This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

FastMAC: Stochastic Spectral Sampling of Correspondence Graph

Yifei Zhang[2,1], Hao Zhao[2†], Hongyang Li[3], Siheng Chen[4,3]
[1]University of Chinese Academy of Sciences, [2]AIR, Tsinghua University,
[3]Shanghai AI Laboratory, [4]Shanghai Jiao Tong University
zhangyifei21a@mails.ucas.ac.cn, zhaohao@air.tsinghua.edu.cn

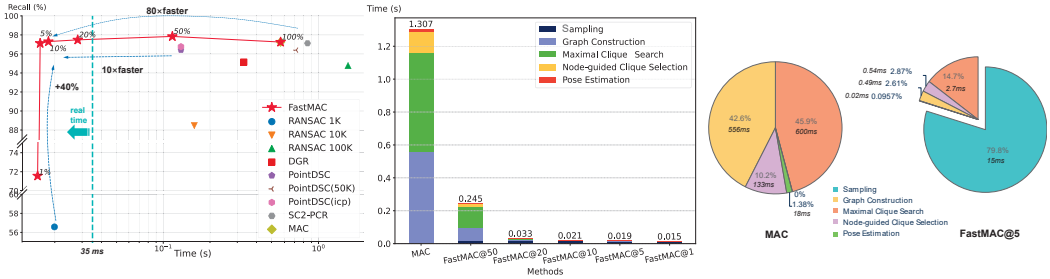

Figure 1. **LEFT:** FastMAC can accelerate MAC [71] by 80 times, while preserving similarly high registration success rate (denoted by registration recall). This is achieved by sampling 5% nodes on the correspondence graph, through a stochastic spectral formulation. Other sampling ratios are also shown, and FastMAC achieves real-time when the ratio is lower than 20%. **MIDDLE:** Time profiling comparison between vanilla MAC and FastMAC with different sampling ratios. FastMAC significantly accelerates all stages of MAC. **RIGHT:** A detailed runtime breakdown for each component in MAC and FastMAC. Maximal clique search is no longer a bottleneck.

Abstract

3D correspondence, i.e., a pair of 3D points, is a fundamental concept in computer vision. A set of 3D correspondences, when equipped with compatibility edges, forms a correspondence graph. This graph is a critical component in several state-of-the-art 3D point cloud registration approaches, e.g., the one based on maximal cliques (MAC). However, its properties have not been well understood. So we present the first study that introduces graph signal processing into the domain of correspondence graph. We exploit the generalized degree signal on correspondence graph and pursue sampling strategies that preserve high-frequency components of this signal. To address time-consuming singular value decomposition in deterministic sampling, we resort to a stochastic approximate sampling strategy. As such, the core of our method is the stochastic spectral sampling of correspondence graph. As an application, we build a complete 3D registration algorithm termed as FastMAC, that reaches real-time speed while leading to little to none performance drop. Through extensive experiments, we validate that FastMAC works for both indoor and outdoor benchmarks. For example, FastMAC can accelerate MAC by 80 times while maintaining high registration success rate on KITTI. Codes are publicly available at https://github.com/Forrest-110/FastMAC.

†Corresponding author.

1. Introduction

Correspondence is one of the most fundamental computer vision concepts, since it encodes important geometric relationships such as multi-view transformation (2D-2D [37] or 3D-3D [59][75][74] correspondence) or single-view projection (2D-3D correspondence [56][64]). 3D correspondence, which is by definition a pair of matched 3D points, plays an important role in 3D registration [72] and downstream applications like SLAM [8], 3D reconstruction [20][35] and 3D scene understanding [12][25][24][60]. While the community has studied 3D correspondence for a long time, 3D correspondence graph is not yet well-understood.

In this graph, each vertex is a 3D correspondence, and the edge connectivity is usually defined according to the compatibility between two correspondences. For example, if a certain compatibility metric is higher than a threshold, an edge is active between two correspondences. This graph is indeed a critical component in state-of-the-art 3D registration methods like MAC [71]. MAC, as our baseline, searches for maximal cliques on this graph and estimates relative poses using compatible correspondences.

However, MAC can be quite slow with numerous input correspondences, with a single registration can cost more than one second, as shown in Figure. 1. This makes it far from deployment for real-time applications such as SLAM. A natural idea is to downsample this graph for efficiency. We introduce the framework of graph signal processing [10] to achieve this goal. Specifically, we use the generalized degree, *i.e.*, the weighted edge sum, of each node in the graph as the graph signal and preserves its high frequency component with a graph filter, which is the Laplacian matrix in our work. We derive an optimal sampling strategy that best preserves the filtered high-frequency signal of interest.

High-frequency. The principle is that we need to sample nodes with fast change of generalized degree over the correspondence graph (short as *high-frequency nodes* for brevity), since they are better suited for maximal clique search. The intuition is three-fold: (1) Every maximal clique must contain one high-frequency node; (2) Maximal cliques tend to contain three or more high-frequency nodes; (3) Cut-points, which are always high-frequency nodes, are contained in more than one maximal clique.

Stochastic Sampling. There exist well-established deterministic sampling methods to recover a certain signal of interest on a graph [10]. But they involve iterative singular value decomposition operations which are time-consuming thus contradict the goal of accelerating MAC. As such, we derive a stochastic approximated sampling strategy that takes constant time w.r.t. sample number.

To summarize, we make these contributions: (1) For the first time, we introduce graph signal processing to 3D correspondence graph. (2) We propose a stochastic sampling strategy that selects high-frequency nodes on a correspondence graph. (3) We develop a full registration pipeline FastMAC upon the proposed sampling method, which is suited for maximal clique search. (4) FastMAC achieves trivial performance drop on indoor and outdoor datasets, while achieving 80 times acceleration to a real-time level.

2. Related Works

3D point cloud registration is important to many real-world problems, including pose estimation [42], SLAM [47], and 3D reconstruction [21]. However, globally optimal 3D point cloud registration is very challenging and existing works can be summarized into six primary categories:

Maximum Consensus (MC) is a widely used robust optimization objective in point cloud registration [41]. Compared with other robust optimization objectives such as Truncated-Least-Squares (TLS), MC has its advantage of being superior under certain circumstances[15]. However, MC may produce an error-prone estimate depending on the input conditions [9]. Moreover, directly solving Maximum Consensus entails an NP-hard computational complexity, which has been confirmed by prior research [16].

Stochastic Techniques. To address the complexity of solving the NP-hard Maximum Consensus problem, stochastic techniques have been proposed and Random Sample Consensus (RANSAC) is a well-known one among them. Numerous extensions and enhancements to RANSAC have been devised, aiming to improve its efficiency [61][19], accuracy [62], and robustness [32][58]. However, it is still essential to recognize that the convergence speed of RANSAC exhibits an exponential relationship with the rate of outliers in the dataset [7].

Branch-and-Bound. The Branch-and-Bound (BnB) [4][43][73] algorithm stands as a fundamental technique in optimization and search problems, for registration. It can explore and assess all solution possibilities systematically and intelligently remove less promising ones, ensuring an optimal solution. Still, it is important to note that BnB exhibits an exponential complexity concerning the problem size and the presence of outliers within the dataset [7].

Mixed Integer Program. BnB has been extended with Mixed-Integer Programming (MIP) [34] to speed up computation. But MIP itself also demonstrates that the incorporation of Linear-Matrix-Inequality Constraints significantly expedites computational processes [57]. Several avenues of exploration has been done including TEASER++ [66], Fast-Global-Registration (FGR) [76] and other works [65]. Nevertheless, it is noteworthy that the computational time still exhibits sensitivity to both the outlier rate and problem size, underscoring the need for further improvements.

Simultaneous Pose and Correspondence (SPC) methods represent another prominent paradigm within the field of point cloud registration, with the pioneering work of the Iterative Closest Point (ICP) algorithm as a cornerstone [5]. Over time, several robust extensions of the ICP method have been introduced [27][31][14][39][54]. SPC methods are often lauded for their swiftness and precision, yet they do exhibit a notable vulnerability to local minima, a limitation that has been acknowledged [47]. Though global SPC methods like Go-ICP [67] have been proposed, it is important to note that many global methods in the SPC paradigm still rely on Branch-and-Bound (BnB) techniques [28][6].

Consistency Graph-based Methods. Recently, several approaches based on consistency graph are proposed for point cloud registration, emphasizing the encoding of consistency among pairs of correspondences through the utilization of invariants [22][40][33][68] structured within a graph framework. Researchers have sought to enhance the efficiency of the search for these maximum cliques through the introduction of more efficient search algorithms [46] and various relaxations to the maximum clique problem [38][55]. Previous research [50] has explored sampling of correspondences, but we address the problem from the perspective of graph signal processing for the first time.

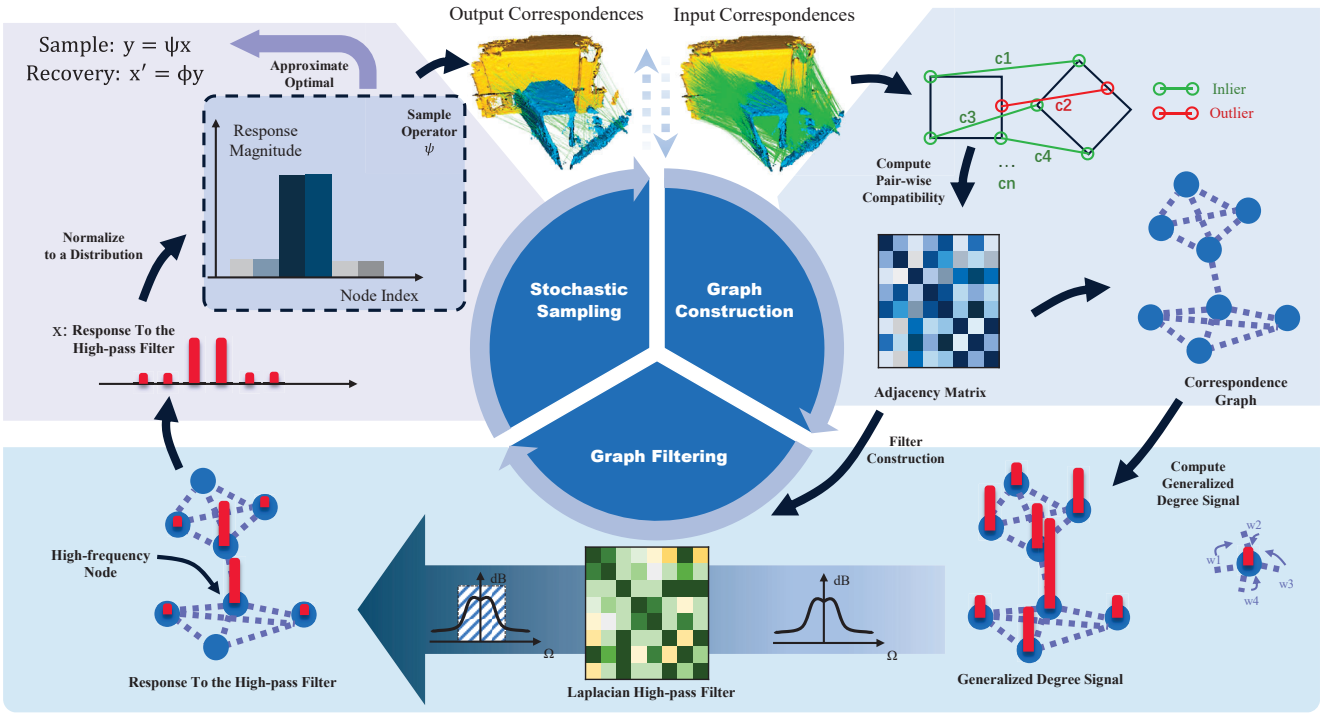

Figure 2. **Pipeline of FastMAC.** In the **top-right** panel, we show the the procedure of constructing a correspondence graph from input correspondences. The graph is mathematically represented by an adjacency matrix. High values in this matrix mean high compatibility between two correspondences. In the **bottom** panel, we define generalized degree signal on the graph as the aggregation of compatibility scores on edges connecting with a node. We pass the signal through a Laplacian high-pass graph filter (constructed from the adjacency matrix) to get its high-frequency component. As mentioned in the text, after filtering, nodes with high response are named as *high-frequency nodes*. In the **top-left** panel, we derive a stochastic sampling strategy in which the sampling probability of a node is proportional to the response magnitude. This sampling strategy is a fast approximation of the optimal (but slow) deterministic sampling strategy that recovers a signal of interest. **Lastly** but not shown in this figure, we use the MAC registration algorithm on output correspondences.

3. Methods

Our goal is fast and accurate 3D point cloud registration and our method is based upon the recently published maximal clique (MAC) method [71]. It has four steps as shown in the time profiling of Fig. 1 and Table. 1. (1) **Graph Construction**, which builds a correspondence graph on the input correspondences. We inherit this step as shown in the top-right panel of Fig. 2. (2) **Maximal Clique Search**, which finds all maximal cliques in the graph as the name implies. (3) **Node-guided Clique Selection**, which reduces the number of maxmimal clique candidates and finally (4) **Pose Estimation**, which evaluates pose hypothesis generated in each clique and chooses the best one as the output pose.

The key intuition behind MAC is to loosen the previous maximum clique constraint [36], and use more maximal clique candidates to generate potentially accurate pose hypotheses. However it is very slow when there are many input correspondences and **Maximal Clique Search** is the biggest bottleneck due to its exponential complexity. Hence, we aim to design a sampling module that reduces graph size without sacrificing the maximal clique registration performance. Our sampling module is shown in the bottom and top-left panels of Fig. 2. It is inserted into the first and second step of MAC, which means the output correspondences are input into the remaining three steps of MAC. That is the difference between MAC and FastMAC.

How to achieve this graph down sampling? There are widely used modules like random sampling and farthest point sampling [48]. But as shown later in Fig. 4, Fig. 5 and Fig. 6, they perform poor for MAC acceleration. So we resort to the graph signal processing theory [52][53][10][11]. Due to page limit, its basics are presented in Appendix A.1.

First, as shown in the top-right panel of Fig. 2, we construct a correspondence graph G_{corr} and the adjacency matrix W_{SOG}, in which SOG means second order graph, following MAC [71]. Due to page limit, the details are presented in Appendix A.4. It is noteworthy that the value in W_{SOG} means compatibility between two correspondences.

Generalized Degree Signal. In order to exploit the graph signal processing theory, we need to define a signal on the correspondence graph. The normal degree signal for

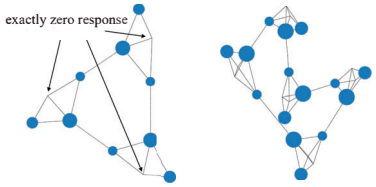

Figure 3. Response of the degree signal to a high-pass filter on connected-caveman graphs. Node size represents the response magnitude. This shows why high-pass filter is suited for MAC.

a node is the number of edges it connects. In a weighted graph, the generalized degree signal for a node is defined as the sum of edge weight it connects. In the following sections, we will explain why this signal can help us construct a graph filter that is suited for maximal clique search.

3.1. Graph Filtering: Key Insight

After constructing a correspondence graph, we go into the second part of Fig. 2. Our objective is to extract the high-frequency component of the generalized degree signal (degree later for brevity), allowing us to sample the nodes in the graph where the degree undergoes rapid changes. This section will highlight the reasons for doing so. We start by analysing the frequency of the degree distribution, and in particular its relation to cliques, as shown in Fig. 3.

In order to explore the relationship between the degree signal frequency and cliques, we investigated the response of the degree signal to a high-pass filter, which is implemented with Laplacian matrix as mentioned above, on crafted (not generated from realistic data) connected caveman graphs[63]. Note that our focus is on the degree frequency, a local feature determined solely by the neighboring nodes. By prioritizing this local feature, we can simplify other types of graphs into the connected caveman graph, which will be explained later.

As illustrated in Fig. 3, nodes with high response exhibit certain properties. If we consider each clique as a community, then: 1.In every community, there must exist a node to generate strong response. 2.There are a sufficient number of nodes within a community that can elicit a strong response. 3.Nodes with significant response lie on the periphery of each community and are susceptible to consituting cut points. They have links not only with nodes within their respective community but also with nodes in other ones.

These properties prompt the idea of sampling the high-frequency nodes. Maximal Clique registration process involves searching all maximal cliques in the correspondence graph, generating hypothesis for each maximal clique and selecting the best one. Suppose the output samples consist of high-frequency nodes, then: 1. since such nodes must exist in every community, they can cover nearly every maximal clique. 2. A sufficient number of samples in each clique guarantees the ability to generate a hypothesis. 3. Considering link between nodes represents compatibility, the selected correspondences are not only compatible with the correspondences within their own clique, but also with some others, indicating that these correspondences are more reliable and thus generating better hypotheses.

Now we explain why the features seen in the connected caveman graph can apply to other graphs. In a typical graph, cliques are either mutually connected or not. Mutually connected cliques maintain similar local properties of a connected caveman graph, whereas isolated cliques exhibit distinct features. Nevertheless, isolated cliques are rare and often negligible in the scenario of graph-based registration.

3.2. Graph Filtering: Formulation

High-pass. Guided by the aforementioned insight, we propose to selectively sample the high-frequency nodes by first formulating the graph filter. There are three typical graph filters: high-pass, low-pass and all-pass. A simple design of high-pass filter is a Haar-like high-pass graph filter:

$$\mathcal{H} = I - \mathcal{A} = V \begin{bmatrix} 1-\lambda_1 & 0 & \dots & 0 \\ 0 & 1-\lambda_2 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1-\lambda_N \end{bmatrix} V^{-1} \tag{1}$$

where $\mathcal{A}$ is a normalized graph shift as defined in Appendix A.1., V and λ_i are the corresponding eigenvectors and eigenvalues. Note that $\lambda_{\max} = 1$ and if we order λ_i in a descending order, we have $1-\lambda_i \leq 1-\lambda_{i+1}$, indicating low frequency response attenuates and high frequency response amplifies. A detailed interpretation is given in [11].

Low-pass. The opposite of this is a the Haar-like low-pass graph filter, that is

$$\begin{aligned} \mathcal{H} &= I + \frac{1}{|\lambda_{\max}|}\mathcal{A} \\ &= V \begin{bmatrix} 1+\frac{\lambda_1}{|\lambda_{\max}|} & 0 & \dots & 0 \\ 0 & 1+\frac{\lambda_2}{|\lambda_{\max}|} & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1+\frac{\lambda_N}{|\lambda_{\max}|} \end{bmatrix} V^{-1} \end{aligned} \tag{2}$$

All-pass. An all-pass graph filter is simple: $\mathcal{H} = I$. The all-pass filter keeps all information of the degree signal and intuitively samples those nodes with large degrees.

For correspondence graph. When it comes to our filter, we first compute the generalized degree signal $s = [s_1, s_2, \dots, s_N]^T \in \mathbb{C}^{N\times 1}$ where $s_i = \sum_j W_{\text{SOG}_{\text{ij}}}$, and N is the size of correspondence set. Then a high-pass graph filter is adopted to filter the high frequency information of s. For G_{corr}, we define the high-pass graph filter as:

$$\mathcal{H} = \text{Diag}(s) - W_{\text{SOG}}, \tag{3}$$

or the Laplacian Matrix of the correspondence graph. In the graph vertex domain, the output for a signal $X = (x_i)$, $(\mathcal{H}X)_i = s_i x_i - \sum_{j \in \mathcal{N}_i} W_{\mathrm{SOG_{ij}}} x_j$ reflects the difference between a node and the combination of its neighbors.

Then we have the response of signal s corresponding to $\mathcal{H}$ as $f = \mathcal{H}s$. We further compute the response magnitude: $\pi_i = ||f_i||_2^2$ which quantifies the energy of the signal on each node after high-pass graph filtering. It reflects how much information we know about a signal value on the node from its neighbors in the graph.

3.3. Stochastic Sampling

Sampling operator definition. After obtaining the response magnitude of each node to the graph filter as shown in Fig. 2, we perform sampling based on this response magnitude. Suppose we aim to sample m components of a graph signal $x = \mathcal{H}s \in \mathbb{C}^n$ to produce a sampled signal $y = x_{\mathcal{M}} \in \mathbb{C}^m$, where $\mathcal{M}$ is the set of sampled indices. The sampling operator Ψ is defined as a linear mapping from $\mathbb{C}^n$ to $\mathbb{C}^m$, $\Psi_{ij} = \delta_{j,\mathcal{M}_i}$ and the interpolation operator Φ is defined as a linear mapping from $\mathbb{C}^m$ to $\mathbb{C}^n$:

$$y = x_{\mathcal{M}} = \Psi x, \tag{4a}$$

$$x' = \Phi y = \Phi \Psi x, \tag{4b}$$

where $x' \in \mathbb{C}^n$ is the recovery of the original signal. A properly designed sampling operator Ψ aims to minimize the reconstruction error $||x - x'||$.

Non-stochastic methods attempt to create a well-designed deterministic sampling operator Ψ. [10] finds the optimal sampling operator:

$$\Psi_{\mathrm{opt}} = \arg\max_{\Psi} \sigma_{\min}(\Psi V_{(\mathrm{K})}), \tag{5}$$

where $\sigma_{\min}$ means the smallest singular value and $V_{(\mathrm{K})}$ represents the independent columns in the eigenvectors V of the graph shift $\mathcal{A}$. In practice, a greedy algorithm [10] is used to find an approximate solution. It maintains M, a set of rows of $V_{(\mathrm{K})}$, and loops to find another row r in $V_{(\mathrm{K})}$ to maximize $\sigma_{\min}$ of the matrix formed by $M + \{r\}$ until $|M|$ meets the termination condition. However, it is extremely slow when processing large matrices, as it involves a number of SVD decompositions with a total complexity of $O(MN^3 + M^3N)$ where M is the sample size and N is the original size. Proof will be given in 5.2.

Stochastic sampling. By contrast, we adopt a stochastic strategy. We consider π_i fetched from *Graph Filtering* as a sampling distribution and apply probability sampling on the initial correspondence set, resulting in a sampled set denoted as C_{sampled}. π_i approximates the sampling operator Ψ and it is optimal in terms of minimizing the reconstruction error, according to proof in [11], and is much faster, which will be proved in 5.2. A detailed proof of optimality is given in Appendix.A.5.

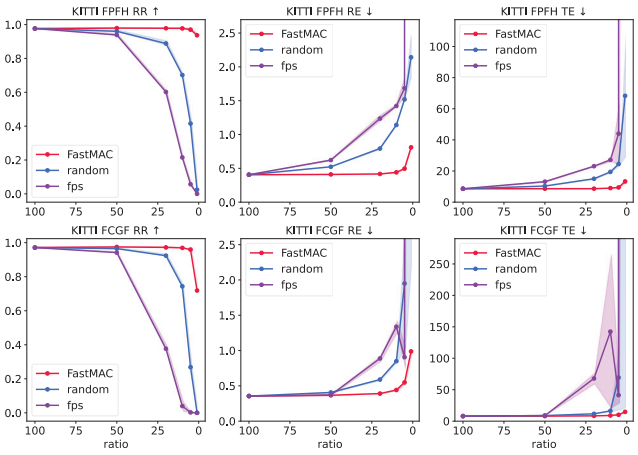

Figure 4. Sampling performance on KITTI. Each column represents a metric in TE,RE and RR and each row represents a setting composed of datasets and descriptors. Shaded areas represent variance from multiple runs.

4. Experiments

For information about datasets, evaluation metrics and implementation details, please refer to the Appendix A.6.

4.1. Time-Accuracy Trade-off Comparison

We perform an extensive comparison in Fig. 1. The correspondence based registration methods are presented for comparison. All methods are tested on the KITTI dataset with FCGF as the correspondence generation descriptor.

Fig. 1 demonstrates RR performance of different methods. **Our FastMAC can outperform all other methods even with a sample ratio as low as 5%.** It runs nearly **80 times** faster than methods with comparable RR performance, and achieves a **40% higher** RR when compared to methods that are almost as fast as it. Moreover, when sampling ratio declines to 20%, our method runs at *real-time* level, with a single registration requires less than 35ms.

4.2. Sampling Strategy Comparison

We also compare our method with different sampling strategies. This demonstrates that the MAC itself is still sensitive to the number of correspondences, thus showing that our method is superior and suited for Maximal Clique registration. The sampling strategies for comparison are Random Sampling and Furthest Point Sampling(FPS). Notably, a correspondence is not a traditional 3D point and we define their distance as the euclidean distance in 6D space. Both FPFH and FCGF descriptor are tested.

Results on KITTI Dataset: Fig. 4 shows results of FPFH and FCGF settings on KITTI Dataset. Our method maintains a consistent RR, RE and TE when sampling ratio declines from 100% to 5% and only becomes slightly

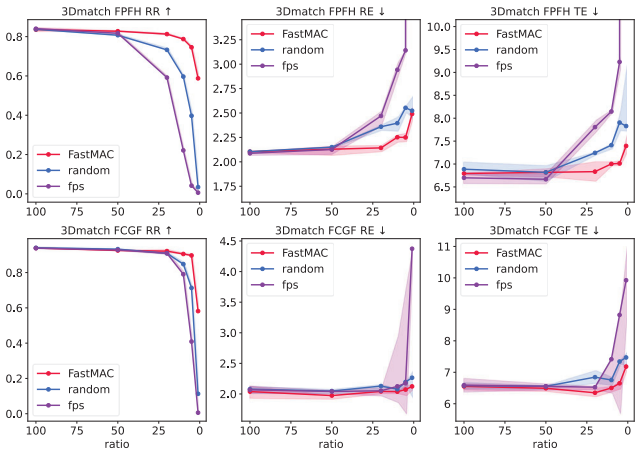

Figure 5. Sampling performance on 3DMatch. Each column represents a metric in TE,RE and RR and each row represents a setting composed of datasets and descriptors.

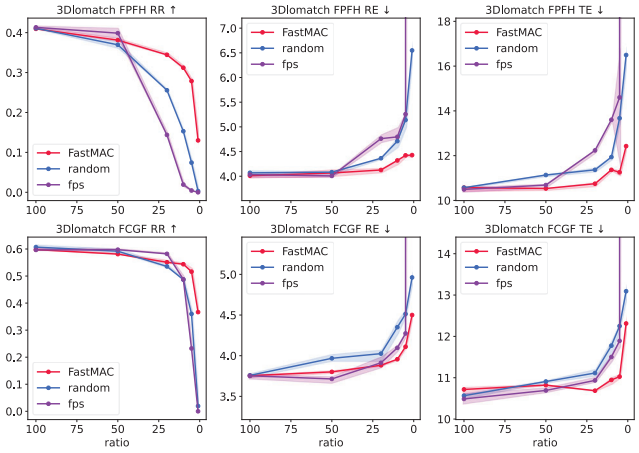

Figure 6. Sampling performance on 3DLoMatch. Each column represents a metric in TE,RE and RR and each row represents a setting composed of datasets and descriptors.

worse at the 1%. By comparison, Random Sampling and FPS strategy show rapid deterioration in performance. It's worth noticing that FPS behaves even worse than Random strategy, suggesting that a 6-dimensional vector space with Euclidean distance is ill-suited for the correspondence set. Another notable point is that the FPS performance shows dramatic fluctuations, indicating lack of robustness.

Results on 3DMatch Dataset: As shown in Fig. 5, our method still works well. Though the Random Strategy performs closely to our method on RE and TE, it has a much lower success rate. This means Random Strategy has a high performance on only a few point cloud pairs, which limits its usage in challenging real-world problems.

Results on 3DLoMatch Dataset: As in Fig. 6, RR of 3DLoMatch drops faster than 3DMatch when the sample ratio decreases for all three methods, due to low-overlap of this dataset. Still, our method significantly surpasses the performance of the other two methods, considering their almost-zero success rate at 1% sample ratio.

method	ratio	Sampling	GC	MCS	NCS	PE	Total
MAC	100	0	749.057	61.390	132.809	18.056	961.312
FastMAC	50	15.13	129.404	26.845	2.225	5.039	178.643
	20	15.02	14.945	4.794	0.408	0.903	36.070
	10	14.97	3.369	1.113	0.070	0.256	19.778
	5	14.99	0.971	0.267	0.005	0.067	16.300
	1	15.06	0.056	0.001	0.001	0.019	15.137

Table 1. Time profiling for 3DMatch Dataset with FCGF descriptor. Vanilla MAC is compared with our FastMAC. Time consumption is measured in milliseconds. GC: Graph Construction; MCS: Maximal Clique Search; NCS: Node-guided Clique Selection; PE: Pose Estimation

	FPFH			FCGF			Time(s)
	RR(%)	RE(°)	TE(cm)	RR(%)	RE(°)	TE(cm)	
FGR[76]	5.23	0.86	43.84	89.54	0.46	25.72	9.350
TEASER++[66]	91.17	1.03	17.98	94.96	0.38	13.69	0.070
RANSAC-4M[23]	74.41	1.55	30.20	80.36	0.73	26.79	52.40
CG-SAC[68]	74.23	0.73	14.02	83.24	0.56	22.96	2.140
SC2-PCR[13]	96.40	**0.41**	8.00	97.12	0.41	9.71	0.850
DGR[18]	77.12	1.64	33.10	94.90	0.34	21.70	0.330
PointDSC[2]	96.40	0.38	8.35	96.40	0.61	13.42	0.130
MAC[71]	97.66	**0.41**	8.61	97.25	**0.36**	8.00	0.570
FastMAC@50	97.84	**0.41**	8.61	**97.84**	**0.36**	**7.98**	0.114
FastMAC@20 (**real-time**)	**98.02**	**0.41**	8.64	97.48	0.38	8.20	0.028

Table 2. Comparison with baseline methods on KITTI Dataset. The best and second-to-best results of baseline methods are repectively marked in bold and underlined. FastMAC@x refers to our method sampling at x% ratio.

4.3. Time Profiling

To demonstrate FastMAC's time efficiency, we study on-device time profiling to report time consumption. The original MAC is used for comparison.

Fig. 1 depicts our results on KITTI Dataset with FPFH descriptor. On the right side, we demonstrate time profiling of MAC and FastMAC with a 5% sample ratio . For MAC, *Graph Construction* and *Maximal Cliques Search* occupy a major part of time consumption, pushing the total time to over *1* second. Whilst for FastMAC, they are no longer bottlenecks. Middle shows variation of time taken versus the sample ratio. *Sampling* gradually became the dominatant factor, with rest parts being barely time-consuming.

Table 1 further presents our findings on 3DMatch. Acceleration from *Sampling* negates the time spent during the process itself. And its time decrease is essentially consistent across various sample rates, proving its efficiency.

4.4. Comparison to State-of-the-arts

Our method is compared with baseline approaches on the 3DMatch, 3DLomatch, and KITTI datasets, and the outcomes can be found in Tables 2, 3 and 4. When sampling

	FPFH			FCGF			Time(s)
	RR(%)	RE(°)	TE(cm)	RR(%)	RE(°)	TE(cm)	
RANSAC-1M[23]	64.20	4.05	11.35	88.42	3.05	9.42	23.30
RANSAC-4M[23]	66.10	3.95	11.03	91.44	2.69	8.38	95.80
GC-RANSAC[3]	67.65	2.33	6.87	92.05	2.33	7.11	0.450
TEASER++[66]	75.48	2.48	7.31	85.77	2.73	8.66	0.030
SC2-PCR[13]	83.73	2.18	6.70	93.16	2.09	**6.51**	0.920
3DRegNet[44]	26.31	3.75	9.60	77.76	2.74	8.13	0.050
DGR[18]	32.84	2.45	7.53	88.85	2.28	7.02	1.260
PointDSC[2]	72.95	2.18	**6.45**	91.87	2.10	6.54	0.140
MAC[71]	**83.90**	**2.11**	6.80	**93.72**	**2.02**	6.54	0.914
FastMAC@50	82.87	2.15	6.73	92.67	2.00	6.47	0.203
FastMAC@20 (**real-time**)	80.71	2.17	6.81	92.30	2.02	6.52	0.038

Table 3. Comparison with baseline methods on 3DMatch Dataset. The best and second-to-best results of baseline methods are repectively marked in bold and underlined. FastMAC@x refers to our method sampling at x% ratio.

	FPFH			FCGF			Time(s)
	RR(%)	RE(°)	TE(cm)	RR(%)	RE(°)	TE(cm)	
RANSAC-1M[23]	0.67	10.27	15.06	9.77	7.01	14.87	19.60
RANSAC-4M[23]	0.45	10.39	20.03	10.44	6.91	15.14	86.30
TEASER++[66]	35.15	4.38	10.96	46.76	4.12	12.89	0.030
SC2-PCR[13]	38.57	4.03	10.31	58.73	3.80	10.44	0.720
DGR[18]	19.88	5.07	13.53	43.80	4.17	10.82	1.220
PointDSC[2]	20.38	4.04	10.25	56.20	3.87	10.48	0.140
MAC[71]	**40.88**	**3.66**	**9.45**	**59.85**	**3.50**	**9.75**	1.181
FastMAC@50	38.46	4.04	10.47	58.23	3.80	10.81	0.271
FastMAC@20 (**real-time**)	34.31	4.12	10.82	55.25	3.84	10.71	0.051

Table 4. Comparison with baseline methods on 3DLoMatch Dataset. The best and second-to-best results of baseline methods are repectively marked in bold and underlined.

at various ratios, FastMAC reports no significant decrease in performance, remaining competitive with other state-of-the-art methods. This showcases the efficacy of our sampling technique, which accelerates the state-of-the-art MAC method whilst maintaining accuracy.

4.5. Descriptor Robustness

Since our method accepts correspondences as input, it is crucial to demonstrate its ability to work with correspondences generated by different descriptors. We perform extensive experiments with various descriptors, including FPFH[51], FCGF[17], Predator[30], Spinnet[1], Cofinet[69] and Geotransformer[49]. These descriptors are used to generate point-wise features which are subsequently used to obtain correspondences. We adopt KITTI as the dataset and the results are presented in Fig. 7.

For RR metric, stronger descriptors like Geotransformer, Cofinet, Spinnet and Predator exhibit amazing excellence, with their RR remaining unaffected by the sample ratio, while FCGF and FPFH perform slightly worse when sample ratio comes to 1%. For RE and TE, most descriptors behave similarly. These metrics first stabilize and then slowly increase as the sampling rate decreases. It can therefore be concluded that our method is highly robust to the correspondences produced by different descriptors.

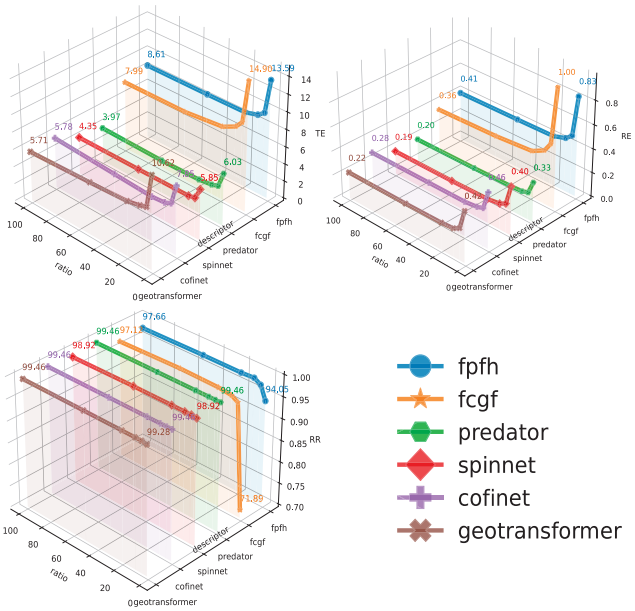

Figure 7. Performance on KITTI with various descriptors.

Ratio	High-pass			All-pass			Low-pass		
	RR(%)	RE (°)	TE(cm)	RR(%)	RE (°)	TE(cm)	RR(%)	RE (°)	TE(cm)
50	**97.66**	**0.368**	**8.016**	97.12	0.368	8.098	97.48	0.368	8.135
20	**97.66**	**0.391**	**8.457**	96.58	0.395	8.672	97.30	0.403	8.648
10	**96.94**	**0.446**	**9.201**	96.94	0.480	9.329	94.96	0.468	9.417
5	**96.04**	**0.525**	**10.038**	93.33	0.583	10.907	90.09	0.616	11.182
1	**71.89**	**0.997**	**14.899**	28.11	1.058	15.217	13.33	3.516	106.211

Table 5. Registration results on KITTI FCGF dataset for comparison between the High-pass, Low-pass and All-pass filters.

5. Ablation Study

In this section, we mainly focus on the analysis of the core parts of our method, i.e, *High-pass Graph Filtering* and *Stochastically Sampling*. Since xyz coordinates of the point cloud can also be a graph signal and has been widely used before [11], we further compare xyz signal with our generalized degree signal to demonstrate our superiority.

5.1. High-pass, Low-pass and All-pass Filter

In this part, we compare our high-pass filter with a low-pass filter and an all-pass filter to demonstrate its efficacy. The filters are implemented as described in 3. For the Haar-like low-pass filter, we choose the graph shift $\mathcal{A}$ to be $D^{-1}W$ where D is the generalized degree matrix and W is the adjacency matrix. Then we have the sample distribution $\pi_i = ||((I + D^{-1}W)s)_i||_2$ in which s is the generalized degree signal. For the all-pass filter, the corresponding sampling strategy is $\pi_i = ||s_i||_2$.

We use these types of filters to sample the correspondences and feed the output into the MAC module. The re-

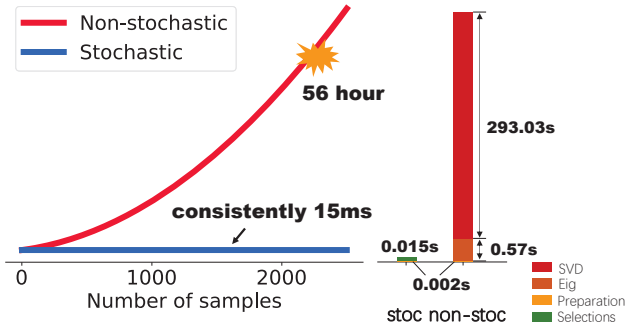

Figure 8. Left figure shows how time increases with number of samples on KITTI FCGF dataset. Right is time profiling when sampling 50 from 5000 correspondences. **Preparation**: Computation of the graph shift and the degree score. **Sampling**: For stochastic method, only 50 selection operations are required according to the given distribution. For non-stochastic method, it includes an eigendecomposition and 50 loops of ∼5000-scale SVD.

sult is shown in Table. 5. The high-pass filter is more effective , providing evidence to support our intuition. In fact, an all-pass filter solely samples out nodes with high degrees, whereas a low-pass filter samples nodes which lie within a clique. If the clique is big in size, low-pass and all-pass filters function in a manner similar to random sampling. That is why these two filters do not function well.

5.2. Stochastic v.s. Non-stochastic

To clarify the efficiency of our stochastic method, we will give both theoretical analysis and experimental analysis.

Theoretical Analysis Since both stochastic and non-stochastic methods involve computing the signal after filtering, we only analyze operations after the computation. Suppose we are sampling M nodes from the original N nodes.

Non-stochastic: $V_{(\mathrm{K})}$ is first required for the graph shift $\mathcal{A}$, which is $O(N^3)$ complexity [45]. Then we enter the greedy algorithm which loops for M times. In loop i, $N-i$ times of computing the smallest singular value is performed, each using $O(N^3+(i+1)^2N)$ [29]. After summing them all, we get the final computation complexity

$$\begin{aligned} &O(N^3)+\sum_{i=0}^{M-1} O(N^3+(i+1)^2N) \\ =&O(MN^3+M^3N) \end{aligned} \tag{6}$$

Stochastic: Our stochastic method simply use the norm of the filtered signal y as a distribution and samples M times from it. The time complexity is $O(M)+O(N)$.

Experimental Analysis We present our experimental results in Fig. 8. The time taken by the stochastic method remains consistently stable at 15 ms as the sample rate increases. By contrast, the non-stochastic method's time consumption increases in a polynomial trend, necessitating two days to obtain 2500 samples! This is largely caused by the frequent, multiple large matrix SVD decompositions.

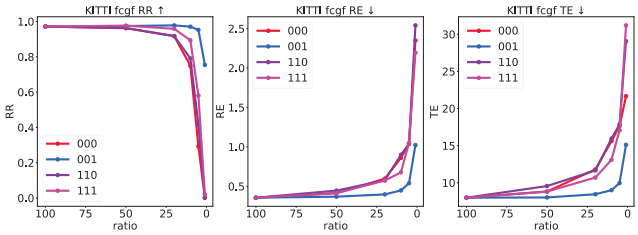

Figure 9. Sampling performance on KITTI FCGF with different degree signal settings. **000**: random sampling. **110**: sample guided by xyz signals from source and target pointcloud. **001**: sample guided by our generalized degree signal. **111**: sample guided by simply adding both signals.

5.3. XYZ Signal v.s. Generalized Degree Signal

In this section, we compare our generalized degree signal with xyz signal which are commonly used in point cloud sampling. As we accept correspondences as input, we denote the source point cloud formed by source points as $Q_{\mathrm{s}} \in \mathbb{C}^{N\times 3}$ and target point cloud $Q_{\mathrm{t}} \in \mathbb{C}^{N\times 3}$. They are considered as graph signals and pass through a high-pass filter which is created by KNN adjacency matrix following [11] to detect their contour points. The response are denoted as $q_{\mathrm{s}} \in \mathbb{C}^{N\times 3}$ and $q_{\mathrm{t}} \in \mathbb{C}^{N\times 3}$. The response magnitude $||q_{\mathrm{s}}||_2 \in \mathbb{C}^N$ and $||q_{\mathrm{t}}||_2 \in \mathbb{C}^N$ are then scaled to 0-1 to form the sampling distribution π_1, π_2.

Fig. 9 demonstrates the results of four settings. The signal of the generalized degree performs the best, while xyz signal does not differ significantly from random sampling. When combined with xyz signal, the performance of generalized degree signal gets worse, indicating that the information within the generalized degree signal is contaminated by xyz signal.

6. Conclusion

In this paper, we propose a stochastic spectral correspondence graph sampling method to discover high-frequency nodes, boosting Maximal Clique Registration to real-time level with little performance loss. Moreover, it's robust to various descriptors, showing its potential for usage in real-time complex applications. Still we have limitations. As shown in Fig. 1, when sample ratio decreases, graph construction gradually becomes a bottleneck. In the future we plan to solve this issue by learning graph prior without building a graph, thus eliminating this cost.

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. 7, 12

[2] 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/CVF Conference on Computer Vision and Pattern Recognition*, pages 15859–15869, 2021. 6, 7, 13, 14

[3] Daniel Barath and Jiří Matas. Graph-cut ransac. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6733–6741, 2018. 7

[4] Jean-Charles Bazin, Yongduek Seo, and Marc Pollefeys. Globally optimal consensus set maximization through rotation search. In *Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part II 11*, pages 539–551. Springer, 2013. 2

[5] 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

[6] Thomas M Breuel. Implementation techniques for geometric branch-and-bound matching methods. *Computer Vision and Image Understanding*, 90(3):258–294, 2003. 2

[7] 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. 2

[8] C Cao, H Zhu, Z Ren, H Choset, and J Zhang. Representation granularity enables time-efficient autonomous exploration in large, complex worlds. *Science Robotics*, 8(80): eadf0970, 2023. 1

[9] Luca Carlone. Estimation contracts for outlier-robust geometric perception. *arXiv preprint arXiv:2208.10521*, 2022. 2

[10] Siheng Chen, Rohan Varma, Aliaksei Sandryhaila, and Jelena Kovačević. Discrete signal processing on graphs: Sampling theory. *IEEE transactions on signal processing*, 63 (24):6510–6523, 2015. 2, 3, 5

[11] Siheng Chen, Dong Tian, Chen Feng, Anthony Vetro, and Jelena Kovačević. Fast resampling of three-dimensional point clouds via graphs. *IEEE Transactions on Signal Processing*, 66(3):666–681, 2017. 3, 4, 5, 7, 8

[12] Xiaoxue Chen, Hao Zhao, Guyue Zhou, and Ya-Qin Zhang. Pq-transformer: Jointly parsing 3d objects and layouts from point clouds. *IEEE Robotics and Automation Letters*, 7(2): 2519–2526, 2022. 1

[13] 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. 6, 7, 13, 14

[14] Dmitry Chetverikov, Dmitry Stepanov, and Pavel Krsek. Robust euclidean alignment of 3d point sets: the trimmed iterative closest point algorithm. *Image and vision computing*, 23 (3):299–309, 2005. 2

[15] Tat-Jun Chin and David Suter. *The maximum consensus problem: recent algorithmic advances*. Springer Nature, 2022. 2

[16] Tat-Jun Chin, Zhipeng Cai, and Frank Neumann. Robust fitting in computer vision: Easy or hard? In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 701–716, 2018. 2

[17] 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. 7, 12, 13

[18] Christopher Choy, Wei Dong, and Vladlen Koltun. Deep global registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2514–2523, 2020. 6, 7, 13, 14

[19] Ondrej Chum and Jiri Matas. Matching with prosac-progressive sample consensus. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, pages 220–226. IEEE, 2005. 2

[20] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312, 1996. 1

[21] Jeffrey Delmerico, Stefan Isler, Reza Sabzevari, and Davide Scaramuzza. A comparison of volumetric information gain metrics for active 3d object reconstruction. *Autonomous Robots*, 42(2):197–208, 2018. 2

[22] Olof Enqvist, Klas Josephson, and Fredrik Kahl. Optimal correspondences from pairwise constraints. In *2009 IEEE 12th international conference on computer vision*, pages 1295–1302. IEEE, 2009. 2

[23] 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. 6, 7, 14

[24] Huan-ang Gao, Beiwen Tian, Pengfei Li, Xiaoxue Chen, Hao Zhao, Guyue Zhou, Yurong Chen, and Hongbin Zha. From semi-supervised to omni-supervised room layout estimation using point clouds. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2803–2810. IEEE, 2023. 1

[25] Huan-ang Gao, Beiwen Tian, Pengfei Li, Hao Zhao, and Guyue Zhou. Dqs3d: Densely-matched quantization-aware semi-supervised 3d detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21905–21915, 2023. 1

[26] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE, 2012. 13

[27] Sébastien Granger and Xavier Pennec. Multi-scale em-icp: A fast and robust approach for surface registration. In *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part IV 7*, pages 418–432. Springer, 2002. 2

[28] Richard I Hartley and Fredrik Kahl. Global optimization through rotation space search. *International Journal of Computer Vision*, 82(1):64–79, 2009. 2

[29] Michael Holmes, Alexander Gray, and Charles Isbell. Fast svd for large-scale matrices. In *Workshop on Efficient Machine Learning at NIPS*, pages 249–252, 2007. 8

[30] 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. 7, 12, 13

[31] Shun'ichi Kaneko, Tomonori Kondo, and Atsushi Miyamoto. Robust matching of 3d contours using iterative closest point algorithm improved by m-estimation. *Pattern Recognition*, 36(9):2041–2047, 2003. 2

[32] Anton Konouchine, Victor Gaganov, and Vladimir Veznevets. Amlesac: A new maximum likelihood robust estimator. *Proc. of Graphicon-2005. Novosibirsk*, pages 93–100, 2005. 2

[33] Marius Leordeanu and Martial Hebert. A spectral technique for correspondence problems using pairwise constraints. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, pages 1482–1489. IEEE, 2005. 2

[34] Hongdong Li. Consensus set maximization with guaranteed global optimality for robust geometry estimation. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1074–1080. IEEE, 2009. 2

[35] Pengfei Li, Ruowen Zhao, Yongliang Shi, Hao Zhao, Jirui Yuan, Guyue Zhou, and Ya-Qin Zhang. Lode: Locally conditioned eikonal implicit scene completion from sparse lidar. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8269–8276. IEEE, 2023. 1

[36] 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. 3

[37] Ce Liu, Jenny Yuen, and Antonio Torralba. Sift flow: Dense correspondence across scenes and its applications. *IEEE transactions on pattern analysis and machine intelligence*, 33(5):978–994, 2010. 1

[38] Parker C Lusk, Kaveh Fathian, and Jonathan P How. Clipper: A graph-theoretic framework for robust data association. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 13828–13834. IEEE, 2021. 2

[39] Lena Maier-Hein, Alfred Michael Franz, Thiago R Dos Santos, Mirko Schmidt, Markus Fangerau, Hans-Peter Meinzer, and J Michael Fitzpatrick. Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error. *IEEE transactions on pattern analysis and machine intelligence*, 34(8):1520–1532, 2011. 2

[40] Joshua G Mangelson, Derrick Dominic, Ryan M Eustice, and Ram Vasudevan. Pairwise consistent measurement set maximization for robust multi-robot map merging. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 2916–2923. IEEE, 2018. 2

[41] Gerard Medioni and Sing Bing Kang. *Emerging topics in computer vision*. Prentice Hall PTR, 2004. 2

[42] Prajval Kumar Murali, Anirvan Dutta, Michael Gentner, Etienne Burdet, Ravinder Dahiya, and Mohsen Kaboli. Active visuo-tactile interactive robotic perception for accurate object pose estimation in dense clutter. *IEEE Robotics and Automation Letters*, 7(2):4686–4693, 2022. 2

[43] Carl Olsson, Fredrik Kahl, and Magnus Oskarsson. Branch-and-bound methods for euclidean registration problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(5):783–794, 2008. 2

[44] 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/CVF conference on computer vision and pattern recognition*, pages 7193–7203, 2020. 7

[45] Victor Y Pan and Zhao Q Chen. The complexity of the matrix eigenproblem. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 507–516, 1999. 8

[46] 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. 2

[47] François Pomerleau, Francis Colas, Roland Siegwart, and Stéphane Magnenat. Comparing icp variants on real-world data sets: Open-source library and experimental protocol. *Autonomous robots*, 34:133–148, 2013. 2

[48] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017. 3

[49] 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. 7

[50] 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. 2

[51] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (fpfh) for 3d registration. In *2009 IEEE international conference on robotics and automation*, pages 3212–3217. IEEE, 2009. 7, 12, 13

[52] Aliaksei Sandryhaila and Jose MF Moura. Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure. *IEEE signal processing magazine*, 31(5):80–90, 2014. 3

[53] Aliaksei Sandryhaila and Jose MF Moura. Discrete signal processing on graphs: Frequency analysis. *IEEE Transactions on Signal Processing*, 62(12):3042–3054, 2014. 3

[54] Aleksandr Segal, Dirk Haehnel, and Sebastian Thrun. Generalized-icp. In *Robotics: science and systems*, page 435. Seattle, WA, 2009. 2

[55] Jingnan Shi, Heng Yang, and Luca Carlone. Robin: a graph-theoretic approach to reject outliers in robust estimation using invariants. In *2021 IEEE International Conference*

on Robotics and Automation (ICRA), pages 13820–13827. IEEE, 2021. 2

[56] Jamie Shotton, Ben Glocker, Christopher Zach, Shahram Izadi, Antonio Criminisi, and Andrew Fitzgibbon. Scene coordinate regression forests for camera relocalization in rgb-d images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2930–2937, 2013. 1

[57] Pablo Speciale, Danda Pani Paudel, Martin R Oswald, Till Kroeger, Luc Van Gool, and Marc Pollefeys. Consensus maximization with linear matrix inequality constraints. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4941–4949, 2017. 2

[58] Lei Sun. Ransic: Fast and highly robust estimation for rotation search and point cloud registration using invariant compatibility. *IEEE Robotics and Automation Letters*, 7(1):143–150, 2021. 2

[59] Zachary Teed and Jia Deng. Raft-3d: Scene flow using rigid-motion embeddings. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8375–8384, 2021. 1

[60] Beiwen Tian, Liyi Luo, Hao Zhao, and Guyue Zhou. Vibus: Data-efficient 3d scene parsing with viewpoint bottleneck and uncertainty-spectrum modeling. *ISPRS Journal of Photogrammetry and Remote Sensing*, 194:302–318, 2022. 1

[61] Ben Tordoff and David W Murray. Guided sampling and consensus for motion estimation. In *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I 7*, pages 82–96. Springer, 2002. 2

[62] Philip H. S. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50:35–61, 2002. 2

[63] Duncan J Watts. Networks, dynamics, and the small-world phenomenon. *American Journal of sociology*, 105(2):493–527, 1999. 4

[64] Xin Wu, Hao Zhao, Shunkai Li, Yingdian Cao, and Hongbin Zha. Sc-wls: Towards interpretable feed-forward camera relocalization. In *European Conference on Computer Vision*, pages 585–601. Springer, 2022. 1

[65] Heng Yang, Pasquale Antonante, Vasileios Tzoumas, and Luca Carlone. Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection. *IEEE Robotics and Automation Letters*, 5(2):1127–1134, 2020. 2

[66] Heng Yang, Jingnan Shi, and Luca Carlone. Teaser: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2):314–333, 2020. 2, 6, 7

[67] 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. 2

[68] 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. 2, 6

[69] 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. 7, 12

[70] 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. 13

[71] Xiyu Zhang, Jiaqi Yang, Shikun Zhang, and Yanning Zhang. 3d registration with maximal cliques. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17745–17754, 2023. 1, 3, 6, 7, 13, 14

[72] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, 13(2):119–152, 1994. 1

[73] Yinqiang Zheng, Shigeki Sugimoto, and Masatoshi Okutomi. Deterministically maximizing feasible subsystem for robust model fitting with unit norm constraint. In *CVPR 2011*, pages 1825–1832. IEEE, 2011. 2

[74] Chengliang Zhong, Peixing You, Xiaoxue Chen, Hao Zhao, Fuchun Sun, Guyue Zhou, Xiaodong Mu, Chuang Gan, and Wenbing Huang. Snake: Shape-aware neural 3d keypoint field. *Advances in Neural Information Processing Systems*, 35:7052–7064, 2022. 1

[75] Chengliang Zhong, Yuhang Zheng, Yupeng Zheng, Hao Zhao, Li Yi, Xiaodong Mu, Ling Wang, Pengfei Li, Guyue Zhou, Chao Yang, et al. 3d implicit transporter for temporally consistent keypoint discovery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3869–3880, 2023. 1

[76] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast global registration. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, pages 766–782. Springer, 2016. 2, 6