Learning Normals of Noisy Points by Local Gradient-Aware Surface Filtering – Supplementary Material –

Qing Li[1] Huifang Feng[2*] Xun Gong[1] Yu-Shen Liu[3]
[1] Southwest Jiaotong University, Chengdu, China
[2] Xihua University, Chengdu, China [3] Tsinghua University, Beijing, China
qingli@swjtu.edu.cn fhf@xhu.edu.cn xgong@swjtu.edu.cn liuyushen@tsinghua.edu.cn

1. More Experimental Results

Normal Estimation. The PGP curve represents the percentage of good point normals (PGP) under a series of given angle thresholds, providing a comprehensive evaluation of normal estimation accuracy across varying levels of precision requirements. In Fig. 1, the PGP curves of oriented normals for unsupervised methods clearly illustrate that our method consistently achieves superior performance across all thresholds and all data categories. This highlights the robustness of our approach in capturing accurate normals, even under challenging scenarios. Similarly, the PGP curves of unoriented normals, as depicted in Fig. 2, further demonstrate the effectiveness of our method, showcasing improved performance over baseline methods in all tested conditions. To validate the practical applicability of our method, we extended the evaluation to real-world scanned data from the ScanNet dataset [1]. This dataset provides complex, noisy, and unstructured 3D data that are captured in real-world indoor scenarios. As shown in Fig. 3, our approach achieves significantly improved results compared to previous methods, with visual examples demonstrating more accurate normal estimations. Overall, the superior performance across synthetic and real-world datasets underlines the robustness, accuracy, and generalizability of our approach for normal estimation in noisy point clouds.

Surface Reconstruction. As illustrated in Fig. 4, we provide a detailed visual comparison of reconstructed surfaces generated from point clouds with varying noise levels. These results highlight our method's ability to maintain surface integrity and recover fine details, even in the presence of significant noise, outperforming baseline approaches in terms of clarity and structural fidelity. Further, in Fig. 5, we visualize reconstructed surfaces from real scanned point clouds of the 3D Scene dataset [11]. The results demonstrate our method's robustness in processing real-world data, effectively handling challenges such as noise, irregular sampling, and complex geometries, producing reconstructions that are both accurate and visually coherent. In Fig. 6, we explore an even more challenging scenario by generating surfaces from wireframe-like point clouds, characterized by extreme sparsity and non-uniform sampling. Remarkably, our approach successfully reconstructs surfaces with consistent geometry and minimal artifacts, demonstrating its adaptability and resilience to highly irregular data distributions. Overall, these visual comparisons confirm the versatility and effectiveness of our method across various challenging datasets, making it well-suited for both synthetic and real-world applications.

Point Cloud Denoising. We compare our approach with baseline methods that rely on supervised training with ground truth labels. In Fig. 7, we present a visual comparison of denoising results on the PointCleanNet dataset [8] for various methods. Our approach demonstrates a superior ability to recover clean and smooth surfaces from noisy point clouds while preserving geometric details, even in challenging cases with significant noise levels. To further evaluate its practical applicability, we also apply our method to real-world scanned data. Fig. 8 showcases denoising results on selected scenes from the Paris-Rue-Madame dataset [9]. Despite not requiring ground truth supervision, our method effectively mitigates noise and produces visually coherent surfaces, demonstrating its robustness and adaptability to real-world scenarios. These results validate the effectiveness of our designed loss functions, which guide the network to accurately infer the clean surfaces from various noisy point clouds. By leveraging statistical reasoning rather than relying on explicit ground truth, our method achieves reliable performance in both synthetic and real-world applications, setting a new standard for unsupervised point cloud denoising. The above results confirm the suitability of our method for real-world applications in 3D vision tasks.

Comparison with NeuralGF on runtime & convergence. Since our method includes more sophisticated losses compared to NeuralGF [3], the optimization incurs a slightly higher per-instance time cost ($\sim +0.4\%$). However, de-

*Corresponding author

Table 1. Ablations for unoriented and oriented normal estimation on the FamousShape dataset. We decompose $\mathcal{L}_d$ into $\mathcal{L}_{ld}$ and $\mathcal{L}_{pd}$.

	Category	Unoriented Normal RMSE							Oriented Normal RMSE						
		Noise				Density		Average	Noise				Density		Average
		None	Low	Medium	High	Stripe	Gradient		None	Low	Medium	High	Stripe	Gradient	
(a)	$\mathcal{L}_{ld}$	25.19	33.08	38.42	46.76	28.80	25.51	32.96	37.51	67.44	84.01	69.70	45.80	35.95	56.73
	$\mathcal{L}_{ld} + \mathcal{L}_{pd} + \mathcal{L}_{sd}$	15.74	23.08	54.64	56.10	14.33	15.53	29.90	22.11	39.46	88.36	92.84	17.97	21.94	47.11
	$\mathcal{L}_{ld} + \mathcal{L}_{pd} + \mathcal{L}_{sd} + \mathcal{L}_v$	16.98	16.60	33.83	41.23	18.85	18.37	24.31	26.56	23.91	55.90	83.79	28.98	27.41	41.09
	$\mathcal{L}_{ld} + \mathcal{L}_{pd} + \mathcal{L}_{sd} + \mathcal{L}_n$	13.41	16.70	31.11	40.48	13.00	12.78	21.25	19.42	19.48	36.03	50.38	14.76	**13.29**	25.56
	$\mathcal{L}_{ld} + \mathcal{L}_{pd} + \mathcal{L}_n + \mathcal{L}_v$	12.00	15.87	30.00	39.30	11.95	13.14	20.38	14.66	18.89	34.96	55.00	15.12	16.45	25.85
	$\mathcal{L}_{ld} + \mathcal{L}_{sd} + \mathcal{L}_n + \mathcal{L}_v$	12.61	15.59	30.19	39.13	12.38	11.69	20.27	16.06	18.30	38.79	48.26	15.70	14.99	25.35
	$\mathcal{L}_{ld} + \mathcal{L}_n + \mathcal{L}_v$	12.28	15.63	29.88	39.26	12.18	12.19	20.24	14.98	18.25	**34.62**	50.82	14.95	21.90	25.92
	$\mathcal{L}_{pd} + \mathcal{L}_n + \mathcal{L}_v$	11.88	15.72	29.97	**39.05**	12.24	12.14	20.17	13.32	18.34	34.87	48.44	15.08	23.17	25.54
	$\mathcal{L}_{sd} + \mathcal{L}_n + \mathcal{L}_v$	12.49	**15.47**	29.94	39.19	12.49	12.11	20.28	15.06	**17.15**	35.25	**48.27**	15.44	16.62	24.63
	$\mathcal{L}_{pd} + \mathcal{L}_{sd} + \mathcal{L}_n + \mathcal{L}_v$	**11.81**	15.73	29.97	39.08	12.01	11.43	20.00	**13.31**	18.36	34.89	48.87	15.15	14.07	24.11
(b)	w/o Aggregation	12.05	16.02	30.04	39.18	11.96	11.51	20.13	13.82	18.53	35.11	49.34	14.46	13.85	24.19
(c)	$\mathcal{K} = 4$	11.89	16.00	30.22	39.51	17.45	**11.17**	21.04	14.80	19.07	35.16	51.31	21.49	13.65	25.91
	$\mathcal{K} = 16$	13.71	15.77	**29.86**	39.32	12.11	13.49	20.71	20.41	18.44	34.90	51.19	**14.14**	17.87	26.16
	Full	11.90	15.84	29.90	39.08	**11.82**	11.36	**19.98**	13.71	18.40	34.97	49.25	14.35	13.76	**24.07**

Table 2. Comparison with NeuralGF on convergence.

Iterations	10,000	20,000
NeuralGF [3]	19.47	18.70
Ours	**19.22**	**17.00**

spite this small overhead, our method can achieve better convergence. In Table 2, we provide the comparison of average oriented-normal RMSE on the PCPNet dataset at different iterations. The richer constraints of our method lead to lower error within the same iterations.

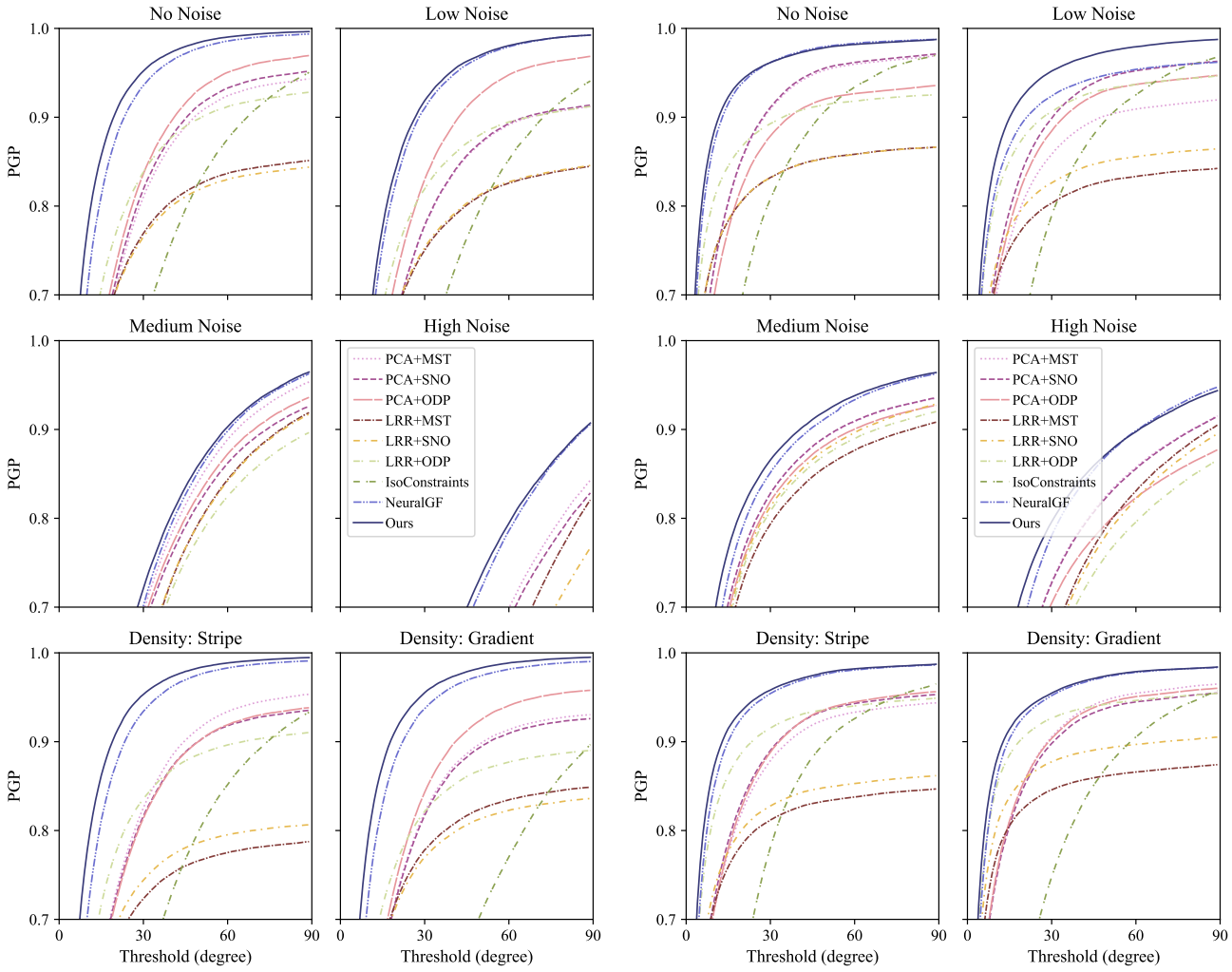
2. Discussion

Hyperparameter tuning in loss. Compared to prior works, we enhance existing losses from a new perspective and introduce novel richer constraints to boost performance. To ensure each loss term contributes meaningfully and no term is orders of magnitude larger or smaller, we let terms with larger raw values receive smaller weight factors, and vice versa. The weights λ_1 and λ_2 are chosen via small grid searches on a held-out subset. We find that varying each weight by $\pm 50\%$ changes the average RMSE by $< 3\%$, showing low sensitivity. For λ_1, we recommend trying $\{0.05, 0.1, 0.3\}$ to choose the value that minimizes normal RMSE. λ_2 are less sensitive and can remain the defaults unless targeting extremely noisy or highly detailed data. In $\mathcal{L}_{sd}$, a $10\times$ factor on the signed distances of surface points forces them to be located on the zero level set and balances surface fitting and denoising. We will release code with default settings for transparency.

Detail preservation and over-smoothing on real data. Our primary focus in this work is on accurately estimating normals from noisy point clouds. In Fig. 8, we demonstrate an extension of our method to point cloud denoising on real-world scans from the Paris-Rue-Madame dataset. While multiple rounds of denoising noticeably improve the visual quality of the surface, fully recovering all fine details remains challenging, particularly because real-world scans often contain a substantial number of outliers. These outliers can skew the surface filtering process, leading to localized over-smoothing even as most noise is removed. Our multi-round filtering on the point cloud denoising task highlights this trade-off: while most outliers and small artifacts are successfully removed, some local regions appear overly smoothed. We acknowledge this limitation and agree that in scenes with many outliers, our method may struggle to balance outlier rejection and structure preservation. We note that this localized over-smoothing is not unique to our approach, some state-of-the-art point cloud denoising methods exhibit similar behavior when aggressively removing noise. To address datasets with heavy outlier contamination, a practical strategy is to incorporate a lightweight preprocessing step that detects and removes outliers before applying the denoising pipeline. Integrating robust outlier filtering or adaptive smoothing strategy into the end-to-end framework is an important direction for future work.

3. Limitation

Our method can be applied to various point cloud processing tasks, as demonstrated in our experiments. However, a limitation of our approach is that the neural network must be optimized individually for each point cloud. As a result, the trained model cannot be directly applied to shapes that were not part of the optimization process. This limitation is similar to some implicit representation methods [5–7, 10] that overfit to individual data, making our method less generalizable compared to pre-trained models like SHS-Net [4] that can be used out-of-the-box for normal estimation on new



(a) FamousShape dataset

(b) PCPNet dataset

Figure 1. Oriented normal PGP curves of unsupervised methods. The Y-axis indicates the percentage of good point normals with errors below the angle thresholds specified on the X-axis. Our method achieves superior results across almost all thresholds.

point cloud data. Therefore, it will incur additional time costs for users to directly start using the proposed method for normal estimation. Future work will focus on addressing this limitation by generalizing the method to handle unseen data, enabling more convenient and broader applicability.

Potential directions for improving the generalization. (1) Front-end feature extractor: mapping objects to a unified metric space using shape-specific features, enabling fast adaptation to diverse geometries. (2) Meta-learning or model-agnostic initialization: pretraining on a small corpus of shapes to warm-start the network for unseen objects.

References

[1] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. ScanNet: Richly-annotated 3D reconstructions of indoor scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5828–5839, 2017. 1

[2] Michael Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Transactions on Graphics*, 32(3): 1–13, 2013. 6

[3] Qing Li, Huifang Feng, Kanle Shi, Yue Gao, Yi Fang, Yu-Shen Liu, and Zhizhong Han. NeuralGF: Unsupervised point normal estimation by learning neural gradient function. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 66006–66019. Curran Associates, Inc., 2023. 1, 2

[4] Qing Li, Huifang Feng, Kanle Shi, Yue Gao, Yi Fang, Yu-Shen Liu, and Zhizhong Han. Learning signed hyper surfaces for oriented point cloud normal estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 46(12):9957–9974, 2024. 2

[5] Shengtao Li, Ge Gao, Yudong Liu, Ming Gu, and Yu-Shen Liu. Implicit filtering for learning neural signed distance

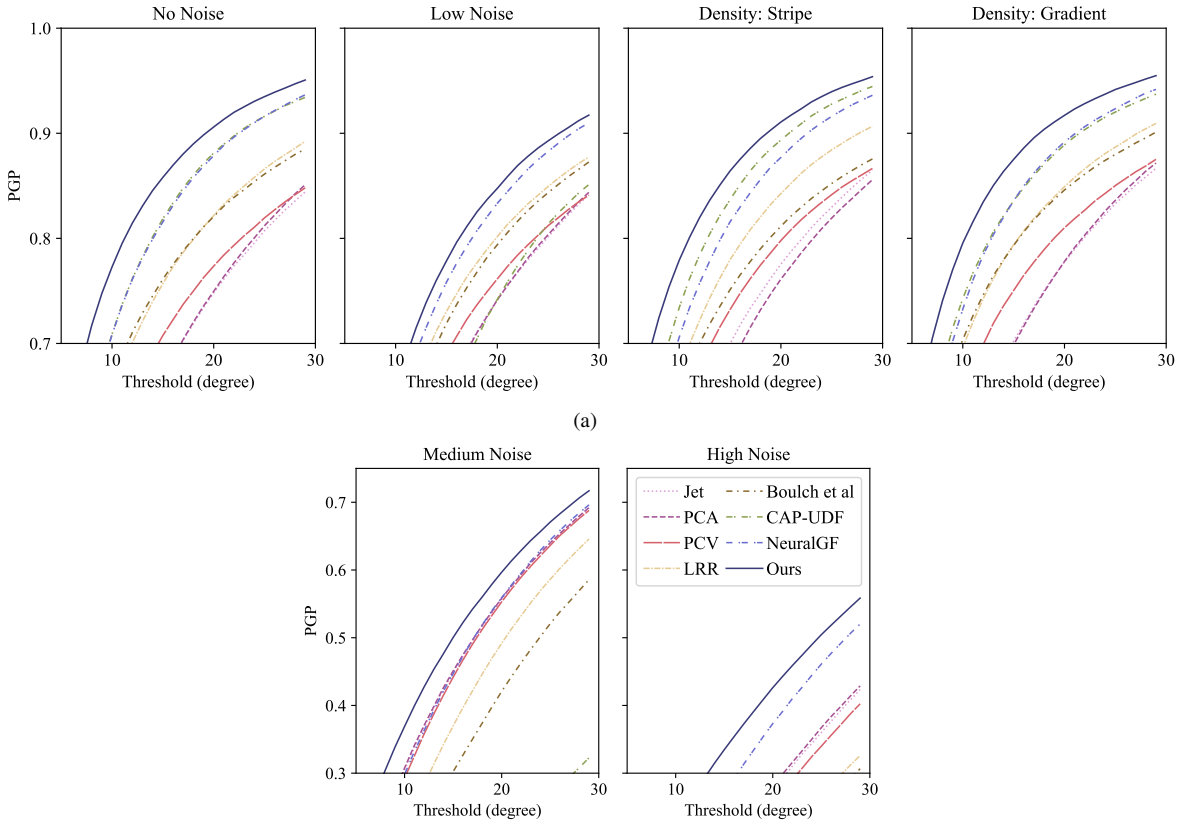


Figure 2. Unoriented normal PGP curves of unsupervised methods on the FamousShape dataset. The Y-axis represents the percentage of good point normals whose errors are below the angle thresholds specified on the X-axis. The scale of the Y-axis in (a) is different from that in (b). Our method consistently outperforms others across all thresholds.

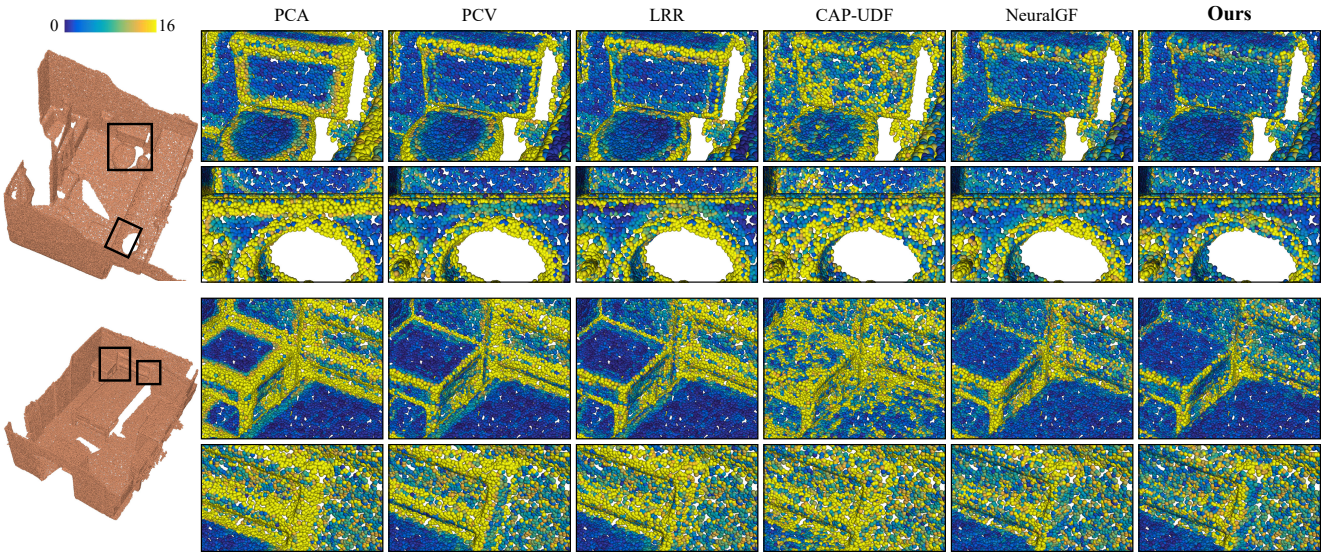


Figure 3. Visual comparison of unoriented normals on real-world point clouds of the ScanNet dataset. Point colors indicate normal errors.

functions from 3D point clouds. In *European Conference on Computer Vision*, pages 234–251. Springer, 2025. 2, 6

[6] Baorui Ma, Zhizhong Han, Yu-Shen Liu, and Matthias Zwicker. Neural-Pull: Learning signed distance functions

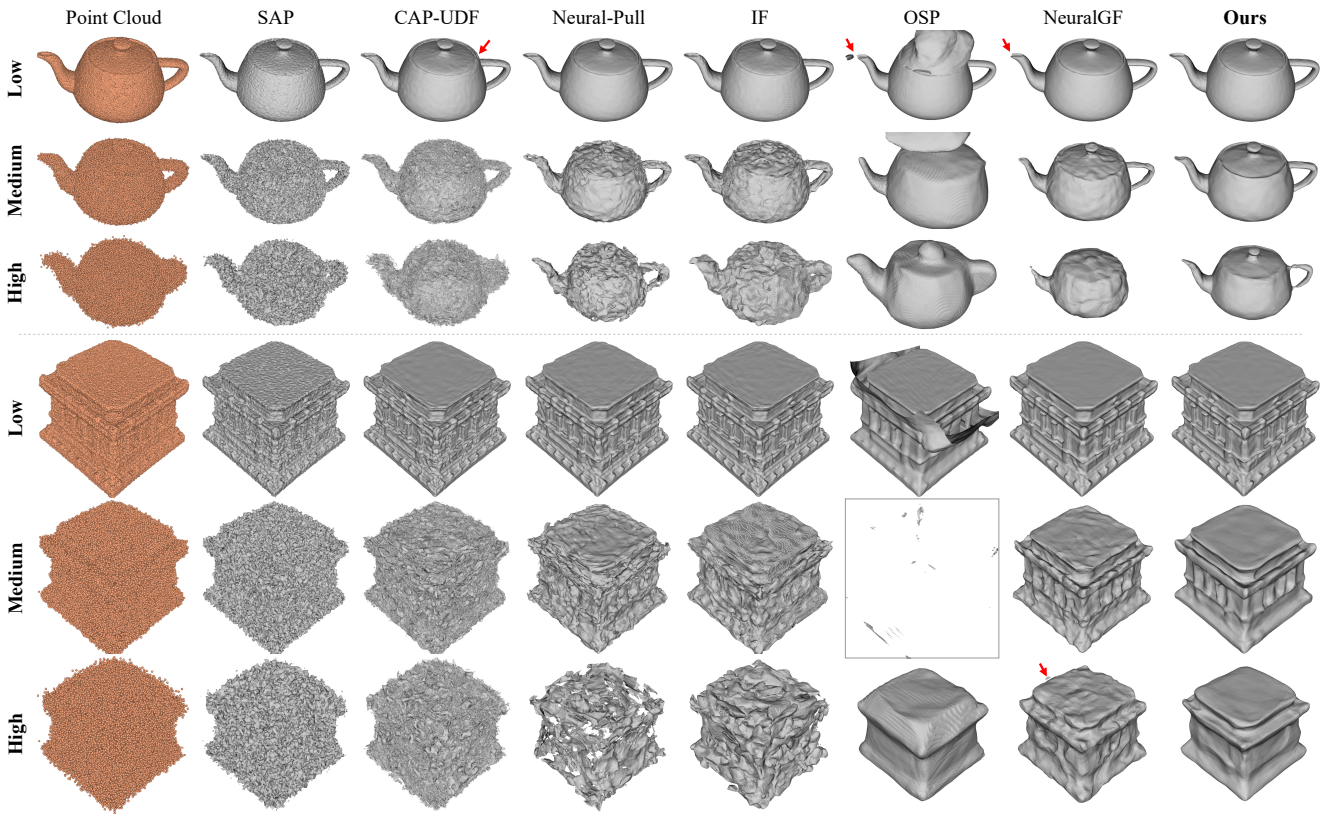


Figure 4. Comparison with implicit representation methods for surface reconstruction. As the noise increases (from low to high), our method becomes more advantageous.

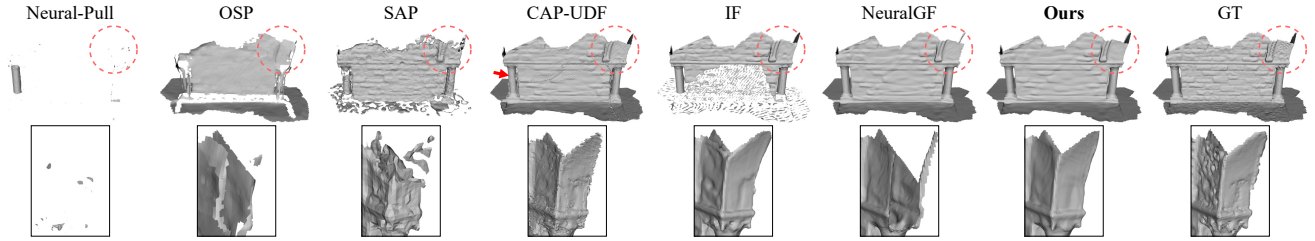


Figure 5. Comparison with implicit representation methods for surface reconstruction on the 3D Scene dataset.

from point clouds by learning to pull space onto surfaces. In *International Conference on Machine Learning*, pages 7246–7257, 2021. 6

[7] Baorui Ma, Yu-Shen Liu, and Zhizhong Han. Reconstructing surfaces for sparse point clouds with on-surface priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6315–6325, 2022. 2, 6

[8] Marie-Julie Rakotosaona, Vittorio La Barbera, Paul Guerrero, Niloy J Mitra, and Maks Ovsjanikov. PointCleanNet: Learning to denoise and remove outliers from dense point clouds. In *Computer Graphics Forum*, pages 185–203. Wiley Online Library, 2020. 1

[9] Andrés Serna, Beatriz Marcotegui, François Goulette, and Jean-Emmanuel Deschaud. Paris-rue-Madame database: a 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods. In *International Conference on Pattern Recognition, Applications and Methods*, 2014. 1

[10] Junsheng Zhou, Baorui Ma, Shujuan Li, Yu-Shen Liu, Yi Fang, and Zhizhong Han. CAP-UDF: Learning unsigned distance functions progressively from raw point clouds with consistency-aware field optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. 2, 6

[11] Qian-Yi Zhou and Vladlen Koltun. Dense scene reconstruction with points of interest. *ACM Transactions on Graphics (TOG)*, 32(4):1–8, 2013. 1

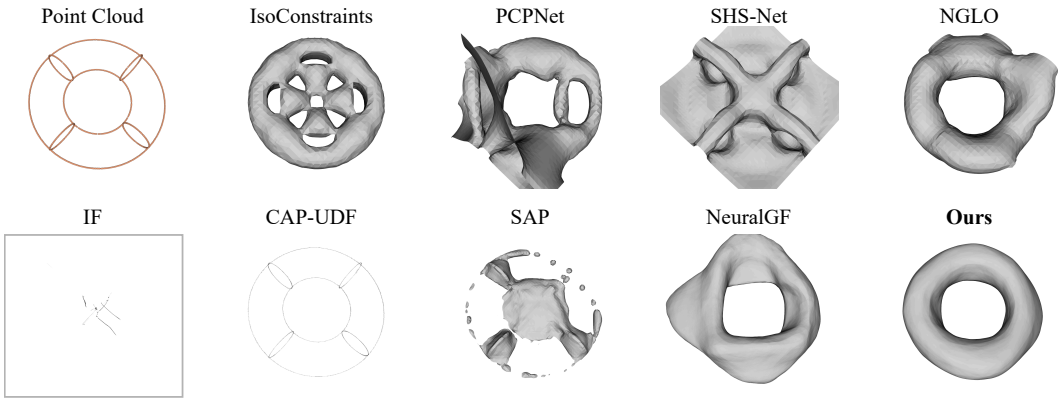


Figure 6. Comparison with normal estimation methods and implicit representation methods for surface reconstruction on a wireframe point cloud of doughnut. The reconstructed surfaces of normal estimation methods in the first row are generated using the Poisson surface reconstruction algorithm [2]. The Neural-Pull [6], OSP [7], IF [5] and CAP-UDF [10] methods fail to generate surfaces.

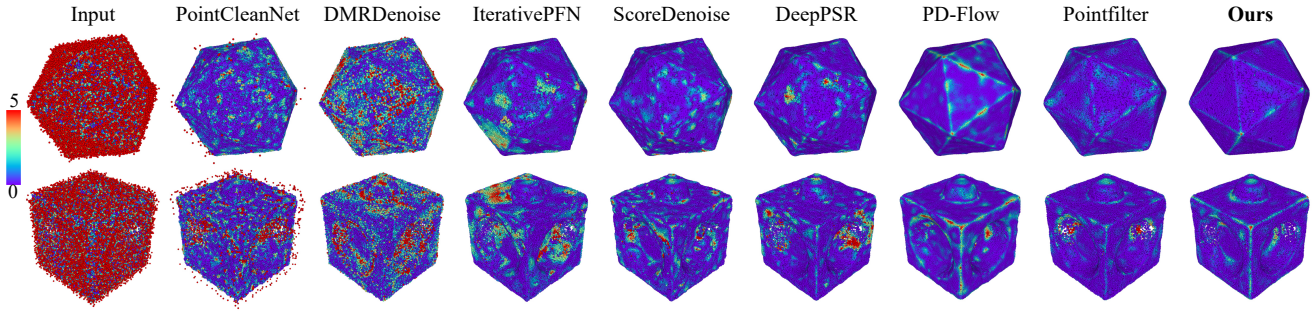


Figure 7. Visual comparison of point cloud denoising. Our method is unsupervised, whereas all baseline methods rely on supervised training. The input consists of 50K-resolution shapes with 3% Gaussian noise from the PointCleanNet dataset. Point colors represent P2M distance error ($\times 10^{-4}$).

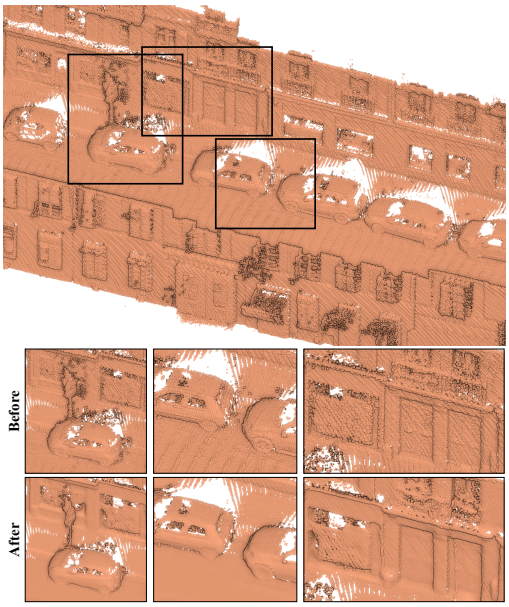


Figure 8. Visualization of point cloud denoising on real-world point clouds of the Paris-rue-Madame dataset. We show the raw data of a street scene in the first row, and compare the local details of the scene before and after denoising using our method in the last two rows.