





P2P-Bridge:3D 点云的扩散桥

去噪

Mathias Vogel<sup>1</sup>,  馆野圭介<sup>2</sup>, 马克·波勒费斯<sup>1,3</sup>,  费德里科·汤巴里<sup>2,4</sup> ,  
玛丽·朱莉·拉科托萨纳\*,<sup>2</sup>, 和 Francis Engelmann<sup>1,2</sup> 

<sup>1</sup> 苏黎世联邦理工学院    <sup>2</sup> 谷歌    <sup>3</sup> 微软    <sup>4</sup> 慕尼黑工业大学

摘要。在这项工作中,我们解决了点云去噪的任务  
使用一种新颖的框架,将扩散薛定谔桥应用于点集等非结构化数据。与之前从点特征或  
学习到的噪声分布预测点位移的工作不同,我们的

方法学习成对点云之间的最优传输计划。在  
在 PU-Net 数据集和真实世界数据集等对象数据集上进行实验  
在 ScanNet++ 和 ARKitScenes 等数据集上,P2P-Bridge 的表现比现有方法好不到哪里去。  
尽管我们的方法证明了  
仅利用点坐标就能得到令人满意的结果,我们证明了  
结合 RGB 信息和逐点等附加功能  
DINOv2 功能进一步改善了结果。代码和预训练  
网络可在<https://github.com/matvogel/P2P-Bridge> 上找到。

关键词:点云去噪、扩散模型、薛定谔桥

1 简介

使用点云来表示 3D 物体和场景[9,24,54]非常普遍  
涉及各个领域,包括 3D 视觉、机器人技术[28, 75]、增强/虚拟  
现实和自动驾驶[26, 59]。最近,3D 扫描设备如  
激光雷达传感器已经越来越受欢迎,并被纳入现成的消费产品中。使用这些手持设备,用户  
可以扫描物体  
或场地。然而,由此产生的点  
由于硬件限制,例如低  
扫描仪分辨率、传感器噪声、有限范围或环境因素,例如  
反射、散射或遮挡可能会对  
依赖于高质量点云的下游任务[14, 53, 62]。为了解决  
针对这一问题,点云去噪已成为一种关键技术,可以减少  
消除扫描物体中的噪点并增强几何细节。  
虽然点云去噪研究取得了显着进展,但  
清理被现实世界的扫描仪噪音破坏的扫描仍然具有挑战性,因为  
需要掌握底层拓扑结构和底层性质  
清洁的表面以及噪音的特点。虽然传统的

<sup>\*</sup> 平等贡献。

2 M.沃格尔等人。

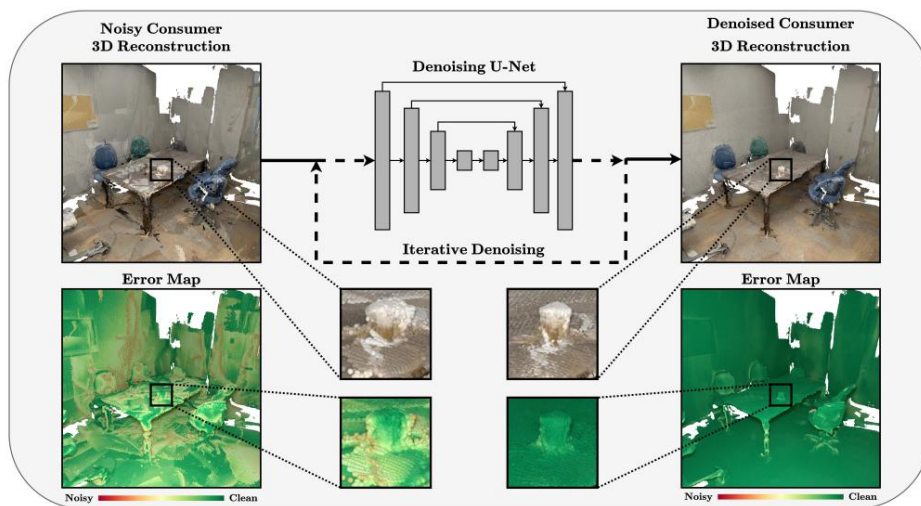


图 1:P2P-Bridge 应用于嘈杂的 LIDAR 扫描的示意图。

点云去噪方法[2, 10, 17, 18, 31, 57, 66, 70, 71]在特定情况下可能表现良好,但它们通常需要对参数或附加点特征 (如法线)进行大量微调,并且往往无法推广到复杂的噪声模式。

由于采用数据驱动方法,深度学习方法[34,36,38,45,68]表现出优于传统方法的性能。一类深度学习方法[19,34]通过首先将点云重新采样为一组粗略的点来解决去噪问题,从而有可能消除高频噪声。它们通过上采样和细化点云来恢复底层干净的面。其他方法[13,45,66]尝试通过回归或逐点位移预测来恢复干净的数据,而 PointCleanNet [45]也结合了异常值去除。

最近,基于分数的[36,68]和基于流的[38]模型通过直接学习噪声分布的分数或概率显示出令人兴奋的结果。

然而,当前处理点云噪声的模型都是在合成噪声 (如各向同性高斯噪声)的假设下训练的。我们的实验表明,这种假设通常不足以对从现成设备获得的真实世界 3D 扫描进行去噪,因为它忽略了异常值聚类、鬼点或边缘眩光等影响[2]。此外,以前的方法通常是在最小化距离度量上进行训练的,而这些距离度量与输入大小的线性关系较差[46],从而抑制了模型架构的扩展,而模型架构的扩展在点特征学习中起着至关重要的作用[44]。最后,最近的模型通常专注于仅从基于点的特征 (如颜色或法线)学习去噪任务,而这些特征没有考虑到底层数据的高级语义属性[41, 61]。

我们建议通过结合从 DINOv2 [41]中提取的逐点特征来利用高级学习特征。

本研究提出了一种基于扩散模型的新型点云去噪监督方法[8, 20]。我们将去噪任务视为薛定谔桥问题[8, 48],并通过训练网络在嘈杂的点云和相应的干净点云之间找到最佳传输方案来解决它,从而使我们的方法可以在任何数据上进行训练,例如室内场景(参见

图1)。我们结合 RGB 和 DINOv2 [41]特征来进一步改进我们的方法。实验表明,我们的方法在合成和真实场景中的表现优于其他最先进的方法。

总而言之,我们的主要贡献是:

- 1.我们提出了 P2P-Bridge (点云到点云桥),这是一种新的点云去噪方法,受薛定谔桥问题启发,将点云去噪表述为可处理的数据到数据扩散过程。
- 2.此外,我们提倡通过结合高级特征(例如 DINOv2)来进行语义信息去噪,以指导去噪过程。

## 2 相关作品

传统的去噪方法大致可分为基于滤波器的方法和基于优化的方法。基于滤波器的方法源自图像和信号处理,假设干净的点云受到高频噪声的破坏。双边滤波器方法[2, 18, 67]适用于去除物体表面的噪声,同时保留尖锐的边缘。基于滤波器的引导方法[17, 31, 57, 70, 71]尝试将局部线性模型拟合到嘈杂的点云中,旨在通过使用点坐标或法线的引导来保留局部细节。

基于图的方法[12, 23, 64]将点云建模为图,以捕捉底层几何结构和点之间的关系。

基于优化的方法包括稀疏重建[11, 27, 39]和基于非局部的点云去噪方法[3, 33, 64]。然而,所有这些传统方法都依赖于手动调整的超参数,这些超参数很难获得,而且通常泛化性不佳。

基于深度学习的方法最近显示出了良好的效果,并且比传统的去噪方法有所改进。PointCleanNet [45]首先去除异常值,然后预测逐点位移矢量以对点云进行去噪。TotalDenoising [19]是一种无监督方法,它将总体最小二乘 [16] 应用于点云等非结构化数据。DMRDenoise [34]使用差分池化的下采样来估计流形,然后从中对点进行重新采样以获得去噪的点云。基于分数匹配的方法[52]学习可处理的噪声分布的分数函数,并使用(动量)梯度上升[36, 68]来预测推理过程中的局部位移。PD-Flow [38]利用正则化流通过在潜在空间中噪声与干净的点云分离来直接估计噪声概率密度函数。I-PFN [49]改进了迭代去噪方法,在训练过程中为每次去噪迭代使用分离的迭代模块。

除了 PD-Flow 之外,所有提到的方法都是在高斯噪声假设下进行训练的,因为很容易生成干净点和噪声点的训练对

4 M.沃格尔等人。

云。然而,正如我们将通过实验证明的那样,这并不一定会转化为现实世界室内场景中的复杂噪声。我们的方法与大多数以前的研究的主要区别在于,我们的方法可以应用于任何一般的数据到数据问题。通过学习特定于数据的噪声特性,我们的方法可以更好地恢复底层的干净数据,去除异常聚类并恢复精细细节。最后,我们的方法使用 DDPM 采样[20] 进行去噪,使其对消融研究显示的去噪步骤数更具鲁棒性。

我们的方法仅需三次函数评估就能显示出良好的效果。

3D 重建涉及使用 2D 图像和深度等附加数据创建真实世界场景的三维表示。它不同于 3D 点云去噪,但可以作为从真实世界场景生成 3D 点云的初始步骤。因此,我们将讨论该领域的一些相关工作。3DMatch [63]是一种使用学习到的体积特征匹配 RGB-D重建的数据驱动方法。RoutedFusion [56]和 Map-Adapt [69]引入了基于机器学习的实时深度图融合方法,并使用神经网络预测基于体素的融合的非线性更新,解决了常见的错误和伪影,尤其是对于薄物体和边缘。NICE-SLAM [74]是一种基于分层网格的 SLAM 方法,使用 RGB-D 数据进行精确的环境重建。它结合了预先训练的模型来增强空间理解,提高了映射和跟踪效率。

NICER-SLAM [73]使用 RGB 输入数据来优化端到端联合映射和跟踪系统,使其能够预测颜色、深度和法线。扭曲和光流损失等额外损失进一步增强了NICER-SLAM 的几何一致性。

### 3 方法

#### 3.1 概述

假设噪声点集  $P = \{x_i\} \in \mathbb{R}^M$  和干净点集  $P = \{x_i\} \in \mathbb{R}^M \times D$  的分布,其中  $D$  是点特征维度,  $N$  和  $M$  分别是噪声点云和干净点云中的点数。我们旨在使用扩散模型[20,35,51]去除噪声点集  $P$  的噪声,并利用其从噪声先验中预测干净数据的性质。扩散模型已成功应用于许多图像生成和翻译任务[30, 40, 47],最近也用于点云生成和完成[25, 37, 55, 65, 65, 72]。

虽然大多数基于扩散的方法以及点云去噪的相关工作[36,45,68]都使用高斯先验,但我们认为,采用数据到数据而不是数据到噪声的方法更适合点云去噪,尤其是在处理特定传感器数据时。使用噪声点集  $P$  的分布作为先验分布使我们的方法能够学习依赖于数据的真实世界噪声特性。然而,对于训练扩散模型,将干净样本扩散到噪声样本的过程通常必须是可处理的,因为扩散模型是经过训练来学习逐步消除噪声的。对于真实世界

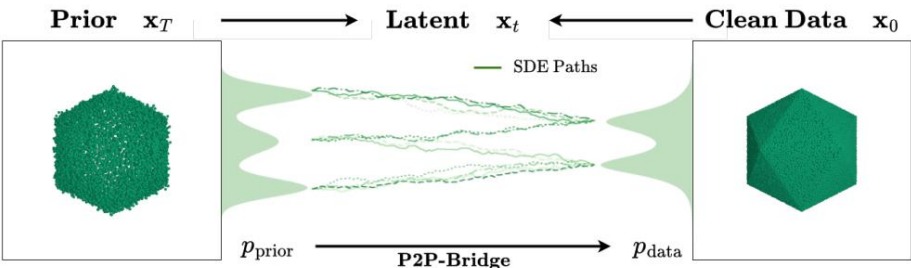


图 2:P2P-Bridge 示意图,将点云去噪建模为反向数据到数据扩散过程。我们的模型可以通过学习干净数据和嘈杂数据之间的桥梁,有效地将嘈杂数据转换为更干净的数据。

数据,无论如何,这个过程都是未知的。模拟干净和嘈杂样本之间的扩散过程的一种方法是使用薛定谔桥 (SB) 。薛定谔桥越来越多地用于图像到图像的翻译[4,8,30]、蛋白质匹配[50]或最近的文本到语音[6]的生成模型中。据我们所知,我们是第一个使用这种方法进行点云去噪的人。

3.2 点云到点云桥

可处理的扩散桥。我们从扩散模型的角度考虑点云去噪问题[20]。通过在  $T$  个时间步长  $\{x_1, \dots, x_T\}$  上生成一个扩散过程,其中  $x_t \in \mathbb{R}^3$ ,来自干净数据  $x_0 \sim p_{\text{data}}$  的样本会扩散到噪声样本  $x_T \sim p_{\text{prior}}$  中。在点云去噪的情况下,先验分布对应于噪声点集  $P$  上的分布 (参见图2) 。考虑描述此过程的参考路径测度  $p_{\text{ref}}(x_0 : T)$ ,我们的目标是找到一个过程  $p(x_0 : T)$ ,使得  $p(x_0) = p_{\text{data}}$  并且

页  $p(x_T) = p_{\text{prior}}$ 最小化  $p_{\text{ref}}$ 和  $p$ 之间的 Kullback-Leibler 散度这个问题也称为薛定谔桥 (SB) 问题 [29, 48]。它可以用正向和反向随机微分方程(SDE) 来描述,定义为

(1)

其中  $W_t$ 是维纳过程,  $f$  是矢量值函数,称为漂移,  $g$  是标量值项,称为扩散系数。项  $\nabla \log \Psi_t(x_t)$ 和  $\nabla \log \Psi_t(x_t)$ 是附加非线性漂移项,用于求解以下耦合偏微分方程 (PDE)

(2)

6 M.沃格尔等人。

使得  $\Psi_0 \Psi_0 = \text{pdata}$ ,  $\Psi^T \Psi = \text{pprior}$ 。Chen 等人 [4] 表明, 等式 (1) 是基于分数的生成模型 (SGM) [51] 到非线性过程的推广。直接求解微分方程组并不切实际, 而且计算成本高昂。然而, 最近的研究 [6, 30] 在处理成对边界数据的假设下引入了简化的易处理框架, 即  $p(x_0, x_T) = \text{pdata}(x_0) \text{pprior}(x_T | x_0)$ 。在点云背景下, 这意味着噪声点云上的分布被建模为干净点集 ( $\text{pdata}(x_0)$ ) 和噪声 ( $\text{pprior}(x_T | x_0)$ ) 的联合分布。

当边界数据为 Diracs ( $\delta x_0, \delta x_T$ ) 混合, 且  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  和  $(t)$  时, 可以证明 [6] 使用线性扩散方案  $g$  等式 (1) 的解析形式为

$$\frac{1}{\sigma^2} \left( \frac{1}{\sigma^2} \int_0^t \sigma^2(\tau) d\tau + \sigma^2(t) \right) \exp \left( -\frac{1}{2\sigma^2(t)} \left\| x_T - x_0 \right\|^2 \right) \quad (3)$$

和

$$\frac{1}{\sigma^2} \left( \frac{1}{\sigma^2} \int_0^t \sigma^2(\tau) d\tau + \sigma^2(t) \right) \exp \left( -\frac{1}{2\sigma^2(t)} \left\| x_T - x_0 \right\|^2 \right) \quad (4)$$

其中  $\sigma^2 = \int_0^t \sigma^2(\tau) d\tau$  和  $\sigma^2(t) = \int_0^t \sigma^2(\tau) d\tau$ 。这简化了方程 (2) 并使其完全易于处理。我们可以参数化一个网络  $\theta$ , 该网络预测在时间步  $t$  添加到  $x_0$  的噪声, 从而使用噪声预测损失得出噪声样本  $x_t$  :

$$\frac{1}{\sigma^2} \left( \frac{1}{\sigma^2} \int_0^t \sigma^2(\tau) d\tau + \sigma^2(t) \right) \exp \left( -\frac{1}{2\sigma^2(t)} \left\| x_T - x_0 \right\|^2 \right) \quad (5)$$

在推理过程中, 我们可以使用 DDPM 采样进行迭代采样 [20]

$$\frac{1}{\sigma^2} \left( \frac{1}{\sigma^2} \int_0^t \sigma^2(\tau) d\tau + \sigma^2(t) \right) \exp \left( -\frac{1}{2\sigma^2(t)} \left\| x_T - x_0 \right\|^2 \right) \quad (6)$$

因为只要  $x_0$  接近实际  $x_0$ , 就会产生相同的 SB 路径边际密度 [6, 30]。但是, 从  $\mu_t$  采样描述了两个点云之间的插值。此操作对于具有固定域的数据 (例如图像数据) 来说很简单, 其中像素附加到静态网格。然而, 对于无序点云, 它定义不明确, 并且依赖于适当的距离度量 [5]。

无序点集之间的有意义插值。我们使用 PointMixup [5] 中的最短路径插值方法来描述后验均值  $\mu_t$  的路径。最短路径插值试图找到一个最佳分配, 使点云  $x_T$  中每个点到  $x_0$  中最近邻点的平均距离最小化。假设噪声点集和干净点集都包含  $N$  个点, 则分配问题定义为

$$\frac{1}{\sigma^2} \left( \frac{1}{\sigma^2} \int_0^t \sigma^2(\tau) d\tau + \sigma^2(t) \right) \exp \left( -\frac{1}{2\sigma^2(t)} \left\| x_T - x_0 \right\|^2 \right) \quad (7)$$

其中  $\Phi = \{1, \dots, N\} \rightarrow \{1, \dots, N\}$  是  $x_T$  和  $x_0$  中点之间可能的双射分配集。使用最短路径插值类似于在成本为平方时在两个点集之间寻找最佳运输计划

测地线距离[42]。最短路径插值所得路径对应于当桥梁随机性消失时方程(1)后验所取的路径,即 $g(t) \rightarrow 0$  [6, 30, 42],这促使我们选择最近路径插值而不是其他可能的插值方法。降低桥梁的随机性可以有效地将桥梁 SDE 简化为形式为最优传输常微分方程 (OT-ODE)

(8)

实际上,我们只需为数据集中的每个数据对计算一次最佳分配。随后,我们可以使用重新排序干净的点云,以便它们与相应的噪声点云对齐。在训练期间,我们可以对 $x_t$ 进行采样而无需再次解决最佳分配问题,从而实现快速且可扩展的训练。有关最短路径插值的进一步讨论和实验可以在附录中找到。

3.3 实施

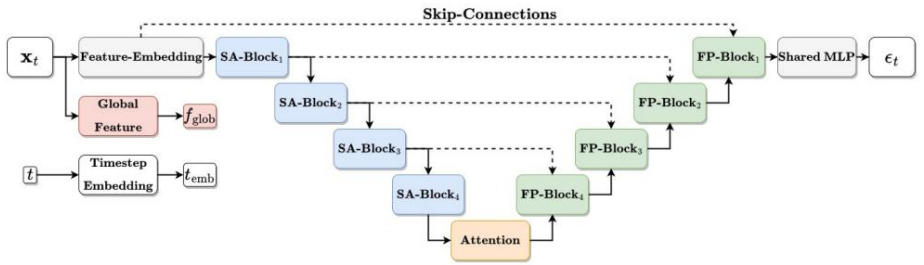


图 3:基于 PointVoxelConvolutions (PVC) [32] 的网络架构。我们改编了 LION [65] 的网络实现,并为其添加了多头全局注意和特征嵌入模块。特征嵌入和最终的共享 MLP 块均使用  $1 \times 1$  卷积实现。

模型架构。我们遵循先前关于点云扩散模型的研究[37,65,72],并使用基于 PointVoxel-CNN (PVCNN) [32] 的模型架构 (参见图3)。PVCNN 是一种 PointNet++ [43] 启发的架构,它使用从发声点云表示中提取的全局特征增强了集合抽象 (SA) 和特征传播 (FP) 块。与 LION [65] 类似,我们结合了挤压激励块[22]和全局特征提取网络。此外,我们使用特征嵌入层,使用  $1 \times 1$  卷积将传入特征映射到更高维度。我们遵循最近关于图像扩散模型的研究[21],并且仅在最低层使用多个注意力来利用注意力。网络以时间步长 $t$ 为条件



8 M.沃格尔等人。

使用正弦位置嵌入和点特征。通过自适应组规范化整合全局特征。输入数据以块的形式处理，从数据集相关的半径球体中采样点，以确保空间中噪声数据和干净数据之间的对应性。

输入特征。手机扫描通常可以获得额外的 RGB 数据。我们提出使用 DINOv2 [41]从原始RGB 特征中提取逐点特征。使用相机姿态和内在函数将逐像素的 DINOv2 特征投影到嘈杂的点云中，从而得到逐点特征。

## 4 实验

### 4.1 数据集

为了进行评估，我们将我们的 P2P-Bridge 与成熟的 3D 对象数据集（描绘单个对象）和室内场景数据集进行比较。对于对象级去噪，我们使用 PU-Net 数据集[60]，其中包含 40 个用于训练的对象和 20 个用于评估的对象。我们使用 PC-Net 数据集[45]提供另外 10 个对象仅用于模型评估。对于这两个对象数据集，我们遵循常用的做法并使用各向同性高斯模拟噪声。

与之前的工作不同，我们建议额外对场景级点云进行评估。此设置更接近现实世界的使用场景。为了在现实世界数据集上进行评估，我们选择了室内场景数据集 ScanNet++ [58]和ARKitScenes [1]，因为它们提供了成对的干净和嘈杂点云数据以及姿势数据。ScanNet++ 包含 330 个室内场景，其中每个场景都有一系列由手持式 LIDAR 扫描仪获得的嘈杂深度图和由 Faro 激光扫描仪获得的干净扫描。我们使用ScanNet++ 作者提供的 3D 重建脚本来构建嘈杂点云，然后在其上应用去噪方法。此外，我们还评估了何时使用 3DMatch [63]细化重建。ARKitScenes包含 5047 个不同室内场所的扫描，其中嘈杂扫描是使用 Apple ARKit 表面重建获得的。干净扫描由 Faro 激光扫描仪获取。

更多详细信息请参阅补充材料。

### 4.2 评估指标

我们使用倒角距离 (CD) 和点到网格 (P2M) 距离作为定量评估指标。CD 测量预测点云  $P = \{x_i \in \mathbb{R}^3\}_{i=1}^N$  和干净点云  $P = \{x_j \in \mathbb{R}^3\}_{j=1}^M$

作为

$$\text{CD} = \frac{1}{N} \sum_{i=1}^N \min_{j \in \{1, \dots, M\}} \|x_i - x_j\| \quad (9)$$

其中  $NN$  是最近邻函数。第一项( $\rightarrow$ CD)大约为描述噪声点到地面真实表面的平均距离,并且



第二项 (←CD)鼓励均匀覆盖。点到网格 (P2M)

距离定义为

(10)

其中 $d(x, f)$ 是测量点 $x$ 到面 $f$ 的距离的函数。因此,第一项描述面到点的距离 (F2P),第二项对应于点到面的距离 (P2F)。为了计算对象级指标,我们按照 ScoreDenoise [36]将预测和真实值居中并缩放到单位球面。

4.3 实验细节

我们在 PU-Net 数据集上训练我们的模型,以去除带有高斯噪声的人工噪声对象,这遵循了以前的研究。对于 ScanNet++ 和 ARKitScenes,我们训练所有基于深度学习的去噪方法 (包括我们的方法),批处理大小为 32,最多 100,000 步。我们使用以前研究的公开代码库中提供的训练参数和模型权重[34,36,38,68]。补充材料中提供了其他实验细节。

4.4 对象比较

我们使用传统方法 (例如 Bilateral [10]或 GLR [64])以及基于深度学习的方法 (包括 PC-Net [45]、DMR [34]、ScoreDenoise [36]、MAG [68]和 PD-Flow [38])定量评估了我们的方法。对于评估,我们针对稀疏 (10k 点)和密集 (10k 点)物体选择高斯噪声水平,范围从物体边界框对角线的 1% 到 3%。表1显示,在大多数噪声设置中,我们的方法优于以前基于优化的方法和基于深度学习的方法。在 1% 的设置中,我们的方法表现仅次于 PD-Flow,而对于更高的噪声水平,我们看到测量精度与以前的方法相比显著提高。正如 PC-Net 数据集上的结果所示,我们的方法似乎也能更好地适应看不见的物体。请注意,这些结果仅通过三个去噪步骤即可实现。

图4定性比较了我们的方法和最新的基于深度学习的方法对受 3% 各向同性高斯噪声破坏的对象进行去噪的效果。与以前的研究相比,我们的方法似乎可以产生更少噪声和更平滑的结果。

补充材料中提供了使用不同噪声类型和运行时间的额外定性结果和实验。

4.5 室内场景对比

为了进一步研究我们的方法和以前方法的去噪能力,我们对应用于大规模室内场景重建的方法进行了评估。这种设置带来了额外的挑战和噪声源

10 M.沃格尔等人。

表 1:对象级分数。我们展示了倒角距离 (CD) 和点 2-PU-Net (顶部)和 PC-Net (底部)数据集上的网格距离 (P2M)。分数为乘以 104 。如果可能,基线分数取自[36]和[68],否则,我们使用公开的权重和测试脚本来评估测试数据由[36]提供。

点数	10 103 (稀疏)						50 103 (密集)					
高斯噪声	1%		2%		3%		1%		2%		3%	
方法	光盘	P2M	光盘	P2M	光盘	P2M	光盘	P2M	光盘	P2M	光盘	P2M
双侧[10]	3.65	1.34	5.01	2.02	7.00	3.56	0.88	0.23	2.38	1.39	6.30	4.73
PCNet [45]	3.52	1.15	7.47	3.97	13.1	8.74	1.05	0.35	1.45	0.61	2.29	1.29
DMRDenoise [34]	4.48	1.72	4.98	2.12	5.89	2.85	1.16	0.47	1.57	0.80	2.43	1.53
GLR [64]	2.96	1.05	3.77	1.31	4.91	2.11	0.70	0.16	1.59	0.83	3.84	2.71
降噪得分[36]	2.52	0.46	3.69	1.07	4.71	1.94	0.72	0.15	1.29	0.57	1.93	1.04
磁力[68]	2.50	0.46	3.63	1.05	4.69	1.92	0.71	0.15	1.29	0.56	1.93	1.05
PD流量[38]	2.13	0.38	3.25	1.01	5.19	2.52	0.65	0.16	1.42	0.78	3.90	2.86
I-PFN [49]	2.31	0.37	3.43	0.9	5.49	2.5	0.66	0.12	1.05	0.43	2.54	1.65
P2P-Bridge (我们的)	2.28	0.39	3.20	0.81	3.99	1.42	0.59	0.09	0.90	0.32	1.56	0.84
双侧[10]	4.32	1.35	6.17	1.65	8.30	2.39	1.17	0.20	2.50	0.63	6.08	2.19
个人计算机网络[45]	3.85	1.22	8.75	3.04	14.5	5.87	1.29	0.29	1.91	0.51	3.25	1.08
DMRDenoise [34]	6.60	2.15	7.15	2.24	8.09	2.49	1.57	0.35	2.01	0.49	2.99	0.86
总LR [64]	3.40	0.96	5.27	1.15	7.25	1.67	0.96	0.13	2.02	0.42	4.50	1.31
降噪分数[36]	3.37	0.83	5.13	1.20	6.78	1.94	1.07	0.18	1.66	0.35	2.49	0.66
弹匣[68]	3.37	0.83	5.13	1.19	7.24	1.94	1.07	0.18	1.66	0.35	3.56	1.15
PD-流[38]	3.24	0.62	4.62	0.92	6.61	1.62	0.97	0.15	1.80	0.40	4.28	1.37
I-PFN [49]	3.05	0.72	4.95	1.16	7.39	2.21	0.99	0.14	1.43	0.27	3.03	0.86
P2P-Bridge (我们的)	2.88	0.63	4.47	0.89	5.58	1.29	0.92	0.12	1.35	0.24	2.12	0.49

比如异常值聚类或表面增厚效应[2] ,并展示了扩展到大输入的方法。

在 ScanNet++ 上,我们评估了作者提供的噪声点云重建的所有模型。这些重建是通过过滤获得的

根据深度图与 Faro 激光深度的一致性,然后通过使用全局优化的姿势进行投影,无需进一步融合方法[58]。此外,我们评估了所有方法在重建结果上的

通过在预过滤的深度图上应用 3DMatch [63] ,使用全局优化姿势。在 ARKitScenes 上,我们直接将这方法应用于 ARKit

作者提供的重建是 3D 对象检测子集的一部分。

有关这些重建的更多详细信息,请参阅相应的论文[1]。表2显示我们的方法主要使用 RGB+DINO 特征取得了最好的结果,其次是我们的方法只使用RGB或坐标特征。我们定性地比较了图5 中表现最佳的方法,包括噪声和 Faro 点云。我们使用颜色渐变来表示

预测点和真实点之间的距离,范围从绿色变为红色,分别表示低距离和高距离。 ScoreDenoise 以及所有在高斯噪声假设下训练的其他方法会受到影响

来自类似图案的工件。由于内存限制,所有深度学习方法对大点云进行去噪。对于在合成噪声下训练的方法,