The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W8-2024
8th International ISPRS Workshop LowCost 3D - Sensors, Algorithms, Applications, 12–13 December 2024, Brescia, Italy

Joint neural denoising and consolidation for portable handheld laser scan

Tian Zhang*, Sagi Filin

Mapping and Geo-Information Engineering, Technion – Israel Institute of Technology, Haifa, Israel
(tianz, filin)@technion.ac.il

Keywords: Point cloud denoising, Handheld laser scanners, Deep learning, Uniform point cloud

Abstract

Mobile and handheld laser scanners document scenes in an economical manner, but the data they acquire are often noisy, of low resolution, unevenly distributed, and feature voids within the scanned scene. These characteristics challenge such applications as feature extraction and 3D modeling when processing the raw pointset. To date, point cloud denoising and consolidation (address of distribution and void regions) have been treated independently despite their complementary nature and their mutual dependence on the underlying surface representation. We argue that if treated jointly, richer shape context features can be learned and an improved enhancement framework can be derived. Accordingly, we formulate the shape context description as a joint contribution by both denoising and consolidation, within an end-to-end framework. To this end, we introduce densely packed graph convolution layers to extract contextual information, allowing to query points offset to the underlying surface and to compensate for the structural loss. We demonstrate how the commonly used L_2-driven loss functions generate non-smooth output and volume shrinkage, and alleviate this by ones that mitigate the noisy outcome, repair voids, and improve point density distributions. Performance analysis on benchmark datasets demonstrates how we outperform state-of-the-art solutions, produce high-fidelity outcomes, and improve reconstruction-based tasks in real-world setups.

1. Introduction

Low-cost mobile and handheld portable laser scanners (MLS & PLS) provide means for rapid 3D modeling of large-scale sites, offering an economical solution for numerous modeling-related challenges (Conti et al., 2024; Tupinambá-Simões et al., 2024). Notwithstanding, the resulting point clouds offer a sparse and noisy representation of the underlying 3D surface and feature an excessive amount of outliers, compared to their stationary terrestrial laser scanner (TLS) counterparts (Wang and Feng, 2016; Ren et al., 2022; Li et al., 2024). In addition, due to data acquisition characteristics, the pointsets are usually unevenly sampled and feature void regions within the scanned scene (Zhang and Filin, 2022; Xia et al., 2023). This reduced data quality seriously hinders the development of downstream tasks, affecting applications such as primitive estimation, semantic interpretation, rendering, and architectural modeling (Cui et al., 2019; Xu et al., 2022; Antova, 2024).

The literature shows that point cloud denoising and consolidation (addressing uneven distribution and existence of void region), are traditionally approached as two independent tasks, where the former involves computing a per point offset to the actual surface, and the latter predicting coordinates of missing points in sparse sampled regions and voids. Considering denoising, common solutions introduced the underlying surface as a constraint in forms of the, L_1-median operator, graph Laplacian framework, and surface proximity-based weighting (Huang et al., 2013; Sun et al., 2015; Digne and De Franchis, 2017; Mattei and Castrodad, 2017; Dinesh et al., 2020). More recently deep learning (DL) based approaches predicted the displacement through an Euclidean multi-scale feature concatenation, or by learned global features via pooling (Pistilli et al., 2020; Rakotosaona et al., 2020; Luo and Hu, 2021; Edirimuni et al., 2023b). Analysis shows that the learned surface description neither handles high noise levels nor the presence of outliers, though attempts have been made to adapt to unknown noise levels (e.g., Pistilli et al., 2020). The reliance on Euclidean measures also led to failure near void regions, or when the point density varied (e.g., Liu et al., 2023). To address uneven density and voids, DL-based solutions proposed concatenation and replication of multi-scale per point features followed by point coordinate prediction, or by predicting coarse anchor points in the missing region followed by detail refinement to best-fit ground truth (Yu et al., 2018; Yuan et al., 2018; Li et al., 2021; Cui et al., 2023). While proving successful in some applications, they required noise-free data input and corresponding class label priors (Yu et al., 2021). They were also challenged by local structural variations of missing regions, often creating spurious content and noisy output.

In practice, real-world data, as MLS & PLS provide, are noisy and consist of high outlier rates, all while featuring voids and uneven point distribution. Our analysis demonstrates how standalone denoising models lead to volume shrinkage around incomplete regions, only to exacerbate voids and structural incompleteness (Sec. 4). Conversely, using data consolidation models often leaves residual noise, while a sequential application of the two tends to introduce further artifacts (Sec. 4). Addressing this, we propose in this paper to treat these two facets simultaneously as a single task. Performance analysis shows how our framework sets state-of-the-art results in denoising and consolidation when tested on benchmark data and applied to real-world scans. The contributions of our framework are: *i*) a joint denoising and consolidation neural framework that also filters outliers. As the paper demonstrates both tasks can be learned jointly through a single shape context representation; *ii*) improved loss modeling, which emphasizes smooth and complete outcomes rather than only predicting a per-point operation; *iii*) computational effective formulation that reduces the feature extraction and loss evaluation computational complexities; and *iv*) improved outcome compared to an individual application.

* Corresponding author

2. Related Work

Pistilli et al. (2020) were of the first to study DL-based 3D point cloud denoising solutions. The authors utilized a dynamic graph convolution network (DGCNN, Wang et al., 2019) followed by the prediction of the noise-free coordinates. Minimization of an L_2 loss with respect to the ground truth, has led to a blur of geometric details. Rakotosaona et al. (2020) denoised each point independently and detected outliers using the PointNet architecture (Qi et al., 2017). Offsets to noisy points were predicted iteratively, using a similar loss as in Pistilli et al. (2020). Realizing that the predicted offsets are imprecise, Luo and Hu (2021) modeled the noise distribution and identified the clean surface as the region of highest density. During training, the network learned a direction per point towards high-density regions, but during testing the displacement was iterative. Hence, inexact vector predictions led to slow convergence and long inference time. For faster convergence, Edirimuni et al. (2023b) replicated their denoising module four times where weights were learned per module. Their network was DGCNN based, with each intermediate result being supervised by ground truth contaminated by reduced levels of noise until becoming noise-free. Here again, corrections were performed independently through minimization of the L_2 norm. Edirimuni et al. (2023a) used the PointNet architecture to jointly denoise points and estimate normals. Their approach was patch-based, where identical features were learned for all central points of all patches under different noise levels. This was followed by a multilayer perceptron (MLP) predicting the normal and filtered coordinates per central point. The prediction was, again, L_2-based, and training was slow as each point was processed independently. To preserve sharp features, Liu et al. (2023) added precomputed normals by the principle component analysis (PCA) to the input data. Here again, a patch-based network was utilized, where the normals helped identify relevant neighbors for the center point during feature extraction. This was followed by normal correction and point displacement for the center point. This patch-based approach is memory-intensive and sensitive to the quality of the estimated normals. Unreliable computed normals, as with data featuring excessive levels of noise, led this framework to fail in securing clean results.

A common, though implicit, assumption of these denoising frameworks is of an even point density. Hence, they tend to suffer from volume shrinkage and distortion when facing missing regions in the data (Edirimuni et al., 2023b). To enhance the point density distribution and fill in voids in sparse scans, upsampling and completion methods have been proposed for point cloud consolidation. Yu et al. (2018) concatenated per-point features at multiple Euclidean scales and replicated them r times to predict a dense pointset. A combination of an L_2 loss with respect to the ground truth, and a point cloud uniformity loss, ensured dense and even output, but the correlation between the replicated features has led to an over-smooth outcome and loss of details. Li et al. (2021) used DGCNN based prediction strategy where a coarse-upsampled pointset ensured surface coverage, which was then refined by predicting offsets to better fit the ground truth. This framework required noise-free input with a small number of voids, making it less useful with actual data. To address structural loss for shapes from the same class, Yuan et al. (2018) generated a coarse pointset as an initial shape approximation, and refined it by deforming a local 2D grid around each coarse point to produce a detailed output. Supervised by the Chamfer distance to the ground truth, both the coarse and refined predictions discarded the original observations and retained noise. Liu et al. (2020) utilized 16 surface decoders to complete separate parts of the shape, combined their prediction through minimum density sampling, and concatenated it with the original pointset. A minimization of the Earth mover's distance (EMD) with respect to ground truth data facilitated noise reduction, but the model's reliance on a fixed number of decoders limited its capacity to handle broader scenes. Yu et al. (2021) introduced a transformer to improve the completion feature extraction, where a sparse set of points from the original pointset and its features, learned through graph convolution, acted as the transformer input. Following Yuan et al. (2018), the network predicted a coarse pointset and refined it by learning local grid deformation. Though effective on synthetic datasets, the dependence on paired complete ground truth for training impeded their application to real-world scans. Cui et al. (2023) addressed this dependence by introducing a self-supervised training framework that as input masked random regions of partial scans and predicted the complete scan. To ensure smoothness, it derived normals from the completion prediction and minimized the cosine distance between neighbors. The training was memory-intensive due to the PCA related computations of the local covariance matrices.

To sum up, the review shows that denoising approaches tend to ignore the effect of data incompleteness and uneven distribution. Ignoring their presence leads to volume shrinkage and structural incompleteness (Luo and Hu, 2021; Edirimuni et al., 2023b). Conversely, using current consolidation methods requires prior information about object parts and class labels to complete missing observations, which are always challenging to obtain (Yuan et al., 2018; Yu et al., 2021). Though using noise-free data as input, these consolidation applications create spurious structures and non-smooth output, limiting their application to P- and MLS scans which are noisy and sparse (Li et al., 2021; Cui et al., 2023). It can be assumed that both enhancement methods can apply sequentially, but we show that such a framework is not optimal. Therefore, we propose to learn an underlying surface representation that jointly treats outlier, noise, and data consolidation in an end-to-end framework.

3. Methodology

Our modeling considers the following point cloud formation representation:

$$P' = \{p_i + d_i\}_{p_i \in P} \cup O, \tag{1}$$

where P' is the observed noisy point cloud, P is the set of perfect surface samples, such that $p_i \in P$ lies on the scanned surface, $d_i \in D$ is the offset from the surface due to ranging noise, and O is the set of outlying points. To account for shape incompleteness due to the acquisition process, we assume P' features a structural loss. Our proposed framework aims to yield a clean surface, P, through denoising and data consolidation that yields added points to complete the shape.

3.1 Point Representation Encoder

To share information between the noise and outlier treatment and the data consolidation, we build upon our earlier dynamic graph convolution-based framework (Zhang and Filin, 2022), which models long-range contextual relationships and extracts shape description per point. The shape context description is obtained through a stack of densely connected graph convolution layers, forming a shared encoder (Fig. 1). Our dense block

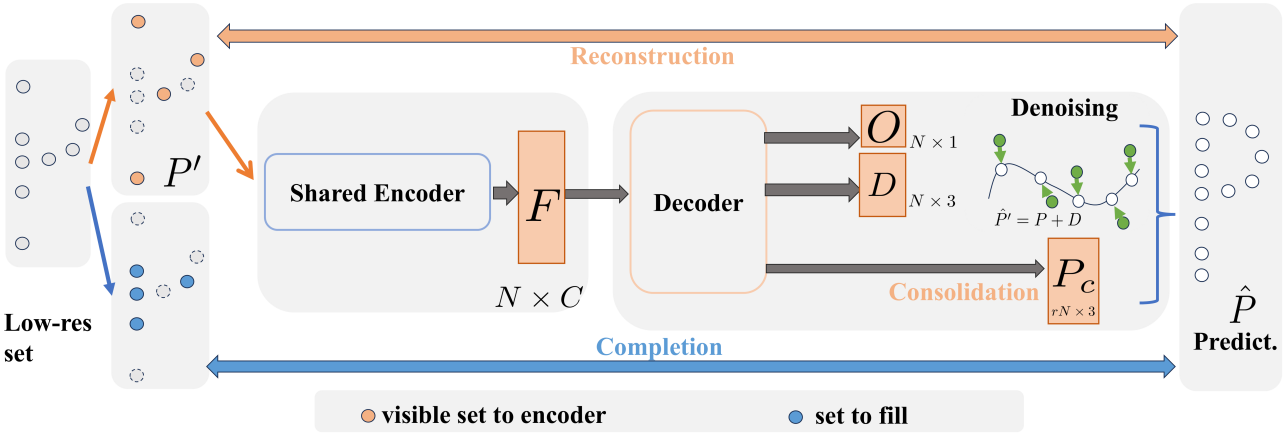

Figure 1. Overview of the proposed joint denoising and consolidation network for mobile laser scan.

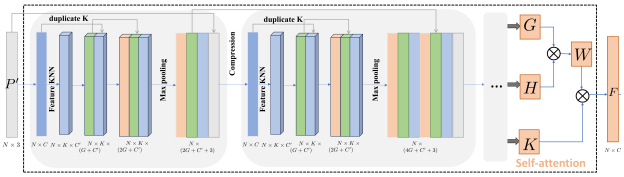

Figure 2. Point representation encoder.

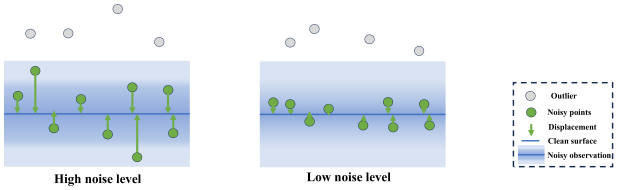

Figure 3. Outlier detection and denoising module.

convolves a point feature, f_i, with its neighbors, $f_j \in \mathcal{N}_i$,

$$e_{ijm} = \text{ReLU}\left(\theta_m \cdot (f_j - f_i) + \phi_m \cdot f_i\right), \tag{2}$$

where $\mathcal{N}_i$ is a set of neighbors of the i-th point, and θ_m and ϕ_m are MLPs. The final activation x_{im} at the m-th layer is:

$$x_{im} = \max_{j \in \mathcal{N}_i} \text{MLP}\left(\text{MLP}(e_{ijm}) \,\|\, e_{ijm}\right), \tag{3}$$

To avoid vanishing gradient and over-fit of local connections due to stacking within blocks, each MLP's output is passed to all subsequent MLPs via skip-connections (Fig. 2). Also, between blocks, point features produced by each block are fed as an input, again, through skip-connections, to all blocks that follow. The skip-connections ensure gradient flow and because of the compression performed at each block end, they significantly reduce the model size (Fig. 2). For global context and over-fitting avoidance, a self-attention unit, which refines the densely connected blocks' output is added. The learned output per point feature (shape context) is denoted F_i.

3.2 Denoising and Outlier Filtering

Using an MLP whose input is F_i, we classify each point as either an outlier or inlier, $\text{O}_i = \text{MLP}\left(F_i\right)$, where O_i is an indicator. F_i is concurrently used to estimate offsets for the outlier-free set, $\tilde{P}' = P' \setminus O$, by predicting displacements, d_i, to the true surface (Fig. 3):

$$d_i = \text{Score}\left(F_i\right), \tag{4}$$

where $\text{Score}(\cdot)$ is an MLP as in Luo and Hu (2021). Using d_i, we compute a noise-free pointset, $\tilde{P}$. To ensure consistent performance, we add during training to the clean sampled points varying levels of random Gaussian, Laplace, and non-uniform noise. Notably, the tendency of volume shrinkage around voids is addressed by the consolidation module which we now describe.

3.3 Data Consolidation

To consolidate the data by addressing unevenly sampled regions and voids, our decoder learns a mapping from the local 2D unit grid, $[0, 1]^2$, centered around each data point, to the 3D surface. This mapping is performed by an MLP, which mimics the morphing of the 2D square into a local 3D surface. This is done by creating r replica of F_i, and concatenating independently, the position of the 2D unit grid points, g^j_{xy}, $j = [1, r]$:

$$P^{ij}_c = \text{MLP}\left(F_i || g^j_{xy}\right), \quad \forall j = [1, r] \tag{5}$$

where P^{ij}_c is a predicted point, and $P^i_c \in P_c$ is the set of completion predictions around each point. This upsampling is supplemented by a designated loss (Sec. 3.4) responsible for the point repositioning to consolidate the data. Note that the locality of our grid morphing facilitates using only a single MLP, rather than the conventional two folding operations (Liu et al., 2020; Cui et al., 2023).

Existing consolidation models train their network by pairing scans from a fixed set of views of the ground truth data. This setting cannot represent the actual structural loss from the sampling process (Yu et al., 2021). Here, we adapt to information loss under different views, by randomly dropping observations from the noisy shape to obtain the input P' to our network. We use the farthest point sampling (FPS) to sample M points as region centers $\mathcal{C} = \{c_i\}^M_{i=1}$ from the complete shape. Then, we gather the k-nearest Euclidean neighbors for each point in $\mathcal{C}$ to obtain a region disk $\mathcal{G}_i = \{p | p \in \mathcal{N}^{c_i}_k\}$, where $\mathcal{N}^{c_i}_k$ denotes the set of neighbors for c_i. We partition the set of disks into two groups, G^{rec}, G^{com}, with ratio $r_1 : r_2$, where G^{rec} is the observable region for the network, and G^{com} is the group of points

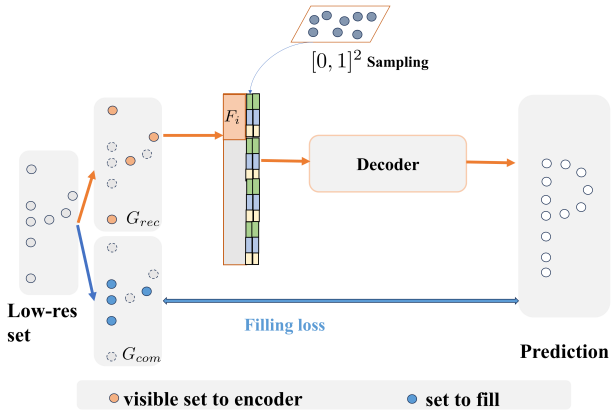

Figure 4. Our data consolidation module.

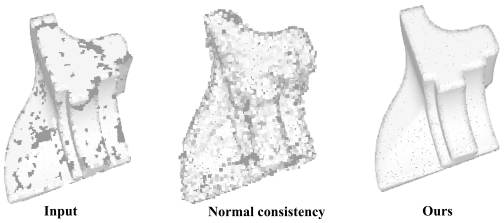

Figure 5. Our loss compared with the P2C.

to complete. We force the network to generate a consolidated shape P_c that preserves regions in G^{com}.

3.4 Loss Design

Our predicted, noise-free consolidated shape, is the union of the denoised point set, $\tilde{P}$, and the consolidation prediction, P_c, $\hat{P} = \tilde{P} \cup P_c$. To enable $\hat{P}$ to approximate the underlying surface, it is customary to employ the Chamfer distance (CD) as a metric, but the CD tends to generate a noisy output (Liu et al., 2020; Cui et al., 2023). Therefore, we introduce the EMD to measure the network prediction:

$$L_{EMD}(\hat{P}, P) = min_{\phi:\hat{P}\to P} \frac{1}{|\hat{P}|} \sum_{\hat{p}_i \in \hat{P}} ||\hat{p}_i - \phi(\hat{p}_i)||, \quad (6)$$

where, P, is the ground truth, and ϕ is a bijective mapping function from the network prediction, $\hat{P}$, to P. As the computation of bijective mapping is iterative, we also introduce here an efficient approximation algorithm to obtain ϕ by solving the linear assignment problem (Liu et al., 2020). The EMD forces the output to have the same density distribution as the ground truth and is thus more discriminative to the local details and the density distribution. To measure the performance of the outlier detection, we use a cross-entropy loss that evaluates the labels in O:

$$L_O = CrossEntropy(O, y), \quad (7)$$

where y is the ground truth labels set. To secure the smoothness of $\hat{P}$, we opt to minimize the Dirichlet energy, which measures the stretch induced by the performed mapping, hence securing a smooth map as the functional minimizer. The Dirichlet energy minimization expression in the current discrete form becomes,

$$L_D(\hat{P}) = \sum_{(\hat{p}_i, \hat{p}_j \in \xi(\hat{P}))} ||\hat{p}_i - \hat{p}_j||^2 \quad (8)$$

where $\xi(\hat{P})$ is the set of edges on $\hat{P}$ from the k-nearest-neighbors graph. It differs from the normal consistency evaluation (Cui et al., 2023), which is slow due to the cubic complexity of normal evaluation and also exhibits instability (c.f. Fig. 5). Compared to normals derived from second-order measures, our minimization is based on first-order ones, less sensitive to prediction noise. Our final loss is a linear combination of the individual losses $L = L_O + L_{EMD} + L_D$.

4. Results

4.1 Dataset and Implementation

We test our model performance by the: PU-Net (PU, 40 shapes, Yu et al., 2018) and PointCleanNet (PC, 10 shapes, Rakotosaona et al., 2020) benchmark datasets. We use the PU 20 shapes train subset and apply the Poisson disk to sample points from the meshes with resolutions ranging from 10K to 50K points (Yu et al., 2018). Before being fed into the model the point clouds are partitioned into 1K points patches (as in e.g., Luo and Hu, 2021; Liu et al., 2023). To introduce structural loss, we partition the input shape into non-overlapping disks and randomly drop 20% of them. Following the convention, during training the data is contaminated with Gaussian random noise with standard deviations (std.) between $0.5 - 2\%$ of the shape's bounding sphere radius. During testing, the Gaussian noise std. ranged from $1 - 3\%$ of the shape's bounding sphere radius.

Our real-world datasets consist of the widely used MLS Paris-Rue-Madame database (Serna et al., 2014) that features complex structures, outliers, and platform motion-induced noise. As no ground truth reference data is unavailable, only qualitative results are presented, in similarity to, e.g., Luo and Hu (2021); Liu et al. (2023) and Edirimuni et al. (2023b). In addition a dataset featuring complex architectural and richly decorative elements, and acquired by the GeoSLAM ZEB-REVO handheld scanner, was used for testing. It is noisy with an excessive amount of outliers, of low-density, and uneven point distribution.

Baselines and Metrics Our model is compared with state-of-the-art quality enhancement approaches including the graph Laplacian-based denoising (GLR, Dinesh et al., 2020), the learning-based model, Score-denoise (Score, Luo and Hu, 2021), the completion model, P2C (Cui et al., 2023), and the sequential application of the two models (Score + P2C). To measure the model's performance, we use the CD and point-to-mesh distance (P2M), two common quantitative metrics. As the point clouds size vary, we normalize the denoised results, before assessment, into the unit sphere.

4.2 Model Analysis

Evaluation of our network performance against the baselines demonstrates robustness even at the unseen noise levels of 3% (Fig. 6 & Table 1). Visual inspection of the complex curved structure of significant voids, e.g., the icosahedron and camel models, shows how the underlying structure becomes discontinuous, almost indistinguishable under heavy noise. For missing data and under this level of noise, the GLR collapsed and failed to provide denoising output. Quantitatively, our CD/P2M values are: $3.870/1.646$ on the PU dataset, representing $38.7\%/42.8\%$ improvement over Score

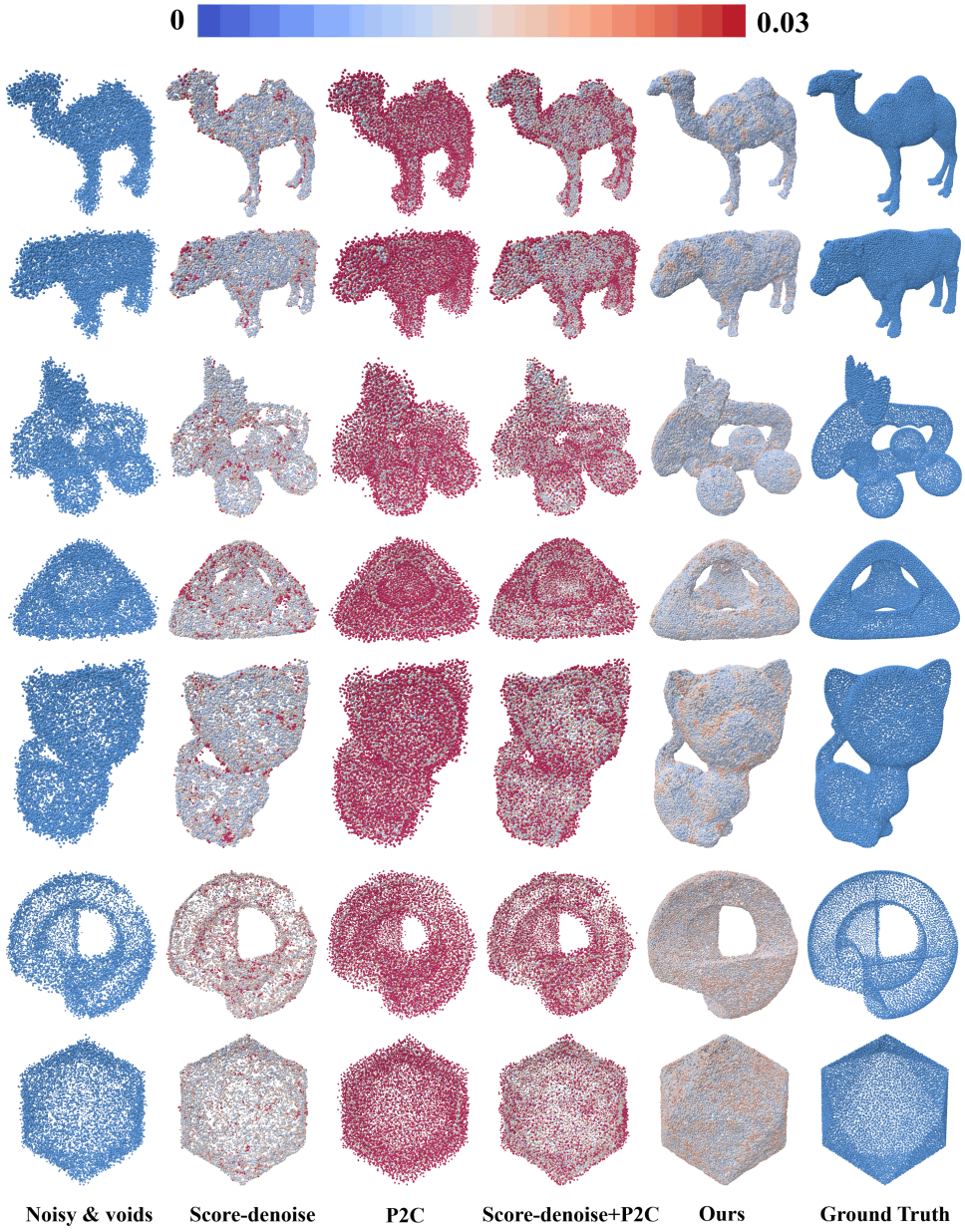

Figure 6. Data quality enhancement of the PU & PC sets. Note our improved density distribution also revealing underlying surface.

Dataset	Metric	Bilateral	Jet-denoising	MRPCA	GLR	Score-denoise	P2C	Score & P2C	Ours
PU	CD ($\times 10^4$)	6.998	6.262	5.009	-	6.313	34.883	12.954	**3.870**
	P2M ($\times 10^4$)	3.557	2.921	1.963	-	2.879	30.091	9.411	**1.646**
PC	CD ($\times 10^4$)	8.296	7.650	6.502	-	5.908	37.398	15.808	**3.811**
	P2M ($\times 10^4$)	2.393	2.227	1.676	-	2.311	23.892	7.962	**1.614**

Table 1. Comparison to state-of-the-art methods (c.f. text for references), CD and P2M values are factored by 10^4.

(6.313/2.879), and 70.5%/82.9% over their sequential application (12.954/9.411). We conjecture that both approaches either treat noise components as high-frequency features and thereby leading to errors, or cannot compensate for the structural loss on noisy input. As Fig. (6) shows, the Score, and P2C, failed to repair voids and recover the underlying geometry, resulting in significant residual noise. Even the sequential application, Score + P2C, led to a completely noisy outcome. Notably, applying the reverse order, P2C + Score, has led to complete failure, due to the noisy data that the P2C could not handle. In contrast, ours is designed to robustly capture geometric structures across varying noise levels and voids, effectively producing state-of-the-art results that closely match the ground truth. Similar improvements were recorded on the PC dataset when testing our network (Table 1).

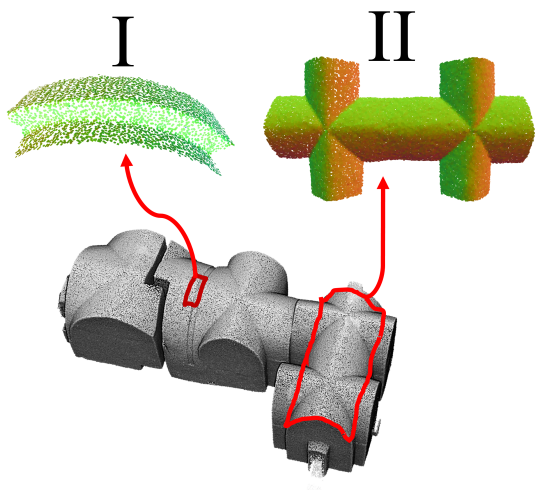

Figure 7. PLS datasets and study subsets colored by normals.

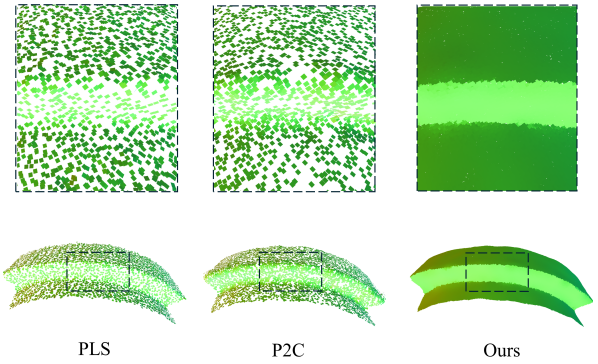

Figure 8. Consolidation results on subset #I.

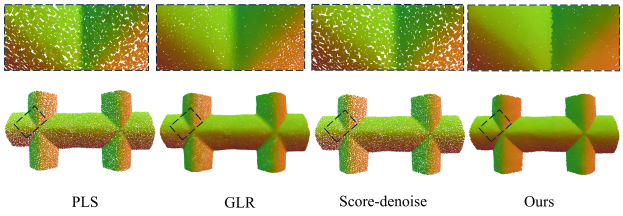

Figure 9. Consolidation results, subset #II. Coloring by normals.

4.3 MLS and PLS Denoising

We demonstrate our network performance on two subsets of the low quality PLS data (Fig. 7). Note the uniform point distribution with the high fidelity surface representation it generates on subset #I. Compared to the P2C – state-of-the-art consolidation model, the improvement is evident. This improvement stems from the joint shape context representation of the underlying surface, embodying both denoising and completion facets. Testing on the symmetric groin vault, a combination of the complex curved surface and their intersections (Fig. 7, subset #II), the GLR and Score either over-smoothed the data or created voids after denoising. Ours, in contrast, denoised the pointset while revealing the underlying surface features (Fig. 9). Solely minimizing the L_2 loss, as in the GLR and Score, has led to blurred details and artifacts. Our joint enhancement form not only denoised the data but also consolidated it in sparse areas,

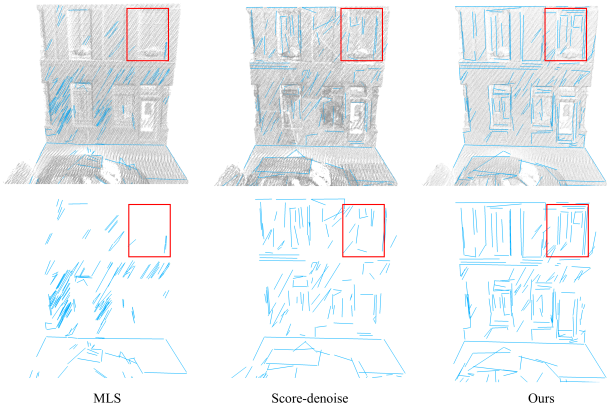

Figure 10. Line-based reconstruction using raw scan, score-denoise, and our results.

effectively recovering surface features and generating a uniform point cloud distribution.

Contribution to downstream application Using the MLS benchmark, we tested the contribution of our network for the line-based reconstruction model by Lu et al. (2019). Applying the model directly to raw MLS data led to erroneous reconstruction that mirrored the scanning pattern rather than the underlying geometric features (Fig. 10). Applying Score partially restored some geometric entities. However, the overall reconstructed structure echoed the significant noise levels in the data. Our results demonstrate how we follow closely the definition of building outlines, and enhance the quality of the reconstruction (Fig. 10). This is so due to more even density distribution and recovering finer details through an effective consolidation process.

5. Conclusions

Neural point cloud denoising and consolidation approaches present promising solutions for enhancing data obtained by low-cost scanners. In this paper, we presented a novel learning-based framework that jointly addresses noise, uneven density distribution, and structural voids to improve overall data quality for downstream applications. Through this joint formulation, our network effectively captures robust surface descriptors across varying noise levels, density fluctuations, and voids. Through dynamic graph convolution, the network learns intricate shape contexts, enabling noise-free structure recovery by applying corrective offsets to noisy observations. For consolidation, the network compensates for potential structural loss by a local 2D-3D grid deformation around each center point, effectively filling voids. Our enhanced loss formulation promotes smoothness by minimizing the Dirichlet energy and improves density distribution by the EMD loss. Experiments on reference benchmark datasets have confirmed the efficacy of our model in both denoising and consolidation, also showing substantial quality improvements of mobile and handheld laser scans. We have demonstrated enhanced performance in the line-based reconstruction application, where our model attenuates the point distributions, typical of mobile scans, yielding significant improvements in the retrieval of geometric line layouts.

References

Antova, G., 2024. Portable laser scanning solutions for 3D modelling of large buildings. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 13–19.

Conti, A., Pagliaricci, G., Bonora, V., Tucci, G., 2024. A comparison between terrestrial laser scanning and hand-held mobile mapping for the documentation of built heritage. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 141–147.

Cui, R., Qiu, S., Anwar, S., Liu, J., Xing, C., Zhang, J., Barnes, N., 2023. P2c: Self-supervised point cloud completion from single partial clouds. *Proc. of ICCV*, 14351–14360.

Cui, Y., Li, Q., Yang, B., Xiao, W., Chen, C., Dong, Z., 2019. Automatic 3-D reconstruction of indoor environment with mobile laser scanning point clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(8), 3117–3130.

Digne, J., De Franchis, C., 2017. The bilateral filter for point clouds. *Image Processing On Line*, 7, 278–287.

Dinesh, C., Cheung, G., Bajić, I. V., 2020. Point cloud denoising via feature graph laplacian regularization. *IEEE Transactions on Image Processing*, 29, 4143–4158.

Edirimuni, D., Lu, X., Li, G., Robles-Kelly, A., 2023a. Contrastive learning for joint normal estimation and point cloud filtering. *IEEE Transactions on Visualization and Computer Graphics*.

Edirimuni, D., Lu, X., Shao, Z., Li, G., Robles-Kelly, A., He, Y., 2023b. Iterativepfn: True iterative point cloud filtering. *Proc. of CVPR*, 13530–13539.

Huang, H., Wu, S., Gong, M., Cohen-Or, D., Ascher, U., Zhang, H., 2013. Edge-aware point set resampling. *ACM transactions on graphics*, 32(1), 1–12.

Li, R., Li, X., Heng, P.-A., Fu, C.-W., 2021. Point cloud upsampling via disentangled refinement. *Proc. of CVPR*, 344–353.

Li, X., Lu, J., Ding, H., Sun, C., Zhou, J. T., Chee, Y. M., 2024. Pointcvar: Risk-optimized outlier removal for robust 3d point cloud classification. *Proc. of the AAAI Conference on Artificial Intelligence*, 38(19), 21340–21348.

Lipman, Y., Cohen-Or, D., Levin, D., Tal-Ezer, H., 2007. Parameterization-free projection for geometry reconstruction. *ACM Transactions on Graphics*, 26(3), 22–es.

Liu, M., Sheng, L., Yang, S., Shao, J., Hu, S.-M., 2020. Morphing and sampling network for dense point cloud completion. *Proc. of AAAI*, 34(7), 11596–11603.

Liu, Z., Zhao, Y., Zhan, S., Liu, Y., Chen, R., He, Y., 2023. PCDNF: Revisiting learning-based point cloud denoising via joint normal filtering. *IEEE Transactions on Visualization and Computer Graphics*.

Lu, X., Liu, Y., Li, K., 2019. Fast 3D line segment detection from unorganized point cloud. *arXiv preprint arXiv:1901.02532*.

Luo, S., Hu, W., 2021. Score-based point cloud denoising. *Proc. of ICCV*, 4583–4592.

Mattei, E., Castrodad, A., 2017. Point cloud denoising via moving rpca. *Computer Graphics Forum*, 36(8), Wiley Online Library, 123–137.

Pistilli, F., Fracastoro, G., Valsesia, D., Magli, E., 2020. Learning graph-convolutional representations for point cloud denoising. *ECCV*, Springer, 103–118.

Qi, C. R., Su, H., Mo, K., Guibas, L. J., 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. *Proc. of CVPR*, 652–660.

Rakotosaona, M.-J., La Barbera, V., Guerrero, P., Mitra, N. J., Ovsjanikov, M., 2020. Pointcleannet: Learning to denoise and remove outliers from dense point clouds. *Computer graphics forum*, 39(1), Wiley Online Library, 185–203.

Ren, J., Pan, L., Liu, Z., 2022. Benchmarking and analyzing point cloud classification under corruptions. *International Conference on Machine Learning*, PMLR, 18559–18575.

Serna, A., Marcotegui, B., Goulette, F., Deschaud, J.-E., 2014. Paris-rue-madame database: A 3d mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods. *4th ICPRAM*.

Sun, Y., Schaefer, S., Wang, W., 2015. Denoising point sets via L0 minimization. *Computer Aided Geometric Design*, 35, 2–15.

Tupinambá-Simões, F., Pascual, A., Guerra-Hernández, J., Ordóñez, C., de Conto, T., Bravo, F., 2024. Accuracy of tree mapping based on hand-held laser scanning comparing leaf-on and leaf-off conditions in mixed forests. *Journal of Forestry Research*, 35(1), 93.

Wang, Y., Feng, H.-Y., 2016. Effects of scanning orientation on outlier formation in 3D laser scanning of reflective surfaces. *Optics and Lasers in Engineering*, 81, 35–45.

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., Solomon, J. M., 2019. Dynamic graph cnn for learning on point clouds. *Acm Transactions On Graphics*, 38(5), 1–12.

Xia, Z., Liu, Y., Li, X., Zhu, X., Ma, Y., Li, Y., Hou, Y., Qiao, Y., 2023. Scpnet: Semantic scene completion on point cloud. *Proc. of CVPR*, 17642–17651.

Xu, L., Gong, J., Na, J., Yang, Y., Tan, Z., Pfeifer, N., Zheng, S., 2022. Shield tunnel convergence diameter detection based on self-driven mobile laser scanning. *Remote Sensing*, 14(3), 767.

Yu, L., Li, X., Fu, C.-W., Cohen-Or, D., Heng, P.-A., 2018. Pu-net: Point cloud upsampling network. *Proc. of CVPR*, 2790–2799.

Yu, X., Rao, Y., Wang, Z., Liu, Z., Lu, J., Zhou, J., 2021. Pointr: Diverse point cloud completion with geometry-aware transformers. *Proc. of ICCV*, 12498–12507.

Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M., 2018. Pcn: Point completion network. *IEEE 3DV*, 728–737.

Zhang, T., Filin, S., 2022. Deep-Learning-Based Point Cloud Upsampling of Natural Entities and Scenes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 321–327.