3D Registration with Maximal Cliques

Xiyu Zhang   Jiaqi Yang*   Shikun Zhang   Yanning Zhang

School of Computer Science, Northwestern Polytechnical University, China

{2426988253, zhangshikun}@mail.nwpu.edu.cn; {jqyang, ynzhang}@nwpu.edu.cn

Abstract

As a fundamental problem in computer vision, 3D point cloud registration (PCR) aims to seek the optimal pose to align a point cloud pair. In this paper, we present a 3D registration method with maximal cliques (MAC). The key insight is to loosen the previous maximum clique constraint, and mine more local consensus information in a graph for accurate pose hypotheses generation: 1) A compatibility graph is constructed to render the affinity relationship between initial correspondences. 2) We search for maximal cliques in the graph, each of which represents a consensus set. We perform node-guided clique selection then, where each node corresponds to the maximal clique with the greatest graph weight. 3) Transformation hypotheses are computed for the selected cliques by the SVD algorithm and the best hypothesis is used to perform registration. Extensive experiments on U3M, 3DMatch, 3DLoMatch and KITTI demonstrate that MAC effectively increases registration accuracy, outperforms various state-of-the-art methods and boosts the performance of deep-learned methods. MAC combined with deep-learned methods achieves state-of-the-art registration recall of 95.7% / 78.9% on 3DMatch / 3DLoMatch.

1. Introduction

Point cloud registration (PCR) is an important and fundamental problem in 3D computer vision and has a wide range of applications in localization [13], 3D object detection [17] and 3D reconstruction [25]. Given two 3D scans of the same object (or scene), the goal of PCR is to estimate a six-degree-of-freedom (6-DoF) pose transformation that accurately aligns the two input point clouds. Using point-to-point feature correspondences is a popular and robust solution to the PCR problem. However, due to the limitations of existing 3D keypoint detectors & descriptors, the limited overlap between point clouds and data noise, corre-

*Corresponding author.
Code will be available at https://github.com/zhangxy0517/3D-Registration-with-Maximal-Cliques.
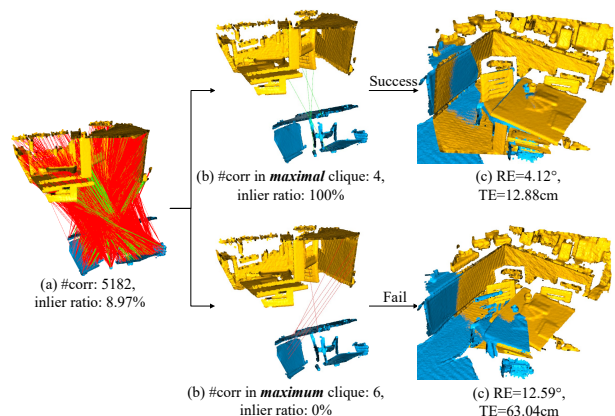

Figure 1. Comparison of **maximal** and **maximum** cliques on a low overlapping point cloud pair. Maximal cliques (MAC) effectively choose the optimal 6-DoF transformation hypothesis with low rotation error (RE) and translation error (TE) for two point clouds with a low inlier ratio, while the maximum clique fails in this case.

spondences generated by feature matching usually contain outliers, resulting in great challenges to accurate 3D registration.

The problem of 3D registration by handling correspondences with outliers has been studied for decades. We classify them into geometric-only and deep-learned methods. For geometric-only methods [5, 6, 21, 30, 31, 38–41], random sample consensus (RANSAC) and its variants perform an iterative sampling strategy for registration. Although RANSAC-based methods are simple and efficient, their performance is highly vulnerable when the outlier rate increases, and it requires a large number of iterations to obtain acceptable results. Also, a series of global registration methods based on branch-and-bound (BnB) are proposed to search the 6D parameter space and obtain the optimal global solution. The main weakness of these methods is the high computational complexity, especially when the correspondence set is of a large magnitude and has an extremely high outlier rate. For deep-learned methods, some [1–4, 9, 10, 14, 16, 18, 19, 27, 35] focus on improving

one module in the registration process, such as investigating more discriminate keypoint feature descriptors or more effective correspondence selection techniques, while the others [22, 29, 43] focus on registration in an end-to-end manner. However, deep-learned based methods require a large amount of data for training and usually lack generalization on different datasets. At present, it is still very challenging to achieve accurate registrations in the presence of heavy outliers and in cross-dataset conditions.

In this paper, we propose a geometric-only 3D registration method based on maximal cliques (MAC). The key insight is to loosen the previous maximum clique constraint, and mine more local consensus information in a graph to generate accurate pose hypotheses. We first model the initial correspondence set as a compatibility graph, where each node represents a single correspondence and each edge between two nodes indicates a pair of compatible correspondences. Second, we search for maximal cliques in the graph and then use node-guided clique filtering to match each graph node with the appropriate maximal clique containing it. *Compared with the maximum clique, MAC is a looser constraint and is able to mine more local information in a graph.* This helps us to achieve plenty of correct hypotheses from a graph. Finally, transformation hypotheses are computed for the selected cliques by the SVD algorithm. The best hypothesis is selected to perform registration using popular hypothesis evaluation metrics in the RANSAC family. To summarize, our main contributions are as follows:

- We introduce a hypothesis generation method named MAC. Our MAC method is able to mine more local information in a graph, compared with the previous maximum clique constraint. We demonstrate that hypotheses generated by MAC are of high accuracy even in the presence of heavy outliers.

- Based on MAC, we present a novel PCR method, which achieves state-of-the-art performance on U3M, 3DMatch, 3DLoMatch and KITTI datasets. Notably, our geometric-only MAC method outperforms several state-of-the-art deep learning methods [3, 9, 19, 27]. MAC can also be inserted as a module into multiple deep-learned frameworks [1, 10, 18, 29, 43] to boost their performance. MAC combined with GeoTransformer achieves the state-of-the-art registration recall of **95.7% / 78.9%** on 3DMatch / 3DLoMatch.

2. Related Work

2.1. Geometric-only PCR Methods

Various geometric-only methods [6, 8, 20, 36, 45] have been proposed recently. Typically, RANSAC and its variants [5, 13, 30, 31, 38–40] remain the dominant approaches to the problem of estimating a 6-DoF pose from correspondences. RANSAC iteratively samples correspondences from the initial set, generating and evaluating geometric estimations for each subset until a satisfactory solution is obtained. Efficient and robust evaluation metrics are extremely important for using RANSAC to achieve accurate registration. To address the current problems of time-consuming and noise-sensitive evaluation metrics, [40] analyzes the contribution of inliers and outliers during the computation and proposed several metrics that can effectively improve the registration performance of RANSAC. A large number of variants have also been proposed to achieve further improvement. For example, Rusu *et al*. [31] presented the simple consensus-based initial alignment (SAC-IA) method, which samples correspondences spread out on the point cloud and leverages the Huber penalty for evaluation. Graph cut RANSAC (GC-RANSAC) [5] uses the graph-cut algorithm before model re-fitting in the local optimization step. Compatibility-guided sample consensus (CG-SAC) [30] additionally considers the normal information of key points during the sampling process. Yang *et al*. [39] proposed the sample consensus by sampling compatibility triangles (SAC-COT) method, which generates estimations by ranking and sampling ternary loops from the compatibility graph. Although many previous efforts have been made, these methods suffer from low time efficiency and limited accuracy in cases with high outlier rates.

A series of globally optimal methods based on BnB have been proposed recently. Yang *et al*. [41] proposed globally optimal ICP (GO-ICP), which rationalizes the planning of ICP update tasks at different stages, and its biggest advantage is that it minimizes the local optimum. Bustos and Chin [6] presented guaranteed outlier removal (GORE), which calculates the tight lower bound and tight upper bound for each correspondence and reduces the size of correspondence set by rejecting true outliers. Motivated by GORE, Li [21] proposed a polynomial time outlier removal method, which seeks the tight lower and upper bound by calculating the costs of correspondence matrix (CM) and augmented correspondence matrix (ACM). However, BnB techniques are sensitive to the cardinality of the input and are time-consuming for large-scale inputs.

2.2. Deep-learned PCR Methods

In addition to geometric-only methods, recent works also adopt deep learning techniques to perform PCR. Some methods aim to detect more repeatable keypoints [4, 18] and extract more descriptive features [1, 10]. FCGF [10] computes the features in a single pass through a fully convolutional neural network without keypoint detection. D3Feat [4] uses a fully convolutional network to obtain local information of point clouds and a joint learning framework to achieve 3D local feature detection and description.
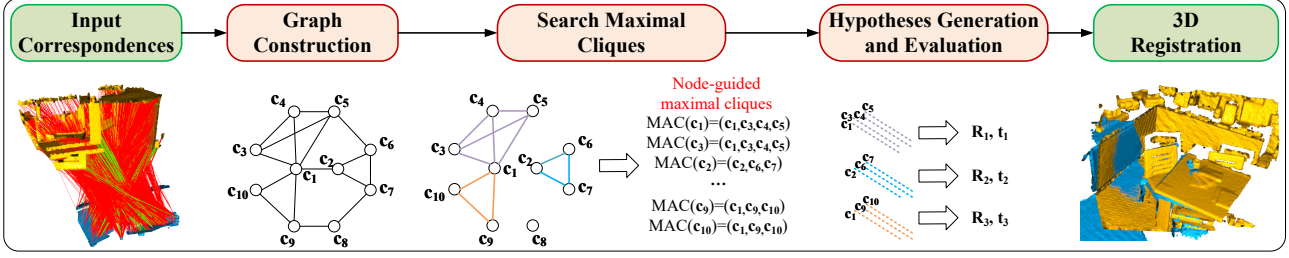
Figure 2. **Pipeline of MAC**. **1.** Construct a graph for the initial correspondence set. **2.** Select a set of maximal cliques from the graph as the consistent sets. **3.** Generate and evaluate the hypotheses according to the consistent sets. **4.** Select the best hypothesis to perform 3D registration.

Predator [18] applies an attention mechanism to extract salient points in overlapping regions of the point clouds, thus achieving robust registration in the presence of low overlap rates. Spinnet [1] extracts local features which are rotationally invariant and sufficiently informative to enable accurate registration. Some methods [3, 9, 14, 27] focus on efficiently distinguishing correspondences as inliers and outliers. Deep global registration (DGR) [9] and 3DReg-Net [27] classify a given correspondence by training end-to-end neural networks and using operators such as sparse convolution and point-by-point MLP. PointDSC [3] explicitly explores spatial consistency for removing outlier correspondences and 3D point cloud registration. Fu *et al*. [14] proposed a registration framework that utilizes deep graph matching (RGM) that can find robust and accurate point-to-point correspondences. More recently, several methods [29, 43] follow the detection-free methods and estimate the transformation in an end-to-end way. CoFiNet [43] extracts correspondences from coarse to fine without keypoint detection. GeoTransformer [29] learns geometric features for robust superpoint matching and is robust in low-overlap cases and invariant to rigid transformation.

While deep learning techniques have demonstrated a great potential for PCR, these methods require a large amount of training data and their generalization is not always promising. By contrast, MAC does not require any training data and achieves more advanced performance than several deep-learned methods. Moreover, MAC can be served as a drop-on module in deep learning frameworks to boost their performance.

3. MAC

3.1. Problem Formulation

For two point clouds $\mathbf{P}^s$ and $\mathbf{P}^t$ to be aligned, we first extract local features for them using geometric or learned descriptors. Let $\mathbf{p}^s$ and $\mathbf{p}^t$ denote the points in the $\mathbf{P}^s$ and $\mathbf{P}^t$, respectively. An initial correspondence set $\mathbf{C}_{initial} = \{\mathbf{c}\}$ is formed by matching feature descriptors,

where $\mathbf{c} = (\mathbf{p}^s, \mathbf{p}^t)$. MAC estimates the 6-DoF pose transformation between $\mathbf{P}^s$ and $\mathbf{P}^t$ from $\mathbf{C}_{initial}$.

Our method is technically very simple, and its pipeline is shown in Fig. 2.

3.2. Graph Construction

The graph space can more accurately depict the affinity relationship between correspondences than the Euclidean space. Therefore, we model the initial correspondences as a compatibility graph, where correspondences are represented by nodes and edges link nodes that are geometrically compatible. Here, we consider two approaches to construct a compatibility graph.

- **First Order Graph.** The first order graph (FOG) is constructed based on the rigid distance constraint between the correspondence pair $(\mathbf{c}_i, \mathbf{c}_j)$, which can be quantitatively measured as:

$$S_{dist}(\mathbf{c}_i, \mathbf{c}_j) = \left| \left\| \mathbf{p}_i^s - \mathbf{p}_j^s \right\| - \left\| \mathbf{p}_i^t - \mathbf{p}_j^t \right\| \right|. \quad (1)$$

The compatibility score between $\mathbf{c}_i$ and $\mathbf{c}_j$ is given as:

$$S_{cmp}(\mathbf{c}_i, \mathbf{c}_j) = \exp(-\frac{S_{dist}(\mathbf{c}_i, \mathbf{c}_j)^2}{2d_{cmp}^2}), \quad (2)$$

where d_{cmp} is a distance parameter. Notably, if $S_{cmp}(\mathbf{c}_i, \mathbf{c}_j)$ is greater than a threshold t_{cmp}, $\mathbf{c}_i$ and $\mathbf{c}_j$ form an edge $\mathbf{e}_{ij}$ and $S_{cmp}(\mathbf{c}_i, \mathbf{c}_j)$ is the weight of $\mathbf{e}_{ij}$, otherwise $S_{cmp}(\mathbf{c}_i, \mathbf{c}_j)$ will be set to 0. Since the compatibility graph is undirected, the weight matrix $\mathbf{W}_{FOG}$ is symmetric.

- **Second Order Graph.** The previous study [8] proposes a second order compatibility measure, which relates to the number of commonly compatible correspondences in the global set. The second order graph (SOG) evolves from FOG. The weight matrix $\mathbf{W}_{SOG}$ can be calculated as:

$$\mathbf{W}_{SOG} = \mathbf{W}_{FOG} \odot (\mathbf{W}_{FOG} \times \mathbf{W}_{FOG}), \quad (3)$$

where $\odot$ represents the element-wise product between two matrices.

Both graph construction methods can adapt to our frameworks. Compared with FOG, 1) SOG has stricter edge construction conditions and a higher degree of compatibility with adjacent nodes; 2) SOG is sparser, which facilitates a more rapid search of cliques. In Sec. 4.5, we experimentally compare FOG and SOG in our MAC framework.

3.3. Search Maximal Cliques

Given an undirected graph $G = (\mathbf{V}, \mathbf{E})$, clique $C = (\mathbf{V}', \mathbf{E}')$, $\mathbf{V}' \subseteq \mathbf{V}$, $\mathbf{E}' \subseteq \mathbf{E}$ is a subset of G, in which any two nodes are connected by edges. A maximal clique is a clique that cannot be extended by adding any nodes. In particular, the maximal clique with the most nodes is the maximum clique of a graph.

Searching for Maximal cliques. To generate hypotheses, RANSAC-based methods repeatedly take random samples from the correspondence set. Nevertheless, they fail to fully mine the affinity relationships between correspondences. Theoretically, inliers would form cliques in the graph, because inliers are usually geometrically compatible with each other. Previous works [23,24,28,36] focus on searching for maximum cliques in the graph, however, the maximum clique is a very tight constraint that only focuses on the global consensus information in a graph. Instead, we loosen the constraint and leverage maximal cliques to mine more local graph information.

By using the $igraph\_maximal\_cliques$ function in the igraph[1] C++ library, which makes use of a modified Bron-Kerbosch algorithm [12], the search of maximal cliques can be very efficient. The process's worst time complexity is $\mathcal{O}(d(n-d)3^{(d/3)})$, where d is the degeneracy of the graph. Note that d is typically small in our problem because the graph is usually sparse when dealing with point cloud correspondences.

Node-guided Clique Selection. After executing the maximal clique searching procedure, we obtain the maximal clique set $MAC_{initial}$. In practice, $MAC_{initial}$ usually contains tens of thousands of maximal cliques, which will make it very time-consuming if we consider all maximal cliques. We introduce a node-guided clique selection method in this section to reduce $|MAC_{initial}|$. First, we calculate the weight for each clique in $MAC_{initial}$. Given a clique $C_i = (\mathbf{V}_i, \mathbf{E}_i)$, the weight w_{C_i} is calculated as:

$$w_{C_i} = \sum_{e_j \in \mathbf{E}_i} w_{e_j}, \qquad (4)$$

where w_{e_j} represents the weight of edge e_j in $\mathbf{W}_{SOG}$. A node may be included by multiple maximal cliques and we

---

[1]https://igraph.org

only retain the one with the greatest weight for that node. Then, duplicated cliques are removed from the rest, obtaining $MAC_{selected}$. The motivation behind this is to use information about the local geometric structure around graph nodes to find the best consistent set of corresponding nodes. It is clear that the number of maximal cliques $|MAC_{selected}|$ will not exceed $|\mathbf{V}|$. We could send these maximal cliques directly to the following stages for 3D registration. However, when $|\mathbf{V}|$ is quite large, the number of retained maximal cliques can still be very large. Here, we propose several techniques to further filter the maximal cliques.

- **Normal consistency.** In the maximal cliques, we find that the normal consistency is satisfied between each correspondence. Given two correspondences $\mathbf{c}_i = (\mathbf{p}_i^s, \mathbf{p}_i^t)$, $\mathbf{c}_j = (\mathbf{p}_j^s, \mathbf{p}_j^t)$ and the normal vectors $\mathbf{n}_i^s, \mathbf{n}_j^s, \mathbf{n}_i^t, \mathbf{n}_j^t$ at the four points, the angular difference $\alpha_{ij}^s = \angle(\mathbf{n}_i^s, \mathbf{n}_j^s)$, $\alpha_{ij}^t = \angle(\mathbf{n}_i^t, \mathbf{n}_j^t)$ between the normal vectors can be calculated then. The following inequality ought to hold if $\mathbf{c}_i$ and $\mathbf{c}_j$ are normal consistent:

$$\left| \sin\alpha_{ij}^s - \sin\alpha_{ij}^t \right| < t_\alpha, \qquad (5)$$

where t_α is a threshold for determining whether the angular differences are similar.

- **Clique ranking.** We organize $MAC_{selected}$ in a descending order using the clique's weight w_{C_i}. The top-K ones are supposed to be more likely to produce correct hypotheses. This makes it flexible to control the number of hypotheses.

These techniques' experimental analysis is presented in Sec. 4.5.

3.4. Hypothesis Generation and Evaluation

Each maximal clique filtered from the previous step represents a consistent set of correspondences. By applying the SVD algorithm to each consistency set, we can obtain a set of 6-DoF pose hypotheses.

- **Instance-equal SVD.** Transformation estimation of correspondences is often implemented with SVD. Instance-equal means that the weights of all correspondences are equal.

- **Weighted SVD.** Assigning weights to correspondences is commonly adopted by recent PCR methods [8,9,27,29]. Correspondence weights can be derived by solving the eigenvectors of a compatibility matrix constructed for a compatibility graph. Here, we take the primary eigenvalues of $\mathbf{W}_{SOG}$ as correspondence weights.

The final goal of MAC is to estimate the optimal 6-DoF rigid transformation (composed of a rotation pose $\mathbf{R}^* \in SO(3)$ and a translation pose $\mathbf{t}^* \in \mathbb{R}^3$) that maximizes the objective function as follow:

$$(\mathbf{R}^*, \mathbf{t}^*) = \arg\max_{\mathbf{R},\mathbf{t}} \sum_{i=1}^{N} s(\mathbf{c}_i), \qquad (6)$$

where $\mathbf{c}_i \in \mathbf{C}_{initial}$, $N = |\mathbf{C}_{initial}|$, and $s(\mathbf{c}_i)$ represents the score of $\mathbf{c}_i$. We consider several RANSAC hypothesis evaluation metrics here [40], including mean average error (MAE), mean square error (MSE) and inlier count. Their behaviors will be experimentally compared in Sec. 4.5. The best hypothesis is taken to perform 3D registration then.

4. Experiments

4.1. Experimental Setup

Datasets. We consider four datasets, i.e, the object-scale dataset U3M [26], the scene-scale indoor datasets 3DMatch [44] & 3DLoMatch [18], and the scene-scale outdoor dataset KITTI [15]. U3M has 496 point cloud pairs. 3DLoMatch is the subset of 3DMatch, where the overlap rate of the point cloud pairs ranges from 10% to 30%, which is very challenging. For KITTI, we follow [3,8] and obtain 555 pairs of point clouds for testing.

Evaluation Criteria. We follow [39] that employs the root mean square error (RMSE) metric to evaluate the 3D point cloud registration performance on the U3M object-scale dataset. In addition, we employ the rotation error (RE) and translation error (TE) to evaluate the registration results on the scene-scale dataset. By referring to the settings in [9], the registration is considered successful when the RE $\leq 15°$, TE ≤ 30 cm on 3DMatch & 3DLoMatch datasets, and RE $\leq 5°$, TE ≤ 60 cm on KITTI dataset. We define a dataset's registration accuracy as the ratio of success cases to the number of point cloud pairs to be registered.

Implementation Details. Our method is implemented in C++ based on the point cloud library (PCL) [32] and igraph library. For U3M, we use the Harris3D (H3D) [33] keypoint detector and the signatures of histograms of orientation (SHOT) [34] descriptor for initial correspondence generation as in [42]. For 3DMatch and 3DLoMatch datasets, we use the fast point features histograms (FPFH) [31] descriptor and fully convolutional geometric features (FCGF) [10] descriptor to generate the initial correspondence set. The main steps in the comparative experimental sections are SOG construction, searching node-guided maximal cliques, hypotheses generation by instance-equal SVD and evaluation by MAE. Default values for compatibility threshold t_{cmp} and distance parameter d_{cmp} mentioned in Sec. 3.2 are 0.99 and 10 pr respectively; if input matches exceed 5000, t_{cmp} is set to 0.999 to reduce computation. Here, 'pr' is
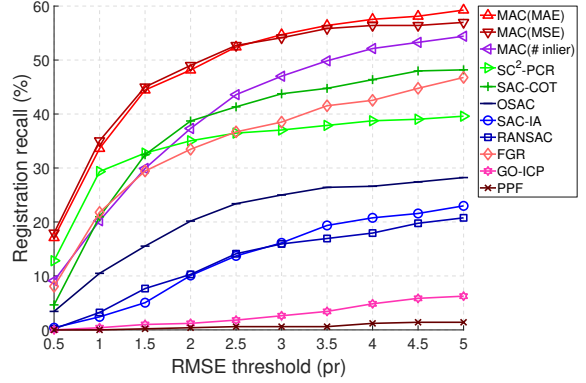

Figure 3. Registration performance of tested point cloud registration methods on U3M.

	FPFH			FCGF		
	RR(%)	RE(°)	TE(cm)	RR(%)	RE(°)	TE(cm)
i) Traditional						
SM [20]	55.88	2.94	8.15	86.57	2.29	7.07
FGR [45]	40.91	4.96	10.25	78.93	2.90	8.41
RANSAC-1M [13]	64.20	4.05	11.35	88.42	3.05	9.42
RANSAC-4M [13]	66.10	3.95	11.03	91.44	2.69	8.38
GC-RANSAC [5]	67.65	2.33	6.87	92.05	2.33	7.11
TEASER++ [36]	75.48	2.48	7.31	85.77	2.73	8.66
CG-SAC [30]	78.00	2.40	6.89	87.52	2.42	7.66
SC²-PCR [8]	83.73	2.18	6.70	93.16	2.09	6.51
ii) Deep learned						
3DRegNet [27]	26.31	3.75	9.60	77.76	2.74	8.13
DGR [9]	32.84	2.45	7.53	88.85	2.28	7.02
DHVR [19]	67.10	2.78	7.84	91.93	2.25	7.08
PointDSC [3]	72.95	2.18	6.45	91.87	2.10	6.54
MAC	**84.10**	**1.96**	**6.18**	**93.72**	**1.89**	**6.03**

Table 1. Registration results on 3DMatch dataset.

a distance unit called point cloud resolution [42]. Normal vectors are calculated using the NormalEstmation class of PCL with the 20 nearest neighboring points. When searching maximal cliques, the lower bound on clique size is set to 3 with no upper bound defined. All experiments were implemented with an Intel 12700H CPU and 32 GB RAM.

4.2. Results on U3M Dataset

We perform an extensive comparison in Fig. 3. Here, the following methods are tested, including SAC-COT [39], OSAC [37], SAC-IA [31], RANSAC [13], SC²-PCR [8], FGR [45], GO-ICP [41], and PPF [11], where the former four are RANSAC-based methods. The RMSE threshold is varied from 0.5 pr to 5 pr with a step of 0.5 pr.

The results indicate that MAC performs best and significantly outperforms all tested RANSAC-fashion estimators, such as SAC-COT, OSAC, SAC-IA, and RANSAC. The registration performance of MAC based on the MAE evaluation metric is the best on U3M.

	FPFH			FCGF		
	RR(%)	RE(°)	TE(cm)	RR(%)	RE(°)	TE(cm)
i) Traditional						
RANSAC-1M [13]	0.67	10.27	15.06	9.77	7.01	14.87
RANSAC-4M [13]	0.45	10.39	20.03	10.44	6.91	15.14
TEASER++ [36]	35.15	4.38	10.96	46.76	4.12	12.89
SC^2-PCR [8]	<u>38.57</u>	<u>4.03</u>	10.31	<u>58.73</u>	<u>3.80</u>	<u>10.44</u>
ii) Deep learned						
DGR [9]	19.88	5.07	13.53	43.80	4.17	10.82
PointDSC [3]	20.38	4.04	<u>10.25</u>	56.20	3.87	10.48
MAC	**40.88**	**3.66**	**9.45**	**59.85**	**3.50**	**9.75**

Table 2. Registration results on 3DLoMatch dataset.

# Samples	3DMatch RR(%)					3DLoMatch RR(%)				
	5000	2500	1000	500	250	5000	2500	1000	500	250
FCGF [10]	85.1	84.7	83.3	81.6	71.4	40.1	41.7	38.2	35.4	26.8
SpinNet [1]	88.6	86.6	85.5	83.5	70.2	59.8	54.9	48.3	39.8	26.8
Predator [18]	89.0	89.9	90.6	88.5	86.6	59.8	61.2	62.4	60.8	58.1
CoFiNet [43]	89.3	88.9	88.4	87.4	87.0	67.5	66.2	64.2	63.1	61.0
GeoTransformer [29]	92.0	91.8	91.8	91.4	91.2	75.0	74.8	74.2	74.1	73.5
FCGF+MAC	91.3	92.2	91.6	90.4	85.6	57.2	56.0	52.6	42.4	32.1
	6.2↑	7.5↑	8.3↑	8.8↑	14.2↑	17.1↑	14.3↑	14.4↑	7.0↑	5.3↑
SpinNet+MAC	95.3	95.1	93.3	91.4	81.2	72.8	69.9	59.2	54.8	32.1
	6.7↑	8.5↑	7.8↑	7.9↑	11.0↑	13.0↑	15.0↑	10.9↑	15.0↑	5.3↑
Predator+MAC	94.6	94.4	94.0	93.5	92.3	70.9	70.4	69.8	67.2	64.1
	5.6↑	4.5↑	3.4↑	5.0↑	5.7↑	11.1↑	9.2↑	7.4↑	6.4↑	6.0↑
CoFiNet+MAC	94.1	94.4	94.5	93.8	92.7	71.6	71.5	70.6	69.2	68.1
	4.8↑	5.5↑	6.1↑	6.4↑	5.7↑	4.1↑	5.3↑	6.4↑	6.1↑	7.1↑
GeoTransformer+MAC	**95.7**	**95.7**	**95.2**	**95.3**	**94.6**	**78.9**	**78.7**	**78.2**	**77.7**	**76.6**
	3.7↑	3.9↑	3.4↑	3.9↑	3.4↑	3.9↑	3.9↑	4.0↑	3.6↑	3.1↑

Table 3. Performance boosting for deep-learned methods when combined with MAC.

4.3. Results on 3DMatch & 3DLoMatch Datasets

PCR methods comparison. Both geometric-only and deep-learned methods are considered for comparison, including SM [20], FGR [45], RANSAC [13], TEASER++ [36], CG-SAC [30], SC^2-PCR [8], 3DRegNet [27], DGR [9], DHVR [19] and PointDSC [3]. Results are shown in Tables 1 and 2.

The following conclusions can be made: 1) regardless of which descriptor is used, MAC outperforms all compared methods on both 3DMatch and 3DLoMatch datasets, indicating its strong ability to register indoor scene point clouds; 2) even compared with deep-learned methods, MAC still achieves better performance without any data training; 3) in addition to the registration recall (RR) metric, MAC achieves the best RE and TE metrics. This indicates that registrations by MAC are very accurate and MAC is able to align low overlapping data.

Boosting deep-learned methods with MAC. Several kinds of state-of-the-art deep-learned methods are integrated with MAC for evaluation. The considered methods are FCGF [10], SpinNet [1], Predator [18], CoFiNet [43] and Geo-Transformer [29]. Each method is tested under a different number of samples, which refer to the number of sampled points or correspondences. Results are reported in Table 3.

Remarkably, MAC dramatically improves the registration recall under all tested methods on both 3DMatch and 3DLoMatch datasets. Notably, the performance of Spin-Net, Predator and CoFiNet after boosting by MAC exceeds

	FPFH			FCGF		
	RR(%)	RE(°)	TE(cm)	RR(%)	RE(°)	TE(cm)
i) Traditional						
FGR [45]	5.23	0.86	43.84	89.54	0.46	25.72
TEASER++ [36]	91.17	1.03	17.98	94.96	0.38	**13.69**
RANSAC [13]	74.41	1.55	30.20	80.36	0.73	26.79
CG-SAC [30]	74.23	0.73	14.02	83.24	0.56	22.96
SC^2-PCR [8]	<u>99.28</u>	<u>0.39</u>	8.68	**97.84**	**0.33**	20.58
ii) Deep learned						
DGR [9]	77.12	1.64	33.10	96.90	<u>0.34</u>	21.70
PointDSC [3]	<u>98.92</u>	**0.38**	**8.35**	**97.84**	**0.33**	20.32
MAC	**99.46**	0.40	<u>8.46</u>	**97.84**	<u>0.34</u>	<u>19.34</u>

Table 4. Registration results on KITTI dataset.

that of GeoTransformer. MAC working with GeoTransformer achieves state-of-the-art registration recall of **95.7%** / **78.9%** on 3DMatch / 3DLoMatch. The results suggest that: 1) MAC can greatly boost existing deep-learned methods; 2) MAC is not sensitive to the number of samples.

4.4. Results on KITTI Dataset

In Table 4, the results of DGR [9], PointDSC [3], TEASER++ [36], RANSAC [13], CG-SAC [30], SC^2-PCR [8] and MAC are reported for comparison.

As shown by the table, in terms of the registration recall performance, MAC presents the best and is tied for the best results with FPFH and FCGF descriptor settings, respectively. MAC also has a lower TE than the state-of-the-art geometric-only method SC^2-PCR. Note that outdoor point clouds are significantly sparse and non-uniformly distributed. The registration experiments on the object, indoor scene, and outdoor scene datasets consistently verify that MAC holds good generalization ability in different application contexts.

4.5. Analysis Experiments

In this section, we perform ablation studies and analysis experiments on both 3DMatch and 3DLoMatch datasets. We progressively experiment with the techniques proposed in Sec. 3, and the results are shown in Table 5. The quality of generated hypotheses is analyzed in Table 6. The performance upper bound is studied in Table 7. Table 8 presents the time efficiency analysis of MAC.

Performing feature matching selection. Before 3D registration, a popular way is to perform outlier rejection to reduce the correspondence set. Here we employ geometric consistency (GC) [7], which is independent of the feature space and associates the largest consistent cluster relating to the compatibility among correspondences.

By comparing Row 1 and 2 of Table 5, GC has a negative impact on MAC performance, potentially due to that some inliers are also removed in this process. This demonstrates that MAC can still perform well even if the initial correspondence set is directly utilized as input without any filtering.

		FOG	SOG	GC	MC	NG	NC	CR	SVD	W-SVD	MAE	MSE	# inlier	RR(%)		RE(°)		TE(cm)	
FPFH	1)		✓			✓			✓				✓	83.86	39.14	2.17	4.01	6.51	9.94
	2)		✓	✓		✓			✓				✓	77.02	26.61	2.10	3.83	6.19	9.49
	3)	✓				✓			✓				✓	82.26	39.02	2.12	3.98	6.43	9.89
	4)		✓						✓				✓	83.49	38.91	2.22	4.11	6.65	10.05
	5)		✓			✓				✓			✓	83.67	38.85	2.15	4.03	6.53	9.82
	6)		✓			✓			✓		✓			**84.10**	**40.88**	**1.96**	**3.66**	**6.18**	**9.45**
	7)		✓			✓			✓			✓		82.93	39.98	1.95	3.66	6.12	9.48
	8)		✓			✓	✓		✓				✓	82.44	38.46	2.16	3.97	6.41	9.85
	9)		✓		✓	✓			✓				✓	74.06	31.11	2.08	3.89	6.17	9.82
	10) Top100		✓			✓		✓	✓				✓	82.01	37.79	2.13	4.02	6.42	9.82
	11) Top200		✓			✓		✓	✓				✓	83.18	38.85	2.16	4.08	6.55	9.91
	12) Top500		✓			✓		✓	✓				✓	83.06	38.85	2.14	4.03	6.47	9.81
	13) Top1000		✓			✓		✓	✓				✓	83.30	38.91	2.16	4.05	6.53	9.84
	14) Top2000		✓			✓		✓	✓				✓	83.36	38.79	2.14	4.02	6.52	9.78
FCGF	1)		✓			✓			✓				✓	93.41	59.80	2.04	3.78	6.33	10.16
	2)		✓	✓		✓			✓				✓	91.68	49.97	1.99	3.64	6.23	9.90
	3)	✓				✓			✓				✓	93.35	59.24	2.04	3.67	6.28	9.99
	4)		✓						✓				✓	92.91	59.07	2.06	3.88	6.33	10.20
	5)		✓			✓				✓			✓	93.16	59.46	2.04	3.76	6.26	10.00
	6)		✓			✓			✓		✓			**93.72**	**59.85**	**1.89**	**3.50**	**6.03**	**9.75**
	7)		✓			✓			✓			✓		93.59	59.01	1.86	3.49	6.00	9.61
	8)		✓			✓	✓		✓				✓	93.28	59.63	2.02	3.73	6.24	9.98
	9)		✓		✓	✓			✓				✓	87.86	49.35	2.00	3.61	6.09	9.60
	10) Top100		✓			✓		✓	✓				✓	92.42	57.44	2.00	3.75	6.21	10.00
	11) Top200		✓			✓		✓	✓				✓	93.22	57.83	2.01	3.75	6.29	10.06
	12) Top500		✓			✓		✓	✓				✓	93.22	58.90	2.02	3.78	6.33	10.02
	13) Top1000		✓			✓		✓	✓				✓	93.35	59.40	2.05	3.78	6.32	10.18
	14) Top2000		✓			✓		✓	✓				✓	93.35	59.52	2.04	3.78	6.33	10.19

Table 5. Analysis experiments on 3DMatch / 3DLoMatch. **FOG**: First order compatibility graph. **SOG**: Second order compatibility graph. **GC**: Use geometric consistency to preliminarily perform outlier rejection. **MC**: Search the *maximum* clique instead of *maximal* cliques. **NG**: Node-guided clique selection. **NC**: Normal consistency. **CR**: Clique ranking. **W-SVD**: Weighted SVD.

Graph construction choices. We test the performance of MAC by using different graph construction approaches.

As shown in Row 1 and 3 of Table 5, the registration recall obtained by using SOG is 1.6% higher than using FOG when combined with FPFH, and 0.06% higher when combined with FCGF on 3DMatch. Also, the registration recall obtained by using SOG is 0.12% higher than using FOG when combined with FPFH, and 0.56% higher when combined with FCGF on 3DLoMatch. Therefore, SOG is more suitable for MAC. Detailed analyzing descriptions can be found in the supplementary.

Maximum or maximal clique. To justify the advantages of maximal cliques, we change the search strategy of MAC to the maximum cliques and test the registration performance.

As shown in Row 1 and 9 in Table 5, applying maximal cliques surpasses maximum by 9.8% when combined with FPFH, and 5.55% higher when combined with FCGF on 3DMatch. Besides, the registration recall obtained by using maximal cliques is 8.03% higher than using the maximum cliques when combined with FPFH and 10.45% higher when combined with FCGF on 3DLoMatch. There are several reasons for this: 1) maximal cliques include the maximum cliques and additionally consider local graph constraints, so the search for maximal cliques can make use of both local and global information in the compatibility graph; 2) the maximum clique is a very tight constraint which requires maximizing the number of mutually compatible correspondences, but it does not guarantee the optimal result.

Node-guided clique selection. We compare the performance with and without node-guided (NG) clique selection for maximal cliques search.

Comparing Row 1 and 4 in Table 5, using NG achieves a recall improvement of 0.37% when combined with FPFH, and 0.5% improvement when combined with FCGF on 3DMatch. Also, using NG achieves a recall improvement of 0.23% with FPFH and 0.73% improvement with FCGF on 3DLoMatch. It is worth noting that while NG improves recall, the mean RE and mean TE are also decreasing. For example, NG reduces the mean RE by 0.1°and the mean TE by 0.11 cm with FPFH on 3DLoMatch. NG effectively reduces the number of calculations in the subsequent steps and promises accurate hypotheses.

Different approaches for clique filtering. We test the effectiveness of the two filtering methods, normal consistency and clique ranking.

1) Normal consistency: comparing Row 1 and 8 in Table 5, NC slightly degrades MAC's performance. *2) Clique ranking:* Row 10 to 14 demonstrate that the registration recall tends to increase as K increases, suggesting that larger K yields a subset of cliques that generate more correct hypotheses. Remarkably, setting K to 100 can already achieve outstanding performance.

Employing instance-equal or weighted SVD. The comparisons of instance-equal and weighted SVD are shown in Rows 1 and 5 of Table 5.

# hypotheses	3DMatch				3DLoMatch			
	RANSAC		MAC		RANSAC		MAC	
	FCGF	FPFH	FCGF	FPFH	FCGF	FPFH	FCGF	FPFH
100	10.45	0.76	61.94	50.67	1.25	0.05	30.47	12.22
200	20.76	1.50	119.20	89.27	2.52	0.09	55.57	17.59
500	51.74	3.68	269.06	162.41	6.21	0.21	109.32	23.32
1000	103.65	7.39	456.18	217.32	12.43	0.41	156.11	26.02
2000	208.24	14.90	669.32	254.13	24.80	0.81	202.12	29.31

Table 6. Comparison of the number of correct hypotheses generated by MAC and RANSAC on 3DMatch and 3DLoMatch.

	3DMatch RR(%)	3DLoMatch RR(%)
MAC-1	98.46	91.24
MAC-5	97.10	83.32
MAC-10	96.43	77.93
MAC-20	94.70	70.47
MAC-50	91.13	56.37
MAC-origin	93.72	59.85

Table 7. Registration recall on 3DMatch with FCGF setting based on judging MAC's hypotheses. MAC-n: a point cloud pair is considered alignable if at least n hypotheses are correct.

Weighted SVD is slightly inferior to instance-equal SVD. This suggests that samples in MACs are already very consistent, indicating no additional weighting strategies are required.

Varying hypothesis evaluation metrics. Here we compare three evaluation metrics, including MAE, MSE and inlier count, for MAC hypothesis evaluation.

As shown in Row 1, 6 and 7, MAC with MAE achieves the best performance. In Table 5, MAE achieves a recall improvement of 0.24% when combined with FPFH, and 0.31% improvement when combined with FCGF on 3DMatch compared with the commonly used inlier count metric. Also, MAE has a 1.74% improvement when combined with FPFH, and 0.05% when combined with FCGF on 3DLoMatch compared with inlier count. MAE is also very effective in reducing RE and TE. For instance, MAE reduces the mean RE by 0.35°and the mean TE by 0.49 cm with FPFH on 3DLoMatch.

Comparison with RANSAC hypotheses. We evaluate the quality of the generated hypotheses by comparing the hypotheses from RANSAC and MAC with the ground truth transformation. The results are shown in Table 6.

Compared to RANSAC, which randomly selects correspondences and generates hypotheses from the correspondence set without geometric constraints, MAC effectively generates more convincing hypotheses from maximal cliques in the compatibility graph, which fully exploits the consensus information in the graph.

The performance upper bound of MAC. Given an ideal hypothesis evaluation metric, allowing a point cloud pair can be aligned as long as correct hypotheses can be generated. This can test the performance upper bound of MAC. We vary the judging threshold for the number of generated correct hypotheses and report the results in Table 7.

# correspondences	Graph Construction	Search Maximal Cliques	Node-guided Clique Selection	Pose Estimation	Total
250	1.03 (14.55%)	5.24 (74.01%)	0.58 (8.19%)	0.23 (3.25%)	7.08
500	4.07 (17.54%)	15.67 (67.51%)	3.12 (13.44%)	0.35 (1.51%)	23.21
1000	16.90 (29.85%)	36.60 (64.65%)	1.88 (3.32%)	1.23 (2.18%)	56.61
2500	153.92 (53.29%)	104.03 (36.02%)	4.97 (1.72%)	25.93 (8.97%)	288.85
5000	887.03 (27.16%)	1579.61 (48.37%)	65.40 (2.00%)	733.38 (22.47%)	3265.42

Table 8. Average consumed time (ms) per point cloud pair on the 3DMatch dataset. Predator is used for generating correspondences.

Impressively, MAC-1 achieves registration recalls of **98.46% / 91.24%** on 3DMatch / 3DLoMatch. This indicates that even on low overlapping datasets, MAC is able to produce correct hypotheses for most point cloud pairs. In addition, we can deduce that MAC's performance can be further improved with better hypothesis evaluation metrics.

Time consumption of MAC. We employ Predator [18] to generate correspondences with different magnitudes to test the time performance of MAC. The time consumption is reported in Table 8.

The following observations can be made. 1) In general, MAC can complete 3D registration in only tens of milliseconds when the number of correspondences is smaller than 1000. Even with an input with 2500 correspondences, the time consumption is about 0.29 seconds. Note that MAC is implemented on the CPU only. 2) As the number of correspondences increases from 250 to 2500, there is an increase in time cost for graph construction due to $\mathbf{W}_{SOG}$ computation taking more time. 3) When the number of correspondences reaches 5000, there is a large rise in the time cost of MAC's registration. The significant increase in the input size makes the search for maximal cliques more time-consuming. However, MAC is not sensitive to the cardinality of the input correspondence set, as verified in Table 3. Hence, using sparse inputs for MAC can produce outstanding performance while making registration efficient.

5. Conclusion

In this paper, we presented MAC to solve PCR by using the maximal clique constraint to generate precise pose hypotheses from correspondences. Our method achieves state-of-the-art performance on all tested datasets and can adapt to deep-learned methods to boost their performance.

Limitation. As shown in Table 7 and Table 1, MAC produces accurate hypotheses but may fail to find them. In the future, we plan to develop a more convincing hypothesis evaluation technique utilizing semantic information.

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. 1, 2, 3, 6

[2] Yasuhiro Aoki, Hunter Goforth, Rangaprasad Arun Srivatsan, and Simon Lucey. Pointnetlk: Robust & efficient point cloud registration using pointnet. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7163–7172, 2019. 1

[3] Xuyang Bai, Zixin Luo, Lei Zhou, Hongkai Chen, Lei Li, Zeyu Hu, Hongbo Fu, and Chiew-Lan Tai. Pointdsc: Robust point cloud registration using deep spatial consistency. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 15859–15869. IEEE, 2021. 1, 2, 3, 5, 6

[4] 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. 1, 2

[5] Daniel Barath and Jiří Matas. Graph-cut ransac. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6733–6741, 2018. 1, 2, 5

[6] Alvaro Parra Bustos and Tat-Jun Chin. Guaranteed outlier removal for point cloud registration with correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12):2868–2882, 2017. 1, 2

[7] Hui Chen and Bir Bhanu. 3d free-form object recognition in range images using local surface patches. *Pattern Recognition Letters*, 28(10):1252–1262, 2007. 6

[8] Zhi Chen, Kun Sun, Fan Yang, and Wenbing Tao. Sc2-pcr: A second order spatial compatibility for efficient and robust point cloud registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13221–13231, 2022. 2, 3, 4, 5, 6

[9] Christopher Choy, Wei Dong, and Vladlen Koltun. Deep global registration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2514–2523. IEEE, 2020. 1, 2, 3, 4, 5, 6

[10] 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. 1, 2, 5, 6

[11] Bertram Drost, Markus Ulrich, Nassir Navab, and Slobodan Ilic. Model globally, match locally: Efficient and robust 3d object recognition. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 998–1005. IEEE, 2010. 5

[12] David Eppstein, Maarten Löffler, and Darren Strash. Listing all maximal cliques in sparse graphs in near-optimal time. In *International Symposium on Algorithms and Computation*, pages 403–414. Springer, 2010. 4

[13] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. 1, 2, 5, 6

[14] Kexue Fu, Shaolei Liu, Xiaoyuan Luo, and Manning Wang. Robust point cloud registration framework based on deep graph matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8893–8902, 2021. 1, 3

[15] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3354–3361. IEEE, 2012. 5

[16] Zan Gojcic, Caifa Zhou, Jan D Wegner, and Andreas Wieser. The perfect match: 3d point cloud matching with smoothed densities. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5545–5554, 2019. 1

[17] Yulan Guo, Mohammed Bennamoun, Ferdous Sohel, Min Lu, and Jianwei Wan. 3d object recognition in cluttered scenes with local surface features: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(11):2270–2287, 2014. 1

[18] 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. 1, 2, 3, 5, 6, 8

[19] Junha Lee, Seungwook Kim, Minsu Cho, and Jaesik Park. Deep hough voting for robust global registration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15994–16003, 2021. 1, 2, 5, 6

[20] Marius Leordeanu and Martial Hebert. A spectral technique for correspondence problems using pairwise constraints. 2005. 2, 5, 6

[21] Jiayuan Li. A practical o (n2) outlier removal method for point cloud registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. 1, 2

[22] 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

[23] Muyuan Lin, Varun Murali, and Sertac Karaman. A planted clique perspective on hypothesis pruning. *IEEE Robotics and Automation Letters*, 7(2):5167–5174, 2022. 4

[24] Yu-Kai Lin, Wen-Chieh Lin, and Chieh-Chih Wang. K-closest points and maximum clique pruning for efficient and effective 3-d laser scan matching. *IEEE Robotics and Automation Letters*, 7(2):1471–1477, 2022. 4

[25] Ajmal S Mian, Mohammed Bennamoun, and Robyn A Owens. Automatic correspondence for 3d modeling: an extensive review. *International Journal of Shape Modeling*, 11(02):253–291, 2005. 1

[26] Ajmal S Mian, Mohammed Bennamoun, and Robyn A Owens. A novel representation and feature matching algorithm for automatic pairwise registration of range images. *International Journal of Computer Vision*, 66(1):19–40, 2006. 5

[27] G Dias Pais, Srikumar Ramalingam, Venu Madhav Govindu, Jacinto C Nascimento, Rama Chellappa, and Pedro Miraldo. 3dregnet: A deep neural network for 3d point registration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7193–7203. IEEE, 2020. 1, 2, 3, 4, 5, 6

[28] Alvaro Parra, Tat-Jun Chin, Frank Neumann, Tobias Friedrich, and Maximilian Katzmann. A practical maximum clique algorithm for matching with pairwise constraints. *arXiv preprint arXiv:1902.01534*, 2019. 4

[29] 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. 2, 3, 4, 6

[30] Siwen Quan and Jiaqi Yang. Compatibility-guided sampling consensus for 3-d point cloud registration. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10):7380–7392, 2020. 1, 2, 5, 6

[31] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (fpfh) for 3d registration. In *IEEE International Conference on Robotics and Automation*, pages 3212–3217. IEEE, 2009. 1, 2, 5

[32] Radu Bogdan Rusu and Steve Cousins. 3d is here: Point cloud library (pcl). In *IEEE International Conference on Robotics and Automation*, pages 1–4. IEEE, 2011. 5

[33] Ivan Sipiran and Benjamin Bustos. Harris 3d: a robust extension of the harris operator for interest point detection on 3d meshes. *The Visual Computer*, 27(11):963–976, 2011. 5

[34] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In *European Conference on Computer Vision*, pages 356–369. Springer, 2010. 5

[35] Haiping Wang, Yuan Liu, Zhen Dong, and Wenping Wang. You only hypothesize once: Point cloud registration with rotation-equivariant descriptors. In *Proceedings of the ACM International Conference on Multimedia*, pages 1630–1641, 2022. 1

[36] Heng Yang, Jingnan Shi, and Luca Carlone. Teaser: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2):314–333, 2020. 2, 4, 5, 6

[37] Jiaqi Yang, Zhiguo Cao, and Qian Zhang. A fast and robust local descriptor for 3d point cloud registration. *Information Sciences*, 346:163–179, 2016. 5

[38] Jiaqi Yang, Jiahao Chen, Siwen Quan, Wei Wang, and Yanning Zhang. Correspondence selection with loose-tight geometric voting for 3d point cloud registration. *IEEE Transactions on Geoscience and Remote Sensing*, 2022. 1, 2

[39] Jiaqi Yang, Zhiqiang Huang, Siwen Quan, Zhaoshuai Qi, and Yanning Zhang. Sac-cot: Sample consensus by sampling compatibility triangles in graphs for 3-d point cloud registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–15, 2021. 1, 2, 5

[40] Jiaqi Yang, Zhiqiang Huang, Siwen Quan, Qian Zhang, Yanning Zhang, and Zhiguo Cao. Toward efficient and robust metrics for ransac hypotheses and 3d rigid registration. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(2):893–906, 2021. 1, 2, 5

[41] Jiaolong Yang, Hongdong Li, Dylan Campbell, and Yunde Jia. Go-icp: A globally optimal solution to 3d icp point-set registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(11):2241–2254, 2015. 1, 2, 5

[42] Jiaqi Yang, Yang Xiao, Zhiguo Cao, and Weidong Yang. Ranking 3d feature correspondences via consistency voting. *Pattern Recognition Letters*, 117:1–8, 2019. 5

[43] 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, 2021. 2, 3, 6

[44] 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. 5

[45] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast global registration. In *European Conference on Computer Vision*, pages 766–782. Springer, 2016. 2, 5, 6