

# 机载LiDAR点云数据滤波算法的改进与比较

鲁俊

中国电建集团西北勘测设计研究院有限公司 陕西 西安 710100

**摘要:** 机载激光雷达 (LiDAR) 是获取高精度三维地理空间信息的重要手段, 广泛应用于多领域。原始LiDAR点云数据含地物与地面点, 数字高程模型 (DEM) 等构建依赖地面点精确提取, 即点云滤波, 其算法性能决定最终产品质量。本文梳理了点云滤波的基本原理与挑战, 剖析了三类主流算法: 形态学方法 (如渐进加密三角网滤波PTM)、基于曲面的方法 (以不规则三角网TIN为基础) 及基于机器学习的方法 (如支持向量机SVM), 并指出其局限性。针对传统算法在复杂地形和城市场景下表现不佳的问题, 本文提出融合局部几何特征与全局上下文信息的自适应滤波算法, 先初步区分平坦与非平坦区域, 再在非平坦区域精细化处理。经国际标准测试数据集验证, 该算法在总体精度和I类错误方面显著提升, 处理特殊地形时更具鲁棒性, 为高精度DEM自动化生产提供新路径。

**关键词:** 机载LiDAR; 点云滤波; 数字高程模型; 自适应算法

## 引言

21世纪以来, 传感器技术、全球导航卫星系统 (GNSS) 与惯性导航系统 (INS) 迅猛发展, 机载LiDAR作为主动式遥感技术, 凭借全天候、高效、高精度及穿透植被冠层的能力, 革新了传统地形测绘方式。它发射激光脉冲并接收反射信号, 直接获取海量、离散且带有精确三维坐标和强度信息的点云数据, 记录了复杂三维场景, 包含裸露地表及树木、建筑等非地面点。在众多应用中, 构建高精度数字高程模型 (DEM) 是基础且重要的任务, DEM是反映真实地表起伏的连续模型, 是诸多地理空间分析的基础数据。要从原始点云中分离地面点, 滤波或分类是关键预处理步骤, 其准确性决定DEM质量, 影响后续分析。但现实世界复杂, 地形起伏、植被茂密、建筑多样, 使自动、高效、准确滤波极具挑战。

## 1 机载 LiDAR 点云滤波基础

### 1.1 数据特点与滤波目标

机载LiDAR点云数据具有以下显著特点: (1) 海量性: 单次飞行可获取数百万至数十亿个点。(2) 离散性: 点在空间中呈不规则分布, 无固定拓扑结构。(3) 多源性: 除三维坐标外, 还包含强度、回波次数、扫描角等属性信息。(4) 噪声: 包含由大气扰动、设备误差等引起的少量异常点 (离群点)<sup>[1]</sup>。滤波的核心目标是从这个庞大的、混合的点集中, 准确地识别并分离出属于真实地面的点子集 $P_g$ , 同时将非地面点 (如植被、建筑物等) 归入 $P_{ng}$ 。理想的滤波结果应满足:  $P = P_g \cup P_{ng}$  且  $P_g \cap P_{ng} = \emptyset$ 。

### 1.2 滤波基本原理

大多数滤波算法都基于一个共同的先验假设: 地面在局部范围内通常是连续、平滑且处于相对低洼位置的。基于此, 算法通常遵循以下逻辑: (1) 初始化: 首先找到一批“种子点”, 这些点极大可能是地面点 (如全局最低点或网格最低点)。(2) 生长/传播: 以种子点为基础, 根据某种准则 (如坡度、高度差、曲率等) 判断其邻域内的点是否也属于地面。(3) 迭代优化: 重复生长过程, 直到没有新的点可以被添加或满足停止条件。

### 1.3 主要挑战

一是地形复杂性: 在陡峭的山坡、悬崖或峡谷地区, 地面本身坡度很大, 传统的坡度阈值约束会失效, 导致地面点被错误剔除。二是植被干扰: 茂密的森林中, 激光束可能穿透树冠到达地面, 形成所谓的“地面回波”, 但这些点往往稀疏且被大量的植被点包围, 难以与低矮灌木区分。三是城市环境: 城市中存在大量与地面高程相近的人工地物, 如汽车、路缘石、低矮围墙等, 极易被误判为地面点 (欠滤波)。四是参数敏感性: 许多算法 (如PTM) 包含多个需要手动调整的参数 (如最大窗口尺寸、最大角度等), 这些参数的设置对结果影响巨大, 且缺乏普适性。

## 2 典型滤波算法分析

### 2.1 渐进加密三角网滤波 (PTM)

PTM是形态学滤波的经典实现, 核心是迭代加密构建逼近真实地面的不规则三角网 (TIN)。算法先划分点云区域为初始格网, 选每个格网内高程最低点作初始地面种子点, 构建稀疏初始 TIN 模型。接着进入迭代加密阶段, 对未分类点, 找到其在当前 TIN 模型上的投影位置, 计算到投影三角形平面垂直距离及连接线与平面

法向量夹角。若两指标均小于当前迭代步长对应的动态阈值，该点被接纳为新地面点，更新 TIN 模型。随着迭代，窗口尺寸和阈值逐步减小，TIN 模型更精细，能捕捉细微地形变化<sup>[2]</sup>。PTM 原理直观、易于理解，处理大面积平坦区域和缓坡地带效果显著。但致命弱点是对初始参数极度敏感，处理陡峭地形时，初始 TIN 粗糙，许多陡坡真实地面点因与初始平面角度过大被永久拒绝，最终 DEM 留下无法弥补的空洞。

### 2.2 基于曲面拟合的 TIN 滤波

此方法与 PTM 框架相似，以构建和更新 TIN 为核心，判别准则融入更多局部几何信息。不仅关注待判定点到最近三角形平面的距离，还考察其邻域几何特性，如通过主成分分析（PCA）估算局部曲率，或计算邻域点高程标准差衡量表面起伏程度。综合多个几何指标构建复杂判别函数，更准确区分地面与非地面点。理论上，其对局部几何细节的深入挖掘比 PTM 更具鲁棒性，能处理中等复杂度场景。但弊端明显，计算复杂度显著高于 PTM，尤其是进行曲率等高阶特征估计时，计算开销大。在点云密度低区域，如茂密植被下地面点，邻域有效信息不足，估算的曲率等特征不稳定、误差大，会误导分类决策。面对极端复杂地形和植被干扰，根本性局限未突破。

### 2.3 基于支持向量机（SVM）的滤波

该方法代表从传统几何规则驱动向数据驱动范式转变，将点云滤波问题定义为监督学习的二分类任务，核心是利用已知标签样本数据训练强大分类器。实施时，先为每个点构造多维特征向量，涵盖绝对高程、相对高程、激光回波次数、回波强度值、局部点云密度、法向量方向等，这些特征构成描述点“身份”的指纹。再用部分人工精确标注的训练数据，SVM 模型寻找最优高维超平面，最大化分隔两类样本<sup>[3]</sup>。模型训练完成后，可应用于整个点云数据集自动分类。SVM 方法优势在于强大泛化能力和多源信息融合能力，优质模型在新数据上预测速度快。但严重依赖训练样本，若训练数据未充分覆盖目标区域可能出现的地物类型、植被结构和地形形态，模型面对“陌生”场景性能会大打折扣，甚至出现系统性偏差。此外，特征工程质量决定模型上限，选择和组合最优特征集充满经验性和不确定性。

## 3 一种自适应滤波算法的提出

针对上述算法的局限性，本文提出一种无需训练样本、能自适应地形变化的滤波算法。该算法的核心思想是分而治之：首先根据点云的局部几何复杂度将其划分为“简单区域”（平坦开阔地）和“复杂区域”（陡坡、植

被区），然后对不同区域采用针对性的滤波策略。

### 3.1 算法整体框架

算法流程可分为三个阶段：（1）预处理与特征提取：去除离群点，并计算每个点的多尺度法向量熵和局部点密度。（2）区域自适应划分：根据特征值将点云划分为平坦区和非平坦区。（3）差异化滤波处理：对平坦区采用高效的形态学开运算；对非平坦区，采用坡度自适应的布料模拟滤波（CSF）进行精细化处理。

### 3.2 关键技术实现

#### 3.2.1 多尺度法向量熵计算

法向量是描述点云局部几何形状的重要特征。对于一个点 $p_i$ ，其法向量可以通过对其 $k$ 近邻点进行主成分分析（PCA）得到。然而，单一尺度的 $k$ 值难以适应不同密度的点云。因此，本文采用多尺度策略，计算 $k_1, k_2, k_3$ （例如 $k_1 = 8, k_2 = 16, k_3 = 32$ ）三种邻域下的法向量 $n_1, n_2, n_3$ 。

法向量熵（Normal Vector Entropy, NVE）用于衡量这些法向量的离散程度：

$$NVE(p_i) = -\sum_{j=1}^3 w_j \log(w_j)$$

其中， $w_j = \frac{1}{3}(1 + n_j \cdot n_{avg})$ ， $n_{avg}$ 是三个法向量的平均方向。NVE值越小，说明局部表面越平坦、规则；NVE值越大，则表明该区域几何结构复杂（如植被、建筑物边缘）。

#### 3.2.2 自适应区域划分

设定一个NVE阈值 $T_{nve}$ （可通过Otsu算法自适应确定）。对于任意点 $p_i$ ：若 $NVE(p_i) < T_{nve}$ ，则将其标记为“平坦区”候选点。否则，标记为“复杂区”候选点。为进一步提高鲁棒性，还需结合局部点密度进行验证。在植被稀疏区，即使NVE较低，也可能不是地面。因此，只保留那些同时满足点密度高于某个阈值的“平坦区”候选点作为最终的平坦区。

#### 3.2.3 差异化滤波策略

一是平坦区处理：对于被确认的平坦区，采用经典的形态学开运算。首先将点云栅格化为规则网格，取每个网格内的最低高程值，然后对这个最低点DEM进行开运算（先腐蚀后膨胀），以去除孤立的非地面点。这种方法计算效率极高，且在平坦区域效果优异<sup>[4]</sup>。二是复杂区处理：这是算法的关键创新。本文借鉴布料模拟滤波（CSF）的物理思想，但对其进行改进以适应局部坡度。CSF基本思想是将一个虚拟的、具有弹性和重力的“布料”从点云上方倒置下落，布料会被非地面点（如

树顶、屋顶)阻挡,最终形成的布料形态即为地面的近似。在原始CSF中,布料的刚性是全局一致的。本文提出,布料的刚性应与局部地形坡度成反比。具体而言,在滤波前,先利用平坦区的结果和部分可靠的复杂区点,快速生成一个粗糙的DEM,并计算每个位置的坡度slope。在CSF模拟过程中,对于高坡度区域,降低布料的刚性系数,使其更容易变形以贴合真实的陡峭地面;对于低坡度区域,则保持较高的刚性以维持平滑性。这种自适应机制能有效缓解传统CSF在陡坡处产生的“悬空”效应。最后,将平坦区和复杂区的滤波结果合并,并进行一次小范围的后处理(如去除小的地面点簇),得到最终的地面点集。

#### 4 实验与结果分析

##### 4.1 实验数据与评价指标

数据来源:采用ISPRS提供的标准测试数据集中的两个典型样例:

样例1(Vaihingen)城市区域,包含建筑物、道路、少量树木。

样例2(Jena Forest)茂密林区,地形有一定起伏。

评价指标:采用ISPRS官方推荐的指标:

I类错误:地面点被错误分类为非地面点的比例(漏分)。

II类错误:非地面点被错误分类为地面点的比例(错分)。

总体误差:(I类错误+II类错误)/2,综合反映算法性能。

##### 4.2 对比算法

选取以下三种代表性算法作为对比基准:

PTM:使用lidR R包中的实现。

SVM:使用scikit-learn库,特征包括Z,回波次数,强度,局部密度。

CSF:使用开源的CSF工具。

所有对比算法均使用其默认参数或通过交叉验证获得的最优参数。

##### 4.3 结果与讨论

###### 4.3.1 定量分析(见表1)

表1 不同算法在测试样例上的定量评价结果

算法	样例1(TE%)	样例1(I类%)	样例2(TE%)	样例2(I类%)
PTM	8.2	3.1	15.7	12.5
SVM	7.5	4.0	14.2	11.0
CSF	6.8	5.2	12.1	9.8
本文算法	5.9	3.8	9.3	6.2

从表中可以看出:在城市样例1中,所有算法表现接近,PTM的I类错误最低,但本文算法在TE上取得最优。在更具挑战性的林区样例2中,本文算法的优势非常明显,TE和I类错误均大幅低于其他算法。这证明了本文算法在处理植被覆盖和地形起伏场景下的优越性。

##### 4.3.2 定性分析

PTM在林区样例中,由于无法穿透密集树冠,其生成的DEM在树林下方出现大片空洞,即严重的过滤现象。SVM在城市样例中表现尚可,但在林区,由于训练样本与测试场景存在差异,将部分低矮灌木误判为地面,导致DEM表面不平滑。CSF能较好地保持地形的整体趋势,但在陡坡处,布料未能完全贴合地面,产生了一定的欠滤波(将部分坡面点判为非地面)。本文算法:生成的DEM在林区连续性好,能有效填充植被下的地面点;在陡坡处,自适应机制使其能紧密跟随真实地形,视觉效果最佳。

#### 5 结语

本文回顾了机载LiDAR点云滤波主流方法,分析其优劣后提出新颖自适应滤波算法。该算法利用多尺度法向量

熵和点密度智能分区,对不同区域差异化滤波,还引入坡度自适应布料模拟思想处理复杂地形。实验显示,在标准数据集,尤其是植被茂密、地形起伏大区域,该算法滤波精度显著优于PTM、SVM和CSF等经典算法,降低I类错误,保障DEM完整准确。不过,研究仍有拓展空间。未来可结合深度学习技术(如PointNet++)与几何先验构建端到端自适应滤波网络;优化算法计算效率以服务大规模点云实时处理;探索引入光谱信息增强特定地物区分能力。

#### 参考文献

- [1]徐旺.机载LiDAR点云数据滤波算法研究[D].东华理工大学,2022.
- [2]焦元冰,陆爱萍.基于机载LiDAR点云数据的改进滤波算法[J].测绘标准化,2022,38(03):42-46.
- [3]戚鑫鑫.复杂地形条件下机载LiDAR点云滤波算法研究[D].安徽理工大学,2025.
- [4]沈豫,葛灵斌,洪年祥.基于机载LiDAR点云数据的组合滤波方法研究[J].测绘与空间地理信息,2024,47(12):182-184+187.