

图 4:在 3% 各向同性高斯噪声下,在 PU-Net 数据集上对我们的 P2P-Bridge 与最近基于深度学习的点云去噪方法进行定性比较。

这导致了一种情况,即每个patch边界上的点被假定为异常值,并集中在这些patch周围的人工边界上。当使用没有随机性的朗之万采样 (如在ScoreDenoise和MAG中)时,这种效果会进一步增强,导致点崩溃 [52]。我们假设我们的方法不太容易受到patch伪影的影响,原因有二。首先,我们不是在高斯噪声假设下进行训练的,这使得我们的方法在区分真实物体边界和由于基于patch的处理而出现的边界方面更为稳健。此外,我们不是简单地在patch上累积预测,然后进行最远点采样,而是对噪声云中每个点的预测点坐标取平均值。虽然PD-Flow表现出圆形patch伪影,但它不会因为在现实世界噪声上训练而受到patch边界点崩溃的影响。由于与ScanNet++相比,ARKitScenes (参见图5)的噪声扫描中的细节较少,因此去噪通常不太明显。尽管如此,与其他方法相比,我们的方法可以生成更清晰的物体边缘和更平滑的表面。还有一些不完整的物体可见,这是任何方法都无法完成的。为了解决这个问题,未来的研究可以结合点云补全策略[15,37]来进一步改善结果。

12 M.沃格尔等人。

表 2:室内场景得分。定量点云去噪比较
ScanNet++ [58]和 ARKitScenes [1]测试集中的场景。Faro 扫描仪点
云层作为地面实况的参考。→CD描述了
噪声点到地面真实表面和←CD描述从
真实点指向噪声点。类似地,P2F 是点到面的距离,
F2P 是人脸到点的距离。ScanNet++ 上的指标乘以 104
ARKitScenes 的因子是 103 。

方法	数据集 输入源	特征	ScanNet++ [58]						ScanNet++ [58]						ARKit场景[1]		
			苹果 LiDAR						苹果 LiDAR + 3DMatch [63]						苹果 LiDAR		
			P2F	F2P	→CD	←CD	CD	P2M	P2F	F2P	→CD	←CD	CD	P2M	→CD	←CD	CD
双侧[10]		XYZ	6.29	140.59	6.66	145.44	73.44	76.05	108.70	18.32	108.89	19.67	64.28	63.51	15.87	70.49	43.18
精确射程[34]		XYZ	6.48	149.99	6.71	159.13	78.24	82.92	99.96	19.61	100.16	21.27	60.71	59.79	10.84	30.51	20.68
降噪分数[36]		XYZ	3.49	128.59	3.72	132.71	68.21	66.04	97.11	18.87	97.31	20.26	58.78	57.99	9.56	30.86	20.21
弹匣[68]		XYZ	76.49	99.05	24.82	99.26	26.69	62.97	61.93	9.57	80.82	20.20					
PD-流[38]		XYZ	3.80	147.49	4.02	151.90	77.96	75.64	85.29	21.00	85.49	22.56	54.02	53.14	9.93	33.82	21.87
I-PN[49]		XYZ	3.80	132.98	4.03	137.21	70.62	68.39	83.99	18.99	84.19	20.43	52.31	51.49	9.19	31.99	20.59
P2P 桥接器 (我们的)		XYZ	2.48	122.23	2.71	126.22	64.46	62.35	50.87	18.69	51.07	20.05	35.56	34.78	9.65	30.64	20.14
P2P 桥 (我们的)	XYZ,RGB	2.47	122.27	2.70	126.26	64.48	62.37	50.40	18.39	50.60	19.73	35.17	34.39	9.65	30.45	20.05	
P2P 桥接 (我们的)	XYZ,RGB,DINO	2.42	122.23	2.65	126.22	64.44	62.33	49.64	18.57	49.84	19.92	34.88	34.11	9.57	30.27	19.92	

4.6 消融研究

我们进行消融研究,以更好地了解网络和
扩散参数设计选择对性能的影响。
扩散模型主干。我们评估不同的扩散模型主干
在 ScanNet++ 的一个子集上。具体来说,我们考虑一种基于
点像素卷积神经网络 PVCNN [32] (参见图3)、GECCO [55]的基于 Transformer 的架构以及基于稀疏卷
积的架构
使用 Minkowski 引擎[7]。由于 Minkowski 引擎不提供
扩散模型架构,我们重建了 DDPM [20]主干架构
仅使用 Minkowski 引擎的构建块。表4显示了结果
性能,有利于 PVCNN 架构。
桥梁设置。我们评估最近邻插值和
PU-Net 数据集上的随机性。表3表明,在没有先前训练的情况下
无序点云数据的对齐会大大降低性能
我们的方法。事实上,如果没有适当的数据对齐,该方法无法
收敛。在训练过程中向插值路径添加随机性,这
相当于不训练 OT-ODE 而是训练 SDE,也会降低性能。我们
推测这是由于噪声扫描中存在强大的先验信息。然而,
对于其他任务,例如点云完成,可以增加额外的随机性
变得必要[30,37]。

表 3:桥梁设置比较
PU-Net 数据集。CD 和 P2M 是多重
增加了 104 个。

OT-ODE 对准 CD P2M		
✓	49.33	44.22
	✓ 2.45	0.73
✓	✓ 2.11	0.65

表 4 扩散模型主干组成
ScanNet++ 上的比较。CD 和 P2M 是
均乘以 104 。

骨干		光盘P2M
明科夫斯[7]	12.79	32.39
设置变压器[55]	10.50	15.46
聚氯乙烯-CNN [32]	9.76	14.33

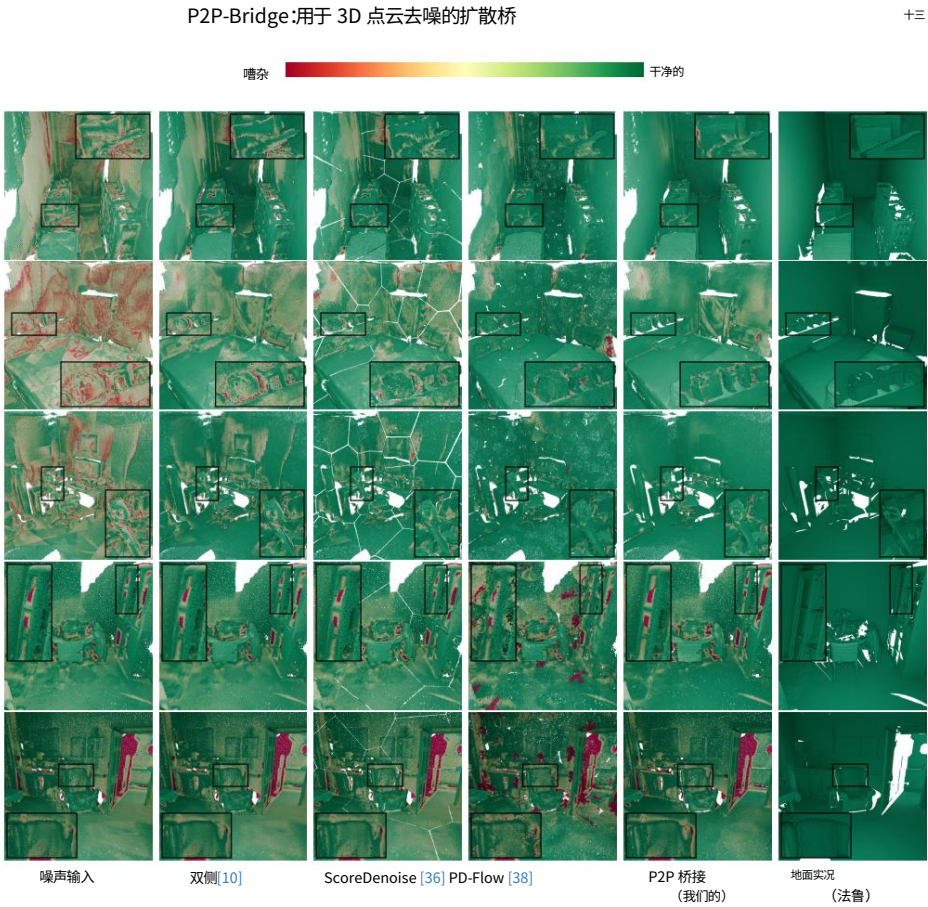


图 5:ScanNet++ [58] (前 3 行)和 ARKitScenes [1]上的定性比较 (底部 2 行)使用嘈杂的 iPhone 扫描作为输入。

模型建筑。我们研究了单个建筑的重要性表 5 中的块及其属性。研究表明,增加块的增加通常会改善结果,其中较浅的差异较大块。由于输入在每个 SA 块之后被下采样,因此较浅的块可以提取更多细粒度的特征,可能解释更大的影响,由于体素卷积和全局特征网络已经提取了粗特征。在其他特征层中,我们看到最大的影响来自特征嵌入。然而,将每层的通道数加倍具有最大的对评估指标的影响。图6显示了指标的相对变化增加推理步骤。只需五到五个步骤即可获得良好的结果十个推理步骤之后,指标似乎趋于稳定。

14 M.沃格尔等人。

表 5:Scan-Net++ 上的网络配置研究。SE 描述了挤压和激励块在[22]中引入的卷积层之后。CD 且 P2M 乘以 104。

基础 PVC 全球		特征	
通道块功能 SE 嵌入 CD (Δ) P2M (Δ)			
32	1222 ✓ ✓ 2122 ✓ ✓ 2212	✓	9.41 (+1.14) 13.78 (+4.08)
32	✓ ✓	✓	9.45 (+1.18) 13.82 (+4.12)
32		✓	9.33 (+1.06) 13.70 (+4.00)
32	2221 ✓ ✓ 2222 ✓ ✓	✓	9.32 (+1.05) 13.70 (+4.00)
32		✓	9.26 (+0.99) 13.67 (+3.97)
32	2222 ✓ 2222 ✓	✓	9.36 (+1.09) 13.75 (+4.05)
32	2222 ✓ ✓	✓	9.33 (+1.06) 13.71 (+4.01)
32			9.76 (+1.49) 14.33 (+4.61)
64	2222 ✓ ✓ 2322 ✓ ✓	✓	8.31 (+0.04) 9.72 (+0.02)
64		✓	8.27 (---) 9.70 (---)

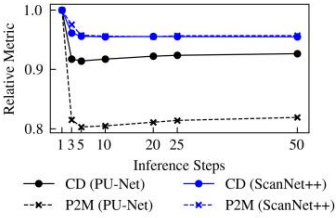


图 6:相对改善 CD 和 P2M 的采样步骤。好的指标可以只需 5 到 10 步骤,从而实现快速推理。

5 结论与讨论

在本文中,我们提出了基于点云去噪框架 P2P-Bridge 扩散薛定谔桥。它通过学习点之间的最佳传输路径,将去噪任务作为数据到数据的扩散问题来解决

集。我们激发了在应用扩散桥时对数据对齐的需求通过绘制最优运输计划之间的相似性来点云数据和最短路径点云插值并实证证明其效率这种方法。我们将我们的方法应用于单个对象数据集以及大规模室内点云,并通过大量实验表明我们的方法在单个物体和大点云上的表现优于以前的研究数据。最后,我们展示了其他基于图像的特征,例如 RGB 信息,以及逐点高级特征,例如 DINOv2 特征,进一步提高成果。

致谢 Francis Engelmann 部分得到了 ETH AI 的支持中心博士后研究奖学金和苏黎世联邦理工学院职业种子奖。

参考

1. Baruch, G., Chen, Z., Dehghan, A., Dimry, T., Feigin, Y., Fu, P., Gebauer, T., Joffe, B., Kurz, D., Schwartz, A. 等人: ARKitscenes - 一种使用移动 RGB-d 数据进行 3d 室内场景理解的多样化真实世界数据集。在: 国际神经信息处理系统会议 (NeurIPS) (2021) 8, 10, 12, 13
2. Chen, H., Shen, J.: 用于计算机辅助设计、工程和制造的点云数据去噪。在: 计算机工程 (2018) 2, 3, 10
3. Chen, H., Wei, M., Sun, Y., Xie, X., Wang, J.: 通过具有图约束的低秩恢复进行多补丁协作点云去噪。引自: IEEE 可视化与计算机图形学交易 (2019) 3
4. Chen, T., Liu, GH., Theodorou, EA.: 使用前向-后向 SDE 理论对薛定谔桥进行似然训练 (2022) 5, 6
5. Chen, Y., Hu, VT., Gavves, E., Mensink, T., Mettes, P., Yang, P., Snoek, CG.: PointMixup: 点云增强。在: 欧洲计算机视觉会议 (ECCV) (2020) 6
6. Chen, Z., He, G., Zheng, K., Tan, X., Zhu, J.: 薛定谔桥节拍扩散文本转语音合成模型 (2023) 5, 6, 7
7. Choy, C., Gwak, J., Savarese, S.: 4D 时空卷积网络: Minkowski 卷积神经网络。在: IEEE 计算机视觉与模式识别会议论文集 (2019) 12
8. De Bortoli, V., Thornton, J., Heng, J., Doucet, A.: 扩散薛定谔桥在基于分数的生成模型中的应用。在: 国际神经信息处理系统会议 (NeurIPS) (2021) 3, 5
9. Delitzas, A., Takmaz, A., Tombari, F., Sumner, R., Pollefeys, M., Engelmann, F.: SceneFun3D: 3D 场景中的细粒度功能和可用性理解。在: 国际计算机视觉和模式识别会议 (CVPR) (2024) 1
10. Digne, J., de Franchis, C.: 点云的双边滤波器。在: 图像处理在线 (2017) 2, 9, 10, 12, 13
11. Digne, J., Valette, S., Chaine, R.: 通过局部形状探测实现稀疏几何表示。在: IEEE 可视化与计算机图形学交易 (2017) 3
12. Duan, C., Chen, S., Kovacevic, J.: 加权多投影: 使用切线平面进行 3D 点云去噪。在: 2018 年 IEEE 全球信号和信息处理会议 (GlobalSIP) (2018) 3
13. Duan, C., Chen, S., Kovacevic, J.: 通过基于深度神经网络的局部表面估计进行 3D 点云去噪 (2019) 2
14. Engelmann, F., Manhardt, F., Niemeyer, M., Tateno, K., Tombari, F.: OpenNeRF: 具有像素级特征和渲染新视图的开放集 3D 神经场景分割。在: 国际学习表征会议 (ICLR) (2024) 1
15. Fei, B., Yang, W., Chen, WM., Li, Z., Li, Y., Ma, T., Hu, X., Ma, L.: 基于深度学习的 3d 点云完成处理和分析的综合回顾。出处: IEEE 智能交通系统汇刊 (2022) 11
16. Golub, GH., Van Loan, CF.: 总体最小二乘问题的分析。出处: SIAM 数值分析杂志 (1980) 3
17. 韩晓峰, 金建生, 王明军, 姜伟: 引导式三维点云过滤。多媒体工具与应用 (2018) 2, 3

16 M.沃格尔等人。

18.韩晓峰、金建生、王明军、姜伟、高玲、肖玲:三维点云滤波算法综述。信号处理:图像通信 (2017) 2、3

19. Hermosilla, P., Ritschel, T., Ropinski, T.:全面去噪:3D 点云清理的无监督学习。在:IEEE/CVF 国际计算机视觉会议论文集 (2019) 2、3 20. Ho, J., Jain, A., Abbeel, P.:去噪扩散概率模型。在:国际神经信息处理系统会议 (NeurIPS) (2020) 3、4、5、6、12

21. Hoogeboom, E., Heek, J., Salimans, T.:简单扩散:用于高分辨率图像的端到端扩散。在:国际机器学习会议 (2023) 7
22. Hu, J., Shen, L., Sun, G.:挤压和激励网络。在: IEEE 计算机视觉和模式识别会议论文集 (2018) 7、14 23. Hu, W., Gao, X., Cheung, G., Guo, Z.:用于 3d 点云去噪的特征图学习。引自:IEEE 信号处理学报 (2020) 3 24. Huang, R., Peng, S., Takmaz, A., Tombari, F., Pollefeys, M., Song, S., Huang, G., Engelmann, F.: Segment3D:无需手动标签即可学习细粒度类别无关 3D 分割。欧洲计算机视觉会议 (ECCV) (2024) 1

25. Kasten, Y., Rahamim, O., Chechik, G.:使用预训练的文本到图像扩散模型进行点云补全。在:国际神经信息处理系统会议 (NeurIPS) (2024) 4 26. Kreuzberg, L., Zulfikar, I., Mahadevan, S., Engelmann, F., Leibe, B.:4D-StOP:使用时空对象提议生成和聚合对 4D LiDAR 进行全景分割。在: ECCVW (2022) 1 27. Leal, E., Sanchez-Torres, G., Branch, J.W.:基于稀疏正则化的点云去噪和尖锐特征增强方法。于:传感器 (2020) 3 28. Lemke, O., Bauer, Z., Zurbrugg, R., Pollefeys, M., Engelmann, F., Blum, H.:Spot-Compose :用于点云中开放词汇对象检索和抽屉操作的框架。于:ICRA 2024 移动操作和具身智能第二届研讨会(2024) 1

29. Léonard, C.:薛定谔问题及其与最优传输的一些联系的调查。在:arXiv 预印本 arXiv:1308.0215 (2013) 5 30. Liu, GH, Vahdat, A., Huang, DA, Theodorou, EA, Nie, W., Anandkumar, A.:I2SB :图像到图像薛定谔桥。在:国际机器学习会议(ICML) (2023) 4、5、6、7、12 31. Liu, Z., Xiao, X., Zhong, S., Wang, W., Li, Y., Zhang, L., Xie, Z.:用于点云去噪的特征保留框架。在:计算机辅助设计 (2020) 2、3

32. Liu, Z., Tang, H., Lin, Y., Han, S.:用于高效 3D 深度学习的点体素 CNN (2019) 7, 12

33. Lu, X., Schaefer, S., Luo, J., Ma, L., He, Y.:3D 几何滤波的低秩矩阵近似。在:IEEE 可视化与计算机图形学学报(2020) 3

34. Luo, S., Hu, W.:用于点云去噪的可微分流形重建 (2020) 2、3、9、10、12

35. Luo, S., Hu, W.:用于 3D 点云生成的扩散概率模型。在:国际计算机视觉与模式识别会议 (CVPR) (2021) 4

36. Luo, S., Hu, W.:基于分数的点云去噪。引自:国际计算机视觉会议(ICCV) (2021) 2、3、4、9、10、11、12、13

37. Lyu, Z., Kong, Z., XU, X., Pan, L., Lin, D.: 用于 3D 点云补全的条件点扩散-细化范式。在:国际学习表征会议 (ICLR) (2022) 4.7.11.12
38. Mao, A., Du, Z., Wen, YH., Xuan, J., Liu, YJ.: PD-Flow: 具有正则化流的点云去噪框架。在:欧洲计算机视觉会议 (ECCV) (2022) 2.3.9.10.11.12.13 39. Mattei, E., Castrodad, A.: 通过移动 RPCA 实现点云去噪。在:计算机图形学论坛 (2017) 3 40. Nichol, A., Dhariwal, P.: 改进的去噪扩散概率模型。在:国际机器学习会议 (ICML) (2021) 4 41. Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A. 等人: DINOv2: 无需监督即可学习稳健的视觉特征 (2023) 2, 3.8 42. Peyré, G., Cuturi, M.: 计算最优传输。基础和趋势机器学习 (2019) 7
43. Qi, CR., Yi, L., Su, H., Guibas, LJ.: Pointnet++: 度量空间中点集的深度分层特征学习。在:国际神经信息处理系统会议 (NeurIPS) (2017) 7 44. Qian, G., Li, Y., Peng, H., Mai, J., Hammoud, H., Elhoseiny, M., Ghanem, B.: PointNeXt: 通过改进的训练和扩展策略重新审视 PointNet++。参见:国际神经信息处理系统会议 (NeurIPS) (2022) 2
45. Rakotosaona, MJ., La Barbera, V., Guerrero, P., Mitra, NJ., Ovsjanikov, M.: PointCleanNet: 学习从密集点云中去除和移除异常值 (2020) 2.3.4.8.9.10 46. Ravi, N., Reizenstein, J., Novotny, D., Gordon, T., Lo, WY., Johnson, J., Gkioxari, G.: 使用pytorch3d 加速 3D 深度学习。在:arXiv 预印本 arXiv:2007.08501 (2020) 2
47. Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., Norouzi, M.: 调色板: 图像到图像的扩散模型。见:ACM SIGGRAPH (2022) 4 48. Schrödinger, E.: 论电子的相对论理论和量子力学的解释。亨利庞加莱研究所年鉴 (1932) 3, 5 49. de Silva Edirimuni, D., Lu, X., Shao, Z., Li, G., Robles-Kelly, A., He, Y.: 迭代: 真正的迭代点云过滤。见:计算机视觉和模式识别国际会议 (CVPR)。页数13530–13539 (2023) 3, 10, 12 50. Somnath, VR., Pariset, M., Hsieh, YP., Martinez, MR., Krause, A., Bunne, C.: 对齐扩散 Schr/ödinger 桥。见: arXiv 预印本 arXiv:2302.11419 (2023) 5
51. Song, J., Meng, C., Ermon, S.: 去噪扩散隐式模型。在:国际学习表征会议 (ICLR) (2021) 4.6 52. Song, Y., Ermon, S.: 通过估计数据分布梯度进行生成建模。在:国际神经信息处理系统会议 (NeurIPS) (2019) 3.11 53. Takmaz, A., Fedele, E., Sumner, RW., Pollefeys, M., Tombari, F., Engelmann, F.: OpenMask3D: 开放词汇 3D 实例分割。出处:国际神经信息处理系统会议 (NeurIPS) (2023) 1 54. Takmaz, A., Schult, J., Kaftan, I., Akçay, M., Leibe, B., Sumner, R., Engelmann, F., Tang, S.: 使用合成数据对点云中的人体进行 3D 分割。出处:国际计算机视觉会议 (ICCV) (2023) 1
55. Tyszkiewicz, MJ., Fua, P., Trulls, E.: Gecco: 几何条件点扩散模型。arXiv 预印本 arXiv:2303.05916 (2023) 4, 12

18 M.沃格尔等人。

56. Weder, S., Schonberger, J., Pollefeys, M., Oswald, M. R.: RoutedFusion: 学习实时深度图融合。刊于: 国际计算机视觉与模式识别会议 (CVPR) (2020) 4

57. Yadav, S. K., Reitebuch, U., Skrodzki, M., Zimmermann, E., Polthier, K.: 使用正态投票张量和受限二次误差度量进行基于约束的点集去噪。在: 计算机与图形学 (2018) 2, 3

58. Yeshwanth, C., Liu, Y. C., Nießner, M., Dai, A.: ScanNet++: 3D 室内场景的高保真数据集。刊于: 国际计算机视觉会议 (ICCV) (2023) 第8, 10, 12, 13页

59. Yilmaz, K., Schult, J., Nekrasov, A., Leibe, B.: MASK4D: 用于4D 全景分割的 Mask Transformer (2024) 1 60. Yu, L., Li, X., Fu, C. W., Cohen-Or, D., Heng, P. A.: PU-Net: 云上采样网络。在: 国际计算机视觉与模式识别会议 (CVPR) (2018) 8, 10 61. Yue, Y., Das, A., Engelmann, F., Tang, S., Lenssen, J.: 通过3D 感知微调改进 2D 特征表示。欧洲计算机视觉会议 (ECCV) (2024) 2 62. Yue, Y., Kontogianni, T., Schindler, K., Engelmann, F.: 连接点: 使用两级查询进行平面图重建。在: 国际计算机视觉与模式识别会议 (CVPR) (2023) 1 63. Zeng, A., Song, S., Nießner, M., Fisher, M., Xiao, J., Funkhouser, T.: 3DMatch: 从 RGB-D 重建中学习局部几何描述符。引自: CVPR (2017) 4, 8, 10, 12 64. Zeng, J., Cheung, G., Ng, M., Pang, J., Yang, C.: 使用低维流形模型的图拉普拉斯正则化进行3D点云去噪 (2019) 3, 9, 10

65. Zeng, X., Vahdat, A., Williams, F., Gojcic, Z., Litany, O., Fidler, S., Kreis, K.: LION: 用于 3D 形状生成的潜在点扩散模型。在: 国际神经信息处理系统会议 (NeurIPS) (2022) 4, 7 66. Zhang, D., Lu, X., Qin, H., He, Y.: Pointfilter: 通过编码器-解码器建模进行点云过滤。在: IEEE 可视化与计算机图形学学报 (2020) 2

67. Zhang, F., Zhang, C., Yang, H., Zhao, L.: 使用主成分分析和新型双边滤波器进行点云去噪。在: 信号处理 (2019) 3 68. Zhao, Y., Zheng, H., Wang, Z., Luo, J., Lam, E. Y.: 通过梯度场中的动量上升进行点云去噪 (2023) 2, 3, 4, 9, 10, 11, 12 69. Zheng, J., Barath, D., Pollefeys, M., Armeni, I.: 地图自适应: 实时质量自适应语义 3D 地图。arXiv preprint arXiv:2406.05849 (2024) 4 70. Zheng, Y., Li, G., Wu, S., Liu, Y., Gao, Y.: 通过锐化引导点云去噪

特征骨架。在: 《视觉计算机》(2017) 2, 3

71. 郑勇, 李光, 徐晓燕, 吴胜, 聂宇: 点云的滚动法线滤波。引自: 计算机辅助几何设计 (2018) 2, 3 72. 周琳, 杜燕, 吴建军: 通过点-体素扩散实现 3D 形状生成与完成。引自: 国际计算机视觉会议 (ICCV) (2021) 4, 7

73. Zhu, Z., Peng, S., Larsson, V., Cui, Z., Oswald, M. R., Geiger, A., Pollefeys, M.: NICER-SLAM: 用于 RGB SLAM 的神经隐式场景编码。在: 国际3D 视觉会议 (3DV) (2024) 4

74. Zhu, Z., Peng, S., Larsson, V., Xu, W., Bao, H., Cui, Z., Oswald, M. R., Pollefeys, M.: NICE-SLAM: 用于 SLAM 的神经隐式可扩展编码。在: 国际计算机视觉与模式识别会议 (CVPR) (2022) 4

75. Zurbrugg, R., Liu, Y., Engelmann, F., Kumar, S., Hutter, M., Patil, V., Yu, F.: ICGNet: 一种以实例为中心的抓取的统一方法。在: 国际机器人与自动化会议 (ICRA) (2024) [1](#)