

## 单光子激光雷达的时间相关多深度估计

邬 森, 陆 侯, 冒添逸, 何伟基\*, 陈 钱

(南京理工大学 电子工程与光电技术学院, 江苏 南京 210094)

**摘要:** 单光子激光雷达广泛应用于获得三维场景的深度和强度信息。对于多表面目标, 如激光经过半透明表面上时, 一个像素上探测到的回波信号可能包含多个峰。传统方法在低光子或相对较高的背景噪声水平下无法准确估计多深度图像。因此, 提出了一种单光子激光雷达时间相关多深度估计方法。该方法利用信号响应的时相关性, 对点云数据进行多深度快速去噪, 能够从背景噪声中识别每个像素上来自多个表面的信号响应。并基于该信号响应集合的泊松分布模型, 通过全变分正则化引入像素之间的空间相关性, 建立多深度估计成本函数。使用快速收敛的交替方向乘子算法从成本函数中估计深度图像。实验结果表明, 所提方法在距离约为 1 km 处的多深度目标上, 相较于常规方法估计深度图像的均方根误差减少了至少 27.05%, 信号重建误差比提高了至少 18.39%, 同时数据量减少至原来的 4%。证明该方法能够以更小的内存需求和计算复杂度提高单光子激光雷达的多深度图像估计性能。

**关键词:** 单光子激光雷达; 多深度图像估计; 时间相关性; 交替方向乘子算法

**中图分类号:** TN958 **文献标志码:** A **DOI:** 10.3788/IRLA20210885

## Time-correlated multi-depth estimation of Single-photon lidar

Wu Miao, Lu Yu, Mao Tianyi, He Weiji\*, Chen Qian

(School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

**Abstract:** Single-photon lidar has been widely used to obtain depth and intensity information of a three-dimensional scene. For multi-surface targets, such as when the laser transmit through a translucent surface, the echo signal detected on one pixel may contain multiple peaks. Traditional methods cannot accurately estimate multi-depth images under low photon or relatively high background noise levels. Therefore, a time-correlated multi-depth estimation method was introduced. Based on the time correlation of the signal responses, a multi-depth fast denoising method was adopted to point cloud data, and could identify the signal responses of multiple surfaces from background noise on each pixel. Considering the Poisson distribution model of the signal response set, the spatial correlation between pixels was introduced through total variation (TV) regularization to establish a multi-depth estimation cost function. The fast-converging alternating direction method of multipliers (ADMM) was used to estimate the depth image from the cost function. Experimental results on a multi-depth target at a distance of about 1 km show that the root mean square error (RMSE) and signal to reconstruction-error ratio (SRE) of the depth image estimated by the proposed method can be at least 27.05% and 18.39% better than that of other state-of-the-art methods. In addition, the data volume of this method is reduced to 4% of the original. It is proved that this method can effectively improve the multi-depth image estimation of single-photon lidar with

收稿日期:2021-11-24; 修订日期:2022-01-11

基金项目:国家自然科学基金(61875088, 62005128)

作者简介:邬森, 女, 博士生, 主要从事单光子成像技术与应用方面的研究。

导师(通讯作者)简介:何伟基, 男, 研究员, 博士, 主要从事光电成像探测与信号处理方面的研究。

smaller memory requirements and computational complexity.

**Key words:** single-photon lidar; multi-depth image estimation; temporal correlation; ADMM

## 0 引言

三维成像在遥感、水下探索和地形测量等领域取得了广泛的应用<sup>[1-3]</sup>。激光雷达系统采用工作于盖革模式下的雪崩光电二极管 (Geiger-Mode Avalanche Photodiodes, GM-APD), 它具有单光子灵敏度和皮秒级时间分辨率<sup>[4]</sup>。时间相关单光子计数 (Time Correlated Single-Photon Counting, TCSPC) 技术<sup>[5]</sup>为每个像素提供光子计数相对于飞行时间 (Time of flight, TOF) 的直方图, 其中计数的位置和数量分别对应目标的深度和反射率信息, 从而重建出三维场景。

实际应用中, 当激光通过部分反射或部分遮挡 (例如: 窗口或遮蔽物) 的目标时, 单个像素上需要估计多个深度, 以完全重建场景中存在的所有表面。然而, 对于此类目标, 尤其是在弱回波信号和高背景噪声的远距离室外场景中, 精确的深度估计具有挑战性。为了解决上述问题, 除了优化激光雷达探测机制<sup>[6-7]</sup>和硬件系统外<sup>[8-10]</sup>, 学者们已经提出许多有效的后端处理算法来改进深度估计质量和鲁棒性。在假设一个像素仅接受来自一个表面的回波信号时, Rapp 等人<sup>[11]</sup>通过逐像素的时间域开窗来滤除噪声, 并建立超像素来提高低光子水平下算法的鲁棒性。Halimi 等人<sup>[12]</sup>利用了非局部空间相关性, 使用交替方向乘子算法 (Alternating Direction Method of Multipliers, ADMM) 估计深度图像, 虽然这种方法采用非均匀采样策略来降低计算成本, 但它仅从二维图像中进行估计, 而忽略了点云数据时间维度的影响。Lindell 等人<sup>[13]</sup>和 Sun 等人<sup>[14]</sup>采用多传感器融合策略, 将来自相机的二维图像与来自激光雷达的点云数据结合, 使用深度卷积网络来得到精确的深度图像, 但是相机与激光雷达之间配准复杂, 且对于远距离成像来说, 相机的价格昂贵且尺寸大。对一个像素存在多个表面的情况<sup>[15]</sup>, Shin 等人<sup>[16]</sup>建立数据项与 $l_1$ 正则化相结合的成本函数, 并使用改进的迭代阈值收缩算法 (Iterative Shrinkage Thresholding Algorithm, ISTA) 求出最优解, 实现多深度估计。Tachella 等人<sup>[17]</sup>将可逆跳跃马尔可夫链蒙特卡洛方法与空间点过程相结合, 能够精确

地重建出复杂的三维场景, 但计算时间太长。

文中提出了一种新的单光子激光雷达多深度估计方法。首先, 该方法基于信号响应的时域相关性, 使用快速的逐像素时间域开窗法从噪声中提取出多深度的信号, 同时能够极大缩减激光雷达数据时间轴的长度。其次, 采用优化框架, 结合多深度的泊松分布模型与引入目标的空间相关性的全变分 (Total Variation, TV) 正则化项, 建立深度估计成本函数。最后, 使用快速收敛的 ADMM 算法从成本函数中估计出深度图像。

## 1 单光子探测模型

在单光子激光雷达探测过程中, 为避免距离模糊, 一次只允许发射一个激光脉冲, 目标的最大距离应小于  $c/2f$ , 其中  $f$  为激光脉冲重频,  $c$  为真空中的光速。由此, 可以假设目标深度  $z \in [0, c/2f)$  和光子飞行时间  $t = 2z/c \in [0, 1/f)$ 。由像素个数  $N = N_r \times N_c$  和时间单元个数  $T = 1/f/tcube$  定义的激光雷达数据  $s$  表示所有像素上所有时间单元的光子计数总数, 其中  $tcube$  表示时间单元的宽度。像素  $(i, j)$  处第  $t$  个时间单元的光子计数  $s_{i,j,t}$  的泊松分布可以描述为<sup>[18]</sup>:

$$s_{i,j,t} \sim \text{Poisson}(\lambda_{i,j,t}) \quad (1)$$

其中

$$\lambda_{i,j,t} = \sum_{l=1}^L r_{i,j,l} g(t - t_{i,j,l}) + (b_{i,j} + d_{i,j}) \quad (2)$$

式中:  $\lambda_{i,j,t}$  代表平均光子计数;  $g(t)$  表示激光雷达的系统脉冲响应;  $L$  为目标表面数;  $r_{i,j,l}$  为第  $l$  个表面上目标点的反射率;  $t_{i,j,l}$  为与第  $l$  个表面上目标点深度  $z_{i,j,l}$  对应的光子飞行时间;  $b_{i,j}$  和  $d_{i,j}$  分别表示由背景光和探测器热效应引起的噪声光子计数, 对所有时间单元和表面可以假定为常数。

## 2 多深度快速去噪方法

在探测过程中, 激光脉冲的时域展宽、系统的时间抖动、环境背景和探测器暗噪声等多种因素引入了深度估计的误差。尤其是在远距离成像中, 微弱的回波信号往往被强背景噪声所淹没, 无法准确估计深度。

针对单深度目标,已经提出了快速去噪方法(Fast-Denoising method with the Temporal Correlation of Photons, FDTCP)<sup>[9]</sup>,该方法通过逐像素开窗法来分离信号与噪声,基本原理在于噪声响应和信号响应概率分布特征的不同:信号响应主要集中在激光脉冲较窄的半峰全宽 $T_p$ (Full Width at Half Maximum, FWHM)内,而噪声响应则均匀分布在整个时间维度上,与激光脉冲无关,因此可以假设为常数。根据上述信号响应的时相关性,在每个像素的光子计数直方图上得到固定时间区间 $T_w$ ( $T_w$ 介于 $T_p$ 和激光周期 $T_r$ 之间)内光子数之和最多的响应的集合,并认为是信号响应的集合。在理想情况下,时间区间以 $2z_{i,j}/c$ 为中心,宽度为 $T_w$ ,足以捕获大部分或所有的信号响应。FDTCP以单光子探测器对回波信号引发的一个响应为中心,如果其前后脉宽 $T_p$ 范围内发生的响应数大于预设的阈值,它将被视为信号响应。然而,在低光子水平下,该方法很有可能会因连续响应的概率很低而崩溃。Rapp 等人<sup>[11]</sup>在时间维度上滑动一个长度为 $T_w$ 的窗口,以选择响应数最多的集合。对于大小为 $N_r \times N_c \times T$ 的单深度激光雷达数据,该方法的计算复杂度为 $O(N_r \times N_c \times T)$ ,当远距离和高时空分辨率激光雷达成像(即大 $N_r \times N_c \times T$ )时,计算成本急剧增大。

基于之前的工作<sup>[9]</sup>,文中提出多深度快速去噪方法,将逐像素的时间域开窗法拓展到多表面目标,从噪声中提取出多深度的信号,并受二分法的启发,使用快速搜索方法加速其运算速度,每次搜索可以舍弃三分之二的值来缩小范围。具体而言,对每个像素 $(i, j)$ 上时间单元个数为 $T$ 的直方图 $s_{i,j} = \{s_{i,j,t} : 1 \leq t \leq T\}$ 进行以下步骤。

设当前阶段为 $l, 1 \leq l \leq L_{max}$ ,其中, $L_{max}$ 为预设的最大目标表面数。首先,设定搜索初值 $t_{i,j}^{Left} = 1, t_{i,j}^{Right} = T$ ,将直方图划分为三个长度相等的子集,并计算它们的光子计数之和:

$$s_{i,j}^{Left} = \sum_{t=t_{i,j}^{Left}}^{t_{i,j}^{Mid}} s_{i,j,t} \quad (3)$$

$$s_{i,j}^{Mid} = \sum_{t=t_{i,j}^{Mid}}^{t_{i,j}^{Right}} s_{i,j,t} \quad (4)$$

$$s_{i,j}^{Right} = \sum_{t=t_{i,j}^{Mid}}^{t_{i,j}^{Right}} s_{i,j,t} \quad (5)$$

式中:  $t_{i,j}^{mLeft} = t_{i,j}^{Left} + (t_{i,j}^{Right} - t_{i,j}^{Left})/4$ ;  $t_{i,j}^{Mid} = t_{i,j}^{Left} + (t_{i,j}^{Right} - t_{i,j}^{Left})/2$ ,  $t_{i,j}^{mRight} = t_{i,j}^{Left} + (t_{i,j}^{Right} - t_{i,j}^{Left}) \times 3/4$ 。比较  $s_{i,j}^{Left}$ 、 $s_{i,j}^{Mid}$ 、 $s_{i,j}^{Right}$ ,保留光子计数之和最大的子集(若两个子集计数相等且大于另外一个子集,则任取两个计数相等的子集中的一个),记为 $\tilde{s}_{i,j}^l$ ,并使用 $\tilde{s}_{i,j}^l$ 的第一个和最后一个时间单元更新搜索值 $t_{i,j}^{Left}$ 、 $t_{i,j}^{Right}$ 。由此,  $\tilde{s}_{i,j}^l = \{s_{i,j,t} : t_{i,j}^{Left} \leq t \leq t_{i,j}^{Right}\}$ ,  $T_m^l(i, j) = t_{i,j}^{Left}$ ,用于索引开窗后集合的第一个时间单元。对长度为 $[t_{i,j}^{Left}, t_{i,j}^{Right}]$ 的 $\tilde{s}_{i,j}^l$ 进行下一次搜索,生成三个新的子集: $s_{i,j}^{Left}$ 、 $s_{i,j}^{Mid}$ 、 $s_{i,j}^{Right}$ 。在每次搜索中都保留光子计数最多的子集,直到满足条件 $t_{i,j}^{Left} + T_w \geq t_{i,j}^{Right}$  or  $T$ ,则搜索停止。得到像素 $(i, j)$ 上光子计数最多的集合 $\tilde{s}_{i,j}^l = \{s_{i,j,t} : T_m^l(i, j) \leq t \leq T_m^l(i, j) + T_w\}$ 。

然后,根据直方图的噪声水平设定阈值 $K$ ,用于分离信号与噪声。如果集合 $\tilde{s}_{i,j}^l$ 的光子计数总和 $\tilde{s}_{sum}^l(i, j) = \sum_{t=T_m^l(i, j)}^{T_m^l(i, j)+T_w} s_{i,j,t}$ 小于阈值 $K$ ,则被判定为噪声响应,忽略该像素的所有计数;反之,则判定为来自第 $l$ 个表面的信号响应,保留区间长度为 $T_w$ 的光子计数:

$$\tilde{s}_{i,j} = \begin{cases} \tilde{s}_{i,j}, \tilde{s}_{sum}^l(i, j) \geq K \\ \emptyset, \tilde{s}_{sum}^l(i, j) < K \end{cases} \quad (6)$$

若 $l = L_{max}$ ,即达到了预设的最大目标表面数,或者任意像素上集合 $\tilde{s}_{i,j}^l$ 的光子计数之和都小于阈值,均被判定为噪声响应,此时 $\tilde{s}_{i,j}^l = \emptyset, \forall 1 \leq i \leq N_r, 1 \leq j \leq N_c$ ,则完成了对所有表面上信号响应的搜索,当前目标表面数为 $L$ ,每个像素 $(i, j)$ 上信号响应集合为 $\tilde{s}_{i,j} = \bigcup_{1 \leq l \leq L} \tilde{s}_{i,j}^l$ 。否则,令 $s_{i,j}(t) = 0$ ,其中 $T_m^l(i, j) \leq t \leq T_m^l(i, j) + T$ ,并令 $l = l + 1$ ,重复上述步骤。

多深度快速去噪方法示例如图 1 所示,其中红色代表均匀分布的噪声响应。阶段为 $l$ 时,得到第 $l$ 个表面上目标的信号响应集合 $\tilde{s}_{i,j}$ ;阶段为 $l+1$ 时,蓝色、粉红色和绿色分别代表三种泊松分布的信号响应,它们的中心分别位于 $t_{i,j,l+1}^1$ 、 $t_{i,j,l+1}^2$ 和 $t_{i,j,l+1}^3$ 。可以看出,快速搜索方法在每次搜索时将集合划分为三个子集,即使信号恰好位于一个子集的边缘,其他两个子集中也始终存在一个子集覆盖绝大部分的信号响应。此外,对于相同大小的数据,单深度的计算复杂度降低至 $O(N_r \times N_c \times \log(T))$ 。最终,输出尺寸为 $N_r \times N_c \times T_w \times L$ 的信号响应集合。

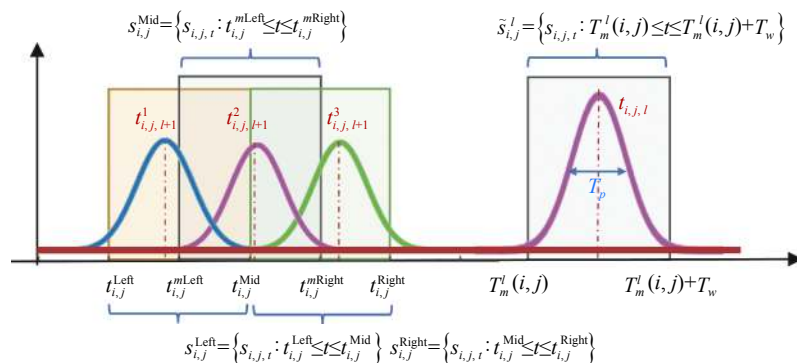


图 1 多深度快速去噪方法示例

Fig.1 Illustrative example of the multi-depth fast-denoising method

### 3 深度估计方法

首先采用优化框架建立成本函数,将深度估计问题转化成最优求解问题。基于第 2 节中获得的信号响应集合的泊松分布模型构造数据项<sup>[20]</sup>,具体公式为:

$$L(t_{n,l}; \tilde{s}_n^l) = - \sum_{t=T_m^l(n)}^{T_m^l(n)+T_w} s_{n,t} \log(g(t-t_{n,l})) \quad (7)$$

式中:  $\tilde{s}_n^l$  为像素  $n$  上的第  $l$  个表面对应的信号响应集合;  $t_{n,l}$  为像素  $n$  上第  $l$  个表面上待求解的深度值,  $1 \leq n \leq N$ 。为了保留边缘的细节特征同时平滑深度图像,基于目标场景的空间相关性引入全变分 TV 正则化模型,建立与数据项结合的深度估计成本函数  $C$ ,具体公式为:

$$C(\hat{\mathbf{t}}) = L(\hat{\mathbf{t}}) + \lambda TV(\hat{\mathbf{t}}) \quad (8)$$

式中:  $\lambda$  为正则化参数;  $\hat{\mathbf{t}} \in \mathbb{R}^{N \times L}$  为多表面深度图,  $\hat{\mathbf{t}} = \{t_{n,l} : 1 \leq n \leq N, 1 \leq l \leq L\}$ ;  $TV(\hat{\mathbf{t}}) = \|\mathbf{D}_h \hat{\mathbf{t}}\|_1 + \|\mathbf{D}_v \hat{\mathbf{t}}\|_1$ , 其中  $\mathbf{D}_h$  和  $\mathbf{D}_v$  分别代表水平和垂直方向的一阶差分。

其次,使用 ADMM 算法从成本函数中估计深度图像。ADMM 算法的内存需求与 6 倍的数据尺寸成正比,在多深度快速去噪方法中将单深度估计的数据尺寸由  $N \times T$  减小到  $N \times T_w$ , 因此,相应的内存需求和计算时间可以显著减少。ADMM 通常解决两个变量的等式约束优化问题,具体如下:

$$\begin{aligned} & \text{minimize } f(\mathbf{u}) + g(\mathbf{v}) \\ & \text{subject to } \mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{v} = \mathbf{c} \end{aligned} \quad (9)$$

式中:  $\mathbf{u} \in \mathbb{R}^{m \times n}$ ;  $\mathbf{v} \in \mathbb{R}^{p \times q}$ ;  $\mathbf{A} \in \mathbb{R}^{d \times m}$   $\mathbf{B} \in \mathbb{R}^{d \times p}$ ;  $\mathbf{c} \in \mathbb{R}^{d \times 1}$ ,  $g$  和

$f$  为适当的闭凸函数。参考优化问题的结构并引入拉格朗日算子,将公式 (7) 的成本函数  $C$  扩展为简化后的增广拉格朗日形式:

$$\begin{aligned} & \text{argmin}_{\hat{\mathbf{t}}, \mathbf{v}, \mathbf{d}} C(\hat{\mathbf{t}}, \mathbf{v}, \mathbf{d}) = \\ & \text{argmin}_{\hat{\mathbf{t}}, \mathbf{v}, \mathbf{d}} \left[ - \sum_{n=1}^N \sum_{l=1}^L \sum_{t=T_m^l(n)}^{T_m^l(n)+T_w} s_{n,t} \log(g(t-t_{n,l})) + \lambda TV(\mathbf{v}) + \right. \\ & \left. (\rho/2) \|\hat{\mathbf{t}} - \mathbf{v} + \mathbf{d}\|_2^2 - (\rho/2) \|\mathbf{d}\|_2^2 \right] \end{aligned} \quad (10)$$

式中:  $\rho > 0$  为惩罚参数;  $\mathbf{v} = \hat{\mathbf{t}}$  为构建的等式约束条件;  $\mathbf{d}$  为拉格朗日算子。利用 ADMM 算法求解公式 (9) 包括下面三个独立的子问题:

$$\begin{aligned} t_{n,l}^{k+1} = & \text{argmin}_{t_{n,l}} \left( - \sum_{t=T_m^l(n)}^{T_m^l(n)+T_w} s_{n,t} \log(g(t-t_{n,l})) + (\rho/2) \right. \\ & \left. \|t_{n,l} - v_{n,l}^k + d_{n,l}^k\|_2^2 \right) \end{aligned} \quad (11)$$

$$\mathbf{v}^{k+1} = \text{argmin}_{\mathbf{v}} \left( \lambda TV(\mathbf{v}) + (\rho/2) \|\hat{\mathbf{t}}^{k+1} - \mathbf{v} + \mathbf{d}^k\|_2^2 \right) \quad (12)$$

$$\mathbf{d}^{k+1} = \mathbf{d}^k + \hat{\mathbf{t}}^{k+1} - \mathbf{v}^{k+1} \quad (13)$$

式中:  $t_{n,l}^{k+1} \in \{T_m^l(n), T_m^l(n) + 1, \dots, T_m^l(n) + T_w\}$ 。因此,可以通过遍历它的范围来求解,由于时域开窗缩减了时间维度的长度,该方案的计算成本是可以接受的。迭代求解三个子问题,当  $\max(\|\hat{\mathbf{t}}^{k+1} - \hat{\mathbf{t}}^k\|_2, \|\mathbf{v}^{k+1} - \mathbf{v}^k\|_2, \|\mathbf{d}^{k+1} - \mathbf{d}^k\|_2) < tol$ , 其中  $tol$  为容限值,停止迭代。最终得到满足精度的解  $\hat{\mathbf{t}}$ , 代表估计的多深度图像。

### 4 仿真结果与分析

#### 4.1 评价指标与比较方法

为了验证所提出方法的性能,通过实验横向对比

了经典的极大似然估计 (Maximum Likelihood Estimation, MLE)、FDTPC<sup>[19]</sup>、Rapp<sup>[11]</sup> 以及针对多深度目标的 ManiPOP<sup>[17]</sup>、SPISTA 方法<sup>[16]</sup> 等。MLE 方法假设系统脉冲近似为高斯分布, 且背景噪声可忽略不计, 在此假设上使用对数匹配滤波, 得到深度估计的具体公式为:

$$t_{i,j}^{ML} = \operatorname{argmax} \sum_{t=1}^T s_{i,j,t} \log(g(t-t_{i,j})) \quad (14)$$

文中使用两个指标来评价实验结果: 信号重建误差比 (Signal to Reconstruction-Error ratio, SRE) 和均方根误差 (Root Mean Square Error, RMSE),  $SRE = 10 \log_{10} (\|\mathbf{x}\|^2 / \|\mathbf{x} - \mathbf{x}_{true}\|^2)$ ,  $RMSE = \sqrt{\|\mathbf{x} - \mathbf{x}_{true}\|^2 / N}$ , 其中  $N$  表示像素总数,  $\mathbf{x}$  和  $\mathbf{x}_{true}$  分别为估计的深度图和参考深度图。

### 4.2 几何形状上结果

第一个仿真数据使用不同大小的几何形状构成的场景<sup>[17]</sup>, 包括  $N_r \times N_c = 99 \times 99$  个像素点和  $T = 4\,500$  个时间单元, 每个时间单元的宽度为 2 ps, 设置激光脉宽  $T_p$  约为 0.8 ns。如图 2(a) 所示, 该场景中不同大小的几何形状在不同位置上重叠, 导致每个像素上的表面数不同, 位于 0~3 范围之内。为了验证所提出方法在低信号光子和高背景噪声环境中的鲁棒性, 图 2 比较了在不同每像素接收平均光子数 (Photon Per Pixel, PPP) 水平和信背比 (Signal-to- Background Ratio, SBR) 下的所提出方法的深度估计结果。对于不同 PPP(11.12、2.23、1.08) 和 SBR(3.40、0.99、0.13) 均使用  $T_w = 100$  的窗口大小。

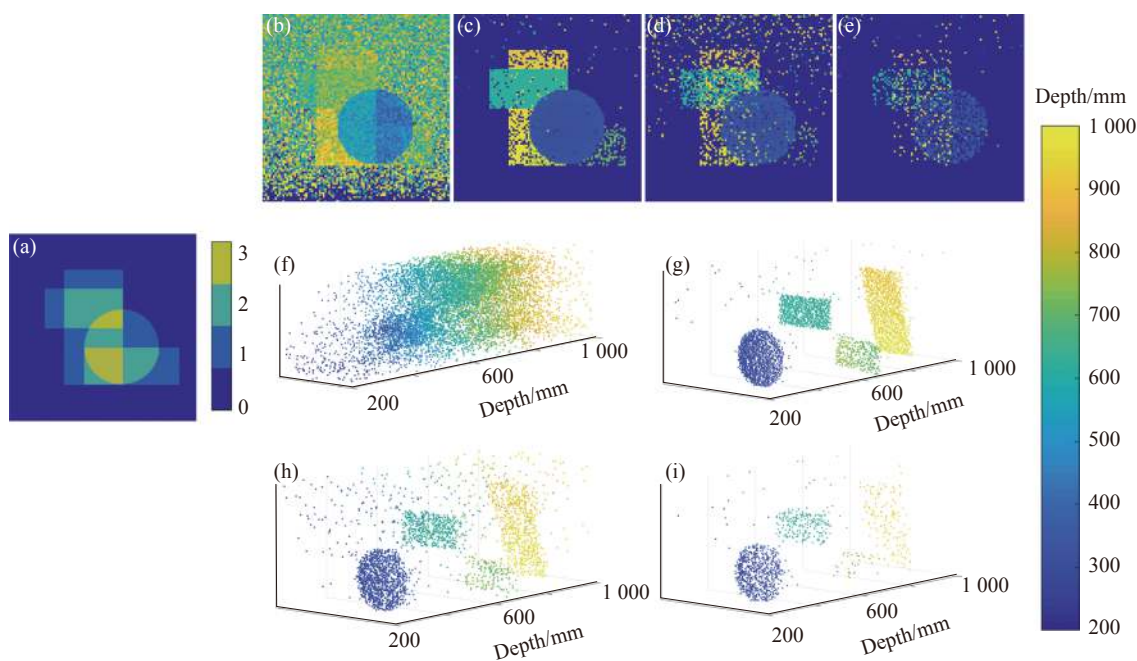


图 2 不同 PPP 和 SBR 下几何形状的深度与点云图。(a) 几何形状上的每像素表面数; (b)  $PPP=11.12$  和  $SBR=3.40$  下 MLE 估计的深度图; (c)~(e)  $PPP=11.12$  和  $SBR=3.40$ 、 $PPP=2.33$  和  $SBR=0.99$ 、 $PPP=1.08$  和  $SBR=0.13$  下所提出方法估计的深度图及 (f)~(i) 对应点云图  
Fig.2 Depth images and 3D point clouds of s of geometric shapes under different PPP and SBR. (a) Surfaces per pixel on geometric shapes; (b) Depth image estimated by MLE under  $PPP=11.12$ ,  $SBR=3.40$ ; (c)-(e) Depth images estimated by the proposed method respectively under  $PPP=11.12$  and  $SBR=3.40$ ,  $PPP=2.33$  and  $SBR=0.99$ ,  $PPP=1.08$  and  $SBR=0.13$ , and (f)-(i) the corresponding 3D point clouds

图 2(b) 和 (f) 显示了 MLE 在  $PPP=11.12$  和  $SBR=3.40$  下的估计结果, 此时目标被噪声所淹没, 且不同表面的深度信息互相干扰, 偏差较大。图 2(c) 和 (g) 显示了所提出方法在同样条件下的结果, 有效地处理了背

景噪声, 同时很好地区分出不同几何形状, 能够有效处理 0~3 个表面数的目标, 恢复出场景绝大部分深度信息。当光子计数水平和信背比变低时, 所提出方法仍能展现出不错的结果, 证明其鲁棒性。

### 4.3 艺术场景上结果

第二个仿真数据使用的是前方放置半透明平面的艺术场景<sup>[17]</sup>, 包括  $N_r \times N_c = 183 \times 121$  个像素点和  $T = 4\ 500$  个时间单元, 每个时间单元的宽度为 2 ps。该场景每个像素由两个表面组成。仿真实验设置激光脉宽  $T_p$  为 90 ps, 半透明平面后  $PPP=6.89$ ,  $SBR=14.57$ 。对于 FDTCP、Rapp 方法和文中所提出方法中的逐像素时域开窗, 均使用  $T_w = 100$  的窗口大小。所有算法根据在调整参数后得到的深度图像、RMSE、SRE 的最佳结果来在试验中选择参数。

图 3 和图 4 分别为艺术场景上不同方法的深度图像与点云图像, 数值结果如表 1 所示。可以看出, 所有方法的深度估计性能都优于 MLE 方法。

FDTCP、Rapp 方法与所提出方法考虑到时间相关性, 使用逐像素时域开窗的方法极大程度降低了背景噪声。FDTCP 和 Rapp 方法针对的是每个像素上只存在一个表面的情况, 无法同时保留 300 mm 处的平面与后面的复杂场景, 导致深度估计的 RMSE 与 SRE 较差。此外, FDTCP 的计算复杂度是最低的, 但由于该算法没有像 Rapp 方法一样考虑到空间先验信息, 它的深度估计精度也是最差的。针对多表面目标, ManiPOP、SPISTA 和文中所提出方法能够有效地处理多深度信息, 具有较低的重建误差。文中所提出方法虽然在深度图像平滑方面不如 ManiPOP, 但从点云图像来看, 该方法在每个像素上保留了更多的深度信息、目标更加完整、轮廓更加清晰, 与其他方法相

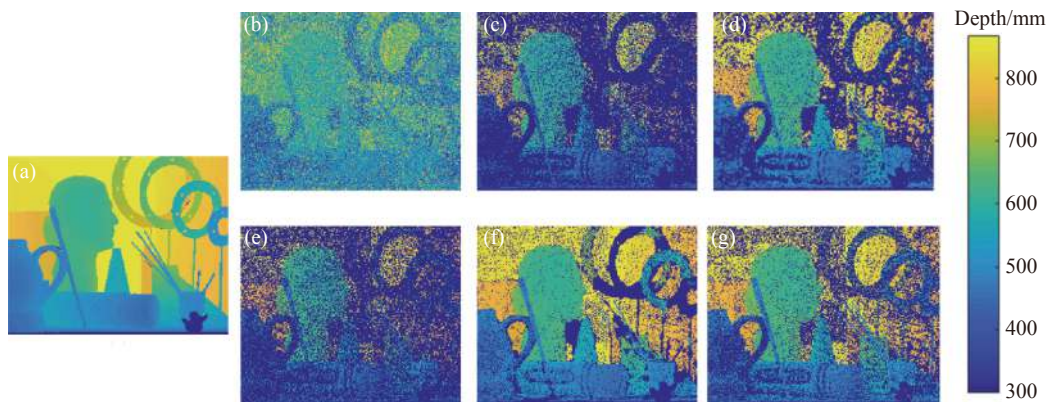


图 3 艺术场景的深度图。(a) 参考深度图; (b)~(g) MLE、FDTCP、Rapp、SPISTA、ManiPOP 和所提出方法估计的深度图

Fig.3 Depth images of the art scene. (a) Reference; (b)-(g) MLE, FDTCP, Rapp, SPISTA, ManiPOP, and proposed method, respectively

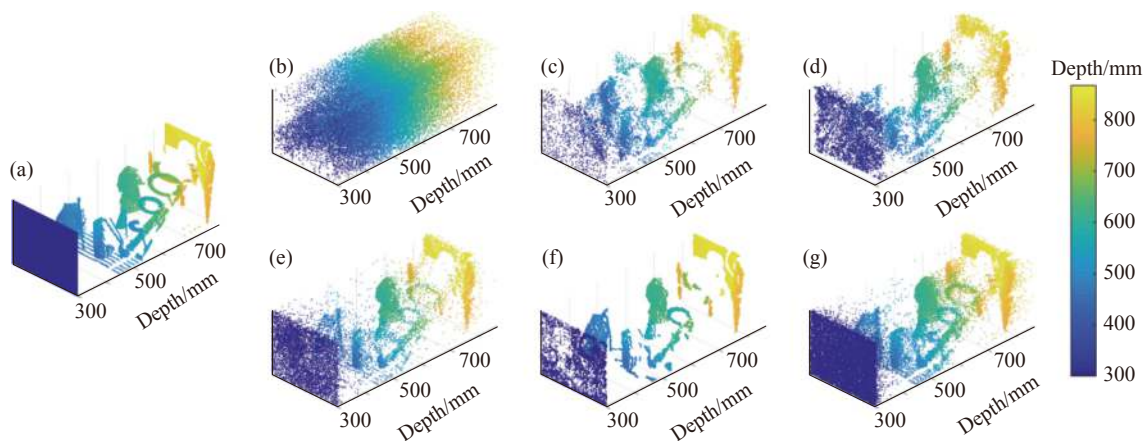


图 4 艺术场景的点云图。(a) 参考点云图; (b)~(g) MLE、FDTCP、Rapp、SPISTA、ManiPOP 和所提出方法估计的点云图

Fig.4 3D point clouds of the art scene. (a) Reference; (b)-(g) MLE, FDTCP, Rapp, SPISTA, ManiPOP, and proposed method, respectively

表 1 艺术场景上不同方法的 RMSE、SRE 和处理时间对比

Tab.1 RMSE, SRE and processing time of different methods on the art scene

Method	RMSE/mm	SRE/dB	Processing time/s
MLE	187.54	11.11	1.12
FDTCP	233.93	12.65	1.56
Rapp	196.41	12.52	154.43
SPISTA	172.14	16.67	197.12
ManiPOP	155.60	13.66	1 103.26
Proposed	87.32	20.27	83.49

比, RMSE 减少了至少 43.88%, SRE 提高了至少 21.60%。该方法经过多深度快速去噪后,数据量由  $183 \times 121 \times 4\ 500$  减小至  $183 \times 121 \times 100 \times 2$ ,  $L = 2$  表示目标的层数,相比于同是处理多深度目标的方法,速度提升 2 倍以上。

## 5 成像实验结果与分析

### 5.1 实验设备

单光子激光雷达系统原理图如图 5 所示<sup>[21]</sup>,光源发射重频为 10 kHz、波长为 1 550 nm、宽度为 3 ns 的激光脉冲,其单脉冲能量为 17.4  $\mu$ J。输出的激光经准直镜准直后进入出射口径为 36 mm 的望远镜,进一步聚焦整形后束散角为 0.1 mrad,经口径为 60 mm 的二维扫描振镜偏转,扫描整个目标场景。单光子探测器 (MPD PDM-IR) 接收目标反射的激光回波和背景光。探测器在门控模式下工作,时间抖动约为 70 ps,门控延迟时间为 30 ns,并与激光源同步。在此模式下,当初始激光脉冲通过收发模块向外传播时,门控系统会保持关闭状态,从而避免探测到不需要的后向散射。TCSPC 模块 (PicoQuant PicoHarp 300) 记录单光子探测器的响应并计算出每个光子的飞行时间,它的最小时间单元为 4 ps。

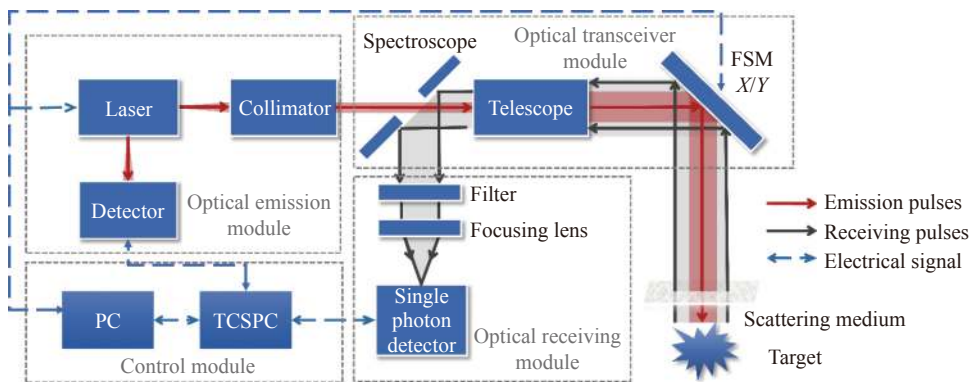


图 5 远距离单光子激光雷达系统原理示意图

Fig.5 Schematic diagram of long-range single-photon lidar system

### 5.2 1 km 处复杂建筑物的结果

如图 6(a) 所示,使用单光子激光雷达系统对距离 1 km 的复杂建筑物进行成像实验。该实验使用  $N_r \times N_c = 200 \times 200$  个像素点和  $T = 10\ 000$  个时间单元,每个时间单元的宽度为 1 ns。在无窗户散射情况下,将采集时间为 50 ms 的实验数据作为参考数据集,并对参考数据集进行以下处理:首先以精确的长度进行距离选通,然后通过设置阈值限制背景噪声的影响,最后使用 MLE 方法得到参考深度图像。关上窗户后,将采集时间为 20 ms 的实验数据作为测试数据。所有像素上的统计直方图如图 6(b) 所示,由于玻

璃的散射较大,来自玻璃的回波光子数比来自复杂建筑物的回波光子数高出一个量级。许多像素在两个建筑物目标的深度边界,因此该场景每个像素至少有两个表面组成。FDTCP、Rapp 方法和文中所提出方法均使用  $T_w = 100$  的窗口大小进行逐像素时域开窗,其中文中所提出方法针对此多表面目标选取深度层数  $L = 4$ 。

图 7 和图 8 分别为复杂建筑物上不同方法的深度图像与点云图像。表 2 对比了复杂建筑物上不同方法的 RMSE、SRE 和处理时间。MLE 和 FDTCP 方法受到来自散射光子的影响,完全无法估计出深度图

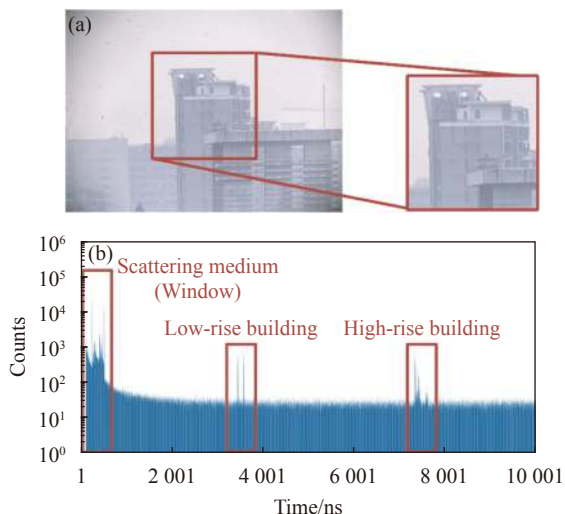


图 6 1 km 处的复杂建筑物。(a) 可见波段图; (b) 测试数据集的直方图  
 Fig.6 Complex buildings at 1 km. (a) Visible-band image; (b) Histogram of test dataset

像, 具有最差的深度估计性能。Rapp 与 ManiPOP 方法虽然能够探测到更多来自目标的光子, 但是 Rapp 方法在单深度逐像素开窗中保留了玻璃深度附近的响应集合, 因此丢失了高建筑物下方的大部分深度信息。而 ManiPOP 方法出现目标深度过平滑的问题, 导致 RMSE 和 SRE 较差。相比之下, 文中所提出方法与 SPISTA 方法获得了明显更好的数值和图像结果。尽管深度缺失较多, 但这两个方法成功地得到了两个建筑物目标的轮廓, 并且将其区分开。所提出方法将逐像素开窗法拓展到多深度目标, 更有效地处理了背景噪声, 并将数据大小由  $200 \times 200 \times 10\,000$  减小至  $200 \times 200 \times 100 \times 4$ , 是原始数据大小的 4%。与其他方法相比, 在估计精度上 RMSE 减少了至少 27.05%, SRE 提高了至少 18.39%, 在处理速度提升 1.21 倍以上。

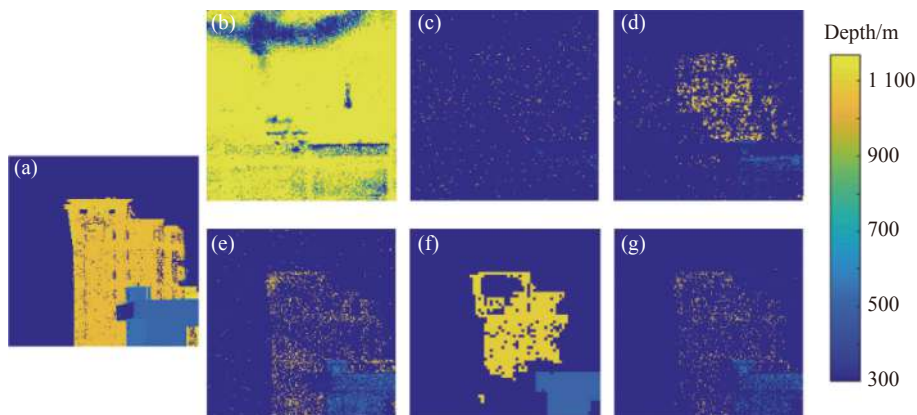


图 7 复杂建筑物的深度图。(a) 参考深度图; (b)-(g) MLE、FDTCP、Rapp、SPISTA、ManiPOP 和所提出方法估计的深度图  
 Fig.7 Depth images of complex buildings. (a) Reference; (b)-(g) MLE, FDTCP, Rapp, SPISTA, ManiPOP, and proposed method, respectively

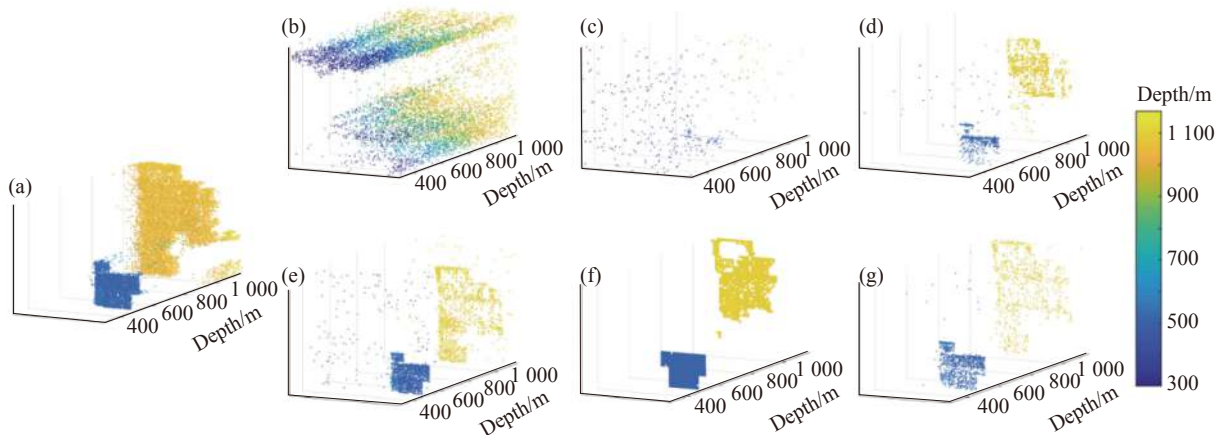


图 8 复杂建筑物的点云图。(a) 参考点云图; (b)-(g) MLE、FDTCP、Rapp、SPISTA、ManiPOP 和所提出方法估计的点云图  
 Fig.8 3D point clouds of complex buildings. (a) Reference; (b)-(g) MLE, FDTCP, Rapp, SPISTA, ManiPOP, and proposed method, respectively



表 2 复杂建筑物上不同方法的 RMSE、SRE 和处理时间对比

Tab.2 RMSE, SRE and processing time of different methods on complex buildings

Method	RMSE/m	SRE/dB	Processing time/s
MLE	751.04	2.16	4.74
FDTCP	876.05	4.77	4.08
Rapp	424.68	15.81	161.75
SPISTA	65.86	30.12	553.88
ManiPOP	584.98	4.32	738.35
Proposed	48.04	35.66	133.30

## 6 结 论

文中提出了一种精确的单光子激光雷达多深度估计方法,该方法强调信号响应的时问相关性。该方法创新性地将逐像素开窗法拓展到多深度目标,进一步使用快速搜索算法加速其计算,显著减小了激光雷达数据大小,相应地降低了后续深度估计方法的内存需求和计算复杂度,同时限制了背景噪声的影响。在深度图像估计中引入了 TV 正则化,使用 ADMM 从成本函数中估计深度图,有效地平滑深度图像并保持了边缘部分的细节。在比较实验中,文中的方法相比于其他深度估计方法在近距离仿真数据集上和户外实验的远距离目标 (~1 km) 均获得了更好的深度估计结果。与其他方法相比,均方根误差减少了至少 27.05%,信号重建误差比提高了至少 18.39%。下一步的工作是在该方法的基础上进一步结合空间相关信息,从而提高深度图像估计质量。

### 参考文献:

[1] Qu Hengkuo, Zhang Qingyuan, Ruan Youtian. Laser radar based on scanning image tracking [J]. *Chinese Optics*, 2012, 5(3): 242-247. (in Chinese)

[2] Zhang Jian, Zhang Lei, Zeng Fei, et al. Development status of airborne 3D imaging lidar systems [J]. *Chinese Optics*, 2011, 4(3): 213-232. (in Chinese)

[3] Wang Zijun, Zhang Yang, Liu Dong, et al. Research on the development of detection satellite technology in the novel multi-beam land and ocean lidar [J]. *Infrared and Laser Engineering*, 2021, 50(7): 20211041. (in Chinese)

[4] He Weiji, Sima Boyu, Cheng Yaojin, et al. Photon counting imaging based on GM-APD [J]. *Optics and Precision Engineering*, 2012, 20(8): 1831-1837. (in Chinese)

[5] O'Connor D V, Phillips D. Time-Related Single Photon

Counting[M]. New York: Academic Press, 2012.

[6] Zheng T, Shen G, Li Z, et al. Frequency-multiplexing photon-counting multi-beam LiDAR [J]. *Photonics Research*, 2019, 7(12): 1381-1385.

[7] Li Z, Wu E, Pang C, et al. Multi-beam single-photon-counting three-dimensional imaging lidar [J]. *Optics Express*, 2017, 25(9): 10189-10195.

[8] Guo Jingjing, Fei Xiaoyan, Ge Peng, et al. High-resolution three-dimensional imaging based on all-fiber photon-counting Lidar system [J]. *Infrared and Laser Engineering*, 2021, 50(7): 20210162. (in Chinese)

[9] Wei Chaojie, Yan Renpeng, Li Xudong, et al. Research progress of sub-nanosecond lasers for 3D imaging lidar [J]. *Optics and Precision Engineering*, 2021, 29(6): 1270-1280. (in Chinese)

[10] Li Z P, Ye J T, Huang X, et al. Single-photon imaging over 200 km [J]. *Optica*, 2021, 8(3): 344-349.

[11] Rapp J, Goyal V K. A few photons among many: Unmixing signal and noise for photon-efficient active imaging [J]. *IEEE Transactions on Computational Imaging*, 2017, 3(3): 445-459.

[12] Chen S, Halimi A, Ren X, et al. Learning non-local spatial correlations to restore sparse 3D single-photon data [J]. *IEEE Transactions on Image Processing*, 2019, 29: 3119-3131.

[13] Lindell D B, O'Toole M, Wetzstein G. Single-photon 3D imaging with deep sensor fusion[J]. *ACM Transactions on Graphics (TOG)*, 2018, 37(4): 113: 1-12.

[14] Sun Z, Lindell D B, Solgaard O, et al. SPADnet: Deep RGB-SPAD sensor fusion assisted by monocular depth estimation [J]. *Optics Express*, 2020, 28(10): 14948-14962.

[15] Tobin R, Halimi A, McCarthy A, et al. Three-dimensional single-photon imaging through obscurants [J]. *Optics Express*, 2019, 27(4): 4590-4611.

[16] Shin D, Xu F, Wong F N C, et al. Computational multi-depth single-photon imaging [J]. *Optics Express*, 2016, 24(3): 1873-1888.

[17] Tachella J, Altmann Y, Ren X, et al. Bayesian 3D reconstruction of complex scenes from single-photon lidar data [J]. *SIAM Journal on Imaging Sciences*, 2019, 12(1): 521-550.

[18] Altmann Y, Ren X, McCarthy A, et al. Lidar waveform-based analysis of depth images constructed using sparse single-photon data [J]. *IEEE Transactions on Image Processing*, 2016, 25(5): 1935-1946.

[19] Feng Z, He W, Fang J, et al. Fast depth imaging denoising with the temporal correlation of photons [J]. *IEEE Photonics Journal*, 2017, 9(5): 1-10.

[20] Chen S, Halimi A, Ren X, et al. Non-local restoration of sparse 3D single-photon data [C]//27th European Signal Processing Conference (EUSIPCO). IEEE, 2019: 1-5.

[21] 何伟基, 黄鹏维, 陈钱, 等. 远距离光子计数三维激光雷达成像系统及其方法: 中国, CN112213737A[P]. 2021-01-12.