

超级激光雷达反射率用于机器人感知

Wei Gao¹, Jie Zhang¹, Mingle Zhao¹, Zhiyuan Zhang², Shu Kong¹, Maani Ghaffari³,
Dezhen Song⁴, Cheng-Zhong Xu¹, and Hui Kong¹

Abstract—传统上, 人类的直觉通常将视觉定义为一种被动光学传感方式, 而主动光学传感通常被视为测量而非视觉的默认模式。然而, 情况现在发生了变化: 传感器技术和数据驱动的模式使主动光学传感能够重新定义视觉的边界, 迎来了主动视觉的新时代。光检测和测距 (LiDAR) 传感器捕捉物体表面的反射率, 这在不同的光照条件下保持不变, 展现出在机器人感知任务中如检测、识别、分割和同时定位与地图构建 (SLAM) 等方面的显著潜力。这些应用通常依赖于密集传感能力, 通常通过高分辨率的昂贵 LiDAR 传感器实现。低成本 LiDAR 面临的一个关键挑战在于扫描数据的稀疏性, 这限制了它们的更广泛应用。为了解决这一限制, 本研究引入了一种创新框架, 用于从稀疏数据生成密集 LiDAR 反射率图像, 利用非重复扫描 LiDAR (NRS-LiDAR) 的独特属性。我们解决了关键挑战, 包括反射率校准以及从静态到动态场景领域的过渡, 促进了真实环境中密集反射率图像的重建。本研究的主要贡献包括一个用于 LiDAR 反射率图像致密化的综合数据集、一个为 NRS-LiDAR 量身定制的致密化网络, 以及使用生成的密集反射率图像的多种应用, 如回路闭合和车道线检测。实验结果验证了所提出方法的有效性, 该方法成功地将计算机视觉技术与 LiDAR 数据处理相结合, 增强了低成本 LiDAR 系统的适用性, 并为机器人主动视觉建立了一个新的范式—激光雷达如同摄像机。数据集和代码将于: 待更新提供。

I. 引言

最近激光雷达技术的进步极大地提高了其可访问性, 这主要是由于传感器尺寸和成本的降低, 从而拓展了其在各种机器人应用中的适用性。除了提供精确的三维空间数据外, 激光雷达系统还能够捕捉物体表面的反射强度, 获取与光照无关的环境信息, 以及丰富的纹理, 以补充纯几何激光雷达传感的局限性。

迄今为止, 许多研究已将 LiDAR 反射强度用于多种应用, 包括闭环检测 [2]–[4]、强度增强的 LiDAR 里程计 [1], [5]–[7]、分割 [8], [9] 和物体检测 [10], [11]。这些应用通常依赖于密集反射图像, 这些图像通常由高分辨率且昂贵的 LiDAR 传感器 (如 Ouster OS1-128 和 Velodyne VLS-128) 生成。相比之下, 低成本的 LiDAR 传感器, 如 Velodyne-16、Velodyne-32 或 Livox MID-360, 通常产生稀疏数据, 从而直接应用为密集反射图像设计的方法时面临显著挑战。

¹Wei Gao, Jie Zhang, Mingle Zhao, Shu Kong, Chengzhong Xu, and Hui Kong are with the State Key Laboratory of Internet of Things for Smart City (SKL-IOTSC), Faculty of Science and Technology, University of Macau, Macau. ({ gw.gao.wei, mc25934, zhao.mingle } @connect.um.edu.mo, { skong, czxu, huikong } @um.edu.mo)

²Zhiyuan Zhang is with the School of Computing and Information Systems, Singapore Management University, Singapore. (zhiyuanzhang@smu.edu.sg)

³Maani Ghaffari is with the Department of Naval Architecture and Marine Engineering and Department of Robotics, University of Michigan, Ann Arbor, MI, USA. (maanigj@umich.edu)

⁴Dezhen Song is with the Department of Robotics, Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), UAE. (dezhen.song@mbzuai.ac.ae)



Fig. 1. 比较所提出方法生成的反射率图像与从高线 LiDAR 数据拼接而成的真实反射率图像的情况。(a) 原始 NRS-LiDAR 输入。(b) 使用所提出的方法加密的 1380×240 反射率图像。(c) 使用高成本的 Ouster 128 线 LiDAR 数据自 [1], [2] 拼接的 1024×128 反射率图像。得益于所提出的网络和 NRS-LiDAR, 该方法在复杂场景中有效消除了线条伪影, 减轻了过曝现象, 生成高清晰度、相机级的反射率图像。

直观上, 从稀疏的 LiDAR 数据生成密集且准确的反射图像具有提高机器人感知任务的巨大潜力, 然而这个领域在现有研究中仍未得到充分探索 [12]。成功解决这一挑战将有效弥合计算机视觉技术与 LiDAR 数据之间的鸿沟, 促进基于相机的方法向 LiDAR 系统的适应。此外, 它将解决传统相机的固有局限性, 如对光照条件的敏感性, 同时增强 LiDAR 技术的纹理表示能力。

创建密集的 LiDAR 反射图像通常依赖于高质量、高分辨率的地面真实数据, 特别是在使用深度神经网络时。然而, 获取此类数据通常需要专门的传感器或配置, 这带来了显著的挑战。在本研究中, 我们利用 NRS-LiDAR 的独特成像特性来生成高质量的反射图像地面真实数据 (图 1)。如图 2 所示, 在静态环境中可以通过扫描积累获取密集的地面真实数据。然而, 当机器人运动时, 获取密集的地面真实数据则不可行。在动态场景中, 必须实时处理稀疏且迅速变化的输入, 从而显著增加了密集重建的复杂性。静态地面真实生成与动态场景处理之间的差异强调了使用低成本 LiDAR 实现密集反射图像的挑战。总之, 我们面临两大难题。第一是在运动条件下获取用于反射图像稠密化的密集地面真实数据。第二个问题是如何在给定地面真实数据的情况下稠密化稀疏反射。为了解决这些问题, 我们提出了一种新方法, 包含以下贡献:

- 1) 我们提出了第一个专门为 LiDAR 反射图像致密化设计的基准数据集。相比于现有的 LiDAR 数据集 [1], [13], 我们的数据集提供了稀疏到稠密的配对真实数据, 支持基于端到端深度学习的致密化基准以及图像级感知任务。
- 2) 我们引入了一种新颖的卷积网络, 专为 NRS-

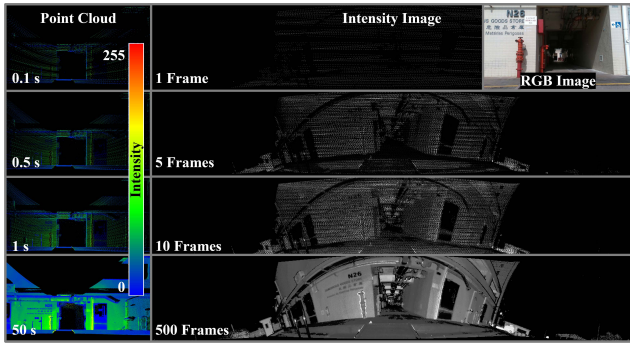


Fig. 2. 非重复扫描机制的示意图。处于静止状态时，LiDAR 会在相同的 FoV 内随着时间的推移逐渐生成更密集的点云，这使得从点云投影得到的反射强度图像也能够达到更高的密度。点云或灰度图像的颜色代表强度值。

LiDAR（例如，Livox MID-360）的反射图像实时密集化而设计。我们的方法有效地解决了两个关键挑战：静止到运动域的差距和 LiDAR 反射图像密集化中的自适应强度校准。与现有方法相比，我们的方法在质量和准确性上取得了显著的改进。

- 3) 我们验证了密集反射图像在闭环检测和车道检测中的有效性，在具有挑战性的条件下取得了显著的改进，并展示了其在增强低成本激光雷达系统实际应用中的潜力。

II. 相关工作

A. 机器人中的激光雷达反射率

在单一基于 LiDAR 的应用中，众多研究主要集中在利用几何信息用于机器人感知 [14] 和定位 [15] 任务。然而，这些方法在几何退化环境中表现出局限性，因为它们无法捕捉纹理信息。为了解决这个问题，研究人员已经探讨了融合摄像机和 LiDAR 数据以利用其互补优势 [16], [17]。不过，这种整合带来了挑战，包括额外的硬件需求、精确校准和时间同步。此外，视觉传感器本身受限于光照条件。一个有前景的替代方案是利用 LiDAR 的内在反射特性。在该领域已取得显著进展。一种将几何和强度数据协同结合的回环检测方法被提出 [18]。此外，反射信息的应用在机器人地形分析中被证明特别有效，能够进行语义分割，以区分道路表面和草地 [19]。

近年来，由高分辨率 LiDAR 传感器生成的密集反射图像被用于增强机器人感知能力 [5], [6]。一个显著的工作是 COIN-LIO [1]，这是一个将 LiDAR 强度图像上的光度误差最小化与几何配准相结合的 LiDAR-惯性里程计框架，用于解决几何退化环境中的挑战。文章 [2] 中介绍了一种混合回环检测方法，通过结合 LiDAR 点云和 LiDAR 强度图像来实现更高的鲁棒性。然而，这些先进的方法依赖于高分辨率、高成本的 LiDAR 传感器，使其在低成本应用中不切实际，并强调了在资源受限场景中需要替代策略。

本研究提出了一种新的方法，通过低成本的 LiDAR 传感器从稀疏反射图像重建高保真度的密集反射图像，并展示了在实际实时机器人应用中的显著潜力。

B. 深度补全

密集的环境感知对于机器人任务至关重要。然而，物理传感器通常仅提供稀疏的测量。因此，将稀疏的感官数据完整化为密集的感知表示是一个关键的研究方向。

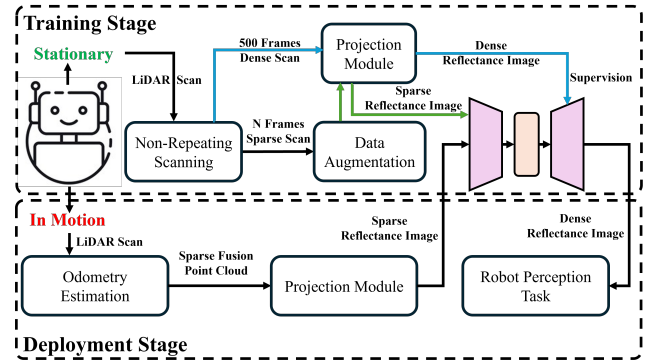


Fig. 3. 系统图：在静止时，机器人使用 NRS-LiDAR 获取稀疏和密集的 LiDAR 数据。稀疏数据被增强以模拟运动特性，而密集数据则用作监督来训练用于运动感知任务的反射率图像密集化网络。

深度补全 [20], [21] 是研究最为广泛的领域之一，这激励我们回顾密集补全方法。深度补全利用来自传感器测量 [22]–[25]、多视图几何 [26], [27] 或数据驱动的先验知识 [28], [29] 的稀疏深度数据，生成密集的深度图像，从而实现导航和操作的连续和密集的感知。

然而，直接将现有的深度补全方法应用于激光雷达反射图像的密集化具有挑战性，这是由于强度测量的复杂性。入射角、表面反射特性以及传感器到目标的距离 [30] 等因素使得学习潜在的强度模式变得复杂，尤其是在可用的公共训练数据有限的情况下。因此，开发专门用于激光雷达反射图像密集化的新型网络是至关重要的。与仅关注几何的深度补全不同，激光雷达反射图像密集化为机器人配备了一个新的高分辨率“眼睛”，从而使得通过激光雷达实现高级视觉感知成为可能，开启了重新想象各种视觉问题的大门。

III. 方法学

运动中的机器人通常需要足够的时间来积累数据，以全面了解其环境。在本节中，我们详细介绍了我们提出的方法以解决这一挑战。如图 3 所示，这项工作的目的是仅依靠机器人在运动期间短时间内收集的稀疏 LiDAR 数据作为输入，生成密集且准确的反射图像作为输出，为机器人感知任务提供全面的环境信息。

A. 数据集构建

1) 数据采集：Super LiDAR Reflectance 数据集是通过使用配备多传感器套装的长颈鹿地面机器人平台获得的：一个 Intel RealSense D435i 相机，以 1280×720 的分辨率捕捉 RGB 图像，一个 Livox MID-360 LiDAR 传感器，具有 360° 水平和 59° 垂直视场 (FoV)，以 10 Hz 的点云生成率运行，以及一个高频 IMU 以 200 Hz 的频率采样。起初，机器人自主导航通过选择的代表环境。到达预定位置后，它保持静止姿态 60 秒，以便累积多个 LiDAR 扫描。同框协调的传感器数据，包括 LiDAR 点云、相机图像和 IMU 测量值，被记录为原始 ROS（机器人操作系统）数据包，并系统地组织成数据集集合。为了确保数据多样性和最小化数据集间的相关性，机器人在记录会话之间的方位和位置被故意调整，从而保持每个数据集实例的独立性和代表性。

2) 生成真实数据：本节利用两种投影方法生成真实的 LiDAR 反射图像：全景球面投影和虚拟相机透视投影。在前一种方法中，LiDAR 点直接映射到

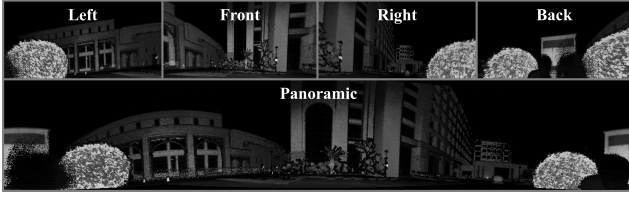


Fig. 4. 当 NRS-LiDAR 在预设位置放置 60 秒时，分别通过虚拟相机投影模式（上方）和全景模式（下方）获得反射率图像。

球面上。特定 3D 点的方位角 $\theta = \arctan 2(y, x)$ 和仰角 $\phi = \arctan 2\left(z, \sqrt{x^2 + y^2}\right)$ 被转换为 2D 坐标: $\text{col} = \left\lfloor \frac{-\theta + 180^\circ}{360^\circ} \cdot W \right\rfloor$ 和 $\text{row} = \left\lfloor \frac{\phi_{\max} - \phi}{\phi_{\max} - \phi_{\min}} \cdot H \right\rfloor$, 其中 W 和 H 是生成图像的宽度和高度。 $[\phi_{\min}, \phi_{\max}]$ 定义了垂直 FoV。强度 I 被分配给对应的像素 (row, col)。这种方法与 LiDAR 传感器的 FoV 对齐, 生成分辨率为 1380×240 像素的图像。

在虚拟摄像机透视投影中, 四个虚拟摄像机被放置在 LiDAR 传感器周围, 将点云投射到四个视图中。使用针孔摄像机模型, 3D 点 (x, y, z) 映射到二维图像平面

$$\text{col} = \left\lfloor -f_x \frac{y}{x} + c_x \right\rfloor, \quad (1)$$

$$\text{row} = \left\lfloor H - 1 - \frac{\arctan 2(z, x) - \phi_{\min}}{\phi_{\max} - \phi_{\min}} \cdot H \right\rfloor, \quad (2)$$

, 其中 $f_x = \frac{W}{2 \cdot \tan(\frac{\text{horizontal FoV}}{2})}$ 、 $f_y = \frac{H}{\phi_{\max} - \phi_{\min}}$ 和 $c_x = \frac{W}{2}$ 。每个视图生成分辨率为 240×480 像素的反射率图像。

这两种投影模式生成的示例反射率图像如图 4 所示。由于 LiDAR 反射率图像的密度随累积扫描次数的变化而变化, 我们使用累积 500 次扫描的反射率图像作为训练过程中的监督地面真值。通过上述数据采集, 我们构建了超级 LiDAR 反射率数据集, 包括 1000 个原始 ROS 数据包 (600K 帧) 和另外 5000 组从稀疏 (1 次扫描) 到密集 (500 次扫描) 配置的反射率图像, 覆盖了多样的城市环境, 如住宅区、街道、停车场和交叉路口。该数据集还包括两个针对密集反射率图像特定应用的测试子集: 一个用于回环检测, 数据来自三个场景的两个时间间隔 (早晨、傍晚), 采用全景球形投影以实现不同视点的鲁棒匹配。另一个子集是为在白天和夜间不同交通场景下的车道检测而设计的, 采用虚拟相机透视投影 (更适合此任务)。

上述的真实监督数据是在激光雷达静止时收集的。然而, 所提出的反射率增密方法通常用于激光雷达运动时的机器人感知任务。在实际的运动模式下, 需要通过里程计估算来融合一定数量的连续扫描, 以构建相对稠密的反射率图像作为网络输入。在融合过程中, 使用过少的扫描 (例如, 仅一两个扫描) 通常会导致数据过于稀疏, 缺乏足够的信息。另一方面, 融合过多的扫描可能引入由里程计估算不准确或场景中动态物体产生的噪声点而导致的误差。为解决在这种情况下动态物体的影响, 我们采用了 [31] 中提出的预处理方法来过滤原始点云数据, 有效地去除由动态物体导致的运动伪影, 并减弱它们对融合过程的不利影响。在我们的案例中, 网络输入是通过使用由 [15] 获得的里程计估算融合 N (一个小数量) 的连续扫描生成的。通过与静态真实训练数据生成的比较, 显然存在静态训练数据与动态推断数据之间的差距, 这可能会在运动场景中降低性能。

如图 ?? 所示, 主要差异可以从两个方面描述。几何失准: 在运动过程中基于里程计的扫描融合会因累积的运动估计误差引入对齐错误, 导致输入不一致。运动引起的分布偏移: 机器人运动改变了数据分布, 导致点密度、空间覆盖和对齐与静态状态相比发生变化。具体来说, 运动减少了扫描之间的重叠, 使得某些区域在 LiDAR 的视野中变得稀疏, 从而进一步增加了适应动态条件的复杂性。

从数学上讲, 静态训练设置中的输入可以表示为

$$\mathbf{X}_{\text{train}} = \mathcal{P} \left(\sum_{i=1}^N \mathbf{R}_i(x, y, z) \right), \quad (3)$$

, 其中 $\mathcal{P}()$ 表示将 3D 点云映射到 2D 反射率图像的投影操作, $\mathbf{R}_i(x, y, z)$ 是第 i 次扫描的原始点云, $\sum$ 表示多次扫描点云的几何合并。相比之下, 当机器人运动时, 训练阶段 (也适用于推理阶段) 的输入表示为

$$\mathbf{X}_{\text{test}} = \mathcal{P} \left(\sum_{i=1}^N \mathbf{T}_i \cdot \mathbf{R}_i(x, y, z) \right), \quad (4)$$

, 其中 $\mathbf{T}_i$ 是从里程计推导出的变换矩阵, 引入了对齐误差。这种错位导致 $\mathbf{X}_{\text{train}}$ 和 $\mathbf{X}_{\text{test}}$ 之间存在差异, 形成域间差距。为了解决这个问题, 我们在训练期间提出了一种数据增强方法, 该方法模拟机器人的运动。

具体来说, 在扫描累积期间, 应用随机旋转和平移到每个扫描, 以模拟运动引起的不一致性。对于序列中的每个扫描, 增强后的输入表示为:

$$\mathbf{X}_{\text{aug}} = \mathcal{P} \left(\sum_{i=1}^N \mathbf{T}_{\text{motion}_j} (\mathbf{R}_i(x, y, z)) + \mathcal{N}(0, \sigma^2) \right), \quad (5)$$

, 其中 $\mathcal{N}(0, \sigma^2)$ 表示均值为 0、方差为 σ^2 的高斯噪声, 加入以模拟传感器噪声和运动变化。变换 $\mathbf{T}_{\text{motion}_j}$ 应用了一个随机预定义的运动模式 j 。为了复制在实际场景中通常观察到的稀疏性, 我们采用随机扫描采样, 其中从初始序列中随机选择 N 扫描, 然后累计为输入数据。此外, 引入里程计的非线性扰动以模拟对齐不良的姿势 (例如, 加速的不连续性、急转弯), 而噪声 IMU 信号模拟传感器退化带来的误差。通过结合这些增强, 所提出的方法提高了网络对齐不良、稀疏和噪声输入的鲁棒性, 有效弥合了静态和动态运动条件之间的域差距。

B. 自适应校准与密集化网络

考虑到卷积神经网络 (CNNs) 能够以计算复杂度近似于输入元素 $\mathcal{O}(n)$ 数量的线性比例有效地提取空间特征, 并在处理大型图像时保持高效, 因此它们被选为我们致密化网络的骨干。相比之下, 虽然 Transformers [32] 具有强大的建模能力, 但其复杂度相对于输入大小为平方级的 $\mathcal{O}(n^2)$, 这显著限制了它们在资源受限场景中的应用, 特别是对于高分辨率图像 [33], [34]。此外, 扩散模型 [35] 尽管在图像复原中表现卓越, 却依赖迭代过程生成最终输出, 导致较长的推理时间 [36], 使它们难以用于实时机器人应用。图 5 展示了我们致密化网络的结构, 该网络采用 U 形架构 [37], 其中提出了一种自适应融合模块 (AFM) 和动态补偿模块 (DCM) 来实现致密化。这些模块解决了两个关键挑战: 1) 自适应融合来自非均匀和稀疏 LiDAR 扫描点的多尺度信息, 并且 2) 缓解随距离和入射角发生的强度衰减。

设 $L_1 \in \mathbb{R}^{H \times W}$ 和 $D_1 \in \mathbb{R}^{H \times W}$ 分别为网络输入的 LiDAR 反射率和深度图像。在反射率补偿步骤中, 在反

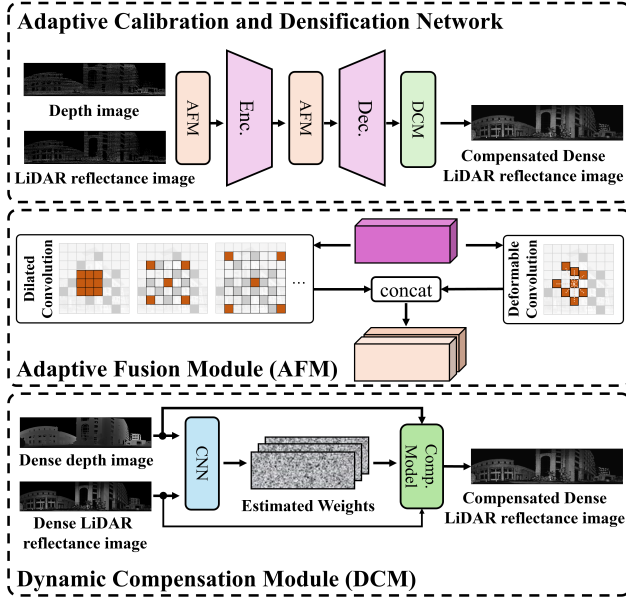


Fig. 5. 提出方法的 LiDAR 反射率密集化框架。该框架包括用于特征提取的自适应特征模块 (AFM)、用于空间重构的编码器-解码器结构, 以及用于校正强度的动态补偿模块 (DCM)。

射率图像重建过程中引入了深度信息。为了最大化利用跨模态交互并减少模型冗余, 首先沿通道维度将 L_1 和 D_1 连接作为网络的输入。然后, 这些输入通过一个整合了 AFM 的编码器-解码器结构进行处理, 以生成稠密反射率图像 $L_2 \in \mathbb{R}^{H \times W}$ 和深度图像 $D_2 \in \mathbb{R}^{H \times W}$ 的初步重建, 其中解码器的第一和第二通道分别是 L_2 和 D_2 。此外, 为了消除由于信号强度衰减引起的伪影, 应用 DCM 来补偿这些效果, 从而得到最终的稠密 LiDAR 反射率图像 $L_3 \in \mathbb{R}^{H \times W}$ 。DCM 采用基于深度学习的动态补偿模型, 以避免手动设计传统方法中所需参数的复杂性, 同时保持端到端重建。

编码器和解码器。编码器有四个堆叠的块, 每个块包含两个顺序连接并带有 ReLU 激活的层, 后接一个平均池化层进行下采样。卷积核设置为大小 3×3 。池化层的池化窗口和步幅都配置为 2, 以便下采样因子为 2。解码器具有对称结构, 包括四个解码器块。在每个解码器块中, 输入特征图通过转置卷积模块进行上采样。随后, 应用两个 3×3 卷积进行局部信息融合。为了增强多尺度特征整合, 我们实施了两个策略: 1) 在编码器和解码器之间插入一个 AFM 模块以促进多尺度空间信息; 2) 类似于 [37], 编码器的特征被并入解码器。来自四个块的相同空间维度的对应特征图被连接起来以整合多层次信息。自适应融合模块。由于 NRS-LiDAR 的扫描特性, 在短时间内捕获的反射图像通常是稀疏且分布不均的。此外, 由于场景中的物体具有多尺度属性并且距离不同, 融合原始 LiDAR 反射数据以补充信息密集化要求模型具备自适应能力以及融合多尺度结构特征的能力。传统的卷积核由于感受野有限, 难以捕捉大区域的信息, 而较大的卷积核会显著增加计算成本。受 [38], [39] 的启发, 我们利用并行膨胀卷积和可变形卷积自适应整合多尺度信息 (图 5)。具体而言, 我们在并行膨胀卷积中采用多种膨胀率以最小化额外计算成本扩展感受野, 从而增强模型多尺度特征提取的能力。同时, 我们引入可变形卷积以便根据局部上下文自适应地融合信息, 从而提升模型

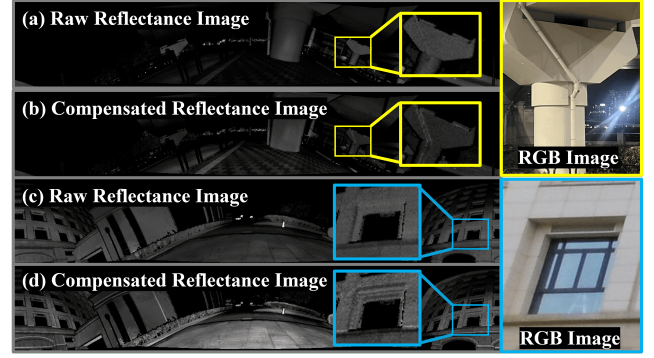


Fig. 6. 原始密集反射率与补偿结果的比较。补偿结果在具有相同表面材料的物体上保持了更一致的反射率值 ((a) 和 (b) 中的黄色方框) 并显示出更清晰的纹理细节 ((c) 和 (d) 中的蓝色方框)。

Method	PSNR / SSIM $\uparrow$	RMSE $\downarrow$	MAE $\downarrow$	Time (ms) $\downarrow$
IP-basic [40]	22.448 / 0.642	0.0672	0.0471	10.6 (CPU)
pNCNN [42]	25.746 / 0.718	0.0537	0.0282	168 (GPU)
Sparse-to-Dense [41]	26.277 / 0.724	0.0505	0.0272	69 (GPU)
Marigold-dc [43]	24.21 / 0.691	0.613	0.354	26140 (GPU)
Ours (w/o AFM)	26.051 / 0.721	0.0521	0.0276	36 (GPU)
Ours (w/ AFM)	27.152 / 0.767	0.0492	0.0247	52 (GPU)

提取详细信息的能力。这些并行卷积的输出随后沿着通道维度连接, 形成 AFM 模块的输出特征。动态补偿模块。LiDAR 强度主要受到三个因素的影响: 光线传播的距离、物体表面的反射率和光的入射角。通常假设物体表面的反射率保持相对一致。因此, 补偿距离和入射角的影响对于强度测量至关重要 [30]。如 [30] 所介绍, 捕获的强度可以表示为

$$I(R, \alpha, \rho) \propto P(R, \alpha, \rho) = \eta(R) \frac{I_e \rho \cos \alpha}{R^2}, \quad (6)$$

其中 I_e 表示激光的发射功率, R 表示从 LiDAR 到物体表面的距离, α 是入射角, ρ 表示物体表面的反射率。考虑近距离效应的项 $\eta(R)$ 可以表示为

$$\eta(R) = 1 - \exp \left\{ -\frac{2r_d^2(R+d)^2}{D^2 S_d^2} \right\}, \quad (7)$$

其中 r_d 是激光探测器的半径, d 是测量距离与真实物体距离之间的偏移量, D 是镜片直径, S_d 是焦距。在近距离时, $\eta(R)$ 很小, 但随着 R 的增长迅速增加, 在较大的 R 时稳定在接近 1.0。重新排列方程 (6) 以用其他变量表示 I_e :

$$I_c \approx I_e = C \frac{I(R, \alpha, \rho) \cdot R^2}{\eta(R) \cdot \rho \cos \alpha}, \quad (8)$$

其中 C 是一个比例常数, 用于系统特定的校准因子。 I_c 表示校正或真实的强度。

在实际操作中, 我们收集了角度和距离受控的相同材料物体的数据, 通过方程 8 拟合所需参数, 并将其应用于原始密集 LiDAR 反射率图像以生成监督补偿结果。补偿结果如图 6 所示。然而, 上述方法进一步受限于数据收集限制和传感器配置, 使得它们不足以捕捉复杂的、非线性的多环境关系。为了解决这个问题, 我们使用了一个简化的参数化补偿模型, 如方程 (9) 和 (10) 所示, 该模型基于方程 (8), 但结合了一个额外的衰减因子, 以全面处理近距离和远距离效果。对于距离补偿, 当 R 值较大时, 距离校正因子 $\frac{R^2}{\eta(R)}$ 变得明显增大。因此, 我们使用结合指数衰减因子的二次多项式进行近似, 以处理

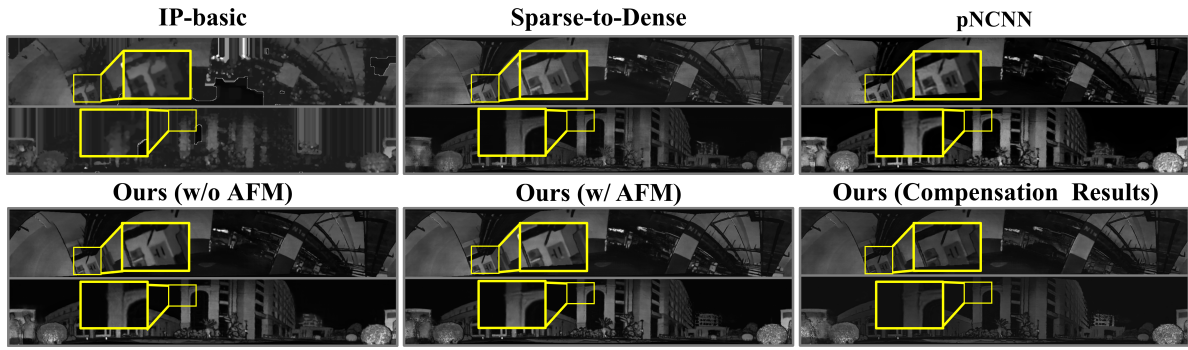


Fig. 7. 在室内和室外场景中, 不同方法的反射率图像扩充的定性比较。依赖非学习技术的 IP-basic [40] 在大孔区域难以恢复结构, 导致输出失真。此外, Sparse-to-Dense [41] 和 pNCNN [42] 在平滑性和边缘保留方面表现出明显的缺陷。我们的 AFM 方法在保留边缘结构和增强细节方面表现优越。补偿结果显示, 相同表面材料 (例如, 相同墙壁) 的强度值保持得更加一致 (如黄色框所示)。

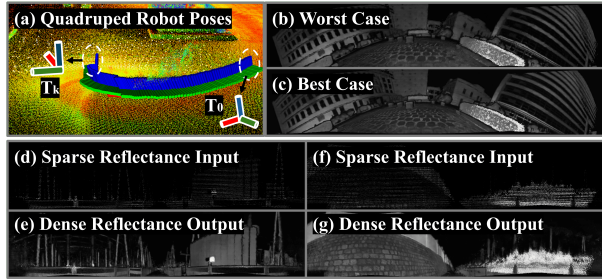


Fig. 8. (a)-(c) 在不平坦地形上使用具有激进四足机器人运动姿态的 5 帧数据进行密集化结果, 其中 (b) 显示了最坏情况 (PSNR = 25.75) 而 (c) 显示了最好情况 (PSNR = 26.82); (d)-(g) 为公共数据集中 [45] 的输入和强度图像密集化输出。

远距离效果。方程

$$g(R) = \frac{(\alpha R^2 + \beta R + \gamma) \cdot \exp(-\lambda R)}{\eta(R)}, \quad (9)$$

, 其中 α 、 β 、 γ 和 λ 是由网络动态预测的可学习参数, 以捕捉强度和距离之间的非线性关系。

为了进行入射角补偿, 强度按照 Lambert 余弦定律 [44] 通过 $\cos(\alpha)$ 进行归一化。通过结合距离和角度补偿, 校正后的强度 I_{com} 表达为

$$I_{\text{com}} = \frac{I_m}{g(R) \cdot \cos(\alpha) \cdot \kappa}, \quad (10)$$

, 其中 I_m 代表测量的强度, 与方程 6 中的 $I(R, \alpha, \rho)$ 相等, 而 κ 是用于考虑表面反射率变化和传感器特定因素的可学习参数。在我们的方法中, 卷积层用于基于局部上下文动态预测每个像素的补偿参数 α 、 β 、 γ 、 λ 和 κ , 距离 R 和入射角 α 都由 D_2 提供。然后将这些参数应用于 I_m 使用方程 (10), 使模型能够进行强度补偿。

我们进行了系列实验, 以评估所提方法的有效性和鲁棒性, 重点研究实时 LiDAR 反射率图像增密及其在回环检测和车道检测中的应用。数据采集和实际实验是使用安装在长颈鹿地面机器人上的 Livox MID-360 进行的, 该机器人配备 NVIDIA RTX 3090 GPU 和 Intel i7-1165G7 CPU。实验涵盖了多种场景, 包括不同的光照条件 (白天、阴天、夜晚) 和城市环境, 以全面评估所提出的方法。

C. LiDAR 反射率加密的评估

为了评估我们提出的致密化网络的性能, 我们进行了比较和消融实验。比较实验对我们的方法生成的致密图像的质量与基于 CNN 的方法和传统方法进行了对比, 分

TABLE II

不同闭环检测方法在不同环境中的定量结果

Env.	Method	Detected Loops	True Positives	False Positives
Env. 1	ISC [46]	53	32 (60 %)	21 (40 %)
	IRIS [18]	36	19 (53 %)	17 (47 %)
	Ours	73	70 (96 %)	3 (4 %)
Env. 2	ISC [46]	88	59 (67 %)	29 (33 %)
	IRIS [18]	96	72 (75 %)	24 (25 %)
	Ours	138	130 (98 %)	8 (2 %)
Env. 3	ISC [46]	152	133 (88 %)	19 (12 %)
	IRIS [18]	162	78 (48 %)	84 (52 %)
	Ours	189	175 (93 %)	14 (7 %)

别进行了定性和定量评估。对于定性评估, 我们将生成的图像与竞争方法的结果进行了视觉比较 (见图 7)。在定量方面, 我们使用标准指标如峰值信噪比 (PSNR)、结构相似性指数 (SSIM)、均方根误差 (RMSE) 以及平均绝对误差 (MAE) 来衡量精度, 同时还评估了推理时间以评估计算效率。由于专门针对反射率致密化的公共数据集或先前工作有限, 我们训练了几个深度完成模型, 并包括传统的插值方法进行比较 [40]–[43]。此外, 我们进行了一个消融研究以分析自适应融合模块 (AFM) 的贡献。表 I 展示了定量结果, 而图 7 展示了我们方法的定性优越性。我们的方法在保持竞争性运行时的同时, 实现了高精度 (最高的 PSNR/SSIM, 最低的 RMSE/MAE)。值得注意的是, 由于 Marigold-dc [43] 是一种为深度完成设计的零样本扩散方法, 其在我们的任务背景下效果较差; 因此, 我们仅提供其定量结果。为了评估我们的方法在机器人运动过程中生成致密反射图像的性能, 我们在机器人受到不平整道路上激烈运动姿态时捕获的数据 (5 帧累积) 上进行了测试。结果的 PSNR 值 (25.75–26.82) 表明了我们方法在对齐噪声和非线性运动伪影方面的鲁棒性 (见图 8)。此外, 我们在补充材料中比较了在不同机器人速度下的实时结果。

尽管我们希望定性地评估我们方法在其他数据集上的泛化能力, 但据我们所知, 我们尚未找到具有密集真值的此类公共数据集。然而, 我们在公共数据集 [45] (没有密集真值) 上展示了定性可视化, 这可以进一步验证我们方法的泛化能力, 如图 8 所示。

D. 在特定应用中的评估

1) 回环闭合检测: 为了验证我们方法的有效性, 我们采用了 [2] 的方法, 但用使用 Livox MID-360 生成的高

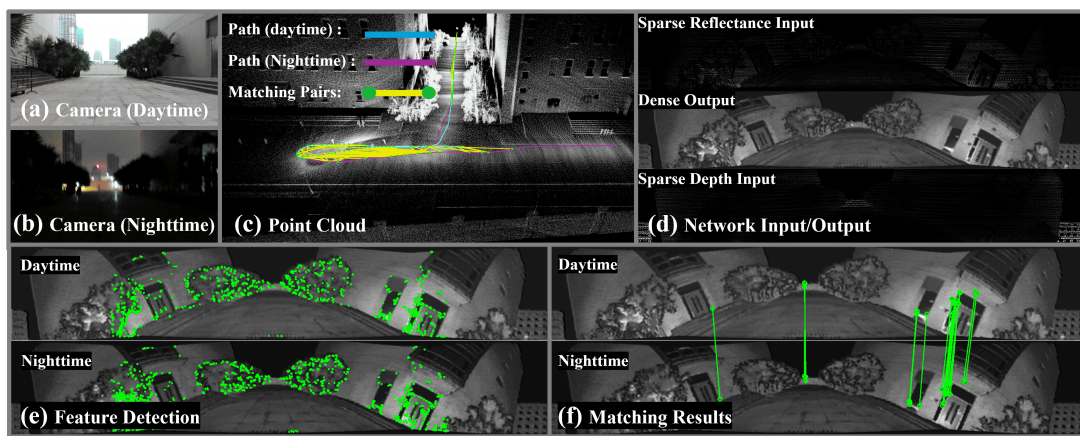


Fig. 9. 回环闭合结果在不同时间点之间进行对比，一个数据集是在白天收集的，另一个是在晚上（全景球面投影）。我们的方法通过视觉回环闭合检测实现了稳健的实时特征匹配，在显著的光照变化下表现出一致的性能。

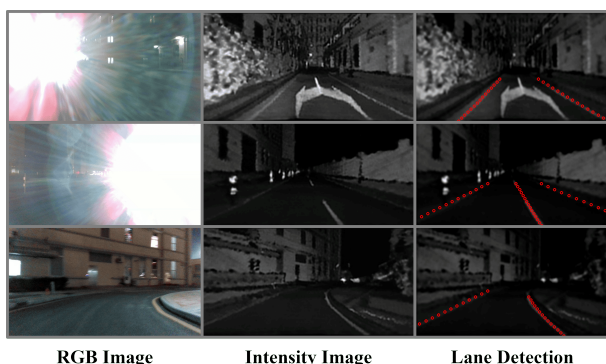


Fig. 10. 顶部：有光污染的单车道场景。中间：来车灯光照明的多车道场景。底部：光照条件不足的弯道路。

密度反射图像替换了由高成本 LiDAR 获取的高分辨率反射数据。具体来说，[2] 从反射图像中提取 ORB 特征描述符 [47]，用于回环检测。这些描述符然后被表示为词袋向量，用于查询回环。在检索到候选匹配后，进行 ORB 描述符匹配，并使用 PnP RANSAC [48] 进行验证以移除离群值。我们在三种不同环境下进行了实验，在白天和夜间两种照明条件下进行（图 9），使用在不同条件下收集的数据来评估我们方法的泛化能力。因为所采用的回环检测方法 [2] 不是基于深度学习的，为了公平比较，我们将我们的回环检测结果与其他非深度学习方法进行对比，包括 [18], [46]，这些方法是具有代表性的基于 LiDAR 描述符的方法。表 II 中的结果表明，我们的方法始终优于竞争方法，即使在环境和光照变化的情况下也能保持可靠的性能。

2) 车道检测：LiDAR 反射率图像通过物质和表面属性的差异揭示了交通车道标记与道路表面之间的强烈对比。通过利用我们方法生成的密集反射率图像，我们在包括对摄像头非常具有挑战性的场景等多种条件下实现了准确的车道检测（见图 10）。车道检测以 LaneATT [49] 作为基线模型进行。该模型在 TuSimple 数据集 [50] 上训练，并在我们的数据集上进行了微调，该数据集由 1,300 张训练图像、300 张验证图像和 200 张测试图像组成，每张图像最多包含四条车道。微调模型使用了 ResNet-34 骨干网络 [51]。检测性能是在具有不同光照条件的三个不同环境中进行评估的。整体检测准确率介于 93.3 % 至 96.1 % 之间，而误检率（FDR）和假阴性率（FNR）则

保持在较低水平，平均值分别为 6.08 % 和 6.22 %。结果展示了我们方法的鲁棒性，特别是在不利光照条件下实现了稳定准确的车道检测（图 10），优于基于相机的方法。

在这项工作中，我们引入了一个新颖的基准数据集和一个框架，用于从稀疏测量生成密集的 LiDAR 反射率图像，解决了域适应、强度校准和反射率恢复等关键挑战。未来的工作将探索将我们的方法推广到重复扫描的 LiDAR 传感器，以及在实时性能与密集化质量之间取得平衡。我们还旨在利用超分辨率反射率图像来增强动态和具有挑战性的环境中的机器人感知。

REFERENCES

- [1] Patrick Pfreundschuh et al. Coin-lio: Complementary intensity-augmented lidar inertial odometry. In *ICRA*, pages 1730–1737, 2024.
- [2] Tixiao Shan et al. Robust place recognition using an imaging lidar. In *ICRA*, pages 5469–5475, 2021.
- [3] Jan Wietrzykowski and Piotr Skrzypczyński. Descriptive power of lidar intensity images for loop closing. In *IROS*, pages 79–85, 2021.
- [4] Luca Di Giammarino et al. Visual place recognition using lidar intensity. In *IROS*, pages 4382–4389, 2021.
- [5] Yanfeng Zhang et al. Ri-lio: Reflectivity image assisted tightly-coupled lidar-inertial odometry. *IEEE RA-L*, 8(3):1802–1809, 2023.
- [6] Wenqiang Du and Giovanni Beltrame. Real-time slam with lidar intensity. In *ICRA*, pages 4164–4170, 2023.
- [7] Mohammad Aldibaja et al. Intensity-based localization on snow-wet roads. *IEEE TII*, 13(5):2369–2378, 2017.
- [8] Christophe Reymann and Simon Lacroix. Improving lidar classification using intensities. In *IROS*, pages 5122–5128, 2015.
- [9] Kasi Viswanath et al. Reflectivity is all you need!: Advancing lidar segmentation. *arXiv preprint*, 2024.
- [10] Alberto Hata and Denis Wolf. Road marking detection using lidar intensity. In *ITSC*, pages 584–589, 2014.
- [11] Xiaolu Li et al. Lidar intensity correction for road marking detection. *Optics and Lasers in Engineering*, 160:107240, 2023.
- [12] Weichen Dai et al. Lidar intensity completion: Fully exploiting sensor messages. *Sensors*, 22(19):7533, 2022.
- [13] Li Li et al. Durlar: High-fidelity lidar dataset for autonomous driving. In *3DV*, pages 1227–1237, 2021.
- [14] Charles R Qi et al. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *CVPR*, pages 652–660, 2017.
- [15] Wei Xu et al. Fast-lio2: Fast direct lidar-inertial odometry. *IEEE T-RO*, 38(4):2053–2073, 2022.

- [16] David Wisth et al. Vilens: Visual, inertial, lidar, and leg odometry for all-terrain legged robots. *IEEE T-RO*, 39(1):309–326, 2022.
- [17] Jiarong Lin et al. R3live: A robust, real-time, rgb-colored, lidar-inertial-visual tightly-coupled state estimation. In *ICRA*, pages 10672–10678, 2022.
- [18] Ying Wang et al. Lidar iris for loop-closure detection. In *IROS*, pages 5769–5775, 2020.
- [19] Wei Gao et al. Active loop closure for osm-guided robotic mapping in large-scale urban environments. In *IROS*, pages 12302–12309, 2024.
- [20] Jonas Uhrig et al. Sparsity invariant cnns. In *3DV*, pages 11–20, 2017.
- [21] Fangchang Ma and Sertac Karaman. Sparse-to-dense: Depth prediction from sparse samples. In *ICRA*, pages 4796–4803, 2018.
- [22] Jiaxiong Qiu et al. Deeplidar: Depth prediction for sparse lidar. In *CVPR*, pages 3313–3322, 2019.
- [23] Kaiyue Lu et al. Depth completion via auxiliary image reconstruction. In *CVPR*, pages 11306–11315, 2020.
- [24] Qingyang Yu et al. Grayscale and normal guided depth completion. In *ICIP*, pages 979–983, 2021.
- [25] Tian Yu Liu et al. Monitored distillation for depth completion. In *ECCV*, pages 35–53, 2022.
- [26] Xingxing Zuo et al. Codevio: Visual-inertial odometry with dense depth. In *ICRA*, pages 14382–14388, 2021.
- [27] Mingle Zhao et al. Dit-slam: Dense visual-inertial slam with implicit depth. *Sensors*, 22(9):3389, 2022.
- [28] Guangkai Xu et al. Towards domain-agnostic depth completion. *Machine Intelligence Research*, 21(4):652–669, 2024.
- [29] Jin-Hwi Park et al. Depth prompting for sensor-agnostic depth estimation. In *CVPR*, pages 9859–9869, 2024.
- [30] Wei Fang et al. Intensity correction of terrestrial laser scanning data. *IEEE TGRS*, 53(2):942–951, 2014.
- [31] Huajie Wu et al. Moving event detection from lidar point streams. *Nature Communications*, 15(1):345, 2024.
- [32] Ashish Vaswani et al. Attention is all you need. *NeurIPS*, 30, 2017.
- [33] Syed Waqas Zamir et al. Restormer: Transformer for high-resolution image restoration. In *CVPR*, pages 5728–5739, 2022.
- [34] Jingyun Liang et al. Swinir: Image restoration using swin transformer. In *ICCV*, pages 1833–1844, 2021.
- [35] Jonathan Ho et al. Denoising diffusion probabilistic models. *NeurIPS*, 33:6840–6851, 2020.
- [36] Xin Li et al. Diffusion models for image restoration and enhancement: A survey. *arXiv preprint*, 2023.
- [37] Olaf Ronneberger et al. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241, 2015.
- [38] Liang-Chieh Chen et al. Deeplab: Semantic image segmentation with deep nets and crfs. *IEEE TPAMI*, 40(4):834–848, 2017.
- [39] Shu Kong et al. Recurrent scene parsing with perspective understanding in the loop. In *CVPR*, pages 956–965, 2018.
- [40] Jason Ku et al. In defense of classical image processing for depth completion. In *CRV*, pages 16–22, 2018.
- [41] Fangchang Ma et al. Self-supervised sparse-to-dense depth completion. In *ICRA*, pages 3288–3295, 2019.
- [42] Abdelrahman Eldesokey et al. Uncertainty-aware cnns for depth completion. In *CVPR*, pages 12014–12023, 2020.
- [43] Massimiliano Viola et al. Marigold-dc: Zero-shot monocular depth completion with guided diffusion. *arXiv preprint arXiv:2412.13389*, 2024.
- [44] Sanna Kaasalainen et al. Incidence angle effects on laser scanner intensity. *Remote Sensing*, 3(10):2207–2221, 2011.
- [45] Yanbo Wang et al. Sfpnet: Sparse focal point network for lidar segmentation. In *ECCV*, 2024.
- [46] Han Wang et al. Intensity scan context: Coding intensity and geometry relations for loop closure detection. In *ICRA*, pages 2095–2101, 2020.
- [47] Ethan Rublee et al. Orb: Alternative to sift or surf. In *ICCV*, pages 2564–2571, 2011.
- [48] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981.
- [49] Lucas Tabelini et al. Keep your eyes on the lane: Real-time attention-guided lane detection. In *CVPR*, pages 294–302, 2021.
- [50] Tusimple lane detection challenge. <https://github.com/TuSimple/tusimple-benchmark>, 2017.
- [51] Kaiming He et al. Deep residual learning for image recognition. In *CVPR*, pages 770–778, 2016.

Super LiDAR Reflectance for Robotic Perception

Supplementary Material

如图 ?? 所示, 对于同一个物体, 由于距离或入射角的差异, LiDAR 光束获得的反射率会有所不同。在实际操作中, 针对相同材料的物体收集多个数据集, 其中距离和角度被分别控制为单一变量, 以拟合它们与反射率的各自关系。然后根据文中的公式获得模型参数。标定结果的一个例子如图 11 所示, 应用我们的补偿方法后, 同一物体的反射率保持一致 (见图 11 (b)), 不像图 11 (a) 中那样, 反射率随入射角或距离变化。此外, 为了验证我们的补偿模型在远距离 (大于 10 米) 下的有效性, 我们分析了同一物体在不同距离和角度下的强度分布, 以及校正后的结果, 如图 12 所示。结果表明, 我们的补偿方法有效地保持了同一物体在不同距离和角度下反射率值的一致性。

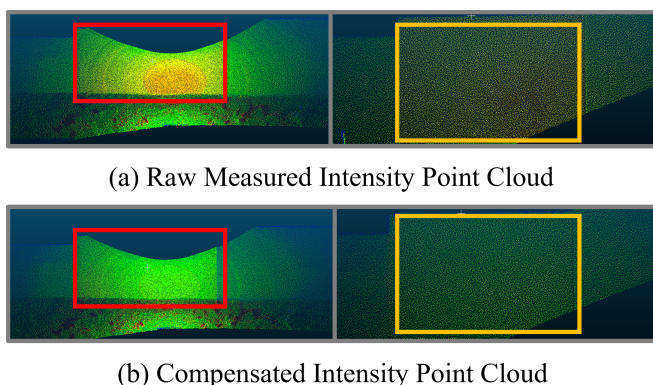


Fig. 11. 对于同一对象 (红色和橙色框), 原始测量显示由于角度和距离的变化导致反射率不一致 (颜色表示反射率值)。补偿之后, 点云中同一对象的反射率保持在一致的范围内。

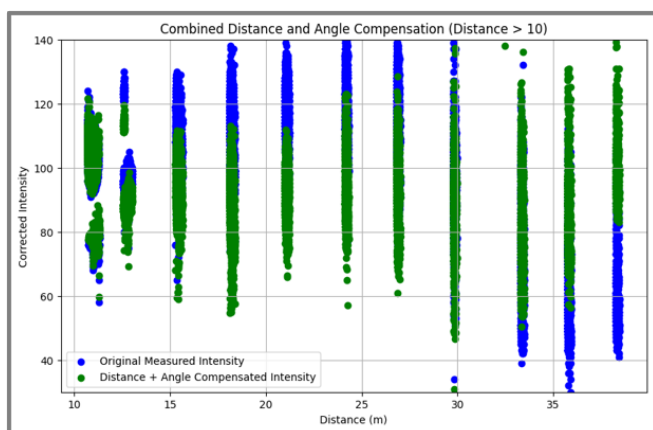


Fig. 12. 反射值分布在不同的距离上。蓝色点代表原始测量的强度, 显示了随着距离的显著变化, 而绿色点代表应用补偿模型后的值, 保持了更为一致的范围。



Fig. 13. 公共数据集 [45] 上的输入和强度图像加密结果

IV. 补充实验

A. 在公共数据集上的测试

正如在第 III-C 节中所述, 我们在一个公共数据集 [45] 上提供了更多定性可视化结果 (没有密集的真实值), 这可以进一步验证我们方法的泛化能力, 如图 13 所示。

B. 帧积累对致密化质量的影响

正如在第 VI 节中所述, 为了证明使用 5 帧累积数据作为输入的合理性, 我们定性和定量地比较了使用单帧和 10 帧累积数据作为输入的致密化结果。如表 III 和图 14 所示, 5 帧输入的致密化性能与 10 帧输入相当, 而单帧输入的重建质量明显较差。这是因为单帧输入通常过于稀疏, 导致物体缺乏足够的信息进行准确的重建。值得注意的是, 这里的评估是基于静态收集数据的致密化质量。虽然 10 帧输入在指标性能上更佳, 但它忽略了机器人运动过程中里程计误差和场景中运动物体的影响。输入数据和生成的图像如图 14 所示。

C. 机器人速度对致密化性能的影响

鉴于我们的方法旨在为机器人运动过程中的感知任务服务, 我们进行了实验, 以评估在相同场景下不同机器人速度的动态致密化结果, 使用回环检测作为测试任务。定量比较结果如表 IV 所示。在图 15 中, 展示了两视

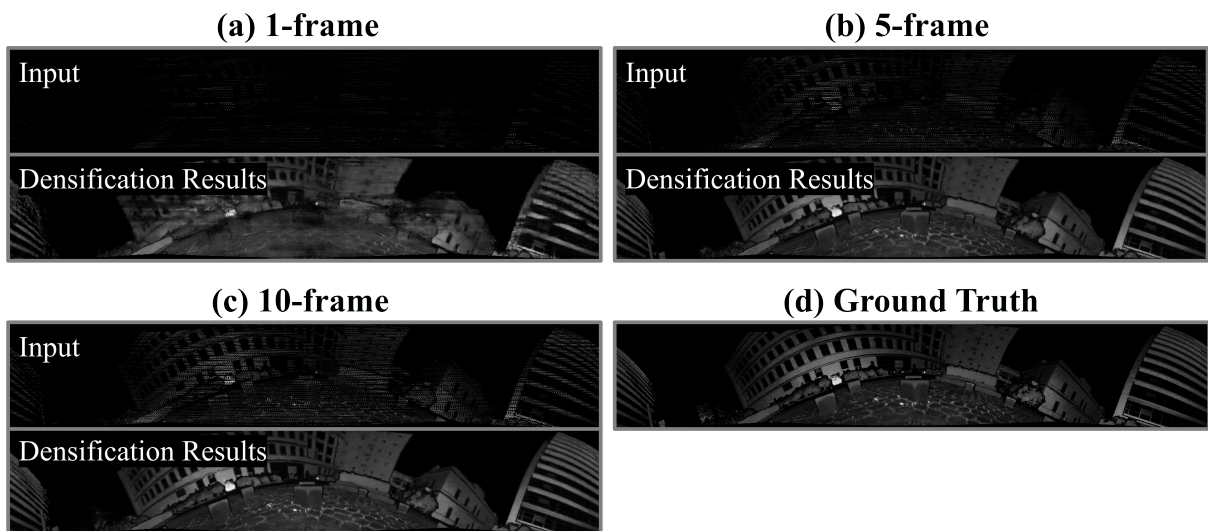


Fig. 14. 从 1 帧、5 帧和 10 帧累积输入生成的密集反射图像的比较

TABLE III

利用不同的指标对 1 帧、5 帧和 10 帧累积输入生成的密集反射率图像进行定量比较

Frames Input	PSNR	SSIM	RMSE	MAE
1-frame	22.303	0.350	0.0802	0.058
5-frame	27.142	0.757	0.0503	0.0263
10-frame	27.239	0.762	0.0496	0.0256

觉回环检测结果，以及不同速度下生成的图像质量和特征点匹配结果。

TABLE IV

不同速度下的回环检测结果

Average Speed (m/s)	Detected Loops	Trajectory Length (m)
0.73	213	193.73
1.52	185	203.51

V. 数据集构建

我们构建了超级激光雷达反射率数据集，其中包括 1000 个原始 ROS 数据包和另外 5000 组反射率图像，配置从稀疏（1 次扫描）到密集（500 次扫描）不等。具体来说，通过全景模式投影，从 1000 个原始 ROS 数据包生成了 1000 张全景反射率图像，其中一些示例结果如图 16 所示。这些图像涵盖了激光雷达传感器的全部视场 (FoV)。同样，通过虚拟相机投影模式，从相同的 1000 个原始 ROS 数据包中生成了 4000 组反射率图像，示例图像如图 17 所示。

此外，该数据集包含两个针对实际任务定制的子集。第一个子集旨在评估生成的全景密集反射图像在闭环检测中的性能。第二个子集则评估生成的虚拟相机密集反射图像在车道检测中的实用性。两个子集的收集都是在机器人运动过程中进行的，旨在验证密集反射图像在增强机器人运动感知方面的效果。

VI. 限制和未来工作

尽管所提出的方法在生成密集 LiDAR 反射图像方面表现出显著进步，但仍然存在一些局限性。首先，由于扫描模式的差异，目前的方法不易泛化到所有 LiDAR 类型，这限制了其在各种硬件设置中的适应性。未来的工作将研究该方法对不同扫描模式的 LiDAR 的泛化和适应性，以解决这一限制。

其次，使用单帧输入有助于缓解静态与动态域差距问题，但会导致数据过于稀疏，这可能影响性能。解决这种权衡是未来研究的重要方向，以进一步增强该方法的鲁棒性和性能，尤其是在动态和具有挑战性的环境中。

此外，本文提出了一种从稀疏输入生成密集 LiDAR 反射图像的新方法，解决了领域适应、强度校准和恢复等挑战。未来，我们将利用生成的超级 LiDAR 反射图像支持更广泛的机器人感知任务，使其能够在动态和复杂的环境中使用。

此外，将探索将相机图像与 LiDAR 反射图像相结合。通过用相机信息为反射图像着色，密集化的图像可以提供更详细的语义理解，弥合 LiDAR 与相机领域之间的差距。这将为基于视觉的任务带来新的可能性，例如语义分割，其中相机基础的方法可以适用于 LiDAR 数据。这些进步将进一步拓宽所提出方法在各种机器人感知场景中的适用性。

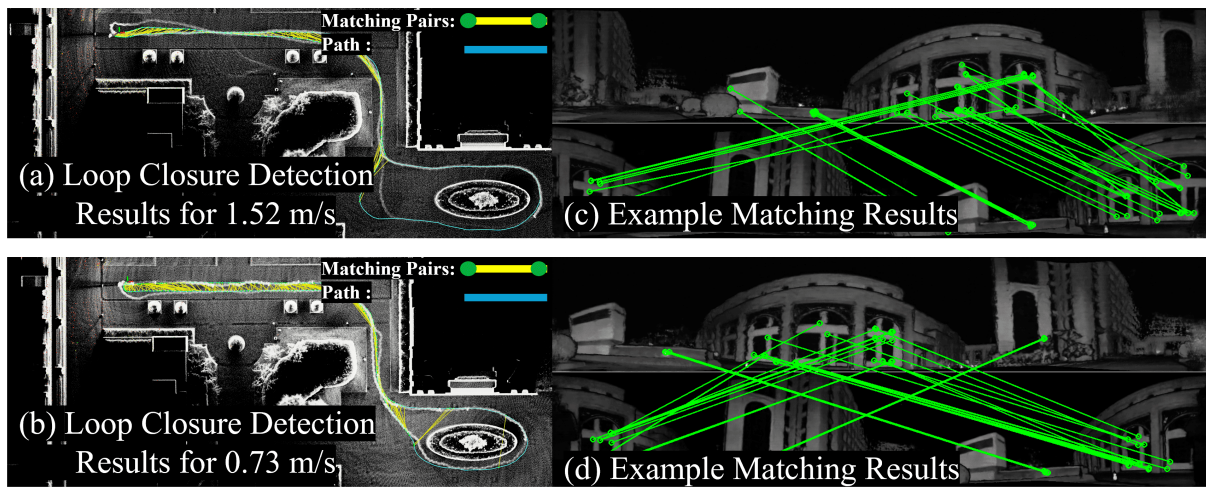


Fig. 15. 同一场景中不同机器人速度下的回环检测和特征点匹配结果。

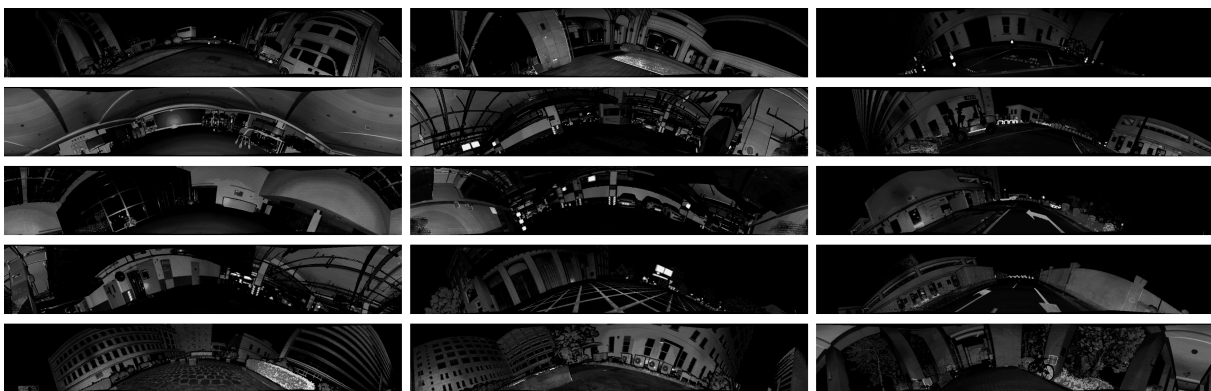


Fig. 16. 利用全景投影生成的全景反射率图像。



Fig. 17. 使用虚拟摄像机投影生成的虚拟摄像机反射图像。