SuP: Sub-cloud Driven Point Cloud Registration

Supplementary Material

7. Overview

The supplementary material includes:
- Evaluation metrics (Sec. 8);
- Additional experiments (Sec. 9);
- Visual comparison results (Sec. 10);

8. Metrics

We follow recent state-of-the-art evaluation protocols [8, 14] and report the Inlier Ratio (IR), Feature Matching Recall (FMR), and Registration Recall (RR) on the color-enhanced 3DMatch benchmark dataset [13, 31]. For fine-grained alignment performance, we further report the Relative Rotation Error (RRE) and Relative Translation Error (RTE) following the protocol in [12], computed over the correctly registered pairs identified by RR.

Inlier Ratio (IR) quantifies the quality of the predicted correspondences. It is defined as the proportion of inlier matches among all predicted matches. A match (x_i, y_i) is considered an inlier if the Euclidean distance between the transformed source point $T_g(x_i)$ and the target point y_i is less than a threshold $\tau_1 = 0.1\,\text{m}$, under the ground-truth transformation T_g:

$$\text{IR} = \frac{1}{|C_p|} \sum_{(x_i, y_i) \in C_p} [\![\|T_g(x_i) - y_i\|_2 < \tau_1]\!], \tag{11}$$

where C_p denotes the set of predicted correspondences, and $[\![\cdot]\!]$ is the Iverson bracket, which evaluates to 1 if the condition holds and 0 otherwise.

Feature Matching Recall (FMR) measures the fraction of point cloud pairs for which the IR exceeds a smaller threshold $\tau_2 = 0.05$. It reflects the potential of feature matching to yield successful registrations:

$$\text{FMR} = \frac{1}{N} \sum_{i=1}^{N} [\![\text{IR}_i > \tau_2]\!], \tag{12}$$

where N is the total number of point cloud pairs in the dataset.

Registration Recall (RR) evaluates the robustness of the overall registration pipeline. It is defined as the fraction of point cloud pairs for which the estimated transformation achieves an error $e_i < 0.2\,\text{m}$:

$$\text{RR} = \frac{1}{N} \sum_{i=1}^{N} (e_i < 0.2m), \tag{13}$$

where N is the number of the samples in the dataset. The transformation error e_i is defined as the Root Mean Square Error (RMSE) of the ground truth correspondence $\mathbb{C}^*$ after applying the estimated transformation:

$$e_i = \sqrt{\frac{1}{|\mathbb{C}^*|} \sum_{(\mathbb{C}_{xi}, \mathbb{C}_{yi}) \in \mathbb{C}^*} \|T_g(\mathbb{C}_{xi}) - \mathbb{C}_{yi}\|_2^2}, \tag{14}$$

where T_g is the ground-truth transformation that aligns the point cloud pair. To further assess fine registration quality, we report Relative Translation Error (RTE) and Relative Rotation Error (RRE) on successfully registered pairs.

RTE is defined as the Euclidean distance between the predicted and ground-truth translation vectors:

$$\text{RTE} = \|t_p - t_g\|_2, \tag{15}$$

where t_p is the predicted translation vector and t_g is the ground truth translation vector. Relative Rotation Error (RRE) measures the differences between the ground truth and the predicted rotation matrices. It is defined as follows:

$$\text{RRE} = \arccos\left(\frac{\text{tr}(R_p^T \cdot R_{gt} - 1)}{2}\right), \tag{16}$$

where $\text{tr}(\cdot)$ is the trace operator. R_p and R_{gt} are the predicted rotation matrix and ground truth rotation matrix, respectively.

9. Additional Experiments

9.1. Number of sub-clouds

Here we fix the subdivision threshold at its default ($\tau = 0.35$) and vary the number of sub-clouds $k \in \{4, 5, 6, 7\}$.

	C3DM			C3DLM		
	RRE (°)	RTE (m)	RR(%)	RRE (°)	RTE (m)	RR(%)
$k = 4$	1.476	0.055	97.4	2.528	0.080	89.6
$k = 5$	1.375	0.053	97.7	2.541	0.075	89.8
$k = 6*$	**1.374**	**0.046**	**97.8**	**2.493**	**0.074**	**90.2**
$k = 7$	1.397	0.054	**97.8**	2.508	0.078	90.0

Table 7. Evaluation results on C3DM and C3DLM with different numbers of sub-clouds, with the top results highlighted in bold. The default value is marked with $*$.

Table 7 shows that too few sub-clouds (e.g. $k = 4$) leads to larger rotation and translation errors, while too many ($k = 7$) offers no additional benefit. At $k = 5$, translation

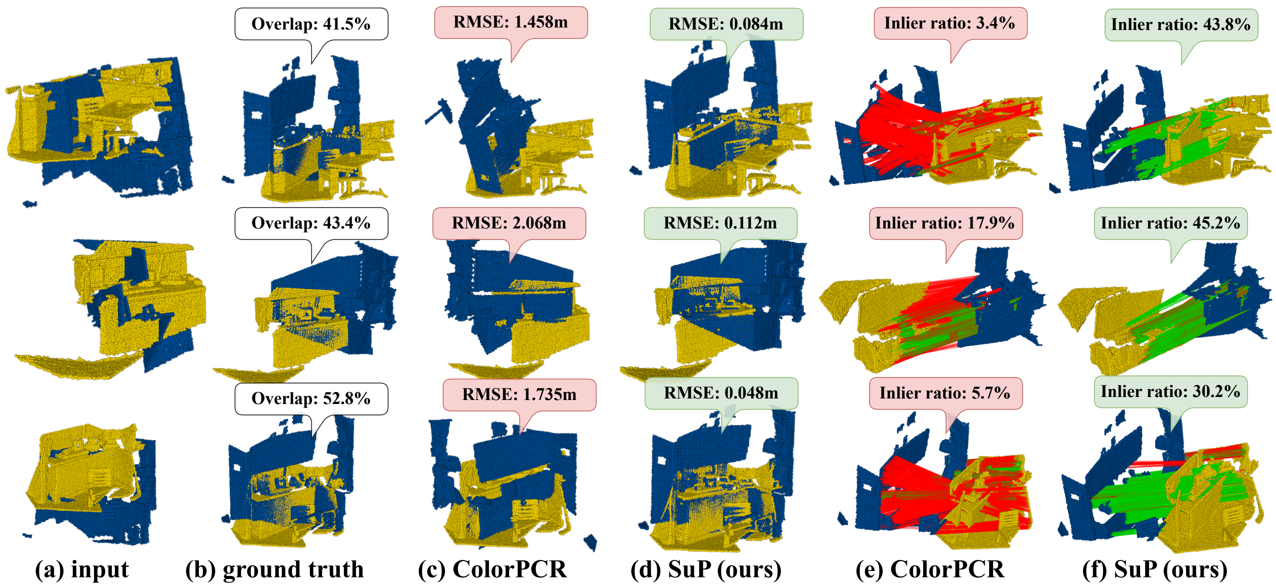

Figure 4. Qualitative comparison of SuP and ColorPCR on low-overlap color-enhanced 3DMatch pairs. ColorPCR produces incorrect alignments due to spurious matches in non-overlapping regions, whereas SuP's sub-cloud strategy can identify the anchor pairs with high-overlap ratios and thus yields more inlier correspondences.

error on C3DM improves (0.053 m) and recall climbs to 97.7%, but C3DLM performance remains suboptimal. The sweet spot is $k = 6$: it achieves the lowest errors (RRE = 1.374°, RTE = 0.046 m) and highest recall (97.8%) on C3DM, and simultaneously the best metrics on C3DLM (RRE = 2.493°, RTE = 0.074 m, RR = 90.2%). We therefore adopt $k = 6$ as our default.

9.2. Anchor-pair selection threshold

Finally, we examine the impact of the feature-consensus weight threshold τ_a, which controls how many candidate anchor pairs are retained for final pose estimation.

	C3DM			C3DLM		
	RRE (°)	RTE (m)	RR(%)	RRE (°)	RTE (m)	RR(%)
$\tau_a = 0.6$	1.430	0.058	97.7	2.581	0.092	89.6
$\tau_a = 0.7$	1.397	0.052	97.6	2.548	0.082	89.9
$\tau_a = 0.8*$	**1.374**	**0.046**	**97.8**	**2.493**	**0.074**	**90.2**
$\tau_a = 0.9$	1.402	0.048	97.5	2.511	0.081	89.9

Table 8. Evaluation results on C3DM and C3DLM with different anchor pair selection thresholds, with the top results highlighted in bold. The default value is marked with $*$.

Results for $\tau_a \in \{0.6, 0.7, 0.8, 0.9\}$ appear in Table 8. A low threshold ($\tau_a = 0.6$) includes noisy correspondences, yielding higher errors (RRE = 1.430°, RTE = 0.058 m), albeit with good recall (97.7%). Raising τ_a to 0.7 reduces translation error and slightly improves consistency. The optimal setting is $\tau_a = 0.8$: it produces the overall lowest errors (RRE = 1.374°, RTE = 0.046 m) and highest recall (97.8%) on C3DM, and similarly optimal results on C3DLM (RRE = 2.493°, RTE = 0.074 m, RR = 90.2%). A too-stringent criterion ($\tau_a = 0.9$) discards useful anchors, degrading performance. Accordingly, we use $\tau_a = 0.8$ by default.

10. Additional Visual Results

We visualize and compare the qualitative registration results between our SuP framework and ColorPCR [13] on the color-enhanced 3DMatch benchmark in Figure 4. For ColorPCR, the majority of correspondences are overwhelmed by false matches, causing true inliers to be hidden amidst the noise. In contrast, our sub-cloud maneuver effectively isolates and highlights the correct correspondences, leading to precise alignment even in regions with repetitive geometry.

Similarly, in Figure 5, we present the results on the color-enhanced 3DLoMatch dataset, which poses more severe low-overlap challenges. ColorPCR again struggles to differentiate true matches from outliers, resulting in a fragmented and unreliable registration. By leveraging localized sub-cloud proposals and our multi-scale post-weighting network, SuP consistently uncovers the underlying inlier structure, producing coherent and accurate transformations across both high- and low-overlap scenarios.

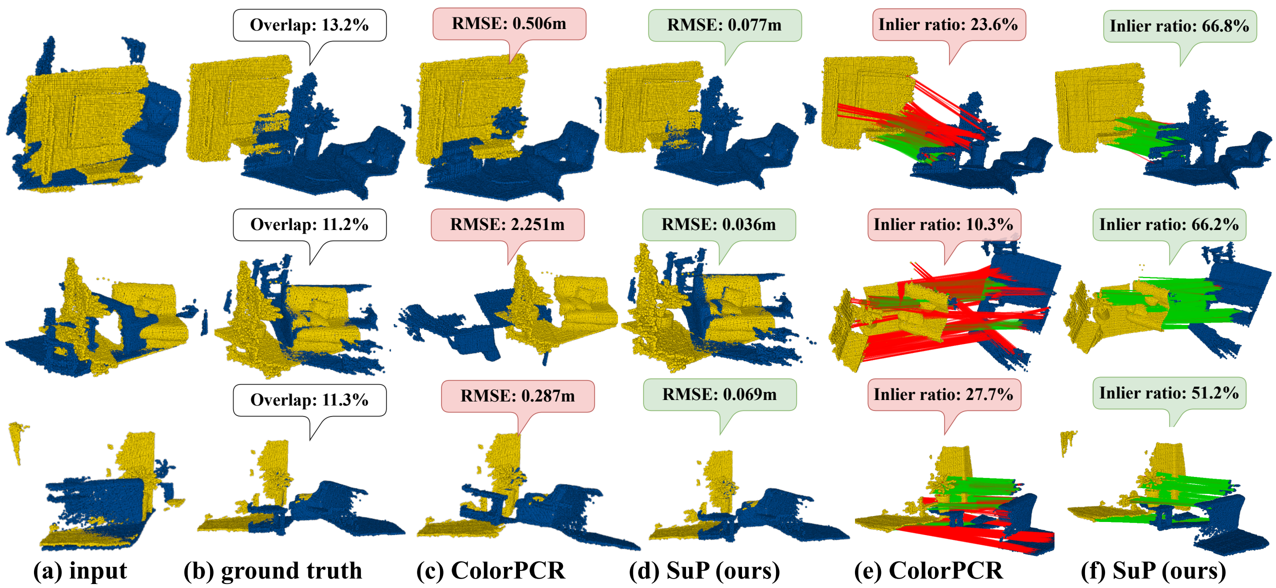

Figure 5. Qualitative comparison of SuP and ColorPCR on challenging, low-overlap pairs from the color-enhanced 3DLoMatch benchmark. We show samples with extremely low overlaps ($< 15\%$). ColorPCR is misled by spurious matches in non-overlapping regions, resulting in incorrect alignment. By contrast, our approach subdivides the point cloud into sub-clouds and the DSAM module reliably identifies high-overlap sub-cloud pairs, and yields accurate inlier correspondences, enabling robust registration even under extreme low-overlap conditions.