近几年提出的主流点云补全网络作为对比方法: PCN^[11], TopNet^[13], VRCNet^[15], SnowflakeNet^[20], ECG^[25], GRNet^[26], AdaPoinTr^[27]和 AnchorFormer^[28], 这些方法均采用原文献开源代码中提供的默认配置在树木补全数据集上进行训练. 将测试集输入到训练完毕的点云补全网络中, 得到的 F 分数和 $CD-I_1$ 分别如表 1 和表 2 所示. 可以看出, 本文网络在 4 个已训练类别和 2 个未训练类别中的树木补全效果均取得最低的 $CD-I_1$ 和最高的 F 分数, 与定量实验结果最优的 AdaPoinTr^[27]相比, 本文网络的平均 $CD-I_1$ 降低 0.62, 平均 F 分数提高 0.04. 表 3 所示为 8 种方法的可视化补全效果.

表 1 不同方法在树木补全数据集上 F 分数对比

方法	F 分数						平均
	Japanese Maple	White Birch	Callistemon	Small Maple	Willow	Weeping Willow	
PCN ^[11]	0.51	0.66	0.53	0.55	0.38	0.38	0.50
TopNet ^[13]	0.37	0.47	0.43	0.37	0.23	0.40	0.38
VRCNet ^[15]	0.86	0.90	0.88	0.84	0.68	0.63	0.80
SnowflakeNet ^[20]	0.85	0.93	0.89	0.85	0.69	0.53	0.79
ECG ^[25]	0.86	0.91	0.87	0.84	0.39	0.56	0.74
GRNet ^[26]	0.66	0.71	0.68	0.62	0.61	0.53	0.64
AdaPoinTr ^[27]	0.88	0.93	0.90	0.85	0.66	0.60	0.80
AnchorFormer ^[28]	0.67	0.77	0.68	0.59	0.42	0.44	0.60
本文	0.91	0.96	0.93	0.90	0.74	0.61	0.84

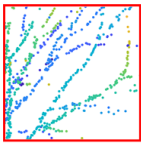
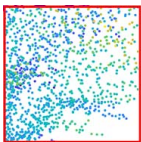
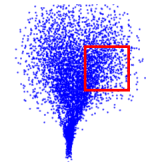
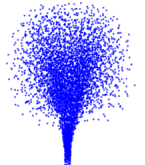
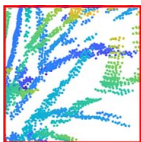
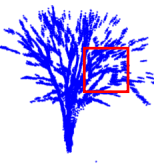
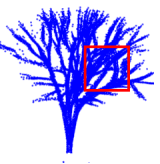
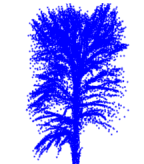
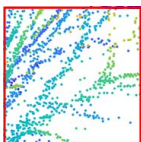
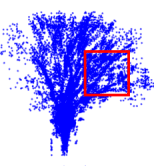
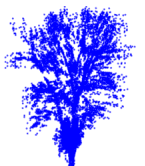
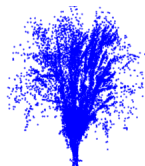
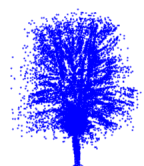
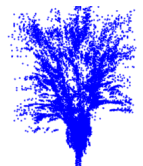
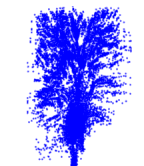
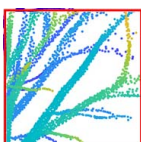
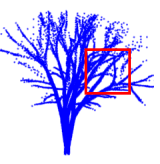
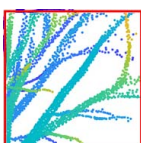
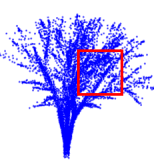
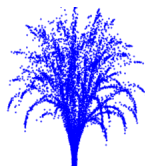
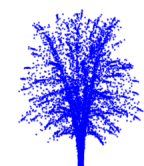
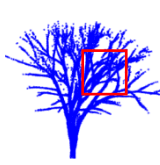
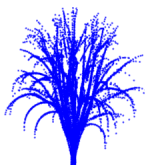
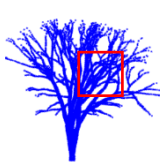
注: 粗体表示最优值.

表 2 不同方法在树木补全数据集上 $CD-I_1$ 对比

方法	$10^{-3} \times CD-I_1$						平均
	Japanese Maple	White Birch	Callistemon	Small Maple	Willow	Weeping Willow	
PCN ^[11]	12.11	10.03	11.77	11.67	15.50	17.46	13.09
TopNet ^[13]	15.56	13.53	13.78	13.94	17.49	15.91	15.04
VRCNet ^[15]	6.45	5.54	6.27	7.41	10.38	13.13	8.20
SnowflakeNet ^[20]	6.48	4.90	5.97	6.70	10.76	14.48	8.22
ECG ^[25]	6.25	5.08	5.95	6.77	12.57	12.41	8.17
GRNet ^[26]	10.73	9.77	9.98	11.19	11.81	14.94	11.40
AdaPoinTr ^[27]	5.49	4.17	5.27	6.31	9.95	12.81	7.33
AnchorFormer ^[28]	9.04	7.59	9.06	10.74	13.89	15.83	11.03
本文	5.24	4.00	4.76	5.61	8.65	12.00	6.71

注: 粗体表示最优值.

表 3 8 种方法树木补全效果对比

方法	细节展示	Japanese Maple	White Birch	Callistemon	Small Maple	Willow	Weeping Willow
输入点云							
PCN ^[11]							
VRCNet ^[15]							
SnowflakeNet ^[20]							
ECG ^[25]							
GRNet ^[26]							
AdaPoinTr ^[27]							
AnchorFormer ^[28]							
本文							
真值							

可以看出, PCN^[11], TopNet^[13]和 GRNet^[26]的补全结果几乎无法辨认树木细节; VRCNet^[15], ECG^[25]和 AnchorFormer^[28]虽然可在一定程度上还原树干和树枝的几何结构, 但存在大量断层或点不均匀等情况; SnowflakeNet^[20]和 AdaPoinTr^[27]的补全结果在总体结构上较为完整和连贯, 但对于树木点云细节的还原度不如本文方法, 尤其是对于未训练的 2 个类别的树木的补全效果较差; 总体上看, 本文网络的补全效果最完整、均匀和平滑, 在 6 个类别的树木上均优于所有对比方法. 图 8 所示为本文网络、SnowflakeNet^[20]和 AdaPoinTr^[27]对未经过训练的 Willow 类的一个树木点云补全效果细节对比. 可以看出, 本文网络生成的补全点云更加完整、连贯、紧凑, 生成的噪声点更少, 整体更接近真值.

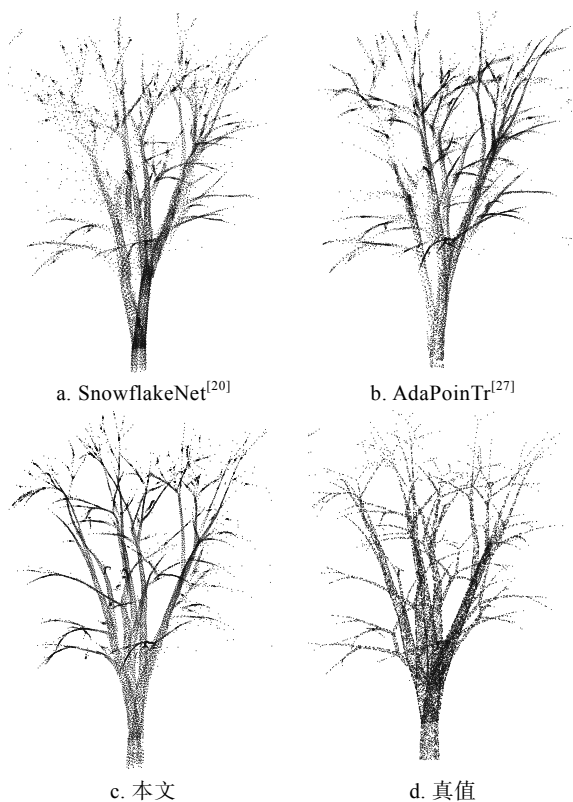


图 8 在树木补全数据集上 3 种方法的细节对比

2.5 消融实验

为了验证本文网络中各部分的有效性, 设置如下消融实验: 将特征提取器替换为 PCN 编码器得到网络 A; 去掉稀疏点云生成器中的 3 个 MAFE-S 模块得到网络 B. 对以上网络在相同配置下进行训练, 并在包含 4 种已训练树木的测试集上进行补全实验, 结果如表 4 所示. 可以看出, 替换掉 A 和 B 后, 平均 $CD-I_1$ 分别上升 0.08 和 0.18, 证明本文网络中各模块对整体的效果提升均有贡献.

表 4 在树木数据集上消融实验的 $CD-I_1$ 对比

网络	Japanese Maple	White Birch	Callistemon	Small Maple	平均
本文	5.34	4.00	4.77	5.63	4.93
A	5.47	4.06	4.87	5.67	5.01
B	5.50	4.16	4.98	5.81	5.11

2.6 真实扫描补全实验

为了验证本文网络在真实树木点云上的补全效果, 截取地面激光雷达扫描的真实场景中的 3 个不同类别缺失树木点云, 并将其输入到本文网络和 2 种对比方法中进行补全, 输出的可视化效果如图 9 所示. 可以看出, 3 种方法均可对缺失和稀疏

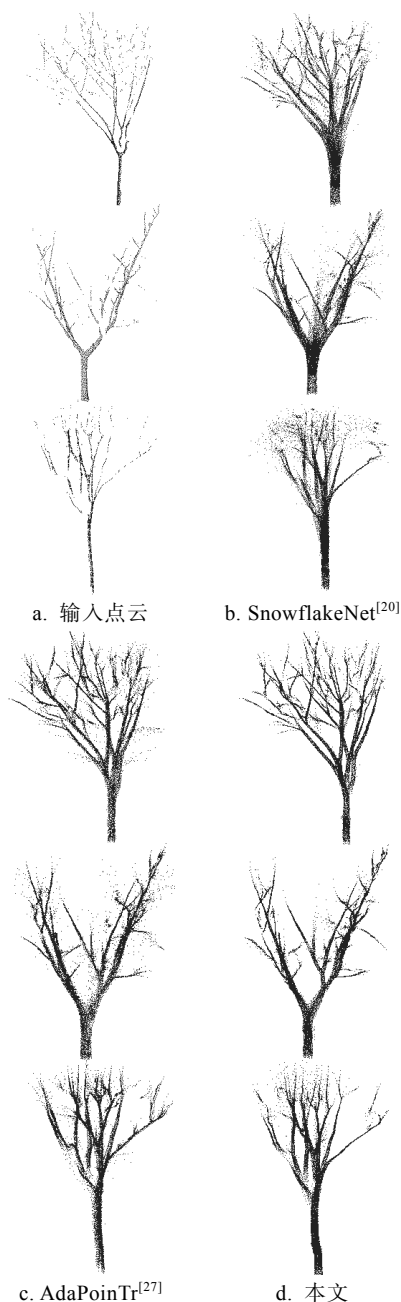


图 9 真实树木点云的补全效果

部分进行一定程度的补全, 缺失较严重的树枝在补全后仍可能存在一些断层; 与其他方法相比, 本文网络可以还原更多细节, 树枝的粗细变化和弯曲更加自然, 且产生的噪点更少. 总体来说, 在真实数据集上的补全效果弱于在虚拟数据集上的补全效果, 原因是虚拟数据集中树木种类和真实补全实验的树木的种类在生长结构上存在一定差异, 且虚拟数据集在模拟真实扫描的方面还存在一些局限, 如树木包含点数较少、树木表面细节相对较少、未模拟出树木不同位置因与扫描仪的距离不同而导致的点云密度变化等.

2.7 重建实验

采用树木三维重建软件 TreeQSM^[29]分别对经过本文网络补全前后的同一棵树木点云进行三维重建, 结果如图10所示. 可以看出, 本文网络对树木的拓扑结构进行了有效的补全, 使更多的树枝细节通过三维重建得到体现.

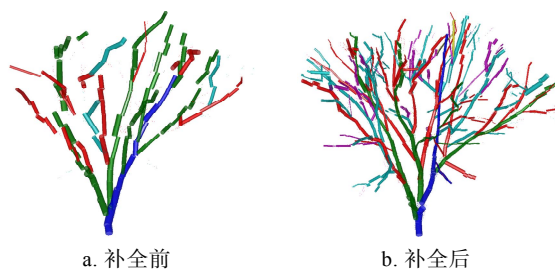


图10 树木点云三维重建效果

3 结 语

本文提出一种基于注意力机制的树木点云补全网络, 并构建了一套树木补全数据集用于训练. 该网络采用编码器-解码器结构, 从粗略到精细地预测输入的缺失点云对应的完整点云; 特征提取器通过引入自注意力机制和交叉注意力机制, 计算输入点云的逐点间的相关性信息和不同尺度点之间的相关性信息, 以提取局部和全局特征; 高密度点云生成器采用 SPD 模块对稀疏点进行上采样, 得到密集且完整的点云. 实验结果表明, 本文网络能够有效地提高对于复杂的缺失树木点云的补全效果. 然而, 本文网络在部分较细的枝条末端出现局部点密度不均匀及连接断层的问题, 说明补全网络对于数据中较为稀疏的部分的处理能力较弱, 仍有继续优化的空间; 此外, 在真实树木点云上的补全实验反映出虚拟数据集的局限, 还需进一步提高虚拟数据和真实数据的相似程度、增加更多不同种类的虚拟树木点云, 并添加一部分真实树木

点云数据. 未来的工作重点是优化补全网络及提高数据集的质量.

参考文献(References):

- [1] Liu Caixia, Wei Mingqiang, Guo Yanwen. 3D point cloud restoration via deep learning: a comprehensive survey[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(12): 1936-1952(in Chinese)
(刘彩霞, 魏明强, 郭延文. 基于深度学习的三维点云修复技术综述[J]. 计算机辅助设计与图形学学报, 2021, 33(12): 1936-1952)
- [2] You Lei, Tang Shouzheng, Song Xinyu. An algorithm of stem surface reconstruction based on tangent plane projection[J]. Scientia Silvae Sinicae, 2016, 52(11): 115-123(in Chinese)
(尤磊, 唐守正, 宋新宇. 基于切平面投影的树干三维表面重建算法[J]. 林业科学, 2016, 52(11): 115-123)
- [3] Guo Qinghua, Liu Jin, Tao Shengli, et al. Perspectives and prospects of LiDAR in forest ecosystem monitoring and modeling[J]. Chinese Science Bulletin, 2014, 59(6): 459-478(in Chinese)
(郭庆华, 刘瑾, 陶胜利, 等. 激光雷达在森林生态系统监测模拟中的应用现状与展望[J]. 科学通报, 2014, 59(6): 459-478)
- [4] Cao W, Wu J Y, Shi Y F, et al. Restoration of individual tree missing point cloud based on local features of point cloud[J]. Remote Sensing, 2022, 14(6): 1346
- [5] Yang Liu, Wu Xiaoqun. 3D shape completion via deep learning: a method survey[J]. Journal of Graphics, 2023, 44(2): 201-215(in Chinese)
(杨柳, 吴晓群. 基于深度学习的三维形状补全研究综述[J]. 图学学报, 2023, 44(2): 201-215)
- [6] Berger M, Tagliasacchi A, Seversky L M, et al. State of the art in surface reconstruction from point clouds[C] //Proceedings of the 35th Annual Conference of the European Association for Computer Graphics. Aire-la-Ville: Eurographics Association Press, 2014: 161-185
- [7] Li Y Y, Dai A, Guibas L, et al. Database-assisted object retrieval for real-time 3D reconstruction[J]. Computer Graphics Forum, 2015, 34(2): 435-446
- [8] Dai A, Qi C R, Nießner M. Shape completion using 3D-encoder-predictor CNNs and shape synthesis[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 6545-6554
- [9] Charles R Q, Su H, Kaichun M, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 77-85
- [10] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc, 2017: 5105-5114
- [11] Yuan W T, Khot T, Held D, et al. PCN: point completion network[C] //Proceedings of the International Conference on 3D

- Vision. Los Alamitos: IEEE Computer Society Press, 2018: 728-737
- [12] Yang Y Q, Feng C, Shen Y R, *et al.* FoldingNet: point cloud auto-encoder via deep grid deformation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 206-215
- [13] Tchapmi L P, Kosaraju V, Rezaatofghi H, *et al.* TopNet: structural point cloud decoder[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 383-392
- [14] Huang Z T, Yu Y K, Xu J W, *et al.* PF-Net: point fractal network for 3D point cloud completion[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 7659-7667
- [15] Pan L, Chen X Y, Cai Z A, *et al.* Variational relational point completion network[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 8520-8529
- [16] Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc, 2017: 6000-6010
- [17] Lu D N, Xie Q, Wei M Q, *et al.* Transformers in 3D point clouds: a survey[OL]. [2024-01-21]. <https://arxiv.org/abs/2205.07417>
- [18] Guo M H, Cai J X, Liu Z N, *et al.* PCT: point cloud transformer[J]. Computational Visual Media, 2021, 7(2): 187-199
- [19] Yu X M, Rao Y M, Wang Z Y, *et al.* PoinTr: diverse point cloud completion with geometry-aware transformers[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2021: 12478-12487
- [20] Xiang P, Wen X, Liu Y S, *et al.* Snowflake point deconvolution for point cloud completion and generation with skip-Transformer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(5): 6320-6338
- [21] Wang J, Cui Y, Guo D Y, *et al.* PointAttN: you only need attention for point cloud completion[C] //Proceedings of the 38th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2022: 5472-5480
- [22] Zhang W X, Zhou H J, Dong Z, *et al.* Point cloud completion via skeleton-detail transformer[J]. IEEE Transactions on Visualization and Computer Graphics, 2023, 29(10): 4229-4242
- [23] Wu Z, Song S, Khosla A, *et al.* 3D ShapeNets: a deep representation for volumetric shapes[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2015: 1912-1920
- [24] Tatarchenko M, Richter S R, Ranftl R, *et al.* What do single-view 3D reconstruction networks learn?[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 3400-3409
- [25] Pan L. ECG: edge-aware point cloud completion with graph convolution[J]. IEEE Robotics and Automation Letters, 2020, 5(3): 4392-4398
- [26] Xie H Z, Yao H X, Zhou S C, *et al.* GRNet: gridding residual network for dense point cloud completion[C] //Proceedings of the 16th European Conference on Computer Vision. Heidelberg: Springer, 2020: 365-381
- [27] Yu X M, Rao Y M, Wang Z Y, *et al.* AdaPoinTr: diverse point cloud completion with adaptive geometry-aware transformers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(12): 14114-14130
- [28] Chen Z K, Long F C, Qiu Z F, *et al.* AnchorFormer: point cloud completion from discriminative nodes[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2023: 13581-13590
- [29] Raunonen P, Kaasalainen M, Åkerblom M, *et al.* Fast automatic precision tree models from terrestrial laser scanner data[J]. Remote Sensing, 2013, 5(2): 491-520