arXiv:2304.01074v1 [cs.RO] 3 Apr 2023

FinderNet：一种无需数据增强、借助规范化的六自由度点云回环检测与闭环技术。

Sudarshan S Harithas[1] ,Gurkirat Singh[1],Aneesh Chavan[1],Sarthak Sharma[1],Suraj Patni[2],Chetan Arora[2],K. Madhava Krishna[1]

摘要——我们聚焦于多智能体场景中基于激光雷达点云的回环检测（或查找）与闭环（LDC）问题。最先进技术（SOTA）会直接生成给定点云的学习嵌入，需要大量数据传输，并且无法适应六自由度（DOF）视点的广泛变化。此外，无结构点云缺乏强先验，导致 LDC 的准确性很低。在这一全新方法中，我们提出利用共同的主导地面，分别对点云进行横滚和俯仰规范化。沿垂直于地平面的轴对规范化后的点云进行离散化，可生成类似数字高程图（DEM）的图像，从而揭示场景中的强空间先验。实验表明，基于此类数字高程图学习嵌入的 LDC 不仅数据效率高，而且比当前最先进水平的方法更加稳健且具有更好的泛化能力。我们报告了在平均精度（用于回环检测）以及相对位姿估计（或回环闭合）的绝对平移误差和旋转误差方面的显著性能提升；测试数据集包括*KITTI*、*GPR*和*Oxford Robot Car* ，并与多种最先进的 LDC 方法进行了比较。我们的编码器技术可将原始点云压缩 830 倍以上。为进一步测试该技术的稳健性，我们创建并开源了一个名为*Lidar-UrbanFly* 数据集（*LUF*）的定制数据集，其中包含由安装在四旋翼无人机上的激光雷达采集的点云。

I. 引言

回环检测与闭环是同步定位与建图（同步定位与建图）处理流程中的关键模块，可减少估计过程中累积的漂移。随着单智能体同步定位与建图系统日趋成熟，研究界正越来越关注多智能体场景和协作式同步定位与建图 [1], [2]。由于在智能体之间传输大型点云并不现实，因此必须谨慎利用有限的带宽，以实现分布式回环检测与闭合（LDC）[2]–[4]。本文着重研究鲁棒且具备高数据效率的分布式 LDC。尽管我们处理流程中的大多数组件适用于通用点云，但我们假设激光雷达是必不可少的传感模态。

这些技术大致分为两类：（**1**）回环检测这类方法通常采用地点识别方法 [5]–[8]检测回环，并使用传统的点云配准算法（如 [9], [10] ）估计查询点云与恢复出的点云之间的相对位姿。（**2**）回环检测与闭合 [11]–[13]等方法在端到端处理流程中执行地点识别，并估计查询点云与检索到的点云之间的相对位姿，而无需采用

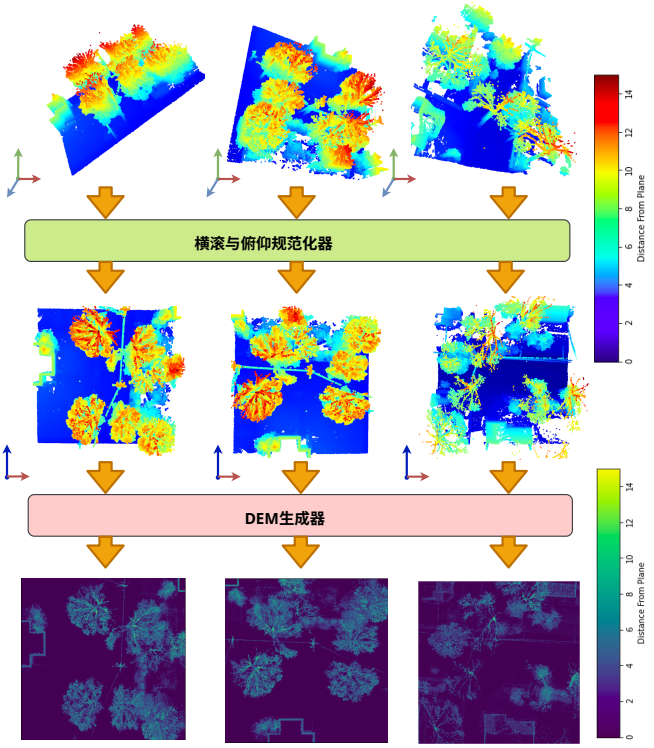


图 1：我们发现，原始激光雷达点云（第一行）缺乏空间结构，难以实现稳健的回环检测与闭环（LDC）。我们先进行局部滚转和俯仰规范化（第二行），再沿 z 轴离散化（第三行），得到类似数字高程图（DEM）的输出，并揭示输入中丰富的场景结构。我们的模型基于此类数字高程图执行 LDC，从而实现较高的数据效率、稳健性，并能泛化至六自由度视点变化。

任何外部点云配准方法。我们的方法属于第二类，在单一处理流程中检测并闭合回环，而无需采用任何外部点云配准方法。

通常，这些方法 [5]–[8], [11] 要么结合特征聚合与数据增强，对输入点云施加随机采样的刚性变换以实现视点不变性，要么通过重叠估计进行 LDC。这种训练过程无法泛化至大范围视点变化。我们采用了一种新颖的方法，将规范表示与可微潜在空间对齐相结合，从几何层面约束系统实现视点不变性。该技术在多个数据集上达到了最先进水平。

我们工作的基石是滚转和俯仰（*RP*）规范化器和一个可微偏航变换器（DYT）。

[1]就职于印度 IIIT 海得拉巴的 RRC[2]就职于印度 IIT 德里大学

[1]项目页面：https://gurkiratsingh.me/FinderNet/

该*RP*规范化器利用主导地平面假设（常见于自动驾驶和无人机应用） [12]–[16] 来补偿两个点云之间的横滚和俯仰。经过横滚和俯仰规范化的点云会转换为数字高程图（DEM），图 1 展示了该过程的直观说明。我们进一步开发了一种可微偏航变换器（DYT），它作用于DEM 的潜在特征嵌入上，以实现偏航不变性，并提供具有六自由度（DOF （SO(3)）相对运动的视角不变回环检测。这与现有技术形成对比，后者仅关注偏航旋转 [12], [15], [17], [18]。

贡献：（1）新颖的处理流程：我们不直接处理天生缺乏结构的原始点云，而是通过横滚与俯仰规范化器（第 III-A 节）将点云转换为规则间隔的 DEM，后续仅需处理偏航。规范化的 DEM 表示具有 CNN 骨干网络可直接处理的结构，从而避开了通常困扰点云表示的等变性问题。尽管 Pointnet [19]及其变体 [20]能够处理等变性问题，但在多种回环检测相关性能指标上的结果表明，本文处理流程优于受 Pointnet 启发的架构 [5]–[7] 。（2）用于视点不变性的规范化流程与可微对齐：不同于以往通过特征聚合 [5]–[7], [11]或重叠估计 [12], [21], 来处理回环检测任务的方法，我们采用规范化与可微对齐流程，在网络中引入几何约束以实现视点不变性，并支持在不使用数据增强的情况下进行训练，同时在多个数据集上达到最先进水平。我们的潜在空间对齐由 DYT 驱动；DYT 是一种新颖的参数估计模块，用于可微网格采样。与 [12], [13] 等方法相比，DYT 能以自监督方式估计相对偏航角，即无需对两个点云之间的相对偏航角进行显式监督。（3）六自由度恢复：不同于以往仅将回环闭合表示为 SE(2) 对齐的方法，本文方法能够恢复两个候选点云之间的六自由度位姿，同时无需数据增强，这得益于处理流程固有的视点不变性。本文框架在 KITTI [22], 、GPR [23] 和 Oxford RobotCar [24] 等多个公开数据集上，在回环检测的既定性能指标方面超越最先进水平。具体而言，该框架在六自由度位姿恢复方面表现最佳，并且在 SE(2) 回环检测任务上优于大多数已有方法。表 I 从概念层面对比了我们的方法与当代技术。其中，DE 指以压缩为手段、并以在压缩空间中执行回环检测或配准为下游目标的方法；此外，这些方法能够从压缩空间中解压缩点云，获得部分或完整形式（DEM 属于部分重建，经过上采样和补全可获得稠密结果）。LD 指执行回环检测任务并使用外部点云的方法

方法	会议	DE	VI	NDA	LD	LDC
[12]	RSS'20	✗	✗	✗	✓	✓(偏航)
[15]	IROS'18	✗	✗	NA	✓	✓(偏航)
[25]	RAL'22	✓	✓	NA	✗	✗
[8]	ICRA'22	✓	✓	✗	✓	✗
[26]	RAL'21	✓	✓	✗	✗	✗
[11]	TRO'22	✗	✓	✗	✓	✓
[13]	IROS'19	✗	✗	✗	✓	✓(偏航)
[7]	CVPR'18	✗	✓	✗	✓	✗
[5]	CVPR'19	✗	✓	✗	✓	✗
[6]	CVPR'21	✗	✓	✗	✓	✗
[27]	ECCV'20	✗	✓	✗	✓	✓
Ours	****'23	✓	✓	✓	✓	✓

表 I：缩略语：**DE：** 通过学习嵌入实现的数据效率，**VI：** 六自由度视点不变性，**NDA：** 无需大量数据增强，**LD：** 回环检测，**LDC：** 联合回环检测与闭环，**NA：** 不适用。

配准方法 [9], [10]以估计相对位姿。**LDC** 是通过单一处理流程联合估计回环及相对位姿的方法。若某方法无需数据增强即可进行学习，则**NDA** 设为 true（或 ✓）；**VI** 则指能够处理 SE(3)运动的方法。

II. 相关工作

手工特征描述符： [15], [16], [28]依靠手工特征描述符提取局部几何信息，并将其聚合为适用于回环检测与闭合的全局描述符。 [15], [16] 假设存在主导地面，以检测并闭合回环。 [15]采用极坐标表示，将地面沿径向和方位角方向划分为多个分箱，并在每个分箱中存储垂直体积内的最大高度。 [16] 以笛卡尔形式将地平面离散为矩形网格单元，并在每个单元中存储点云密度。我们的 DEM 表示是地平面的离散笛卡尔表示，其中每个网格单元存储其中点的最大高度。不过，与 [15], [16],不同，我们进行完整的六自由度估计和回环闭合。

基于学习的回环检测方法：PointNet[19] 提出了一种神经网络模型，可直接处理点云，同时保持置换不变性。*PointNetVLAD* [7] 使用 [19]和 NetVLAD [29] 生成用于地点识别的全局描述符。*PCAN* [5] 使用[19] 作为骨干架构来提取局部特征和相应的注意力图，并结合 [29] 进行特征聚合。然而，这两种方法都使用 PointNet 作为骨干架构，通过多层感知机分别处理每个点，无法捕获局部邻域信息。近期，Retriever [8]借助基于 Perceiver [30] 的机制聚合局部特征，直接在压缩后的特征表示中检测回环。上述方法均使用聚合后的局部特征来计算具有视点不变性的全局描述符，但这类方法需要耗费大量计算的数据增强。我们提出了一种规范化过程，以显式地实现视点不变性。与仅执行回环检测的 [5], [7], [8] 不同，我们的方法还执行回环检测与闭合。

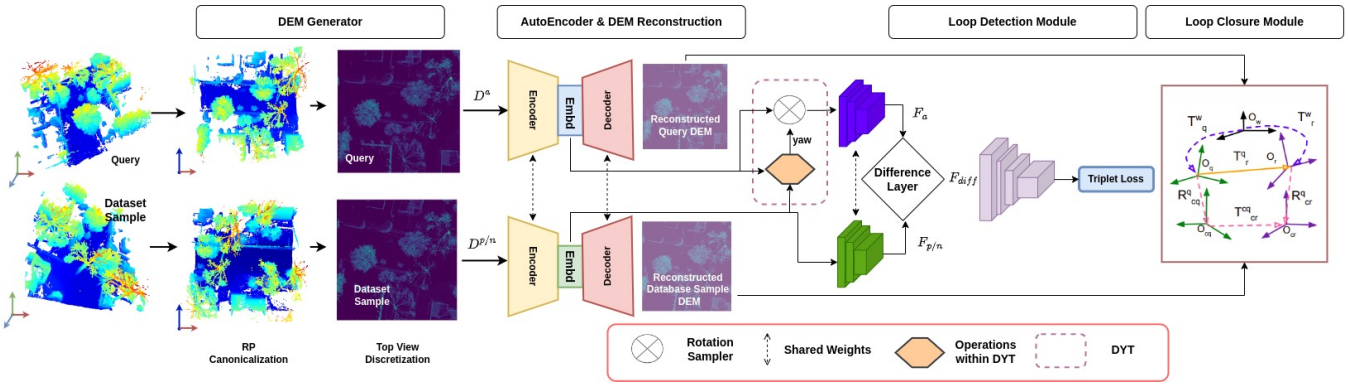


图 2： 该图展示了我们处理流程的概览；最左侧的两个点云分别是输入查询和数据库样本；*DEM*生成器 (第 III-A 节) 会生成点云的离散化俯视图；自动编码器结构会进一步压缩*DEM* (第 III-B 节可微偏航变换器 (DYT)（第 III-C 节）用于偏航对齐；DYT 中的操作包括 *CPC*、极坐标嵌入的水平填充和相关性；每项操作都在第 III-C 节中说明。完整的操作集合以一个橙色六边形表示；这些操作的结果是一个标量偏航值，并将其输入旋转采样器。我们设计了一个网络，利用这些压缩后的嵌入执行回环检测（第 III-D 节）和闭环（第 III-E 节），无需显式解压缩。

基于学习的回环检测与闭合方法：LCDNet [11] 提出了一种端到端可训练系统，采用非平衡最优传输算法来估计两个点云之间的六自由度相对变换。DH3D [27] 使用层次网络聚合局部特征，以获得用于回环检测的全局特征。这两种方法都依赖于对输入点云进行计算开销较大的六自由度数据增强，以实现方向不变性。所提出的框架通过显式的横滚和俯仰规范化，随后进行偏航对齐，从而无需数据增强。与[11], [27], 不同，我们处理的是高度压缩的点云表示，因此适用于多智能体场景中的数据传输。OverlapNet [12] 和 OverlapTransformer [21] 等基于重叠的方法，使用距离图像[31]上的显式重叠信息进行训练。OREOS [13] 提出了两个独立分支：一个用于回环闭合，另一个用于回环检测。与仅估计输入点云之间相对偏航角的 [12], [13],[21] 不同，我们估计完整的六自由度相对位姿。

III. 方法

我们的目标是为三维点云开发一种用于LDC的六自由度视点不变地点识别框架。我们的方法概览如图2所示。我们首先对点云进行规范化，然后将其离散化以获得 DEM 表示（第III-A节）。我们使用一种自动编码器式编码器-解码器网络，学习DEM 的压缩潜在表示（第III-B节）。该潜在表示在各智能体之间传输，用于回环检测与闭合，从而降低数据带宽需求。

为了在压缩空间中实现回环检测的偏航不变性，我们设计了一种可微偏航变换器（DYT），它会变换潜在查询嵌入，使其旋转对齐到数据库样本的潜在嵌入（第III-C节）。DYT 的输出用于回环检测（第III-D节）。检测到回环后，解码器会解压潜在数字高程模型表示，并使用从查询和数据集中解码得到的数字高程模型来估计用于回环闭合的六自由度相对位姿（第III-E节）。

A. 数字高程模型生成

数字高程模型是输入点云的数字表示，可捕捉地形或叠加物体的高程信息。数字高程模型具有强大的表示能力，既能保留边缘和角点等特征丰富区域，又能通过高强度压缩并以高质量恢复来节省带宽。此外，与仅保留偏航信息的距离图像不同，数字高程模型还保留平面平移信息，因此是六自由度点云配准的实用表示。

平面参数化：考虑输入点云 $\mathbf{P}_c$ 、其对应的地平面 $\mathbf{r}_c$以及世界地平面 $\mathbf{r}_w$。我们旨在通过估计它们之间的相对横滚和俯仰（RP），使平面 $\mathbf{r}_c$ 与 $\mathbf{r}_w$ 对齐。我们将输入点云 ($\mathbf{P}_c$) 移至中心，并使用 RANSAC 提取地平面 $\mathbf{r}_c$ 。地平面通过 ($\mathbf{n}_c$, $\mathbf{C}_c$) 参数化，其中 $\mathbf{n}_c \in \mathbf{R}^3$ 是垂直于该平面的单位向量，$\mathbf{C}_c = \{c_i^c \mid i = \{1...n\}\} \in \mathbf{R}^{n\times 3}$ 是点集 $c_i^c \in \mathbf{R}^3$, , 满足 c_i^c 位于 $\mathbf{r_c}$ 上且 $\|c_i^c\| = 1$。世界地平面 $\mathbf{r}_w$ 采用类似方式参数化为 ($\mathbf{n}_w$, $\mathbf{C}_w$)，其中 $n_w = [0, 0, 1]$, 且 $\mathbf{C}_w = \{c_i^w \mid i = \{1...n\}\} \in \mathbf{R}^{n\times 3}$是点集 $c_i^w \in \mathbf{R}^3$, , 满足 c_i^w位于 $\mathbf{r_w}$ 上且 $\|c_i^w\| = 1$。请注意，世界地平面并非通过数据估计得出，而是构造的规范参考平面。横滚 (α) 和俯仰 (β) 的规范化分为两个步骤。首先，通过对齐法向量 $\mathbf{n}_c$ 和 $\mathbf{n}_w$，得到 (α) 和 (β) 的粗略估计。随后，使用迭代最近点算法（ICP）进行更精细的估计。

粗略 RP 规范化：给定来自 RANSAC 的法线 $\mathbf{n}_c = [n_{c,x}, n_{c,y}, n_{c,z}]$ 和 $\mathbf{n}_w = [0, 0, 1]$，我们通过求解以下问题来估计相对横滚 α 和俯仰 β：

$$\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}\cos(\alpha) & 0 & \sin(\alpha)\\ \sin(\beta)\sin(\alpha) & \cos(\beta) & -\cos(\alpha)\sin(\beta)\\ -\cos(\beta)\sin(\alpha) & \sin(\beta) & \cos(\alpha)\cos(\beta)\end{bmatrix}\begin{bmatrix}n_{c,x}\\n_{c,y}\\n_{c,z}\end{bmatrix}.$$

所得的闭式解如下：

$$\alpha = \arctan\left(\frac{-n_{c,x}}{n_{c,z}}\right), \beta = \arctan\left(\frac{n_{c,y}}{n_{c,z}\cos(\alpha) - n_{c,x}\sin(\alpha)}\right)$$

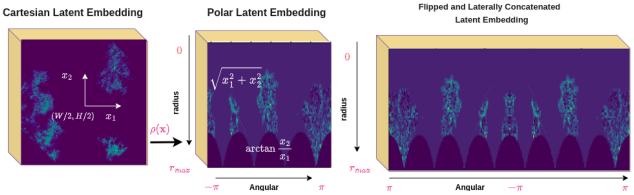


图 3：最左侧图像展示了一个以笛卡尔形式表示的样本 数字高程模型潜在空间。中间的图像描绘了同一嵌入的极坐标形式；右侧图像则是翻转与拼接操作的结果。

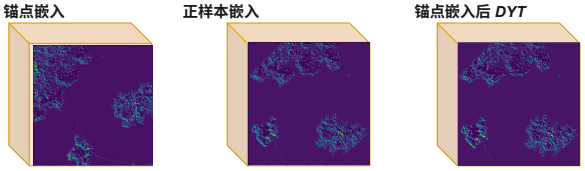


图 4：使用 DYT 对偏航进行对齐的可视化。请注意，锚点和正样本最初并未实现偏航对齐，但经过*DYT* 操作后，两个嵌入实现了对齐。为便于可视化，我们将特征体的第一个通道显示为二值图像。

使用迭代最近点进行精细规范化：我们使用上述对 α 和 β 的粗略估计，并通过迭代最近点对其进行细化（以粗略估计作为初始值），如下所示：

$$R(\alpha,\beta)=\underset{(\alpha,\beta)}{\arg\min}\,||\mathbf{C_w}-R(\alpha,\beta)\mathbf{C_c}||^2$$

俯视图离散化：完成横滚和俯仰规范化后，将点云的俯视图离散化为均匀的 $2D$ 网格单元，以获得点云的数字高程模型（DEM）$_c$。我们定义一个维度为$G_w\times G_h$、分辨率为 d_g的网格 G。根据分辨率 d_g，为每个网格单元 $g_i\in G$ 分配一组点 P_i^g，并赋予其高度值 $h_i^g=\max(h(p)\,|\,p\in P_i^g)$，其中 $h(p)$ 表示点 p的高度，从而将点云转换为数字高程模型。这样的网格表示可以轻松地输入深度网络，以充分利用此类结构信息。

B. 学习压缩潜在表示

我们旨在使用数字高程图 执行LDC 任务，并使用自编码器架构生成数字高程模型的压缩潜在嵌入。 编码器 由依次堆叠的 CNN 层组成，用于特征提取。编码过程充当压缩器，生成特征体（潜在嵌入），与原始数字高程模型 （或其对应的点云）相比，所需的存储和传输内存显著更少。我们用ϕ 表示该嵌入，并且 $\phi\in R^{4\times125\times125}$。回环检测模块旨在使用 ϕ进行视点不变地点识别。然而，对于回环闭合，我们会将 ϕ解码回数字高程模型，然后进行位姿估计。我们的编码器—解码器架构的详细说明见 †项目页面。

C. 可微偏航变换器（*DYT*）用于参数化网格采样

视点不变性是实现稳健地点识别/回环检测的一项重要属性。以往的方法 [11], [13] 尝试通过数据增强实现这一点，即将输入点云按照随机选取的旋转角度进行旋转。然而，这类方法无法推广到复杂序列或视点发生大幅变化的情况。 [11] 指出，仅靠数据增强未必足以实现视点不变性。因此，我们提出一种可微偏航变换器 （DYT）模块，无需显式数据增强即可实现视点不变性。该模块以锚点和正样本（或负样本）DEM 的潜在嵌入（分别记为 ϕ_A、 ϕ_p 和ϕ_n ）为输入，并在输出端返回相对偏航角 $\psi\in R$。随后，它旋转锚点 DEM，使锚点与正样本之间的相对偏航角为零。下文将详细介绍 DYT 内部的操作。

笛卡尔坐标到极坐标转换（CPC）：设$\mathbf{G}$ 为由 ψ 参数化的旋转变换群（属于 SO(2)），使得 T_ψ：$R^d\to R^d$, $\forall T_\psi\in SO(2)$。$\mathbf{G}$ 的典范坐标定义如下：笛卡尔坐标中旋转 T_ψ 表现为典范坐标中的平移 ψ。极坐标系构成该旋转变换群 [32], [33], 的此类典范坐标，可由笛卡尔坐标$\mathbf{x}$：$(x_1,\ x_2)$ 得到，具体如下：

$$\rho(\mathbf{x})=\left(\arctan\frac{x_2}{x_1},\sqrt{x_1^2+x_2^2}\right).\qquad(1)$$

CPC在嵌入张量的每个通道上执行ϕ，从而得到尺寸相同的输出张量。CPC过程如图3的前两列所示。

水平填充极坐标嵌入：如上所述，在笛卡尔坐标中通过偏航旋转相关联的两个嵌入，在转换为极坐标后通过平移相关联。然而，如果我们尝试直接估计平移，估计过程只能在重叠区域之间建立相关性。我们观察到，极坐标潜在嵌入的水平轴位于范围 $[-\pi,\pi]$ 内，并且具有周期性。周期性使我们能够通过复制嵌入并将其翻转（翻转后的嵌入将位于范围 $[\pi,-\pi]$内）来填充嵌入，然后使用翻转版本对嵌入进行水平填充。所得嵌入如图3（最右侧）所示。该操作将潜在嵌入的大小加倍至 $4\times125\times250$，从而允许我们使用完整嵌入进行平移估计。

相关层：将正样本（负样本）的极坐标嵌入填充后，我们尝试通过相关性在其中定位锚点嵌入。我们将该层实现为一个卷积层，以锚点的极坐标潜在嵌入作为卷积核，并在正样本/负样本的水平填充极坐标特征体上执行互相关。由此得到大小为$1\times1\times126$的一维输出。相关层的输出将 360 度旋转划分为 126 个分箱，每个分箱的分辨率约为 2.85 度。我们对相关性得分输出应用 Soft max，将得分向量转换为

不同候选平移对应的概率向量。预测平移乘以 2.85 后，即可转换为预测旋转角。此前已有研究 [12]，使用了与相关层类似的模块，但其设置要求明确的偏航角监督。我们的贡献之一是通过将软偏航角估计作为自监督DYT的一部分来放宽这一要求。如第*KITTI-08*和*LUF* 序列所示，这使得模型在视点变化较大时性能有所提升，详见第V-A节。

旋转采样器：我们根据上一步确定的预测偏航角构造一个旋转矩阵 $R_\psi \in SO(2)$，即ψ。与 [34], 类似，我们使用 R_ψ 对输入特征体进行可微采样，并生成变换后的$\widehat{}$输出特征图，记为 ϕ。该操作在图$\otimes$ 中记为 2。请注意，该操作作用于笛卡尔坐标系中的 $4\times125\times125$ 维嵌入张量。图 4 展示了 DYT 模块的结果。可以看出，在输入 DYT 时，锚点与正样本的方向不同；而经过 DYT 后，它们的方向相同。为便于说明，我们仅展示 $4\times125\times125$ 张量的第一个通道。变换后的锚点张量被送入下一模块进行回环检测。*D.* 回环检测

我们的处理流程通过 RP 规范化器和 DYT 实现旋转不变性。此外，回环检测模块中的差分层提供平移不变性，并衡量两个偏航对齐 DEM 之间的相似度。全卷积网络（CNN）具有平移等变性。我们的回环检测模块由共享的 CNN 层组成，用于提取锚点的特征 $F_a \in R^{H\times W\times C}$以及正样本或负样本 DEM 的特征 $F_{p/n} \in R^{H\times W\times C}$。差分层将两个特征体作为输入，并计算所有像素对之间的绝对差。为实现所有像素对的绝对差，我们首先将 F_a 重塑为 $HW\times1\times C$张量，然后将每个通道中的第一列重复HW次，从而构造平铺张量 $T_a \in R^{HW\times HW\times C}$。数学表示为：
$\forall i \in \{0,1,2,..,H-1\}$ 和$j \in \{0,1,2,..,W-1\}$。

$$T_a(iW+j,k,c) = F_a(i,j,c), \quad \forall k \in [0, HW-1].$$

我们以类似的方式计算 T_p 和 T_n，但还会在最后转置张量的每个通道。这等价于：

$$T_{p/n}(k, iW+j, c) = F_{p/n}(i,j,c), \quad \forall k \in [0, HW-1].$$

T_a 和 $T_{p/n}$ 使我们能够计算所有成对差值：
$F_{\text{diff}} = |T_a - T_{p/n}|$。

差分层会生成一个特征体 F_{diff}，用于量化两个数字高程图之间的共享信息。F_{diff} 经过所提出的CNN 架构处理（详情请见 †项目页面），得到一个标量值，表示两个数字高程图之间的距离。较低的值表示检测到回环。我们使用基于三元组的对比损失来训练所提出的回环检测模块：

$$\mathcal{L}_{\text{triplet}} = \max\left(0, d\left(\widehat{\phi}_a, \phi_p\right) - d\left(\widehat{\phi}_a, \phi_n\right) + \xi\right), \quad (2)$$

其中，$\widehat{\phi}$ 表示 DEM 编码，ϕ 表示偏航对齐的 DEM 编码，d 表示由回环检测模块计算的两种编码之间的距离，ξ 表示三元组损失的边界值。请注意，在反向传播过程中，DEM 编码器同时接收来自自动编码器均方误差损失和上述三元组损失的梯度。而解码器仅使用均方误差损失进行训练。

E. 回环闭合Af
ter 检测

在回环检测过程中，我们估计 SE(3) 刚体变换，以对齐查询点云和检索到的点云。此过程称为回环闭合（或点云配准）。令 P_q 为查询点云，其位姿为 T_q^w，中心位于 O_q。令 P_r 为检索到的点云，中心位于原点 O_r，其位姿为 T_r^w。请参见图 2 最右侧的回环闭合模块。T_q^w 和 T_r^w（以蓝色虚线标示）均处于世界参考坐标系中，且未知。我们的目标是求出相对变换 T_r^q（以黄色实线标示），使 P_q和 P_r对齐。为估计相对 SE(3) 位姿，我们首先估计 P_q 和 P_r之间的相对 SE(2) 变换。之后，将其与最初估计的横滚和俯仰规范化结果结合起来，以获得 SE(3) 位姿估计。

为估计相对 SE(2) 变换，我们对查询和检索到的数字高程模型进行解码，以还原各自的编码。随后，使用 [35], [36] 获取查询与检索到的数字高程模型中的关键点及其对应关系。SE(2) 位姿通过以下优化问题求得

$$\arg\min_{\psi,\mathbf{t}_{cr}^{cq}} \| (\mathbf{R}(\psi)_{cr}^{cq} a_i + \mathbf{t}_{cr}^{cq}) - b_i \|^2. \quad (3)$$

其中，a_i 和 b_i 分别是查询 DEM 和目标 DEM 上的对应点。旋转矩阵 $\mathbf{R}(\psi)_{cr}^{cq} \in SO(2)$由偏航角 ψ 参数化，平移向量记为 $\mathbf{t}_{cr}^{cq} \in R^2$。利用网格分辨率 d_g 将平移向量缩放至度量尺度（参见第 III-A 节中的俯视图离散化）。优化中的偏航角ψ 使用 DYT 模块给出的偏航估计值进行初始化。令 $\mathbf{R}(\alpha_q, \beta_q)_{cq}^{q}$和 $\mathbf{R}(\alpha_r, \beta_r)_{cr}^{r}$（在图 2 的回环闭合模块中以粉色显示）分别表示将查询点云（P_q）和检索点云（P_r）对齐至各自滚转-俯仰补偿坐标系 O_q 和 O_r的旋转矩阵。为估计 SO(3) 旋转矩阵，我们将 $\mathbf{R}_{cq}^{q}$、$\mathbf{R}_{cr}^{cq}$ 和 $\mathbf{R}_{cr}^{q}$组合起来：$\mathbf{R}_r^q = \mathbf{R}_{cq}^q \mathbf{R}_{cr}^{cq} (\mathbf{R}_{cr}^r)^{-1}$。

我们通过结合$\mathbf{t}_r^q$ 与$\mathbf{t}_{cr}^{cq}$ 和 d_r –d_q来获得平移：其中，d_r 和 d_q 是激光雷达与估计地平面之间的距离（该地平面是通过 RANSAC 与地平面参数一起估计得到的）。$\mathbf{t}_r^q = [\mathbf{t}_{cr}^{cq}(0), \mathbf{t}_{cr}^{cq}(1)\ d_r - d_q]$。

IV. 数据集与实现细节

我们使用 PyTorch，并在单块 NVIDIA GeForce GTX 1080 GPU 上进行训练，批量大小为 12，采用 ADAM [37]作为优化器，训练 200 个周期，共计 8 小时。学习率初始化为 4×10^{-4}，每 50 个周期减半。将一个$50m\times50m$ 点云转换为线性缩放的数字高程模型表示，尺寸为 500×500 像素。式（2）中的三元组边界k设为 0.75。与 [5]–[7], [11], [12], 不同，我们不对输入点云进行任何数据增强。

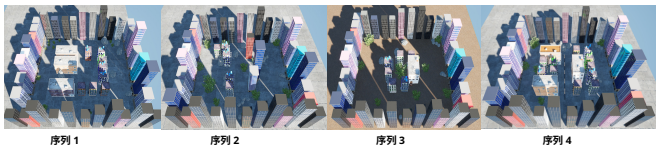


图5：我们创建的 Lidar-UrbanFly 数据集（LUF）环境一览。训练数据：序列（1、2、3）；测试数据：序列 4

为展示数字高程模型揭示底层空间结构的能力，并说明我们的方法对不同密度点云的泛化能力，我们选取了三个公开可用的激光雷达数据集 [22]–[24]。为进一步测试该方法对六自由度运动的鲁棒性，我们使用一架四旋翼无人机构建了一个合成数据集。

（1）KITTI [22]：该数据集包含 11 个序列，与[11] 类似，我们使用 05、06、07、09 进行训练，并使用 00 和 08 进行测试。（2）Oxford RobotCar [24]：该数据集共包含 44 个序列，我们按照 [7] 的建议划分训练集和测试集。（3）Lidar UrbanFly 数据集（LUF）：使用 Unreal Editor [38]，我们创建了一个包含建筑物、树木和不平整道路的自定义环境，用于评估回环检测方法。环境由安装在四旋翼无人机上的 64 通道激光雷达在六自由度运动中扫描得到。如图5所示，我们创建了四个这样的环境，序列（1、2、3）用于训练，序列 4 用于测试。（4）GPR[23]：该数据集共包含 15 个序列，我们使用序列 1、2、3、4、5、6、8、9、11、12 进行训练，并报告序列 10 和 15 上的评估结果。与 [11] 类似，当两个点云之间的真实距离小于 $4m$时，我们将其视为一个回环。此规则使我们能够为训练采样三元组：当两个位姿之间的距离小于$4m$ 时，形成锚点和正样本对；当距离介于 $4m$ 和 $10m$之间时，形成锚点和负样本对。所有与 Oxford Robot Car 相关的结果均可在我们的 †项目页面上查看。

V. 实验与结果

本节通过定量和定性分析展示了我们处理流程的多项功能。具体而言，我们展示了该方法在压缩点云空间中执行六自由度LDC 的高效性。我们通过具有挑战性的场景测试了算法的稳健性，例如视点发生六自由度变化时的回环检测，以及没有任何初始猜测时的回环闭合。此外，我们测量了通过数字高程模型编码过程节省的带宽，并将处理流程与*LIO-SAM* [39] 集成以测量其效果。最后，我们进行了详细的 †消融研究，以测试各个组件的效果。更多信息和演示请访问我们的 †项目页面。*A.* 回环检测结果

我们将*LCDNet* [11]作为基准进行比较。*PointNetVLAD* [3]、*PCAN* [5]和*OverlapNet* [12]是用于回环检测的最先进深度学习方法，但它们对大幅视点变化缺乏鲁棒性。此外，这些方法在训练时使用了大量数据增强，从而限制了泛化能力。我们使用各自作者发布的官方代码和预训练模型。为确保比较公平，对于没有可用预训练模型的数据集，我们会重新训练模型，即重新训练后的模型为 [5]–[7],

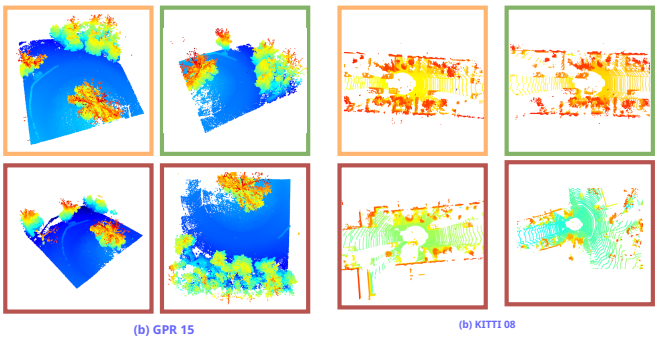


图 6：该图展示了我们的方法在多个序列上的召回率。对于四个序列中的每一个，橙色框（左上）中的点云是查询点云，绿色框（右上）中的点云是排名第一的检索结果。红色框（第二行）中的点云是排名第二和第三的检索结果（从左到右）。该图展示了我们的网络学习空间先验的能力。请注意，所有情况下排名第一的检索结果均正确。

[11], [12]用于*GPR*， [5]–[7] 用于*Kitti*， [11]用于*Oxford RobotCar*；我们还会针对*LUF* 环境重新训练所有模型。此外，我们还将 ScanContext [15]，这一基于手工设计特征的 LDC 方法纳入基准测试。

我们采用精确率-召回率（*PR*）曲线，这是一种评估回环检测的有效指标 [11], [12]。为计算平均精度（*AP*），我们遵循协议*2*，该协议由 [11],提出，并已被证明是一种有效的基准评估指标。以下简要介绍该流程：给定一个查询点云A，我们将其与数据库中的所有点云B进行比较。对于每一对扫描（A，B_i），我们使用第III-D节所述的方法计算它们之间的距离（距离越小，表示相似度越高）。若距离小于固定阈值，则将其视为回环。随后，我们检查激光雷达扫描的真实位姿之间的欧氏距离；若位姿间距小于 $4m$，则视为真正例；若距离大于 $4m$，则视为假正例。通过改变该固定阈值，我们得到多个精确率-召回率值，并据此绘制曲线。

结果见表 II，相应的 PR 曲线如图 7 所示。LCDNet [11] 是 KITTI 数据集上最先进的回环检测方法。在包含相反视角这一挑战性序列 KITTI08 上，我们的表现比 LCDNet 高约 10%。在 KITTI00 序列上，我们的表现位居第二，仅比最先进水平低约 1%。同样，在 GPR 序列（10、15）上，我们的方法均取得最高平均精度，分别比最接近的 LCDNet 高约 11% 和 4.5%。在包含六自由度视角变化的 LUF 数据集上，我们的方法比 LCDNet 高 10%。

在图 6中，我们展示了针对特定查询，由 DEM 召回率最高的 3 个点云，以展示其学习点云底层空间结构的能力。这体现了该网络识别相似空间结构的能力。例如，在 *GPR15* 序列中，输入点云和召回的点云都包含树木等共同结构元素，这表明 DEM已揭示点云的底层几何结构。

方法	KITTI		GPR		*LUF*
	KITTI-00	KITTI-08	GPR-10	GPR-15	Seq-4
LCDNet [11]	**0.89**	0.76	0.82	0.88	0.69
OverlapNet [12]	0.61	0.22	0.75	0.57	NA
PointNetVLAD [7]	0.40	0.39	0.50	0.54	0.67
PCAN [5]	0.46	0.20	0.39	0.20	0.58
ScanContext [15]	0.49	0.20	0.66	0.62	NA
SOE-Net [6]	0.52	0.47	0.74	0.73	0.60
Ours	0.88	**0.84**	**0.91**	**0.92**	**0.76**

表 II：回环检测的平均精度对比。NA：由于 [12] 和 [15] 仅适用于SE(2)运动，因此不适用。

B. 回环闭合评估

本节将所提出的点云配准方法与以下三类算法进行比较：**（1）回环检测与仅偏航估计**方法，例如 ScanContext [15]，和OverlapNet [12]，它们仅估计偏航，而不估计完整的六自由度位姿。**（2）回环检测与六自由度 位姿估计：**LCDNet [11] 是六自由度激光雷达描述符计算任务中的最先进方法，因此我们选择将其作为对比方法。**（3）仅六自由度 相对位姿估计**：我们将该方法与最先进的点云配准技术 TEASER++ [10] 进行比较（该技术不执行回环检测）。我们还将其与经典方法迭代最近点（ICP） [9]进行基准比较。

使用作者开源的官方代码进行基准测试 [10]–[12], [15]，并使用*Open3d* [40]实现迭代最近点。实验结果如表III所示。我们根据平均平移误差（*ATE*）（单位：米）和平均旋转误差（*ARE*）（单位：度）评估我们的方法。我们的方法在*KITTI*数据集（两个序列）、*GPR10*序列和*LUF*数据集上的平均平移误差最低；在*KITTI08*、*GPR10*、*LUF*和*GPR15*上的旋转误差也最低。在*KITTI00*序列上，*FinderNet*的误差比*LCDNet* 高 0.91^0。值得注意的是，诸如 [9]–[11] 之类的方法会处理整个点云；由于闭环检测需要传输大量点云数据，因此不适用于多智能体场景。为改进结果，我们可以采用抗离群点鲁棒ICP公式

方法	KITTI		GPR		LUF
	KITTI-00	KITTI-08	GPR-10	GPR-15	Seq-4
LCDNet [11]	0.77/**1.07**	1.62/3.13	1.44/1.14	**0.50**/4.81	1.82/38.86
OverlapNet* [12]	−/3.6	−/65.29	−/7.85	−/6.25	-
TEASER++ [10]	2.93/16.13	3.24/28.98	2.68/16.87	2.47/20.34	2.05/44.12
ScanContext* [15]	−/1.89	−/3.20	−/4.37	−/4.26	-
迭代最近点 [9]	2.23/9.12	2.31/161.16	2.32/7.81	2.87/8.36	2.15/85.24
Ours	**0.72**/1.98	**1.35**/**2.96**	**0.95**/**0.85**	0.82/**1.14**	**1.78**/**35.15**

表 III：与最先进水平相比的点云配准。结果格式：平移误差（米）/旋转误差（度）。“*”表示仅估计偏航、无法与我们的六自由度方法直接比较的算法。“-”表示不适用，因为 [12]和 [15] 仅适用于 SE(2)运动。

与 LIO-SAM 集成：我们通过将处理流程集成到先进的激光雷达 惯性 同步定位与建图 方法 LIO-SAM 中，测试了处理流程的有效性。LIO-SAM 使用基于因子图的后端优化，已被证明能够提供可靠的状态估计。结果如图 8 所示，与*FinderNet* 相比，使用FinderNet 使均方根误差 降低了*LIO-SAM*的*L2* 基于距离的回环检测 [39]16%。该实验进一步证明了我们的系统可以实时运行。有关实现细节和运行时统计数据，请参阅 †项目页面 。

消融研究：为了评估*RP*规范化

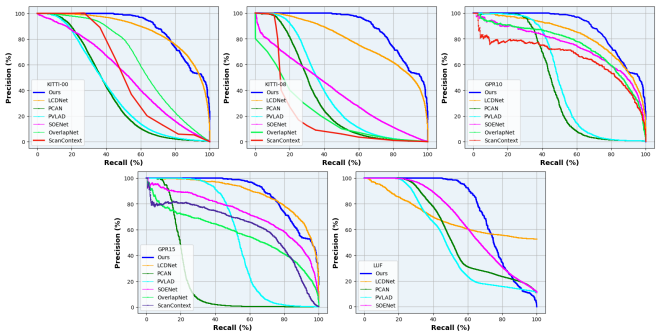


图 7：在 KITTI 和 GPR 数据集以及我们自建的 LUF 数据集上，对比多种回环检测算法。

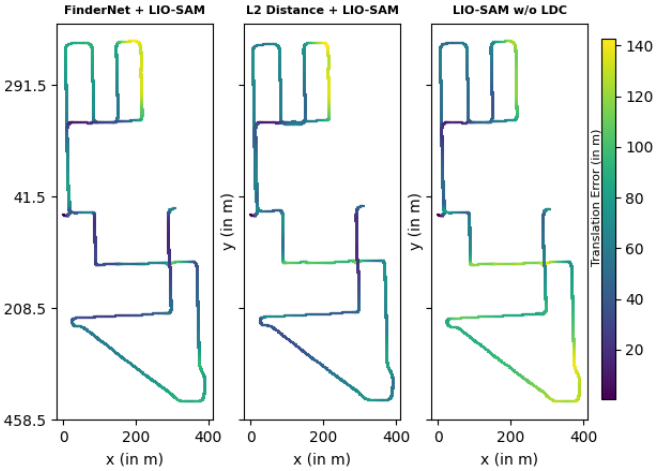


图 8：该图展示了将多种回环检测方法与均方根误差 集成后，在*LIO-SAM* [39] 的 KITTI-08 序列上得到的均方根误差。未使用回环检测的完整轨迹（右图）的均方根误差 为 $48.79m$；使用基于欧氏距离的回环检测（中间图中的完整*LIO-SAM* ）时，观测到均方根误差 为 $35.69m$ 。然而，集成*FinderNet* 后（左图），均方根误差 降至 $29.96m$。

器和DYT的性能，我们进行了消融研究。使用RP规范化器带来了两项关键优势。首先，如LDC 在LUF数据集上的结果（表II和表III）所示，它使我们的方法能够处理处于六自由度运动中的点云。其次，若不使用规范化器，我们就必须依赖诸如 [9], [10]这样的方法来估计六自由度相对位姿。然而，从表III中可以看到，我们的回环闭合处理流程比 [9], [10]具有更高的准确性。此外，我们还单独研究了RP规范化器的性能，并记录了粗RP规范化器的误差为3.249/4.1582（R/P）度，而精细对齐的误差为1.2598/1.352（R/P）度。

为了分析 DYT 的重要性，我们用空间变换器 [34], 替换了它，该变换器用于预测仿射变换的参数。然而，我们发现这种网络的性能较差，在*KITTI 00* [22]，*LUF*和*GPR* [23] 数据集上的平均精度 分别为 0.0551、0.0136 和 0.0818。DYT 能够以 3.12 度的误差估计偏航角。

最后，我们研究了带宽需求，并报告了约830 倍的压缩率，相较于原始点云的大小。我们在 †项目页面上提供了实验流程，以及与完整消融研究相关的其他结果。

VI. 结论与未来工作

我们开发了提出一种新颖的压缩与表示 LiDAR 数据的方法。我们的方法基于所提出的 Matlab 表示，我们的方法利用规范化和 DYT，在不进行数据增强的情况下，实现大旋转角度下的视点不变性。此外，与以往研究不同，我们的方法在训练时不需要数据增强。与最先进水平相比，我们的方法在真实世界和模拟数据集上均表现出显著提升。未来，我们希望在动 y 动态场景。

参考文献

[1] P.-Y. Lajoie、B. Ramtoula、Y. Chang、L. Carlone 和 G. Beltrame，“Door-slam：面向机器人团队的分布式、在线且具备离群点鲁棒性的同步定位与建图”，IEEE《机器人与自动化快报》，第 5 卷，第 2 期，第 1656–1663, 2020 1 页。

[2] P.-Y. Lajoie、B. Ramtoula、F. Wu 和 G. Beltrame，“迈向协作式同步定位与建图：当前研究领域综述”，arXiv 预印本 arXiv: *2108 08325*, 2021 1 。T. Cieslewski 和 D.

[3] Scaramuzza，“使用分布式倒排索引实现高效的去中心化视觉地点识别”，IEEE《机器人与自动化快报》，第 2, 2, 640 647, 2017 1 卷，第

[4] 2, 2, 640 647, 2017 1 期，第–页。——，“利用全图像描述符实现高效的去中心化视觉地点识别”，载于 2017 年国际多机器人与多智能体系统研讨会（MRS）。IEEE, 2017, 78 82 1 第–页。W. Zhang 和 C. Xiao，“Pcan：基于上下文信息的 3d 注意力图学习，用于基于点云的检索”，载于 IEEE/CVF计算机视觉与模式识别会议论文集，

[5] 2019, 12 436 12 445 1, 2, 6, 7 第–页。Y. Xia、Y. Xu、S. Li、R. Wang、J. Du、D. Cremers 和 U. Stilla，“Soe-net：用于点云地点识别的自注意力与方向编码网络”，载于 IEEE/CVF计算机视觉与模式识别会议论文集，2021 年，第 11 348–11 357 页。

[6] 1, 2, 5, 6, 7

[7] M. A. Uy 和 G. H. Lee，“Pointnetvlad：用于大规模地点识别的深度点云检索”，载于 IEEE计算机视觉与模式识别会议论文集，2018 年，第 4470 4479 1, 2, 5, 6, 7 –页。L. Wiesmann、R. Marcuzzi、C. Stachniss 和 J.

[8] Behley，“Retriever：压缩三维地图中的点云检索”，载于 IEEE 国际机器人与自动化会议（ICRA）论文集, 2022 1, 2 。Z.

[9] Zhang，“用于自由形状曲线与曲面配准的迭代点匹配”，《国际计算机视觉期刊》，第 13 卷，第 2 期，第 119 152, 1994 1, 2, 7 –页。H. Yang、J. Shi 和 L.

[10] Carlone，“TEASER：快速且可验证的点云配准”，《IEEE机器人汇刊》，2020 1, 2, 7 。D. Cattaneo、M. Vaghi

[11] 和 A. Valada，“Lcdnet：用于激光雷达同步定位与建图的深度闭环检测与点云配准”，《IEEE机器人汇刊》，2022 1, 2, 3, 4, 5, 6, 7 。X. Chen、T. Labe、A. Milioto、T. R¨ohling、O. Vysotska、A. Haag、J. Behley 和 C. Stachniss，“

[12] OverlapNet：面向激光雷达同步定位与建图的回环闭合”，载于《机器人学：科学与系统》（RSS）论文集，2020 1, 2, 3, 5, 6, 7 。L. Schaupp、M. Burki、R. Dub´e、R. Siegwart 和 C. Cadena，“Oreos：室外场景中三维点云的方向感知识别”，载于

[13] 2019 年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE, 2019, 3255 3261 1, 2, 3, 4 第–页。K.-L. Low，“点到平面ICP曲面配准的线性最小二乘优化”，教堂山，北卡罗来纳大学，第 4 卷，第 10 期，第 1–3 页，2004 年。

[14] [15] G. Kim 和 A. Kim，“Scan Context：三维点云地图中用于地点识别的自我中心空间描述符”，载于 2018 年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE, 2018, 4802 4809 2, 6, 7 第–页。L. Luo、S.-Y. Cao、B. Han、H.-L. Shen 和 J. Li，“Bvmatch：利用鸟瞰图像进行基于激光雷达的地点识别”，IEEE《机器人与自动化快报》，第

[16] 6, 3, 6076 6083, 2021 2 卷，第

[17] 6, 3, 6076 6083, 2021 2 期，第–页。X. Xu、H. Yin、Z. Chen、Y. Li、Y. Wang 和 R. Xiong，“Disco：带方向信息的可微扫描上下文”，IEEE《机器人与自动化快报》，第

[17] 6, 2, 2791 2798, 2021 2 卷，第

[18] 6, 2, 2791 2798, 2021 2 期，第–页。G. Kim、S. Choi 和 A. Kim，“Scan Context：在城市环境中对旋转与横向变化具有鲁棒性的结构化地点识别”，《IEEE机器人汇刊》，2021 年。2

[18] ++ ++

[19] C. R. Qi、H. Su、K. Mo 和 L. J. Guibas，“Pointnet：用于三维分类与分割的深度学习点集网络”，载于 IEEE计算机视觉与模式识别会议论文集，2017, 652 660 2 第–页。C. R. Qi、L. Yi、H. Su 和 L. J. Guibas，“Pointnet++:

[20] 度量空间中的点集深度层次特征学习”，《神经信息处理系统进展》，第 30, 2017 2 卷。J. Ma、J. Zhang、J.

[21] Xu、R. Ai、W. Gu 和 X. Chen，“Overlaptransformer：一种用于基于激光雷达地点识别的高效偏航角不变Transformer网络”，IEEE《机器人与自动化快报》，2022 2, 3 。A. Geiger、P. Lenz 和 R. Urtasun，“我们是否已经准备好实现自动驾

[22] 驶？KITTI视觉基准套件”，载于 2012 年 IEEE计算机视觉与模式识别会议。IEEE, 2012 年，第 3354–3361 2, 6, 7 页。P. Yin、S. Zhao、R. Ge、I. Cisneros、R. Fu、J. Zhang、H. Choset 和 S. Scherer，“Alita：用于长期自主运行的大规模增量数据集”，arXiv 预印本 arXiv: *2205 10737*, 2022 2, 6, 7 。

[23] W. Maddern、G. Pascoe、C. Linegar 和 P. Newman，“年行驶 1000 千米：牛津RobotCar数据集”，《国际机器人研究杂志》，第 36, 1, 3 15, 2017 2, 6 卷，第

[24] 36, 1, 3 15, 2017 2, 6 期，第–页。L. Wiesmann、T.

[25] Guadagnino、I. Vizzo、G. Grisetti、J. Behley 和 C. Stachniss，“Dcpcr：大规模室外环境中的深度压缩点云配准”，IEEE《机器人与自动化快报》，第 7, 3, 6327 6334, 2022 2 卷，第

[25] 7, 3, 6327 6334, 2022 2 期，第–页。L. Wiesmann、A.

[26] Milioto、X. Chen、C. Stachniss 和 J. Behley，“稠密点云地图的深度压缩”，IEEE《机器人与自动化快报》，第 6, 2, 2060 2067, 2021 2 卷，第

[26] 6, 2, 2060 2067, 2021 2 期，第–页。J. Du、R. Wang 和 D. Cremers，“Dh d：用于大规模六自由度重定位的深度层次三维描述符”，载于《欧洲计算机视觉会议》。Springer,

[27] 2020, 744 762 2, 3 第–页。Y.

[27] Wang、Z. Sun、C.-Z. Xu、S. E. Sarma、J. Yang 和 H. Kong，“Lidar Iris：用于闭环检测的激光雷达描述符”，载于 2020 年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE, 2020 年，第 5769 5775 2 –页。R. Arandjelovic、P. Gronat、A.

[28] Torii、T. Pajdla 和 J. Sivic，“Netvlad：用于弱监督地点识别的 CNN 架构”，载于IEEE计算机视觉与模式识别会议论文集，2016, 5297 5307 2 第–页。

[29] A. Jaegle、F. Gimeno、A. Brock、O. Vinyals、A. Zisserman 和 J. Carreira，“Perceiver：利用迭代注意力实现通用感知”，载于国际机器学习会议。PMLR，2021 年，第 4651–4664 页。2[31] I. Bogoslavskyi 和 C. Stachniss，“面向在线运行的稀疏三维激光扫描快速距离图像分割”，载于 2016 年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE, 2016, 163 169 3 第–页。K. S. Tai、P. Bailis 和 G. Valiant，“等

[30] 变Transformer网络”，载于国际机器学习会议。PMLR，2019 年，第 6086–6095 页。4[33] C. Esteves、C. Allen-Blanchette、X. Zhou 和 K. Daniilidis，“极坐标Transformer网络”，arXiv 预印本 arXiv: *1709 01889*, 2017 4 。M. Jaderberg、K. Simonyan、A. Zisserman 等，5, 7

[32] “空间Transformer网络”，《神经信息处理系统进展》，第 28 卷，2015 年。[35] D. DeTone、T. Malisiewicz 和 A. Rabinovich，“Superpoint：自监督兴趣点检测与描述”，载于 IEEE计算机视觉与模式识别会议研讨会论文集，2018, 224 236 5 第–页。P.-E. Sarlin、D. DeTone、T. Malisiewicz 和 A. Rabinovich，“Superglue：利用图神经网络学习特征匹配”，载于IEEE/CVF计算机视觉与模式识别会议论文集，2020, 4938 4947 5 第–页。D. P. Kingma 和 J. Ba，“

[34] Adam：一种随机优化方法”，arXiv 预印本 arXiv: *1412 6980*, 2014 5 。Epic Games，“虚幻引擎”。在线。可用：https://www.unrealengine.com [39] T. Shan、B. Englot、D. Meyers、W. Wang、C. Ratti 和 D. Rus，“Lio-sam：通过平滑与建图实现紧耦合激光雷达惯性里程计”，载于 2020 年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE, 2020, 5135 5142 6, 7 第–页。Q.-Y.

[35] [ ] 6

[36] [37] [38] Zhou、J. Park 和 V. Koltun，“Open d：用于三维数据处理的现代库”，arXiv 预印本 arXiv:1801.09847，2018 年。7

[40] 3