

多通道光子计数激光雷达信息快速提取技术

盛娟^{1,2}, 张冰娜^{1,2*}, 翟景升^{1,2}, 王凤香¹

¹中国科学院上海技术物理研究所, 上海 200083;

²中国科学院大学, 北京 100049

摘要 光子计数激光雷达由于高灵敏度特性使得获取的回波数据存在大量噪声,对数据进行信号鉴别、相关噪声去除是后续进行三维点云数据应用的前提,因此快速提取信号光子具有很大的实际意义。上海技术物理研究所自主研发多通道光子计数激光雷达,对近海面进行数据采集。本实验对局部距离统计去噪算法进行改进,用于预处理后的数据去噪。将改进后算法与改进前算法和两种基于聚类的经典算法进行对比,实验结果表明,所提算法去噪精度为 97.80%,运行时间为 0.46 s,优于改进前算法和两种基于聚类的算法,满足多通道光子计数激光雷达信息快速提取的实际需求。

关键词 遥感; 激光雷达; 光子计数; 点云数据处理; 局部统计分析去噪; 聚类

中图分类号 O436

文献标志码 A

DOI: 10.3788/LOP222064

Fast Extraction Technology of Multichannel Photon Counting LiDAR Information

Sheng Juan^{1,2}, Zhang Bingna^{1,2*}, Zhai Jingsheng^{1,2}, Wang Fengxiang¹

¹Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

²University of Chinese Academy of Sciences, Beijing 100049, China

Abstract Due to the high sensitivity of photon counting LiDAR, there is a lot of noise in the collected echo data. Signal identification and removal of relevant noise are the prerequisites for the subsequent application of three-dimensional point cloud data. As a result, a quick extraction of the signal photons is extremely important in practice. Shanghai Institute of Technical Physics independently developed multichannel photon counting LiDAR to collect data near sea surface. In this experiment, after preprocessing, we first enhanced the local distance statistical denoising technique. The enhanced algorithm was then compared with the original algorithm and two classical clustering-based algorithms. The findings demonstrate that the proposed algorithm has a noise removal accuracy of 97.80% and running time of 0.46 s, which are superior to that of the original algorithm and two clustering-based algorithms; it, thus, meets the practical requirements for the fast extraction of multichannel photon counting LiDAR information.

Key words remote sensing; LiDAR; photon counting; point cloud data processing; local statistical analysis; denoising; clustering

1 引言

激光雷达(LiDAR)能够精确测量目标的距离、方位、速度以及形貌特性,具有探测识别和跟踪目标等能力^[1]。传统线性LiDAR一般采用较大的发射脉冲能量以及较低的飞行高度(较低的条带宽度),以获取足够的激光脉冲回波能量,但这都会使LiDAR数据获取的成本和时间代价^[2]增加,限制传统LiDAR的应用能力与范围。

光子计数LiDAR采用直接脉冲探测方法,通过记录发射激光脉冲和回波光子信号的时间来实现目标距离信息获取^[3-4];同时采用高灵敏度的单光子探测器(盖格型探测器),将线性探测体制下包含大量光子的回波波形探测转换为针对单个回波光子事件的“计数”,降低LiDAR的能量消耗和负载质量,解决了传统LiDAR能量消耗与采集频率之间的主要矛盾。国际上,从林肯实验室到美国航空航天局发射的ICESAT-II上搭载的对地测高LiDAR,已经有大量相当成熟的

收稿日期: 2022-07-13; 修回日期: 2022-07-14; 录用日期: 2022-07-29; 网络首发日期: 2022-08-09

通信作者: hustzbn@126.com

光子计数数据及相应算法研究。国内方面,光子计数 LiDAR 的研究起步较晚,中国科学院上海技术物理研究所最先进行了机载多波束光子计数 LiDAR 设计,并开展了相关外场试验。

由于光子计数 LiDAR 点云数据自身弱信号特性,其极易受到大气散射噪声、太阳辐射噪声等影响,因此信噪比较低,包含大量噪声。由于点云二维剖面中信号和噪声的某些属性,如距离、密度等统计量有明显差异,局部统计分析去噪算法利用统计量分布特征设置阈值剔除噪声。传统局部统计分析算法虽然耗时短,但去噪精度往往有限,本实验在此基础上进行改进,在提升去噪精度的同时在大量噪声中有效快速提取出信号光子点云。

目前,除了局部统计分析去噪算法,国际上主要有栅格图像处理去噪算法和点云聚类分析去噪算法这两类。其中,栅格图像处理去噪算法将剖面点云栅格化为二维图像并采用图像处理算法进行去噪,但点云栅格化过程中会造成有效信息的丢失,且不能充分利用点云的分布特征,使去噪精度降低^[5]。点云聚类分析去噪算法通常利用密度聚类分析将信号点云和噪声点云分成不同“类簇”,从而进行噪声剔除^[6],是一种目前主要的光子计数 LiDAR 去噪算法,在此基础上基于“类簇”的定义和阈值的选取有多种改进。本文采用改进前算法和两种经典的聚类算法与所提算法进行去噪效果与时间两个维度上的对比。

2 基本原理

2.1 数据来源

数据来自上海技术物理研究所自主研发的光子计数 LiDAR 系统,并在烟台开展了船载实验。图 1 是光子计数 LiDAR 探测原理图,激光器在 t_0 时刻向目标物体发射激光脉冲,经目标反射或散射后返回激光脉冲信号,接收系统检测到回波信号并记录下时间 t_1 ,则时间差 $\Delta t = t_1 - t_0$,即激光传播时间。已知激光在大气中传播速度为 c ,则激光器到探测目标的距离 $R = c\Delta t/2$ 。激光器系统接收到数据后交给数据处理系统做后续处理。

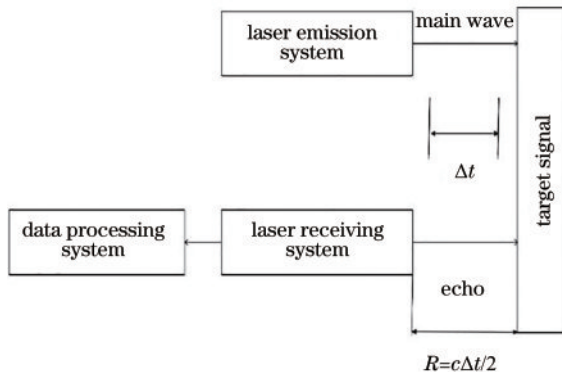


图 1 光子计数 LiDAR 探测原理图

Fig. 1 Schematic diagram of photon counting LiDAR detection

2.2 基于聚类的两种算法

光子计数 LiDAR 数据的信号光子通常比背景光子更密集,基于密度将信号光子与背景光子分离是一种有效的方法。基于聚类的算法原理是每个簇中的点具有比簇外更高的点密度,噪声区域的密度低于集群中的密度,1996 年 Ester 等^[6]最早在论文中进行了详细论述,其是一种被广泛运用的光子计数 LiDAR 去噪算法。实际运用时往往需要定义不同的“类簇”,通过设置合适的阈值去除噪声、提取信号光子。

基于邻域光子数的空间聚类算法(DBSCAN):对于光子点云二维剖面图中的每个点,给定半径的邻域内必须至少包含某个数量的点,即邻域中的密度必须超过某个阈值 D_{min} :

$$D_{min} = \text{mean}(D) + n \times \text{std}(D), \quad (1)$$

式中: $\text{mean}(D)$ 、 $\text{std}(D)$ 分别为频数直方图拟合高斯函数的平均值及方差。邻域的形状由 $p(t_p, h_p)$ 和 $q(t_q, h_q)$ 两点距离函数的选择决定,用 $\text{dist}(p, q)$ 表示。圆形搜索区域如下:

$$\text{dist}(p, q) = \sqrt{(t_p - t_q)^2 + (h_p - h_q)^2}. \quad (2)$$

其中的关键参数为邻域形状大小 $\text{dist}(p, q)$ 和邻域内光子点数阈值 D_{min} 。2015 年, Zhang 等^[7-8]对算法进行了改进,如图 2 所示。由于信号光子沿轨道方向呈线性紧密分布,将搜索区域的形状从圆形变为椭圆形,改进后的 $\text{dist}(p, q)$ 如下:

$$\text{dist}(p, q) = \sqrt{(t_p - t_q)^2/a^2 + (h_p - h_q)^2/b^2}, \quad (3)$$

式中: a 、 b 分别为椭圆长短轴。

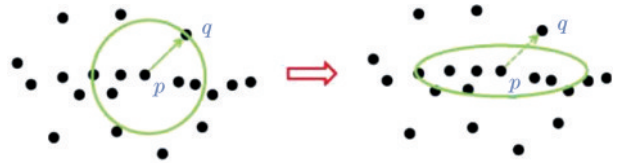


图 2 搜索区域形状更改

Fig. 2 Shape change of searching area

当水平方向上的光子相对于搜索区域中心的权重大于垂直方向上的光子时,水平方向上的连续光子更有可能被分类为表面光子,提高算法实际运用下信号光子提取的准确度。

基于 k 个最近邻数据点的聚类算法(KNN):计算每个光子到其 k 个最近邻光子的平均距离,并对计算结果进行直方图统计。目标物体处光子通常比背景光子更密集,也就是说,目标光子的 k 近邻距离小于背景光子的 k 近邻距离。设置合适的阈值 k_{min} 来提取信号光子。2014 年,夏少波等^[9]采用基于 k 个最近邻数据点的聚类算法,建立每个光子与其 k 个最近邻的累积距离 $\text{dist}(k)$ 的频率直方图,其中,累积距离大于平均距离的点被识别为噪声。此算法在运用时有两个关键参数,最近邻光子数 k 和阈值 k_{min} ,关系如下:

$$k_{\min} = \text{mean}(k) + n \times \text{std}(k)。 \quad (4)$$

2.3 改进的局部统计分析去噪算法(PDF2)

局部统计分析去噪算法按时间段划分光子计数 LiDAR 点云数据,然后将距离划分为小距离段单元,计算每个单元中的光子数并做直方图,选取合适的阈值提取信号光子。2014年, Moussavi 等^[10]建立海拔频率分布直方图,并根据直方图计算标准偏差 $\text{std}(\rho)$,截止阈值设置为 3 倍标准差。2018年, Nie 等^[11]和 Zhu 等^[12]使用 Moussavi 等的方法提取出可能的信号光子。

因此,首先采用局部统计分析去噪算法,对每个通道采集的数据进行分别处理,设置合适的时间和距离单元大小,根据需要,阈值 $\rho_{\min 1}$ 选取为平均值 $\text{mean}_1(\rho)$ 和 n_1 倍标准差 $\text{std}_1(\rho)$ 之和:

$$\rho_{\min 1} = \text{mean}_1(\rho) + n_1 \times \text{std}_1(\rho)。 \quad (5)$$

但这种去噪算法精度不够,去噪结果存在较多簇状噪声。在此基础上进行二次去噪,根据簇状噪声多的问题增大时间/距离单元,同时由于背景噪声已经减少,单元光子数标准差 $\text{std}_2(\rho)$ 增大,因此,在阈值 $\rho_{\min 2}$ 选取上减少了标准差倍数 n_2 :

$$\rho_{\min 2} = \text{mean}_1(\rho) + n_2 \times \text{std}_2(\rho)。 \quad (6)$$

2.4 精度评价指标

由于实际数据无法准确区分信号点云和目标点

云,人工判读是一种常见的精度评价方法。将空间明显离散点标记为噪声,目标物体附近点云标记为信号。为最大限度减少人工判读的主观误差,以 3 次判读的统计平均值作为结果。采用秦磊等^[13]的 F1-score 作为精度评价指标:

$$S_{F1} = 2N_{TP} / (2N_{TP} + N_{FP} + N_{FN}), \quad (7)$$

式中: N_{TP} 为被算法正确识别为信号的个数,即实际为信号点且被识别为信号点的光子个数; N_{FP} 为被算法错误识别为信号的个数,即实际为噪声点但被识别为信号点的光子个数; N_{FN} 为被算法漏掉的信号光子点云,即实际为信号点但被识别为噪声点的光子点云。

3 数据处理过程

3.1 实验数据介绍及预处理

原始数据文件以单列“dat”格式存储,内容包括数据头、编码器数据、有效回波个数、距离数据(单位为 cm)和数据尾。最大回波个数可设为 8、4、2 这 3 档,共 128 通道。本实验采用 8 次回波数据,一次发光的数据量长度为 3226 B,激光重复频率为 20 kHz,采用 0.5 s 内光子计数 LiDAR 采集的激光脉冲和对应的回波光信号数据。数据采集地点为近海面,如图 3(a)所示,图上有明显的船体等目标。

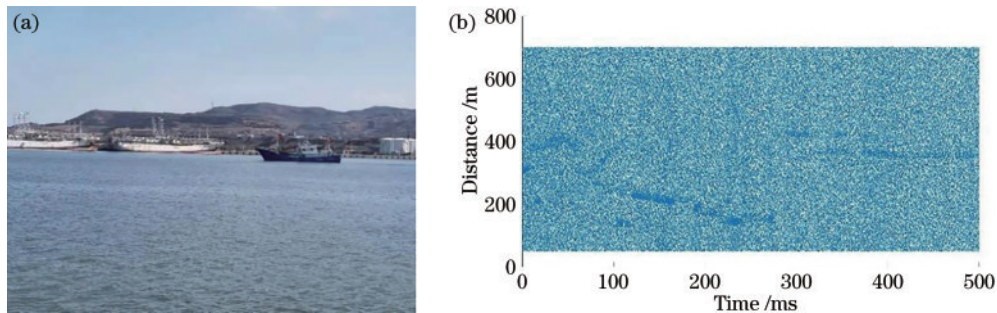


图 3 海面船体照片及点云数据。(a)海面船体;(b)点云数据

Fig. 3 Photos of sea hull and point cloud data. (a) Sea hull; (b) point cloud data

实验数据预处理如下:

1) 解析数据。根据 8 次回波次数可知,一次发光的数据量长度为 3226 B,将多波束扫描仪的原始数据重构为 8×3226 的数组。

2) 提取有效通道数据。将数组每行的有效回波个数列求和得到有效回波次数,去除无效通道的数据。

3) 距离解算。原始一个距离数据 3 B,单位为 cm,解析后以 m 为单位,合并为单个双精度数。

4) 去除过近过远的点,生成坐标对。根据数据的实际采集情况,将 50 m 以内与 1000 m 以外的数据去除掉,减少数据量以减少后续处理的计算量,节省时间。

预处理后的数据二维剖面图如图 3(b)所示。预处理后的多通道光子计数 LiDAR 点云数据特点如下:

1) 每个有效通道的有效回波次数不一致,单光子

数据水平方向上分布不均;

2) 噪声点云很多,信号点云隐藏在大量噪声之下。

3.2 去噪处理

对于基于邻域光子数的空间聚类算法,为增大水平方向上的光子权重,采用椭圆搜索区域。根据式(3),椭圆长轴 a 取 2,短轴 b 取 1。对于邻域内的光子点数阈值 D_{\min} ,式(1)中 n 取 2。算法在 Matlab2021 下 3 次运行平均时间为 17.22 s,提取出的信号光子数为 11317,如图 4(a)所示。

对于基于 k 个最近邻数据点的聚类算法,式(4)中两个关键参数是最近邻光子数 k 和阈值 k_{\min} ,其中, k 取 50(不包括每个光子本身), n 取 -3,即阈值 k_{\min} 为平均值与 3 倍标准差之差。算法在 Matlab2021 下 3 次运行平均时间为 1.78 s,提取出的信号光子数为 13861,如图 4(b)所示。

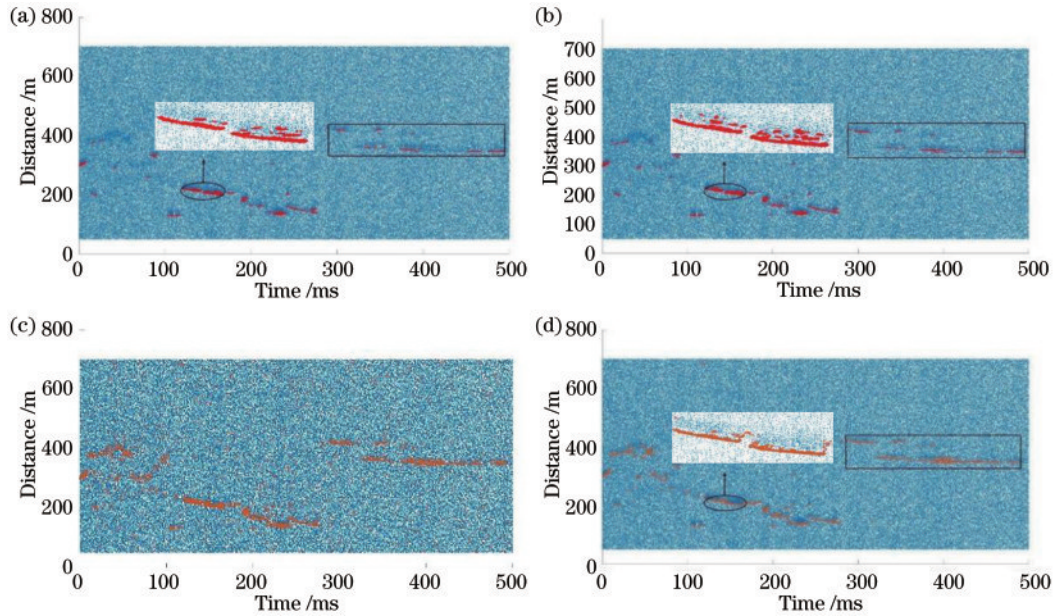


图4 去噪后的点云二维剖面图。(a) DBSCAN; (b) KNN; (c) PDF; (d) PDF2

Fig. 4 2D profile of point cloud after denoising. (a) DBSCAN; (b) KNN; (c) PDF; (d) PDF2

对于基于改进的局部统计分析去噪算法,第一次去噪选取的时间单元为100,距离单元为1,式(5)中 n_1 取5,即阈值 p_{min1} 选取为平均值和5倍标准差之和。算法在Matlab2021下3次运行平均时间为0.37 s,提取出的信号光子数为18361,如图4(c)所示。第二次去噪选取的时间单元为200,距离单元为3,式(6)中 n_2 取2,即阈值 p_{min2} 选取为平均值和2倍标准差之和。算法在Matlab2021下3次运行平均时间为0.46 s,提取出的信号光子数为12166,如图4(d)所示。

4 分析与讨论

由图4可知,传统的局部统计分析去噪算法去噪结果含有过多噪声,去噪效果达不到要求。相较于两种聚类算法,所提改进局部统计分析去噪算法保留了更多的信号光子,去噪后的二维剖面图4(d)相较于图4(a)、(b)右边方框内目标信号光子更加连续,同时对比左边椭圆区域的局部放大图可知,改进的局部统计分析去噪算法将噪声去除得更加干净。

表1 去噪效果对比

Table 1 Comparison of denoising effect

Algorithm	N	N_{TP}	N_{FP}	N_{FN}	F1-score /%	Time /s
DBSCAN	582289	10974	343	1270	93.15	17.22±0.15
KNN	582289	11505	2356	739	88.14	1.78±0.02
PDF	582289	11964	6397	280	78.18	0.37±0.01
PDF2	582289	11937	229	307	97.80	0.46±0.01

根据表1去噪结果可知,去噪前的光子总数(N)为582289,所提改进局部统计分析去噪算法精度评价指标F1-score为97.80%,高于基于聚类的两种算法、远高于改进前算法。对去噪时间进行分析,在同样的运

行环境下,所提改进局部统计分析去噪算法用时0.46 s,与改进前算法的0.37 s相差不大,快于基于聚类的两种算法,具有明显的时间优势,可以快速提取出信号光子。

5 结论

LiDAR是获取三维地理信息的主要途径,光子计数LiDAR解决了传统LiDAR能量消耗与采集频率之间的主要矛盾,同时回波数据存在大量噪声,快速去除噪声提取信号光子具有很大的实际意义。对上海技术物理研究所自主研发的多通道光子计数LiDAR在近海面采集的数据进行了相关处理。首先对采集到的数据进行预处理,基于局部距离统计去噪提出改进的点云去噪算法,对预处理后的数据进行去噪。将改进算法与改进前的局部统计分析去噪算法和两种基于聚类的传统算法进行对比,实验结果表明,根据目视法,所提改进算法精度最高,为97.80%,对噪声的去除效果更好。在去噪时间上,同样的运行环境下,所提算法用时0.46 s,与改进前算法相差不大,大幅度小于基于聚类的两种算法,满足多通道光子计数LiDAR信息快速提取的实际要求。

参考文献

- [1] 龚威, 史硕, 陈博文, 等. 机载高光谱激光雷达成像技术发展与应用[J]. 光学学报, 2022, 42(12): 1200002. Gong W, Shi S, Chen B W, et al. Development and application of airborne hyperspectral LiDAR imaging technology[J]. Acta Optica Sinica, 2022, 42(12): 1200002.
- [2] Swatantran A, Tang H, Barrett T, et al. Rapid, high-resolution forest structure and terrain mapping over large areas using single photon lidar[J]. Scientific Reports,

- 2016, 6: 28277.
- [3] 白一波, 潘康立, 耿林. 激光测距的空间卷积神经网络信号处理研究[J]. 中国激光, 2021, 48(23): 2304001.
Bai Y B, Pan K L, Geng L. Signal processing of spatial convolutional neural network for laser ranging[J]. Chinese Journal of Lasers, 2021, 48(23): 2304001.
- [4] 刘芳华, 贺岩, 罗远, 等. 基于光子计数激光雷达的运动小目标测距测速系统[J]. 中国激光, 2021, 48(13): 1310001.
Liu F H, He Y, Luo Y, et al. Moving target distance and velocity measurement system based on photon-counting lidar[J]. Chinese Journal of Lasers, 2021, 48(13): 1310001.
- [5] Awadallah M, Ghannam S, Abbott L, et al. Active contour models for extracting ground and forest canopy curves from discrete laser altimeter data[C]//Proceedings of 13th International Conference on LiDAR Applications for Assessing Forest Ecosystem, October 9, 2013, Beijing, China. Beijing: Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, 2013.
- [6] Ester M, Kriegel H P, Sandor J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[M]. Menlo Park: AAAI Press, 1996.
- [7] Zhang J S, Kerekes J, Csatho B, et al. A clustering approach for detection of ground in micropulse photon-counting LiDAR altimeter data[C]//2014 IEEE Geoscience and Remote Sensing Symposium, July 13-18, 2014, Quebec City, QC, Canada. New York: IEEE Press, 2014: 177-180.
- [8] Zhang J S, Kerekes J. An adaptive density-based model for extracting surface returns from photon-counting laser altimeter data[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(4): 726-730.
- [9] 夏少波, 王成, 习晓环, 等. ICESat-2 机载试验点云滤波及植被高度反演[J]. 遥感学报, 2014, 18(6): 1199-1207.
Xia S B, Wang C, Xi X H, et al. Point cloud filtering and tree height estimation using airborne experiment data of ICESat-2[J]. Journal of Remote Sensing, 2014, 18(6): 1199-1207.
- [10] Moussavi M S, Abdalati W, Scambos T, et al. Applicability of an automatic surface detection approach to micro-pulse photon-counting lidar altimetry data: implications for canopy height retrieval from future ICESat-2 data[J]. International Journal of Remote Sensing, 2014, 35(13): 5263-5279.
- [11] Nie S, Wang C, Xi X, et al. Estimating the vegetation canopy height using micro-pulse photon-counting LiDAR data[J]. Optics Express, 2018, 26(10): A520-A540.
- [12] Zhu X X, Nie S, Wang C, et al. A ground elevation and vegetation height retrieval algorithm using micro-pulse photon-counting lidar data[J]. Remote Sensing, 2018, 10(12): 1962.
- [13] 秦磊, 邢艳秋, 黄佳鹏, 等. ICESat-2 机载实验光子云数据自适应去噪及分类算法[J]. 遥感学报, 2020, 24(12): 1476-1487.
Qin L, Xing Y Q, Huang J P, et al. Adaptive denoising and classification algorithms for ICESat-2 airborne experimental photon cloud data of 2018[J]. Journal of Remote Sensing, 2020, 24(12): 1476-1487.