Inlier Confidence Calibration for Point Cloud Registration

Supplementary Material

A. Overview

In the supplementary material, we include additional experimental details and present results of further ablation studies. Furthermore, we discuss some future prospects.

B. Code Release

Due to reasons that we cannot discuss without violating anonymity, we are unable to release code at this time. All code necessary to reproduce the results in this paper will be released prior to publication. Our dataset will also be released under a Creative Commons license prior to publication. Our datasets are public, and contain no identifiable personal data or human subjects.

C. Network Configurations

C.1. Implementation Details

We implement and evaluate our model with PyTorch on a NVIDIA RTX 3090 GPU. The network is trained with Adam optimizer for 100 epochs on ModelNet40, Augmented ICL-NUIM and 7Scenes and the batch size is 4. The learning rate starts from 0.0001 and decays when loss has stopped improving. The number of iterations $N_l = 2$ without shared weights among every iteration.

C.2. Dataset

We evaluate the proposed method on synthetic datasets ModelNet40 and Augmented ICL-NUIM, and real-world dataset 7Scenes. ModelNet40 contains 12,308 CAD models of 40 different object categories. Augmented ICL-NUIM consists of 1,478 synthetic model generated by applying data augmentation on original 739 scan pairs. 7Scenes is a subset of 3DMatch and generally used dataset of indoor environment with 7 scenes including Chess, Fires, Heads, Office, Pumpkin, RedKitchen and Stairs.

C.3. Metrics

MIE($\mathbf{R}$) is the geodesic distance in degrees between estimated and ground-truth rotation matrices. It measures the differences between the predicted and the ground-truth rotation matrices. MIE($\mathbf{t}$) is the Euclidean distance between estimated and ground-truth translation vectors. It measures the differences between the predicted and the ground-truth translation vectors. Here, we give the specific calculation formula of MIE($\mathbf{R}$) and MIE($\mathbf{t}$):

$$\mathrm{MIE}(\mathbf{R}) = \arccos\left(\frac{\mathrm{trace}(\mathbf{R}_{gt}^{-1}\mathbf{R}_{est})}{2}\right), \tag{1}$$

$$\mathrm{MIE}(\mathbf{t}) = \|\mathbf{t}_{gt} - \mathbf{t}_{est}\|_2. \tag{2}$$

Algorithm 1 Registration algorithm

Require: Source point cloud $\mathcal{P}$ and reference point cloud $\mathcal{Q}$.
Ensure: Rigid transformation $\{\mathbf{R}, \mathbf{t}\}$.

Initilize rotation matrix $\mathbf{R}_0$ and translation vector $\mathbf{t}_0$.
for $t = 1 : N_l$ **do**
  Construct 1-NN point cloud $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{Q}}$;
  %Finely Initial Correspondences
  Extract local features $\mathcal{F}_{\mathcal{P}}$, $\mathcal{F}_{\mathcal{Q}}$, $\mathcal{F}_{\widehat{\mathcal{P}}}$ and $\mathcal{F}_{\widehat{\mathcal{Q}}}$ with Dynamic Graph CNN for $\mathcal{P}$, $\mathcal{Q}$, $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{Q}}$, respectively;
  Calculate Euclidean distance between the single-point features $\mathbf{D}_{i,j} = \|\mathcal{F}_{\mathbf{p}_i} - \mathcal{F}_{\mathbf{q}_j}\|_2$ and $\widehat{\mathbf{D}}_{i,j} = \|\mathcal{F}_{\widehat{\mathbf{p}}_i} - \mathcal{F}_{\widehat{\mathbf{q}}_j}\|_2$;
  Construct fused neighborhood matching map $\mathbf{M}_{i,j}$ and $\widehat{\mathbf{M}}_{i,j}$;
  Soft Assignment Matrix Optimaization;
  Perform negatively related and exponential strategy and obtain final matching map $\mathbf{F}_{i,j}$;
  %Inlier Confidence Calibration
  Generate reference point cloud copy $\widetilde{\mathcal{Q}}$ via $\mathbf{F}_{i,j}$;
  Construct learnable neighborhood graph;
  Calculate and calibrate the inlier confidence $\mathbf{w}_i$;
  Solve transformation $\{\mathbf{R}_{t-1}, \mathbf{t}_{t-1}\}$ using weighted SVD based on reliable inliers correspondence;
  Update source point cloud $\mathcal{P}$ with $\{\mathbf{R}_{t-1}, \mathbf{t}_{t-1}\}$;
end for

D. Additional Experiments

D.1. Robustness Against Guassion Noise

In this experiment, we further test the robustness against noise by evaluating our model on the unseen categories of the ModelNet40 dataset under different noise levels. As shown in Figure 1, the noises are sampled from $\mathcal{N}(0, \sigma)$ and clipped to $[-1.0, 1.0]$, where the deviation $\sigma \in [0.1, 0.5]$. Our model achieves comparable performance under various noise levels, which supports the conclusion that our method is insensitive to the noise.

D.2. Parameter Sensitivity Analysis

In this experiment, we discuss the optimal number of iteration N_l in the range of 1 to 4 so as to obtain the optimal model. The results are presented in Table 1. We can observe that the most performance gains are the second iter-

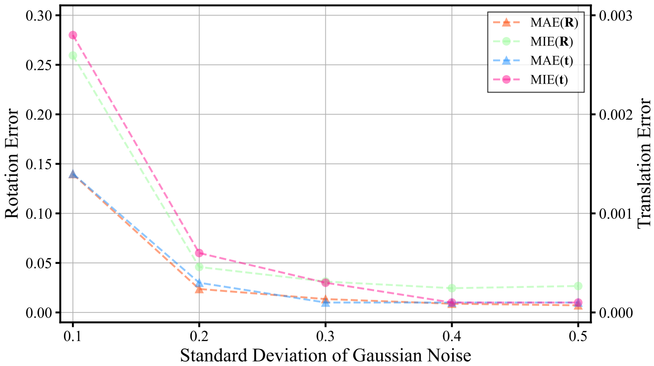

Figure 1. Rotation error and translation error under different noise levels.

ation, and we set number of iteration to 2 for the trade-off between the accuracy in all experiments.

N_l	MAE($\mathbf{R}$)	MAE($\mathbf{t}$)	MIE($\mathbf{R}$)	MIE($\mathbf{t}$)
1	0.0111	0.0001	0.0257	0.0003
2	**0.0012**	**0.0000**	**0.0192**	**0.0000**
3	0.0013	**0.0000**	0.0214	**0.0000**
4	0.0077	0.0001	0.0250	**0.0000**

Table 1. The registration results with different number of iterations. Bold indicates the best performance.

E. Comparison with Previous Work

In this paper, we explore to extend a new definition called inlier confidence calibration (ICC). Distinguished from previous efforts, our work embarks on a novel and unexplored direction, as the pursuit of overlapping regions becomes paramount for the partially overlapping point cloud registration. We approach this challenge from a fresh perspective, focusing on the exploration of reliable inliers and emphasizing their significance in point cloud registration. We argue the proposed ICC concept and soft assignment matrix optimization theorem still contribute valuable insights to the community.

F. Prospects

In this section, we discuss the potential opportunities and challenges of the proposed method. First, some recent works develops several enhanced strategies, e.g., KPConv-FPN backbone for feature extraction and Coarse-to-Finer correspondences refinement. Incorporating these strategies may further improve the performance of the proposed model. Second, this paper puts forth a scalable approach, addressing the issue of the misinterpretation in ICC caused by massive coordinate differences. However, the potential existence of more effective descriptors that further optimize the ICC poses a more promising challenge that requires further investigation and resolution in future research endeavors. Third, the optimization theory and ICC proposed in this paper are intuitively adaptable for integration into the majority of current networks, emerges as a challenge warranting further investigation.