SuP: Sub-cloud Driven Point Cloud Registration

Sheldon Fung[1] Wei Pan[2] Ling Cao[2] Fei Hou[3,4]
Ling Chen[5] Shasha Mao[6] Hongdong Li[7] Xuequan Lu[1*]
[1]University of Western Australia [2]OPT Machine Vision
[3]Key Laboratory of System Software (CAS), Institute of Software, Chinese Academy of Sciences
[4] University of Chinese Academy of Sciences [5]University of Technology Sydney
[6]Xidian University [7]Australian National University
{sheldon.feng, bruce.lu}@uwa.edu.au, {vpan, cling1189}@foxmail.com
houfei@ios.ac.cn, ling.chen@uts.edu.au, ssmao@xidian.edu.cn, hongdong.li@anu.edu.au

Abstract

*While existing point-cloud-registration methods can handle high-overlap scenarios of two point clouds well, they often struggle with low-overlap scenarios due to inevitable geometric/semantic ambiguities in the non-overlapping regions. In this paper, we introduce SuP, a novel framework that reformulates low-overlap registration as a high-overlap sub-cloud pairs (anchor pairs) mining problem. Central to SuP is our Dual-phase Sub-cloud Anchor Mining (DSAM) module, which first subdivides the source and target point clouds into multiple sub-clouds, followed by introducing a dual-phase weighting pipeline: **1)** an efficient overlap-guided prior-weighting scheme (OPS) that leverages feature salience to identify candidate anchor pairs, and **2)** a multi-scale post-weighting network (MPN) that exploits neighborhood feature consensus to further identify anchor pairs. Subsequently, final correspondences are generated through a merge-to-match module using the anchor pairs. To train DSAM, we design an alignment-aware weighting loss that uses on-the-fly alignment errors as supervision. Comprehensive experiments on the color-enhanced 3DMatch and 3DLoMatch demonstrate that SuP significantly outperforms state-of-the-art methods, achieving higher registration recall and more accurate alignment, especially under challenging low-overlap conditions.*

1. Introduction

Recent advances in point cloud registration have enabled high precision alignment when two scans share a substantial amount of common geometry. However, real-world applications such as large-scale mapping and autonomous navigation often encounter scenarios where only a small por-

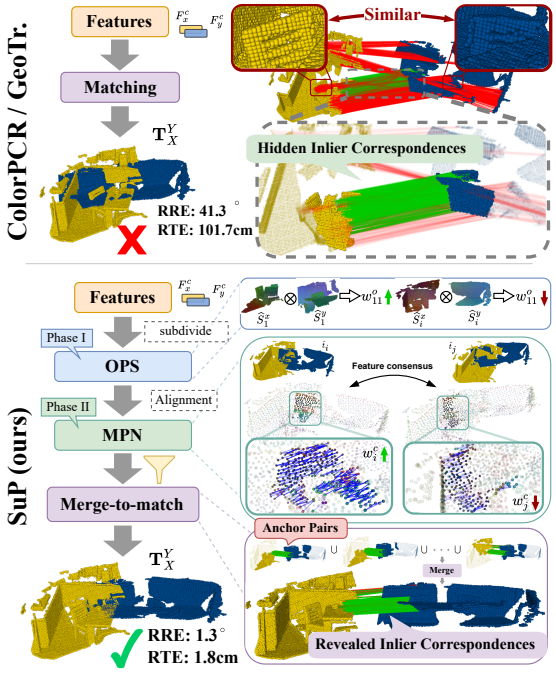

Figure 1. Geometric or semantic similarity in non-overlapping regions (red boxes) often leads to outlier matches that obscure inlier correspondences (dash gray box), causing failures of methods like ColorPCR [13] and GeoTr. [14] under low-overlap conditions. In contrast, our proposed SuP reveals hidden inliers (purple box) by subdividing point clouds into sub-clouds and identifying anchor pairs. The OPS and MPN modules ensure that the selected anchor pairs exhibit high overlap (see blue box) and strong feature consensus (depicted by blue lines in the green box).

*Corresponding author.
Code: https://github.com/SheldonFung98/SuP

tion of the scene overlaps between views. When the shared surface area becomes limited (e.g., below 30%), most classic ICP methods and even modern learning based approaches struggle to establish reliable correspondences, leading to slow convergence, misalignment, or complete failure. This low overlap setting remains a major challenge for robust 3D scene understanding and downstream geometric reasoning.

To mitigate this, pioneering works have attempted to explicitly detect the overlapping region [8, 26] by predicting point-wise overlap weights. Despite its straightforward intuition, direct overlap weights reasoning heavily relies on an underlying consistent feature representation. Recent works [5, 12–14] have abandoned such a scheme and instead focused on implicitly strengthening global feature representations. With the powerful attention [21] layers, these methods are capable of learning transformation-invariant features from the injected intra-cloud geometric encodings [12, 14], semantic features [5], or colors [13]. While these strategies yield improvements, they still face performance degradation as the overlap ratio drops.

The key reason for the failures caused by the lack of overlap is that, despite advanced practices proficiently extracting discriminative features in the overlapping region, the non-overlapping regions still share geometrical/semantical similarities (see Figure 1). This ultimately leads to erroneous correspondence.

In this work, we ask: *instead of wrestling directly with low-overlap point cloud pairs, is it possible to convert (complete) them into high-overlap pairs ?* Our key insight is that by subdividing the source and target point clouds into multiple smaller sub-clouds and performing pairwise matching, we can identify a subset of high-overlap sub-cloud pairs (see Figure 1), denoted as anchor pairs. Yet, since they are hidden among all possible sub-cloud pairs, how we can effectively identify them for robust correspondence generation remains a challenge.

To this end, we propose *SuP*, a **Su**b-cloud Driven **P**oint Cloud Registration framework that reformulates low-overlap registration as an anchor pairs mining problem. It introduces a dual-phase mining pipeline to identify the anchor pairs with consistent features subdivided from source and target point clouds, which are merged and utilized in robust correspondence generation. Concretely, to balance computation complexity and performance, we first aggregate point-wise features from the full point cloud pairs. Subsequently, we propose the Dual-phase Sub-cloud Anchor Mining (DSAM) module to subdivide both source and target into sub-clouds and identify anchor pairs with high-overlap ratios. The two main parts of the proposed DSAM include **1)** an efficient overlap-guided prior-weighting scheme (OPS) that employs feature salience for the preliminary anchor candidate selection, and **2)** a multi-scale post-weighting network (MPN) that exploits neighborhood feature consensus between aligned cross-clouds to identify anchor pairs. With the anchor pairs, we devise a merge-to-match module to generate robust correspondence in a coarse-to-fine manner, with which the final transformation can be estimated. To train the post-weighting network, we design an alignment-aware weighting loss that employs on-the-fly computed alignment errors from anchor pairs as supervision signals. By reframing low-overlap registration as a problem of high-overlap anchor sub-cloud pairs identification, our approach circumvents the ambiguity of non-overlapping regions and yields accurate, robust correspondence for registration.

We evaluate our method on the indoor color-enhanced 3DMatch/3DLoMath [13, 31], and experimental results show that it outperforms state-of-the-art methods. Our contributions are summarized as follows:

- We propose SuP, a sub-cloud driven point cloud registration method that formulates low-overlap registration as a problem of mining high-overlap sub-cloud anchor pairs, yielding robust correspondences for registration.
- We introduce a Dual-phase Sub-cloud Anchor Mining (DSAM) Module that subdivides source and target point clouds into sub-clouds and mines high-overlap anchor pairs via an efficient overlap-guided prior-weighting scheme and a multi-scale post-weighting network.
- We design an alignment-aware weighting loss that employs on-the-fly computed alignment errors from anchor pairs as supervision signals to learn a robust post-weighting network.

2. Related Work

Deep features. ICP-based methods [3, 10, 11, 16] typically rely on point-wise/point-to-plane distances to define correspondences. Although widely adopted, such methods are susceptible to arbitrary poses and low overlap ratios. Recent approaches leverage deep features [1, 6, 8, 12–14, 22, 23, 27, 28] to establish robust point-wise matches to establish reliable correspondences. Wang and Solomon [24] leverage the DGCNN [25] and Transformer [21] network to select correspondences based on feature correlations. Huang et al. [8] employ KPCNN [19] to process dense, large-scale point clouds and explicitly discriminate point-wise overlap scores. However, such an overlap prediction scheme heavily relies on the underlying feature consistency, which becomes less effective on low-overlap samples. More advanced approaches [5, 13, 14, 29] thereby pivot to strengthen the extracted features. With the powerful attention [21] layers, Qin et al. [14] propose to inject intra-cloud geometric encodings to learn transformation-invariant features. Likewise, Li and Harada [12] adopt rotary position encoding to reveal relative position during key-value multiplication in the self-attention process. Inspired by them, more advanced approaches incorporate high-level semantic information, such as universal semantic-aware features [5] and color domain data [13].

Correspondence-based methods. With robust deep features extracted, these paradigms establish point-wise correspondences from feature correlations. Subsequently, correspondences are refined by an outlier estimator, e.g., RANSAC [17]. By using a coarse-to-fine matching scheme, some methods [13, 14, 29] can alternatively use the local-to-global registration (LGR [14]) maneuver to estimate the transformation in a RANSAC-free manner. Since inlier correspondences should exhibit high semantic correlations, Fung et al. [5] propose a semantic-constrained matching scheme where correspondences are restricted to be established only in semantically correlated regions. Another line of research directly predicts the correspondences locations Shen et al. [18], Yew and Lee [26], Yuan et al. [30], and therefore, the transformation matrix can be derived without an outlier estimator. Despite showing high potential in registration performance, these methods often struggle under low-overlap scenarios due to geometric or semantic similarity in non-overlapping regions.

Our method reframes low-overlap registration as a problem of identifying high-overlap sub-cloud anchor pairs. Instead of relying on dense global correspondence estimation, we first identify reliable local regions with substantial overlap and only perform matching within these areas. This design allows our framework to bypass the ambiguity introduced by non-overlapping regions, resulting in more robust and accurate registration, even in challenging settings.

3. Method

Given the source point cloud $\mathbf{X} \in \mathbb{R}^{m\times 3}$ and target point cloud $\mathbf{Y} \in \mathbb{R}^{n\times 3}$ that share partial overlap, our task is to estimate a rigid transformation $\mathbf{T}_X^Y = \{\mathbf{R}, t\}$, with $\mathbf{R} \in SO(3)$ and $t \in \mathbb{R}^3$, which aligns $\mathbf{X}$ to $\mathbf{Y}$. With the ground truth set of correspondences between the source and target $\mathbb{C}^* = \{(x_i, y_i) \mid x_i \in \mathbf{X}, y_i \in \mathbf{Y}, i = 1, \dots, t\}$, an optimal transformation $\mathbf{T}_X^{Y\,*} = \{\mathbf{R}^*, t^*\}$ can be derived by solving the following equation:

$$\mathbf{R}^*, t^* = \min_{\mathbf{R},t} \sum_{(x_i,y_i)\in\mathbb{C}^*} \|\mathbf{R}\cdot x_i + t - y_i\|_2^2. \qquad (1)$$

In practice, $\mathbb{C}^*$ is unknown. Therefore, our method first establishes a set of reliable correspondences, with which the final transformation can be estimated.

3.1. Overview

Figure 2 illustrates the pipeline of our method. We first extract multi-scale features from downsampled point clouds (Sec. 3.2) and enhance them via attention-based aggregation (Sec. 3.3). The proposed Dual-phase Sub-cloud Anchor Mining (DSAM) module (Sec. 3.4) subdivides each point cloud into sub-clouds, selects high-overlap candidates using an overlap-guided prior-weighting scheme (OPS), and estimates preliminary transformations. Subsequently, a multi-scale post-weighting Network (MPN) selects anchor pairs based on the predicted feature consensus weights. Finally, anchor pairs are merged to generate correspondences, from which the final transformation is estimated (Sec. 3.5). The entire pipeline is trained with a combination of contrastive, matching, and alignment-aware losses (Sec. 3.6).

3.2. Multi-scale Feature Extraction

We follow previous work [5, 13, 14] and employ KPConv style [19] to simultaneously down-sample and extract local geometry features hierarchically: $\Phi_x = \kappa(\mathbf{X})$, where $\kappa(\cdot)$ is the KPConv-style network and $\Phi_x = \{(\hat{\mathbf{X}}^i, F_x^i) \mid i = 1, 2, 3, 4\}$ is the multi-scale down-sampled points along with the extracted features (likewise for the source $\mathbf{Y}$). Notably, $\hat{\mathbf{X}}^1$ refers to the points at the densest level while $\hat{\mathbf{X}}^4$ is the coarsest. In our experiments, we employ a color-enhanced KPConv-style backbone [13] by default.

3.3. Attention-based Feature Aggregation

With points and features at the coarsest level $\phi_x^4 \in \Phi_x$, we employ attention-based layers [21] to aggregate robust point-wise features. Specifically, coarse points and features ϕ_x^4 are first fed into a self-attention layer to aggregate global contextual features $F_x^s = \Lambda_s(\phi_x^4, \phi_x^4, \phi_x^4)$, where $\Lambda_s(Q, K, V)$ is the self-attention layer with Q, K, V from the same point cloud. Points $\hat{\mathbf{X}}^4 \in \phi_x^4$ are used for position encoding. In our experiments, we employ a color-guided geometric self-attention [13] by default. Subsequently, we employ a vanilla cross-attention layer [21] to aggregate features conditioned on the other point cloud $F_x^c = \Lambda_c(F_x^s, F_y^s, F_y^s)$, where $\Lambda_c(Q, K, V)$ is the cross-attention layer with Q from the same point cloud and K, V from the other point cloud. Operations on the target point cloud are identical.

3.4. Dual-phase Sub-cloud Anchor Mining

The proposed Dual-phase Sub-cloud Anchor Mining (DSAM) aims to bypass low-overlap registration by reframing the problem into identifying high-overlap sub-cloud pairs (anchor pairs). Instead of estimating a global transformation directly, we first subdivide the point clouds into multiple sub-clouds to achieve high local overlap. Then, we develop an overlap-guided prior-weighting scheme (OPS) to select potentially overlapping sub-cloud pairs as candidates. For each candidate, a preliminary transformation is estimated to align the pair. Finally, a multi-scale post-weighting network (MPN) is designed to predict the feature consensus weights of the aligned candidates, enabling robust anchor pair identification for the final transformation estimation.

Subdivision. Given the coarsest-level points $\hat{\mathbf{X}}^4$, we first pick k well-spread centroids $C^x = \{c_i^x\}_{i=1}^k$ using Furthest Point Sampling (FPS). Around each centroid c_i^x, we form a

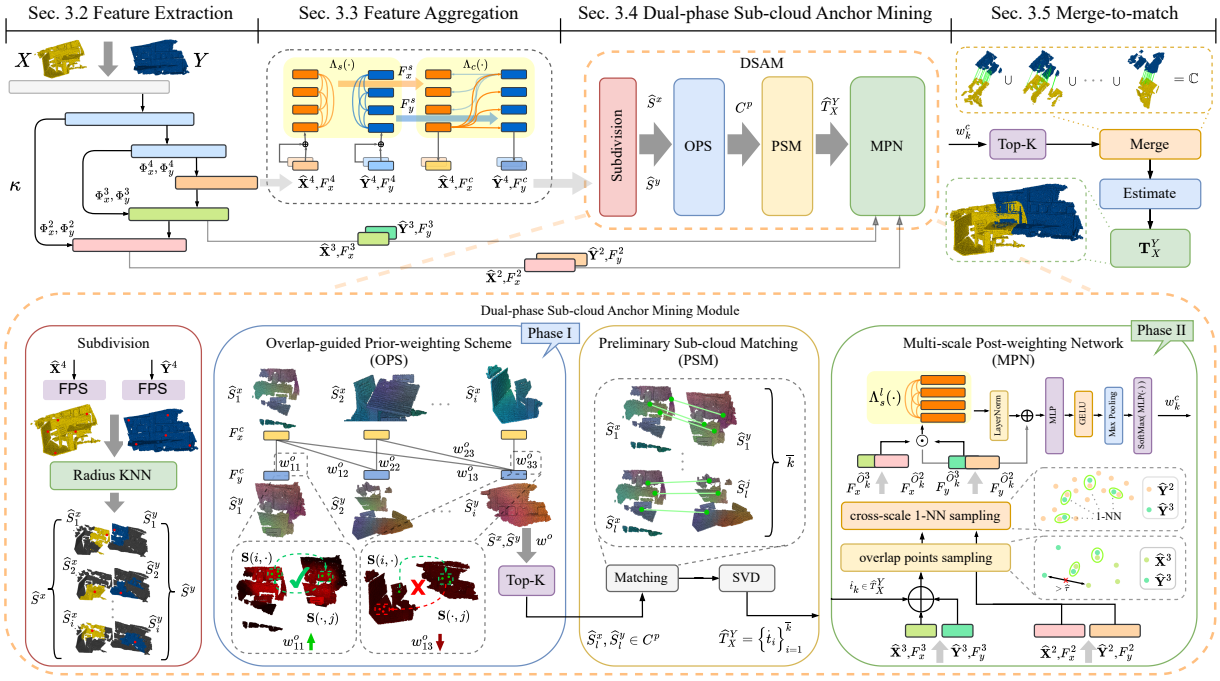

Figure 2. The overview of our proposed SuP. We first obtain point-wise features by feeding the point clouds to the KPConv and the Transformer module. Then each point cloud is subdivided into multiple sub-clouds. We design OPS to pick sub-cloud candidates by estimating the overlap weights among sub-clouds from the source and target. Subsequently, we estimate the poses of the candidates. MPN is proposed to predict feature consensus weights (FCW) from the aligned candidates. The anchor pairs are selected from FCW, and the final transformation is estimated by the merge-to-match module.

sub-cloud $\hat{S}_i^x$ by collecting its K nearest neighbours with a radius constraint $\hat{S}_i^x = \{x_i^4 \mid x_i^4 \in \text{KNN}(c_i^x), \|x_i^4 - c_i^x\|_2 < \tau\}$. The target cloud is subdivided in the same way.

Overlap-guided prior-weighting scheme (OPS). We design an efficient prior-weighting scheme to eliminate non-overlap sub-cloud matches. Concretely, with k sub-clouds from each point cloud, it yields k^2 sub-cloud pairs. Given a sub-cloud pair, we compute a Gaussian correlation matrix $\mathbf{S} \in \mathbb{R}^{|\hat{S}_l^x| \times |\hat{S}_m^y|}$ using their conditioned features: $s_{ij} = exp(-\|\text{norm}(F_{x_i}^c) - \text{norm}(F_{y_j}^c)\|_2^2)$, where norm$(\cdot)$ is the normalization operation. By assuming that two overlapped points tend to share strong mutual feature correlations, the prior weights w_{lm}^o are computed as $w_{lm}^o = \frac{1}{|\mathcal{M}|}\sum_{s\in\mathcal{M}} s$, where $\mathcal{M}$ are the mutual feature correlations:

$$\mathcal{M} = \{s_{ij} \mid s_{ij} \in \Psi_{\hat{k}}(\mathbf{S}(i,\cdot)) \cap \Psi_{\hat{k}}(\mathbf{S}(\cdot,j))\}, \quad (2)$$

where $\mathbf{S}$ is the Gaussian correlation matrix and $\Psi_k(\cdot)$ is a top-k largest operation. Preliminary candidates $C^p = \{(\hat{S}_l^x, \hat{S}_l^y)\}_{l=1}^{\bar{k}}$ are sub-cloud pairs selected by $\Psi_{\bar{k}}(w^o)$.

Preliminary sub-cloud matching (PSM). For each candidate pair $(\hat{S}_l^x, \hat{S}_l^y) \in C^p$, we perform a preliminary transformation estimation for the subsequent multi-scale post-weighting stage. We reuse the Gaussian correlation matrix $\mathbf{S}$ from the prior-weighting pipeline and apply a dual-normalization strategy [14, 15] to further suppress noisy correspondences. Then, the preliminary transformation $\hat{T}_X^Y = \{\hat{T}_i\}_{i=1}^{\bar{k}}$ can be computed using $\hat{S}_l^x$ and $\hat{S}_l^y$ with $\mathbf{S}$ as weights through a weighted SVD.

Multi-scale post-weighting network (MPN). We observe strong feature consensus at cross-cloud overlapped points. Namely, with the well-aligned sub-cloud pairs, their overlapped points (i.e., points at close distances) exhibit high feature similarities, and vise versa. Motivated by this, we propose a learnable post-weighting network that captures multi-scale feature consensus for each aligned candidate. Given $\hat{T}_k \in \hat{T}_X^Y$ for the k-th sub-cloud pair, we formulate the network by first retrieving a set of overlapped points $\hat{\mathcal{O}}_k^3$:

$$\hat{\mathcal{O}}_k^3 = \{(\hat{\mathbf{x}}_i^3, \hat{\mathbf{y}}_j^3) \mid \|\hat{\mathbf{x}}_i^3 - \hat{T}_k(\hat{\mathbf{y}}_j^3)\|_2 < \hat{\tau}, \hat{\mathbf{x}}_i^3 \in \hat{\mathbf{X}}^3, \hat{\mathbf{y}}_j^3 \in \hat{\mathbf{Y}}^3\}, \quad (3)$$

where $\hat{T}_k(\cdot)$ refers to the corresponding rigid transformation. With the coarse overlapped points $\hat{\mathcal{O}}_k^3$, the dense overlapped points $\hat{\mathcal{O}}_k^2$ are located using a cross-scale 1-NN strategy:

$$\hat{\mathcal{O}}_k^2 = \left\{\left(\text{NN}_{\hat{\mathbf{X}}^2}(\hat{\mathbf{x}}_i^3), \text{NN}_{\hat{\mathbf{Y}}^2}(\hat{\mathbf{y}}_j^3)\right) \,\middle|\, (\hat{\mathbf{x}}_i^3, \hat{\mathbf{y}}_j^3) \in \hat{\mathcal{O}}_k^3\right\}, \quad (4)$$

where $\mathrm{NN}_{\mathcal{P}}(p) \triangleq \arg\min_{q\in\mathcal{P}} \|q-p\|_2$. This strategy ensures the order of the constructed overlapped points at different scales while maintaining computational efficiency.

Then, with the established multi-scale overlap points $\hat{\mathcal{O}}_k^3$ and $\hat{\mathcal{O}}_k^2$ consisting of L elements respectively, we compute the affinity features $\mathcal{Z} \in \mathbb{R}^{L\times d_a}$ using the corresponding multi-scale features extracted from KPConv:

$$\mathcal{Z} = [F_x^{\hat{\mathcal{O}}_k^2}, F_x^{\hat{\mathcal{O}}_k^3}] \odot [F_y^{\hat{\mathcal{O}}_k^2}, F_y^{\hat{\mathcal{O}}_k^3}], \tag{5}$$

where $\odot$ is the Hadamard product and $[\cdot,\cdot]$ is the concatenation operator. d_a is the sum of channel numbers of the adopted multi-scale features. We then exploit a light-weight multi-head attention layer with residual connections to learn contextualized affinity features $\hat{\mathcal{Z}}$:

$$\hat{\mathcal{Z}} = \mathcal{Z} + \mathrm{LayerNorm}(\Lambda_s^l(\mathcal{Z})), \tag{6}$$

where Λ_s^l is the light-weight self-attention operation. Subsequently, we employ a projection head to aggregate a feature consensus descriptor $\hat{F}_d \in \mathbb{R}^d$:

$$\hat{F}_d = \mathrm{MaxPooling}(\mathrm{GELU}(\hat{\mathcal{Z}}W_p^T + b)), \tag{7}$$

where $W_p^T \in \mathbb{R}^{d_a\times d}$ is a learnable matrix and b is a learnable vector. We adopt the Gaussian Error Linear Units (GELU) function as activation and a global max-pooling layer for feature aggregation. The final feature consensus weight is computed as $w_k^c = \mathrm{Softmax}(\mathrm{MLP}(\hat{F}_d))$.

With the robust feature consensus weight, our method is able to effectively identify anchor pairs by selecting preliminary candidates with top-k weights constrained by a threshold.

3.5. Merge-to-match

We aim to estimate a final transformation by merging all the inlier coarse correspondences from the anchor pairs generated by DSAM (Sec. 3.4). For the i-th anchor pair $\mathcal{A}_i = \{\hat{S}_j^x, \hat{S}_k^y\}$, we extract the inlier coarse correspondences:

$$\mathbb{C}_i = \{(x_t^4, y_t^4) \mid \|x_t^4 - \hat{T}_{jk}(y_t^4)\|_2 < \tau_m, x_t^4 \in \hat{S}_j^x, y_t^4 \in \hat{S}_k^y\}, \tag{8}$$

where τ_m is an inlier threshold. Then the final set of coarse correspondences $\mathbb{C}$ is derived by merging the inlier correspondences from all anchor pairs: $\mathbb{C} = \bigcup_{i=1}^{N_a} \mathbb{C}_i$, where N_a is the number of anchor pairs.

Coarse correspondences $\mathbb{C}$ are further refined by dense points similar to [13, 14], followed by a RANSAC or LGR outlier estimator.

3.6. Supervision

Our loss function is formulated as $\mathcal{L} = \mathcal{L}_{oc} + \alpha\mathcal{L}_p + \beta\mathcal{L}_a$, where α and β are loss weights, and $\mathcal{L}_{oc}$ is the overlap-aware circle loss and $\mathcal{L}_p$ is a point matching loss [14]. We propose an alignment-aware weighting loss $\mathcal{L}_a$ to facilitate the learning of the proposed MPN.

Alignment-aware weighting loss. Given the ground truth transformation, we first split the multi-scale post weights w^c into positive and negative groups. The positive group ϵ_p comprises anchor pairs with $E_{rmse} < \tau_e$, where E_{rmse} is the transformation RMSE error and τ_e is an error threshold. The negative group is constructed similarly. The alignment-aware weighting loss can be formulated as follows:

$$\mathcal{L}_a = -\frac{1}{|\epsilon_p|}\sum_{p\in\epsilon_p}\lambda_a \log w_p^c - \frac{1}{|\epsilon_n|}\sum_{n\in\epsilon_n}\lambda_a \log(1-w_n^c), \tag{9}$$

where λ_a is an alignment-aware weight that is defined as:

$$\lambda_i^a = \begin{cases} 1-E_{rmse}, & \text{if } \tau_e < E_{rmse} < \tau_e+\delta \\ 1, & \text{otherwise} \end{cases}, \tag{10}$$

where δ is a hyperparameter. Intuitively, δ allows higher feature consensus weight w^c at failure alignments with small errors.

4. Experiments

4.1. Experiment Settings

Implementation. Our SuP framework is implemented in PyTorch and trained end-to-end using vanilla stochastic gradient descent. Each point cloud is subdivided into $k = 6$ sub-clouds, and $\bar{k} = 24$ candidates are selected in the Overlap Prior Scoring (OPS) stage. In the Multi-scale Post-weighting Network (MPN), the overlap threshold $\hat{\tau}$ is set to 0.04, and we select the top 8 sub-cloud pairs with the highest feature consensus weights as anchor pairs. The training starts with a learning rate of 1×10^{-4}, which decays exponentially by a factor of 0.95 after each epoch (i.e., $\eta_t = 1\mathrm{e}{-4}\times0.95^t$). An L_2 weight decay of 1×10^{-6} is applied to all learnable parameters for regularization. Training is distributed across eight NVIDIA RTX 4070 Ti GPUs via DataParallel, with each GPU processing one sample per iteration for an effective batch size of 8. The model is trained for 40 epochs, taking approximately 10 hours to complete on our hardware.

Datasets. Color3DMatch (C3DM) and Color3DLoMatch (C3DLM) [13] both extend the original 3DMatch dataset [31], which consists of 62 scenes: 46 for training, 8 for validation, and 8 for testing. We follow the split convention of earlier work [8], and remove the depth-only scene from the training set [13, 20], resulting in a train set containing $19,499$ point cloud pairs. In C3DM, point-cloud pairs share an overlap greater than 0.3%, whereas C3DLM focuses on more challenging cases with overlaps in the 0.1%–0.3% range.

Metrics. In line with [13, 14], we assess our approach using five key metrics: (1) Inlier Ratio (IR), defined as the proportion of correspondences whose distance under the ground-

truth transform falls below $0.1m$; (2) Feature Matching Recall (FMR), the share of point clouds for which the inlier ratio surpasses 5%; (3) Registration Recall (RR), the fraction of point-cloud pairs whose alignment yields an RMSE under $0.2m$; (4) Relative Rotation Error (RRE), which quantifies angular deviation in degrees; and (5) Relative Translation Error (RTE), reporting translational deviation in centimeters.

#Samples	C3DM 5k	C3DM 2.5k	C3DM 1k	C3DM 500	C3DM 250	C3DLM 5k	C3DLM 2.5k	C3DLM 1k	C3DLM 500	C3DLM 250
Registration Recall %										
FCGF [4]	85.1	84.7	83.3	81.6	71.4	40.1	41.7	38.2	35.4	26.8
D3Feat [2]	81.6	84.5	83.4	82.4	77.9	37.2	42.7	46.9	43.8	39.1
SpinNet [1]	88.6	86.6	85.5	83.5	70.2	59.8	54.9	48.3	39.8	26.8
Predator [7]	89.0	89.9	90.6	88.5	86.6	59.8	61.2	62.4	60.8	58.1
YOHO [22]	90.8	90.3	89.1	88.6	84.5	65.2	65.5	63.2	56.5	48.0
CoFiNet [27]	89.3	88.9	88.4	87.4	87.0	67.5	66.2	64.2	63.1	61.0
GeoTr. [14]	92.0	91.8	91.8	91.4	91.2	75.0	74.8	74.2	74.1	73.5
PEAL [29]	94.6	93.7	93.7	93.9	93.4	81.7	81.2	80.8	80.4	80.1
ColorPCR [13]	96.7	96.5	97.0	96.4	96.5	88.9	88.5	88.1	86.5	85.0
Cross-PCR [32]	94.5	94.2	94.2	94.3	94.0	73.7	73.9	74.1	74.2	74.1
PSReg [9]	95.7	94.9	95.1	95.0	95.2	79.3	79.3	78.7	78.7	78.4
SuP (ours)	**98.1**	**97.7**	**97.5**	**97.2**	**97.4**	**90.4**	**89.5**	**88.9**	**87.9**	**86.2**
Inlier ratio%										
FCGF [4]	56.8	54.1	48.7	42.5	34.1	21.4	20.0	17.2	14.8	11.6
D3Feat [2]	39.0	38.8	40.4	41.5	41.8	13.2	13.1	14.0	14.6	15.0
SpinNet [1]	47.5	44.7	39.4	33.9	27.6	20.5	19.0	16.3	13.8	11.1
Predator [7]	58.0	58.4	57.1	54.1	49.3	26.7	28.1	28.3	27.5	25.8
YOHO [22]	64.4	60.7	55.7	46.4	41.2	25.9	23.3	22.6	18.2	15.0
CoFiNet [27]	49.8	51.2	51.9	52.2	52.2	24.4	25.9	26.7	26.8	26.9
GeoTr. [14]	71.9	75.2	76.0	82.2	85.1	43.5	45.3	46.2	52.9	57.7
PEAL [29]	72.4	79.1	84.1	86.1	87.3	45.0	50.9	57.4	60.3	62.2
ColorPCR [13]	75.0	80.5	84.7	86.5	87.8	51.2	56.6	63.1	66.0	68.0
Cross-PCR [32]	88.7	88.7	88.7	88.7	88.7	65.9	65.9	65.9	65.9	65.9
PSReg [9]	75.8	82.4	87.1	88.9	90.0	49.9	55.5	61.9	64.5	66.3
SuP (ours)	**82.3**	**85.7**	**88.7**	**90.1**	**91.1**	**66.2**	**67.1**	**71.3**	**74.0**	**76.0**
Feature Matching Recall %										
FCGF [4]	97.4	97.3	97.0	96.7	96.6	76.6	75.4	74.2	71.7	67.3
D3Feat [2]	95.6	95.4	94.5	94.1	93.1	67.3	66.7	67.0	66.7	66.5
SpinNet [1]	97.6	97.2	96.8	95.5	94.3	75.3	74.9	72.5	70.0	63.6
Predator [7]	96.6	96.6	96.5	96.3	96.5	78.6	77.4	76.3	75.7	75.3
YOHO [22]	98.2	97.6	97.5	97.7	96.0	79.4	78.1	76.3	73.8	69.1
CoFiNet [27]	98.1	98.3	98.1	98.2	98.3	83.1	83.5	83.3	83.1	82.6
GeoTr. [14]	97.9	97.9	97.9	97.9	97.6	88.3	88.6	88.8	88.6	88.3
PEAL [29]	99.0	99.0	99.1	99.1	98.8	91.7	92.4	92.5	92.9	92.7
ColorPCR [13]	99.5	99.5	99.5	**99.5**	**99.5**	96.5	96.5	**97.0**	**97.0**	**96.7**
Cross-PCR [32]	97.9	97.7	97.9	97.7	97.7	83.5	83.3	83.3	83.1	83.4
PSReg [9]	98.6	98.6	98.6	98.6	98.6	86.4	86.6	87.1	87.4	87.1
SuP (ours)	**99.7**	**99.7**	**99.7**	**99.5**	**99.5**	**96.6**	**96.6**	**97.0**	96.8	**96.7**

Table 1. RANSAC-based evaluation results on C3DM and C3DLM, with the top results highlighted in bold.

4.2. RANSAC Estimator Results

We evaluate the performance of SuP against recent state-of-the-art methods: FCGF [4], D3Feat [2], SpinNet [1], Predator [8], YOHO [22], CoFiNet [27], GeoTransformer [14], PEAL [29], ColorPCR [13], Cross-PCR [32], and PSReg [9] . Following standard protocols, we report results with varying numbers of sampled correspondences and use RANSAC as the transformation estimator for fair comparison. The evaluation metrics include Feature Matching Recall (FMR), Inlier Ratio (IR), and Registration Recall (RR), as shown in Table 1.

SuP consistently outperforms all baselines across all metrics and correspondence sampling levels. For FMR, our method achieves 99.7% on C3DM and up to 97.0% on C3DLM. For IR, SuP reaches 91.1% on C3DM and 76.0% on C3DLM, surpassing the second-best ColorPCR by large margins (3.3% and 8.0% respectively). For RR, SuP further pushes performance to 97.8% on C3DM and 90.2% on C3DLM, setting a new state of the art in both high- and low-overlap scenarios. These results highlight the efficacy of SuP's anchor-based sub-cloud correspondence selection scheme, yielding higher inlier ratios and reliable transformation estimates under challenging low-overlap conditions.

4.3. LGR Estimator Results

To further evaluate the effectiveness of our method in RANSAC-free settings, we compare SuP against recent state-of-the-art methods using the Local-to-Global Registration (LGR) estimator. We report results on both Color3DMatch (C3DM) and Color3DLoMatch (C3DLM) benchmarks in terms of Registration Recall (RR), Relative Rotation Error (RRE), and Relative Translation Error (RTE).

Model	Estimator	Samples	RR(%) C3DM	RR(%) C3DLM
CoFiNet [27]	RANSAC-50k	5000	89.3	67.5
GeoTr. [14]	RANSAC-50k	5000	92.0	75.0
PEAL [29]	RANSAC-50k	5000	94.6	81.7
ColorPCR [13]	RANSAC-50k	5000	96.7	88.9
SuP (ours)	RANSAC-50k	5000	**98.1**	**90.4**
CoFiNet [27]	LGR	all	87.6	64.8
GeoTr. [14]	LGR	all	91.5	74.0
PEAL [29]	LGR	all	94.3	81.2
ColorPCR [13]	LGR	all	96.5	88.3
SuP (ours)	LGR	all	**97.8**	**90.2**

Table 2. Evaluation results on C3DM and C3DLM using LGR estimator, with the top results highlighted in bold.

As shown in Table 2, SuP achieves the highest RR among all evaluated models under the LGR setting, with 97.8% on C3DM and 90.2% on C3DLM. These results are comparable to or even exceed those of RANSAC-based methods, indicating that our sub-cloud isolation strategy produces robust correspondences even without outlier filtering.

In Table 3, we also evaluate the geometric accuracy of the estimated transformations using RRE and RTE. SuP again outperforms all baselines, achieving the lowest rotation and translation errors under LGR: 1.374° / $0.046m$ on C3DM and 2.493° / $0.074m$ on C3DLM. Notably, our performance exceeds that of REGTR and ColorPCR, highlighting SuP's strength in low-overlap scenarios.

These results demonstrate that SuP not only enables accurate registration but also produces geometrically precise transformations under efficient, RANSAC-free estimation.

Model	Estimator	C3DM		C3DLM	
		RRE (°)	RTE (m)	RRE (°)	RTE (m)
Predator [7]	RANSAC-50k	2.029	0.064	3.048	0.093
CoFiNet [27]	RANSAC-50k	2.002	0.064	3.271	0.090
REGTR [26]	weighted-SVD	1.567	0.049	2.827	0.077
GeoTr. [7]	LGR	1.772	0.061	2.849	0.088
PEAL [27]	LGR	1.748	0.062	2.788	0.087
ColorPCR [26]	LGR	1.492	0.048	2.581	0.075
SuP (ours)	LGR	**1.374**	**0.046**	**2.493**	**0.074**

Table 3. Evaluation results of alignment geometric accuracy on C3DM and C3DLM. Top results are highlighted in bold.

4.4. As a Plugin Module

To demonstrate the universality of our approach, we integrate SuP as a post-processing module into several state-of-the-art registration pipelines, including GeoTransformer [14], PEAL [29], and ColorPCR [13]. We enhance each model with our sub-cloud alignment and local consensus refinement, evaluating all configurations using the LGR estimator.

	C3DM			C3DLM		
	RRE (°)	RTE (m)	RR (%)	RRE (°)	RTE (m)	RR (%)
GeoTr.	1.772	0.061	91.5	2.849	0.088	74.0
SuP + GeoTr.	**1.765**	**0.060**	**92.8**	**2.712**	**0.082**	**76.7**
PEAL	1.748	0.062	94.3	2.788	0.087	81.2
SuP + PEAL.	**1.643**	**0.059**	**95.6**	**2.532**	**0.076**	**83.4**
ColorPCR	1.492	0.048	96.5	2.581	0.075	88.3
SuP + ColorPCR.	**1.374**	**0.046**	**97.8**	**2.493**	**0.074**	**90.2**

Table 4. Registration performance of other methods with and without our SuP as a plugin on C3DM and C3DLM, with the top results highlighted in bold.

As shown in Table 4, incorporating SuP consistently improves performance across all baselines on both Color3DMatch and Color3DLoMatch. For instance, integrating SuP into GeoTransformer boosts RR from 91.5% to 92.8% on C3DM and from 74.0% to 76.7% on C3DLM, while slightly reducing both RRE and RTE. Similar gains are observed for PEAL, where the RR increases from 94.3% to 95.6% and from 81.2% to 83.4%, respectively. The largest improvements are observed when SuP is applied to ColorPCR, where the registration recall reaches 97.8% on C3DM and 90.2% on C3DLM, surpassing all other baselines. These improvements are accompanied by lower geometric errors, with RRE and RTE reduced to $1.374°/0.046m$ and $2.493°/0.074m$, respectively. These results clearly indicate that SuP can serve as a general enhancement module for existing methods, yielding consistent and substantial improvements in both recall and registration precision.

4.5. Ablation Study

To evaluate the contribution of each proposed component, we conduct ablation experiments on the Color3DMatch (C3DM) and Color3DLoMatch (C3DLM) datasets. Table 5 reports performance after successively removing our four key modules: the overlap-guided prior-weighting scheme (OPS), feature consensus weights (FCW), multi-scale post-weighting network (MPN), and alignment-aware weighting loss (AWL).

OPS	FCW	MPN	AWL	RR(%)	
				C3DM	C3DLM
-	-	-	-	96.5	88.3
✓	-	-	-	96.6	88.6
✓	✓	-	-	97.0	89.1
✓	✓	✓	-	97.5	89.7
✓	✓	-	✓	97.6	89.2
-	✓	✓	✓	97.7	89.9
✓	✓	✓	✓	**97.8**	**90.2**

Table 5. Ablation results on overlap-guided prior-weighting scheme (OPS), multi-scale post-weighting network (MPN), feature consensus weights (FCW), and alignment-aware weighting loss (AWL). The "-" sign represents the corresponding module being removed.

Starting from a baseline with none of the proposed modules, we observe consistent improvements as each component is introduced. OPS provides a modest gain by emphasizing geometrically overlapped regions. Incorporating FCW further boosts performance by leveraging feature similarity between local correspondences. MPN introduces multi-scale reasoning and significantly enhances robustness under low-overlap settings. AWL contributes additional gains by promoting alignment-consistent hypothesis weighting. Our full model, with all components enabled, achieves the best performance with 97.8% RR on C3DM and 90.2% on C3DLM, demonstrating the complementary effects of each module.

	C3DM			C3DLM		
	RRE (°)	RTE (m)	RR(%)	RRE (°)	RTE (m)	RR(%)
$\tau = 0.15$	**1.361**	0.053	96.4	2.537	0.077	90.0
$\tau = 0.25$	1.384	0.047	97.2	2.495	0.089	89.8
$\tau = 0.35*$	1.374	**0.046**	**97.8**	**2.493**	**0.074**	**90.2**
$\tau = 0.45$	1.417	0.051	97.7	2.498	**0.074**	89.9

Table 6. Evaluation results on C3DM and C3DLM with different sub-cloud subdivision sizes, with the top results highlighted in bold. The default value is marked with $*$.

We also examine the influence of the sub-cloud subdivision threshold τ, which controls the relative size of each sub-cloud. As shown in Table 6, smaller τ values (e.g., 0.15) favor finer granularity but may reduce recall due to insufficient overlap, whereas larger values lead to coarser divisions that miss local details. The best trade-off occurs at $\tau = 0.35$, achieving both minimal translation error (0.046 m) and the highest recall (97.8%) on C3DM, along with strong generalization on C3DLM. We further include two additional ablation studies in the supplementary material.

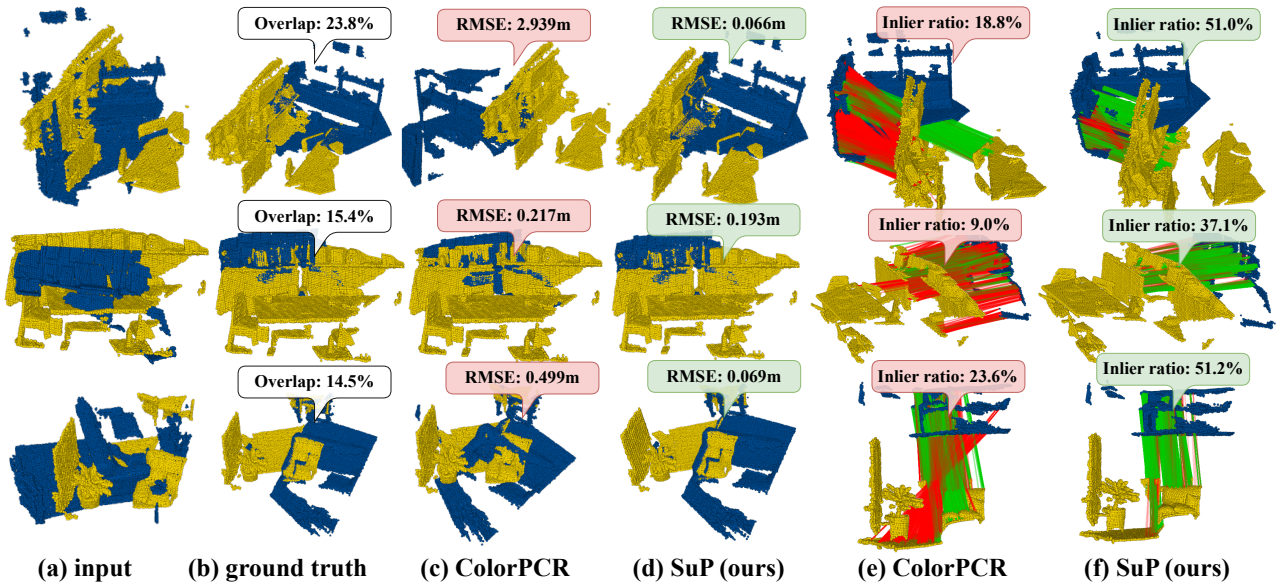

Figure 3. Qualitative comparisons of our SuP and ColorPCR. Each row visualizes a sample from 3DLoMatch with low overlap ratios. ColorPCR fails to estimate accurate transformations due to the outlier correspondences located in the non-overlap region. By leveraging the sub-cloud maneuver, our approach can discard those outlier correspondences and achieve robust performance.

4.6. Qualitative Results

We visualize and compare the qualitative registration results between our approach and ColorPCR [13] in Figure 3. As shown, ColorPCR suffers from a large number of outlier correspondences, especially in non-overlapping or repetitive regions, leading to inaccurate alignments under low-overlap conditions. In contrast, our proposed SuP can effectively reveal reliable inlier correspondences by leveraging the sub-cloud maneuver, where the Dual-phase Sub-cloud Anchor Mining (DSAM) module isolates high-overlap anchor pairs through the overlap-guided prior-weighting scheme (OPS) and further refines them via the multi-scale post-weighting network (MPN). This dual-phase process suppresses spurious matches and enhances feature consensus across aligned sub-clouds, enabling robust and geometrically consistent registration even when the visible overlap ratio is extremely low. These qualitative results confirm that SuP can successfully disentangle inlier structures from ambiguous regions and achieve precise alignment.

5. Limitation and Future Work

Despite the strong performance of our framework across diverse indoor and outdoor datasets, further gains may be achieved by enhancing its robustness under extreme cross-modal variations. In scenarios where scans are captured with significantly different sensing modalities or severe appearance distortions, the extracted features may exhibit subtle inconsistencies that could affect the precision of OPS or the weighting behavior of MPN. While our current design already handles moderate variations well, incorporating lightweight feature normalization or domain adaptation mechanisms could further improve stability under such challenging conditions. We consider this a promising extension rather than a core limitation of the present method.

6. Conclusion

We present SuP, a point cloud registration framework that reframes low-overlap point cloud registration as a problem of mining high-overlap anchor pairs. Our key insight is that by subdividing the source and target point clouds into sub-clouds, we can identify high-overlap sub-cloud pairs to generate robust correspondences. To do so, we propose the Dual-phase Sub-cloud Anchor Mining (DSAM) module, which combines an efficient overlap-guided prior-weighting scheme with a learnable multi-scale post-weighting network that captures cross-cloud feature consensus. These components enable SuP to mine anchor pairs for robust correspondence generation via a coarse-to-fine merge-to-match strategy. By transforming the low-overlap registration problem into high-overlap anchor pair identification, SuP bypasses direct low-overlap registration and significantly improves robustness. Extensive experiments on color-enhanced 3DMatch and 3DLoMatch verify the effectiveness of SuP, achieving state-of-the-art performance, particularly under challenging low-overlap conditions.

References

[1] Sheng Ao, Qingyong Hu, Bo Yang, Andrew Markham, and Yulan Guo. Spinnet: Learning a general surface descriptor for 3d point cloud registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11753–11762, 2021. 2, 6

[2] Xuyang Bai, Zixin Luo, Lei Zhou, Hongbo Fu, Long Quan, and Chiew-Lan Tai. D3feat: Joint learning of dense detection and description of 3d local features. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6359–6367, 2020. 6

[3] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor fusion IV: control paradigms and data structures*, pages 586–606. Spie, 1992. 2

[4] Christopher Choy, Jaesik Park, and Vladlen Koltun. Fully convolutional geometric features. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 8958–8966, 2019. 6

[5] Sheldon Fung, Xuequan Lu, Dasith de Silva Edirimuni, Wei Pan, Xiao Liu, and Hongdong Li. Semreg: Semantics constrained point cloud registration. In *European Conference on Computer Vision*, pages 293–310. Springer, 2024. 2, 3

[6] Sheldon Fung, Wei Pan, Xiao Liu, John Yearwood, Richard Dazeley, and Xuequan Lu. Topformer: topology-aware transformer for point cloud registration. In *International Conference on Computational Visual Media*, pages 112–128. Springer, 2024. 2

[7] Shengyu Huang, Zan Gojcic, Mikhail Usvyatsov, Andreas Wieser, and Konrad Schindler. Predator: Registration of 3d point clouds with low overlap. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 4267–4276, 2021. 6, 7

[8] Shengyu Huang, Zan Gojcic, Mikhail Usvyatsov, Andreas Wieser, and Konrad Schindler. Predator: Registration of 3d point clouds with low overlap. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 4267–4276, 2021. 2, 5, 6

[9] Xiaoshui Huang, Zhou Huang, Yifan Zuo, Yongshun Gong, Chengdong Zhang, Deyang Liu, and Yuming Fang. Psreg: Prior-guided sparse mixture of experts for point cloud registration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3788–3796, 2025. 6

[10] Ji Hoon Joung, Kwang Ho An, Jung Won Kang, Myung Jin Chung, and Wonpil Yu. 3d environment reconstruction using modified color icp algorithm by fusion of a camera and a 3d laser range finder. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3082–3088. IEEE, 2009. 2

[11] Michael Korn, Martin Holzkothen, and Josef Pauli. Color supported generalized-icp. In *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 592–599. IEEE, 2014. 2

[12] Yang Li and Tatsuya Harada. Lepard: Learning partial point cloud matching in rigid and deformable scenes. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5554–5564, 2022. 2

[13] Juncheng Mu, Lin Bie, Shaoyi Du, and Yue Gao. Colorpcr: Color point cloud registration with multi-stage geometric-color fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21061–21070, 2024. 1, 2, 3, 5, 6, 7, 8

[14] Zheng Qin, Hao Yu, Changjian Wang, Yulan Guo, Yuxing Peng, and Kai Xu. Geometric transformer for fast and robust point cloud registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11143–11152, 2022. 1, 2, 3, 4, 5, 6, 7

[15] Ignacio Rocco, Mircea Cimpoi, Relja Arandjelović, Akihiko Torii, Tomas Pajdla, and Josef Sivic. Neighbourhood consensus networks. *Advances in neural information processing systems*, 31, 2018. 4

[16] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the icp algorithm. In *Proceedings third international conference on 3-D digital imaging and modeling*, pages 145–152. IEEE, 2001. 2

[17] Ruwen Schnabel, Roland Wahl, and Reinhard Klein. Efficient ransac for point-cloud shape detection. In *Computer graphics forum*, pages 214–226. Wiley Online Library, 2007. 3

[18] Yaqi Shen, Le Hui, Haobo Jiang, Jin Xie, and Jian Yang. Reliable inlier evaluation for unsupervised point cloud registration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2198–2206, 2022. 3

[19] Hugues Thomas, Charles R Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J Guibas. Kpconv: Flexible and deformable convolution for point clouds. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 6411–6420, 2019. 2, 3

[20] Julien Valentin, Angela Dai, Matthias Nießner, Pushmeet Kohli, Philip Torr, Shahram Izadi, and Cem Keskin. Learning to navigate the energy landscape. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 323–332. IEEE, 2016. 5

[21] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 2, 3

[22] Haiping Wang, Yuan Liu, Zhen Dong, and Wenping Wang. You only hypothesize once: Point cloud registration with rotation-equivariant descriptors. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 1630–1641, 2022. 2, 6

[23] Haiping Wang, Yuan Liu, Qingyong Hu, Bing Wang, Jianguo Chen, Zhen Dong, Yulan Guo, Wenping Wang, and Bisheng Yang. Roreg: Pairwise point cloud registration with oriented descriptors and local rotations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 2

[24] Yue Wang and Justin M Solomon. Deep closest point: Learning representations for point cloud registration. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3523–3532, 2019. 2

[25] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (tog)*, 38(5):1–12, 2019. 2

[26] Zi Jian Yew and Gim Hee Lee. Regtr: End-to-end point cloud correspondences with transformers. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6677–6686, 2022. 2, 3, 7

[27] Hao Yu, Fu Li, Mahdi Saleh, Benjamin Busam, and Slobodan Ilic. Cofinet: Reliable coarse-to-fine correspondences for robust pointcloud registration. *Advances in Neural Information Processing Systems*, 34:23872–23884, 2021. 2, 6, 7

[28] Hao Yu, Zheng Qin, Ji Hou, Mahdi Saleh, Dongsheng Li, Benjamin Busam, and Slobodan Ilic. Rotation-invariant transformer for point cloud matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5384–5393, 2023. 2

[29] Junle Yu, Luwei Ren, Yu Zhang, Wenhui Zhou, Lili Lin, and Guojun Dai. Peal: Prior-embedded explicit attention learning for low-overlap point cloud registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17702–17711, 2023. 2, 3, 6, 7

[30] Yongzhe Yuan, Yue Wu, Xiaolong Fan, Maoguo Gong, Qiguang Miao, and Wenping Ma. Inlier confidence calibration for point cloud registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5312–5321, 2024. 3

[31] Andy Zeng, Shuran Song, Matthias Nießner, Matthew Fisher, Jianxiong Xiao, and Thomas Funkhouser. 3dmatch: Learning local geometric descriptors from rgb-d reconstructions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1802–1811, 2017. 2, 5

[32] Guiyu Zhao, Zhentao Guo, Zewen Du, and Hongbin Ma. Cross-pcr: A robust cross-source point cloud registration framework. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 10403–10411, 2025. 6