QGORE: Quadratic-Time Guaranteed Outlier Removal for Point Cloud Registration

Article *in* IEEE Transactions on Pattern Analysis and Machine Intelligence · April 2023

DOI: 10.1109/TPAMI.2023.3262780

CITATIONS
13

READS
424

4 authors:

Jiayuan Li
Wuhan University
53 PUBLICATIONS **1,189** CITATIONS
SEE PROFILE

Pengcheng Shi
Wuhan University
12 PUBLICATIONS **90** CITATIONS
SEE PROFILE

Qingwu Hu
Wuhan University
109 PUBLICATIONS **1,751** CITATIONS
SEE PROFILE

Yongjun Zhang
Wuhan University
93 PUBLICATIONS **1,156** CITATIONS
SEE PROFILE

QGORE: Quadratic-time Guaranteed Outlier Removal for Point Cloud Registration

Jiayuan Li, Pengcheng Shi, Qingwu Hu, Yongjun Zhang

Abstract—With the development of 3D matching technology, correspondence-based point cloud registration gains more attention. Unfortunately, 3D keypoint techniques inevitably produce a large number of outliers, i.e., outlier rate is often larger than 95%. Guaranteed outlier removal (GORE) [1] has shown very good robustness to extreme outliers. However, the high computational cost (exponential in the worst case) largely limits its usages in practice. In this paper, we propose the first $O(N^2)$ time GORE method, called quadratic-time GORE (QGORE), which preserves the globally optimal solution while largely increases the efficiency. QGORE leverages a simple but effective voting idea via geometric consistency for upper bound estimation, which achieves almost the same tightness as the one in GORE. We also present a one-point RANSAC by exploring "rotation correspondence" for lower bound estimation, which largely reduces the number of iterations of traditional 3-point RANSAC. Further, we propose a l_p-like adaptive estimator for optimization. Extensive experiments show that QGORE achieves the same robustness and optimality as GORE while being 1$\sim$2 orders faster. The source code will be made publicly available.

Index Terms—Point cloud registration, outlier removal, correspondence, 3D matching, robust estimation.

1 Introduction

With the development of high-precision sensor technology such as light detection and ranging (LiDAR), point cloud has become the dominant format to describe the 3D world [2]. However, due to the limitations of field-of-view (FoV) and occlusions, current 3D sensors can only capture a part of the scene in a single scan. Point cloud registration (PCR) is the technique that aligns multiple 3D scans with overlaps to generate a large scene map [3]. Its core problem is to establish an optimal 6D rigid transformation between a scan pair, so that their corresponding parts can be perfectly aligned. PCR has been widely applied in the fields of computer vision, robotics, and photogrammetry, such as 3D reconstruction, 3D localization, pose estimation, etc.

Iterative closest point (ICP) [4] and its variants [5], [6], [7], [8] are the most widely used methods for PCR, which have become the standard approaches for industrial solutions. However, ICP is a local method, whose performance highly relies on the initialization. If the initial guess is not good, ICP is likely to get stuck in a wrong local minimum. Traditional methods usually adopt high reflection targets or high-precision GNSS/INS systems to provide a good coarse alignment, which are labor-intensive or expensive.

To overcome the limitation of ICP-type methods, correspondence-based techniques [1], [3], [9], [10], [11], [12], [13] gain more attention, since they are easier to achieve the global optimality and can largely reduce the sizes of point clouds. Correspondence-based PCR contains two major steps: correspondence establishment and rigid transformation estimation. Benefit from deep learning, 3D feature matching technology has made great progress and many excellent learning-based methods [14], [15], [16], [17] have been proposed for correspondence establishment. However, due to the problems of unorganized, nonhomogeneous density, texture-less, and structural occlusions in point clouds (especially in large-scale LiDAR point clouds), these methods are still prone to produce a large number of outliers (often $> 95\%$). Random sample consensus (RANSAC) [18] is the most straightforward algorithm for robust fitting. However, its required number of iterations grows exponentially with the outlier rate. Moreover, RANSAC-type methods can only obtain approximate solutions [19].

Recently, Bustos and Chin [1] presented a new way to simplify the fitting problem, called guaranteed outlier removal (GORE), which has shown very good robustness to extreme outliers. The basic idea of GORE is to filter out a portion of true outliers before model estimation, so that the inlier rate largely increases and the fitting becomes much easier. However, for the 6D registration problem, it may cost exponential running time due to the usage of branch-and-bound (BnB) [20].

In this paper, we propose a quadratic-time GORE (QGORE) with the following properties: initialization-free, high efficiency, globally optimal, and extremely high robustness. For tight lower bound, we present a one-point RANSAC by exploring the concept of "rotation correspondence". This strategy reduces the number of iterations of 3-point RANSAC by 10^4 times at an outlier rate of 99% in theory. For tight upper bound, QGORE leverages a voting idea based on geometric consistency that only costs an $O(N^2)$ time. As confirmed by experiments, our upper bound achieves almost the same tightness as the one in GORE. QGORE only filters out true outliers and does not recover the registration transformation. Hence, further optimization methods (e.g., BnB or RANSAC) should be adopted after QGORE. To balance the optimality (RANSAC

- *Jiayuan Li, Qingwu Hu, and Yongjun Zhang are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China.*
  Pengcheng Shi is with the School of Computer Science, Wuhan University, Wuhan 430072, China. (Corresponding authors: Qingwu Hu; Yongjun Zhang.)

is not optimal) and efficiency (BnB is slow), we propose a l_p-like adaptive estimator for transformation estimation. A reasonable analysis of its high robustness is also provided according to the influence function and influence interval. Extensive simulated and real experiments demonstrate the superiority of our QGORE.

The contributions of this paper are as follows:

- We propose an $O(N^2)$ time GORE method, called QGORE, which to our knowledge is the first polynomial-time GORE. QGORE proposes a novel voting idea for tight upper bound estimation, which reduces the time complexity of GORE from exponential to quadratic. We also present a lower bound estimation method based on one-point RANSAC, which can be regarded as a good "preprocessor" for GORE-type methods.
- A l_p-like ($p <= 2$) adaptive robust estimator is proposed.

2 Related Work

2.1 3D Keypoint Matching

3D feature matching consists of three major stages:

Feature detection: extract the most distinctive keypoints inside local surfaces of a point cloud by detectors, such as intrinsic shape signatures (ISS) [21], KeypointNet [22], USIP [23], D3feat [24], RSKDD-Net [25], semantic keypoint detection (SKP) [26], etc.

Feature description: descriptors distinguish different keypoints from each other by encoding the 3D local patch of each keypoint into a compact feature vector, e.g., fast point feature histogram (FPFH) [27], RoPS [28], 3DMatch [14], FCGF [29], D3feat [24], SpinNet [16], YOHO [17], etc.

Feature matching: establish one-to-one correspondences according to matching scores between descriptor vectors. Some commonly used strategies are the nearest neighbor, nearest neighbor distance ratio [30], and chi-square test [21].

In the last five years, 3D feature matching technology has made great progress because of the flourish of deep learning. However, these methods are still prone to produce correspondences with extremely high outlier rates, which bring the main challenge for correspondence-based PCR.

2.2 Correspondence-based PCR

RANSAC-family. "Correspondence + RANSAC", perhaps, is the standard method for 2D and 3D robust geometric fitting. RANSAC alternates between minimal set sampling and model estimation until a solution with predefined confidence is obtained. It has many variants, which improve the sampling (e.g., NAPSAC [31], PROSAC [32], Progressive NAPSAC [33]), local optimization (e.g., LORANSAC [34], FLORANSAC [35], GCRANSAC [36]), model testing (e.g., randomized RANSAC [37], [38] and optimal randomized RANSAC [39]), and inlier/outlier threshold (e.g., contrario RANSAC [40], MAGSAC [41], MAGSAC++ [42]) strategies of the original RANSAC. Some methods simultaneously improve multiple aspects of RANSAC, such as USAC [43], USACv20 [44], and VSAC [45]. One common problem of RANSAC-family methods lies in the time complexity, which increases exponentially with the outlier rate or the size of minimal sample set. As 3D matching is much more difficult than its 2D counterpart (e.g., SIFT [30], RIFT [46], LNIFT [47]), cases with extremely high outlier rates often occur in large-scale LiDAR point clouds. Then, traditional 3-point RANSAC-type methods become very slow (dozens of minutes or hours), which prevents their usages in practice.

M-like estimators. M-like robust estimators penalize outliers by assigning small weights to eliminate their influence on the energy function. Traditional estimators such as Cauchy, Welsch, and Tukey cannot handle cases with more than 50% of outliers [48], [19], which limits their usages in correspondence-based PCR. Recently, researchers adapted graduated non-convexity [49], [9], [50] or l_q-norm ($0 < q < 1$) [51] to improve the robustness to high outlier rates. Although these estimators are able to tolerate $70\% \sim 80\%$ of outliers, it is still not enough for real LiDAR registration. To simplify the full 7D/6D registration problem, several methods [52], [10], [13] decompose it into scale, rotation, and translation subproblems. Then, permutation-like strategy, truncated least-squares, RANSAC-variants, and improved robust estimators are introduced for solving these subproblems. These methods have yielded very promising results, i.e., as reported, they can handle up to 99% of outliers. However, they suffer from either the optimality problem or exponential worst-case time.

2.3 Point-based PCR

ICP-family. ICP [4] is a technique that simultaneously establishes correspondences and performs alignment. It is widely used in industrial solutions and has almost become the standard method for PCR. Compared with correspondence-based methods, ICP has one more step for dynamic correspondence update. Many variants are designed to improve the sampling (e.g., Voxel Grid filtering [53] and octree compression [54]), distance metric (e.g., point-to-plane [55], plane-to-plane [56], symmetric point-to-plane [6]), outlier rejection (e.g., trimmed ICP [57], anisotropic ICP [58], sparse ICP [5], robust ICP [7], robust symmetric ICP [8]), efficiency (e.g., EfficientVarICP [59], Anderson-accelerated ICP [60], Fast ICP [7]) of the original ICP algorithm. However, these variants are local methods, which are sensitive to the initializations. They tend to get stuck in local minima when the initial guesses are poor. GO-ICP [61] achieves global optimality by introducing a nested BnB algorithm. However, as aforementioned, the worst-case complexity of BnB is exponential.

Probabilistic-based registration. Probabilistic-based PCR has also attracted widespread attention and becomes an important branch of registration methods. This type of method transfers the Euclidean distance minimization of two point clouds to a probability density function estimation problem. Representative methods include normal distributions transform (NDT) [62], coherent point drift (CPD) [63], [64], Gaussian mixture models (GMM) [65], [66], etc. As pointed out by [67], probabilistic-based methods are more robust to poor initializations compared with ICP-type registration. However, the limitation is that their results are not predictable.

Learning methods. Learning-based registration generally predicts the transformation matrix in an end-to-end

manner. According to [2], learning methods can be categorized into two groups: (1) using regression-based neural networks for registration, e.g., relativeNet [68], deepVCP [69], 3DregNet [70]; (2) combining traditional optimization methods with deep networks, e.g., PointNetLK [71], DeepGMR [72], PointDSC [73], HRegNet [74].

2.4 GORE-type Methods

GORE [1] provides a new way for PCR. It does not directly perform registration on the whole correspondence set $\mathcal{H} = \{\mathcal{I}^*, \mathcal{O}^*\}$, where $\mathcal{I}^*$ and $\mathcal{O}^*$ represent the optimal true inlier set and true outlier set, respectively. The basic idea of GORE is to reject a portion of outliers $\mathcal{O} \subseteq \mathcal{O}^*$ from $\mathcal{H}$ to obtain a small correspondence subset $\mathcal{H}' = \{\mathcal{I}^*, \mathcal{O}^* \backslash \mathcal{O}\}$. GORE achieves this goal based on the bound theorem, which rejects true outliers by the conflict between lower and upper bounds. As shown, GORE only filters out true outliers and $\mathcal{H}'$ preserves the optimal inlier set $\mathcal{I}^*$. It has two important properties: (1) the global optimal solution on $\mathcal{H}'$ is the same as the one on $\mathcal{H}$; (2) the inlier rate of $\mathcal{H}'$ is much higher than the one of $\mathcal{H}$. Thus, GORE can greatly simplify the registration problem without compromising the optimality of the solution. Current GORE methods cost exponential worst-case time complexity due to the usage of BnB. To increase the practicability, we have proposed a polynomial-time variant for bound estimation based on correspondence matrix and augmented correspondence matrix in our previous work [3]. However, this method does not guarantee that the optimality of the solution is not harmed. In this paper, we propose the first $O(N^2)$ time GORE method for PCR.

3 Methodology

3.1 Problem Formulation

Registration Problem. Given a 3D correspondence set $\mathcal{H} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_1^N$ with outliers and noise, the correspondence-based registration satisfies a rigid model,

$$\boldsymbol{y}_i = \mathbf{R}\boldsymbol{x}_i + \boldsymbol{t} + \boldsymbol{o}_i + \boldsymbol{\epsilon}_i \tag{1}$$

where $\mathbf{R} \in SO(3)$ and $\boldsymbol{t}$ are a rotation matrix and a translation vector of the rigid model, respectively. $\boldsymbol{\epsilon}_i$ represents a 3×1 noise vector. $\boldsymbol{o}_i = \mathbf{0}$ if $(\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathcal{I}^*$ (inlier); otherwise, it is a nonzero arbitrary vector. The goal is to recover the optimal $\mathbf{R}^*$ and $\boldsymbol{t}^*$ from $\mathcal{H}$. This robust registration problem can be formulated as a maximum consensus problem or a robust estimation problem.

Maximum consensus. The optimal model $(\mathbf{R}^*, \mathbf{t}^*)$ of **registration problem** corresponds to the optimal inlier set $\mathcal{I}^*$ with the largest size up to a noise bound ε ($\| \boldsymbol{\epsilon}_i \| < \varepsilon, \forall \boldsymbol{\epsilon}_i$). Instead of minimizing the registration error, maximum consensus maximizes the number of inliers,

$$\begin{aligned} \max_{\mathbf{R},\boldsymbol{t},\mathcal{I}\subseteq\mathcal{H}} \ & |\mathcal{I}| \\ s.t. \ & \| \boldsymbol{y}_i - (\mathbf{R}\boldsymbol{x}_i + \boldsymbol{t}) \| < \varepsilon, \ \ \forall (\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathcal{I} \end{aligned} \tag{2}$$

where $\mathcal{I}$ is an inlier set; $|\mathcal{I}|$ is the size of set $\mathcal{I}$; $\| \cdot \|$ represents the l_2-norm.

Robust estimation. Using robust estimators, **registration problem** can be reformulated as minimizing the registration residuals,

$$\min_{\mathbf{R},\boldsymbol{t}} \sum_{i=1}^{N} \rho(r_i) \tag{3}$$

where $r_i = \| \boldsymbol{y}_i - (\mathbf{R}\boldsymbol{x}_i + \boldsymbol{t}) \|$ is the residual of $(\boldsymbol{x}_i, \boldsymbol{y}_i)$; $\rho(\cdot)$ denotes a robust function, which can be the Huber, Cauchy, Tukey functions, etc. Directly optimizing problem (3) is difficult. Fortunately, according to the Black-Rangarajan duality [49], the robust statistics can be viewed as an outlier rejection process and reformulated as,

$$\arg\min_{\mathbf{R},\boldsymbol{t},w_i\in[0,1]} \sum_{i=1}^{N} w_i r_i^2 + \Psi(w_i) \tag{4}$$

where $w(r) = \frac{\psi(r)}{r} = \frac{\partial\rho(r)}{\partial r} \Big/ r$ is a weight function; $\psi(r) = \frac{\partial\rho(r)}{\partial r}$ is an influence function; $\frac{\partial\rho(r)}{\partial r}$ is the derivative of ρ w.r.t. r, and $\Psi(w_i)$ is an outlier penalty function whose expression depends on the function ρ or ψ. Then, problem (4) can be easily optimized via the iterated reweighted least-squares (IRLS) [75].

3.2 Our Quadratic-time GORE (QGORE)

GORE suffers from high computational cost (exponential in the worst case), which largely limits its usages in real applications. Here, we propose the first quadratic-time GORE method, called QGORE. We use the by-product of 3D feature matching, i.e., "rotation correspondence", to simplify traditional 3-point RANSAC into a one-point RANSAC for lower bound estimation, which reduces the number of iterations by $2 \sim 4$ orders. We leverage a simple but efficient and effective voting idea based on geometric consistency for tight upper bound seeking, which achieves the same tightness as the one based on BnB in GORE. Since the calculation of our one-point RANSAC algorithm (abbreviated as 1-ptRS) is independent of the subsequent QGORE pruning, our whole pruning process is actually a 1-ptRS+QGORE pipeline. Hereafter, we use QGORE to represent the pipeline of 1-ptRS+QGORE as long as 1-ptRS has operating conditions (i.e., "rotation correspondence" is available).

3.2.1 GORE Theory Revisit

Instead of directly solving problem (2), GORE [1] reformulates it as a set of subproblems,

$$\max_{(\boldsymbol{x}_k,\boldsymbol{y}_k)\in\mathcal{H}} p_k \tag{5}$$

where $p_k = |\mathcal{I}_k^*| + 1$ represents the biggest inlier size of subproblem $\mathcal{P}_k$,

$$\begin{aligned} \max_{\substack{\mathbf{R}_k,\boldsymbol{t}_k,\\ \mathcal{I}_k\subseteq\mathcal{H}\backslash\{(\boldsymbol{x}_k,\boldsymbol{y}_k)\}}} \ & |\mathcal{I}_k| + 1 \\ s.t. \ & \| \boldsymbol{y}_i - (\mathbf{R}_k\boldsymbol{x}_i + \boldsymbol{t}_k) \| < \varepsilon, \ \ \forall (\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathcal{I}_k \\ & \| \boldsymbol{y}_k - (\mathbf{R}_k\boldsymbol{x}_k + \boldsymbol{t}_k) \| < \varepsilon. \end{aligned} \tag{$\mathcal{P}_k$}$$

Subproblem $\mathcal{P}_k$ is a maximum consensus problem with constraints, where correspondence $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ must obey the model $(\mathbf{R}_k, \boldsymbol{t}_k)$. There are a total of N subproblems. Then,

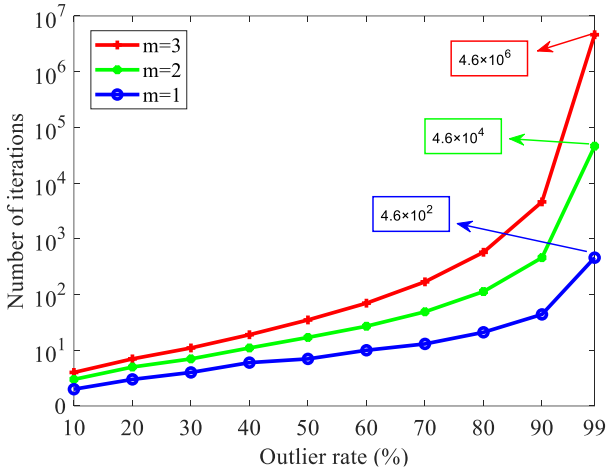


Fig. 1. Number of iterations T required by RANSAC with different subset sizes m and outlier rates η.

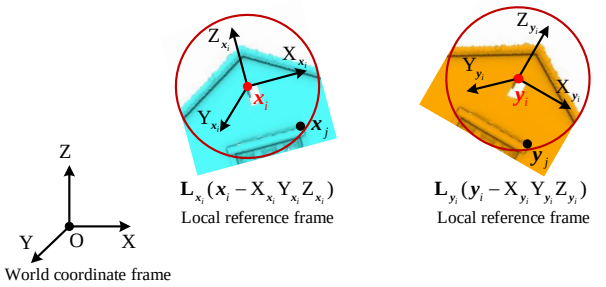


Fig. 2. Local reference frames of $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$.

we can reject true outliers from $\mathcal{O}^*$ based on the following bound confliction theorem.

Theorem 1. *Let $l \leq |\mathcal{I}^*|$ be a lower bound of problem (2) and $u_k \geq p_k$ be an upper bound of problem $\mathcal{P}_k$, if $u_k < l$, then $(\boldsymbol{x}_k, \boldsymbol{y}_k) \in \mathcal{O}^*$ is a true outlier [1].*

The core of GORE is to estimate a tight lower bound l for problem (2) and a tight upper bound u_k for each subproblem $\mathcal{P}_k$.

3.2.2 Lower Bound of QGORE

Any suboptimal solution to the problem (2) provides a reasonable lower bound l [1]. Hence, a popular way that gives rise to l is to use approximate methods such as RANSAC. For the 6D registration problem, the size of minimal subsets sampled by RANSAC is $m = 3$. The number of samples needed to be drawn is calculated as follows,

$$\left\lceil \frac{\log(1-c)}{\log(1-(1-\eta)^m)} \right\rceil \tag{6}$$

where η represents the outlier rate; $\lceil \cdot \rceil$ is a ceiling function; and c is the probability that at least one good minimal subset (only contains inliers) exists in all samples, which is set to 0.99. Fig. 1 displays the number of iterations T required by RANSAC under different outlier rates η and subset sizes m. As shown, T increases exponentially with η and m. For correspondence-based PCR, η is generally very high (often $> 95\%$). Hence, a possible way that can largely improve the efficiency of RANSAC is to reduce m. For example, at an outlier rate of 99%, $T(m = 1)$ is only $\frac{1}{10000}$ of $T(m = 3)$. According to this idea, we propose a one-point RANSAC ($m = 1$).

"Rotation correspondence". Inspired by the concept of "affine correspondence" [76], [77], [78], [79] in 2D image matching, "rotation correspondence" can also be used for the reduction of degree-of-freedoms (DoF) of PCR. In 3D feature description, local reference frame (LRF) is generally used to achieve rotation invariance. Similar to "dominant orientation" in image matching, LRF can be regarded as the relative rotation to the "dominant orientation" of 3D local surface.Ideally, if $(\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathcal{I}^*$, $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ should have the same "dominant orientation". As shown in Fig. 2, $\boldsymbol{x}_i - X_{\boldsymbol{x}_i} Y_{\boldsymbol{x}_i} Z_{\boldsymbol{x}_i}$ and $\boldsymbol{y}_i - X_{\boldsymbol{y}_i} Y_{\boldsymbol{y}_i} Z_{\boldsymbol{y}_i}$ are the LRFs of $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$, respectively. Since LRF has rotation and translation invariance, point $\boldsymbol{x}_j$ in the LRF $\boldsymbol{x}_i - X_{\boldsymbol{x}_i} Y_{\boldsymbol{x}_i} Z_{\boldsymbol{x}_i}$ and point $\boldsymbol{y}_j$ in the LRF $\boldsymbol{y}_i - X_{\boldsymbol{y}_i} Y_{\boldsymbol{y}_i} Z_{\boldsymbol{y}_i}$ should have the same coordinates, where $(\boldsymbol{x}_j, \boldsymbol{y}_j)$ is an inlier. Therefore, we must have,

$$\mathbf{L}_{\boldsymbol{x}_i}(\boldsymbol{x}_j - \boldsymbol{x}_i) = \mathbf{L}_{\boldsymbol{y}_i}(\boldsymbol{y}_j - \boldsymbol{y}_i) \tag{7}$$

where $\mathbf{L}_{\boldsymbol{x}_i}$ and $\mathbf{L}_{\boldsymbol{y}_i}$ are LRF rotation matrices of $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$, respectively. $\mathbf{L}_{\boldsymbol{x}_i}$ ($\mathbf{L}_{\boldsymbol{y}_i}$) is a 3×3 orthogonal matrix formed by three orthogonal unit vectors. Generally, the normal of the local surface centered at a feature is selected as the first axis of LRF and the other two lie in the tangent plane. Considering Eq. (1), we have $(\boldsymbol{y}_j - \boldsymbol{y}_i) = \mathbf{R}(\boldsymbol{x}_j - \boldsymbol{x}_i)$ (noise is not considered for ideal case). Then, we can deduce the rotation between $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$,

$$\mathbf{R}_i = {\mathbf{L}_{\boldsymbol{y}_i}}^T \mathbf{L}_{\boldsymbol{x}_i} \tag{8}$$

The subscript i in $\mathbf{R}$ indicates that each correspondence can obtain a rotation matrix. LRFs are the by-products of many 3D descriptors, which do not introduce any additional computation overhead. Actually, $\mathbf{R}_i$ is also the rotation of $\mathcal{H}$. Then, the 6-DoF rigid registration is simplified as a 3-DoF translation estimation problem and a single correct correspondence gives rise to a solution of $\boldsymbol{t}$.

One-point RANSAC. We adopt the GCRANSAC [36] framework for suboptimal registration. In each iteration, we draw a single sample $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ with an assigned rotation matrix $\mathbf{R}_i$ and calculate a translation vector $\boldsymbol{t}_i$. Then, a consensus set that is consistent with $\boldsymbol{t}_i$ is used for graph-cut based local optimization. The size of largest consensus set is accepted as our lower bound l.

3.2.3 Upper Bound of QGORE

Suppose condition $\mathbb{C}$ gives rise to an upper bound to problem (2), then, we must have that all inliers in $\mathcal{I}^*$ should obey condition $\mathbb{C}$. Moreover, the more outliers conflicted with condition $\mathbb{C}$, the tighter the bound. Our goal is to seek a reasonable condition $\mathbb{C}$.

Geometric consistency. Although geometric consistency has been used for transform decomposition [10], [13] and correspondence grouping [80], [81], [82], it has never been adapted for bound determination. Given two correspondences $(\boldsymbol{x}_i,\ \boldsymbol{y}_i) \in \mathcal{H}$ and $(\boldsymbol{x}_j,\ \boldsymbol{y}_j) \in \mathcal{H}$, we can construct a line correspondence $(\vec{\boldsymbol{x}}_{ij} = \boldsymbol{x}_i - \boldsymbol{x}_j, \vec{\boldsymbol{y}}_{ij} = \boldsymbol{y}_i - \boldsymbol{y}_j)$. According to equation (1), we have,

$$\vec{\boldsymbol{y}}_{ij} = \mathbf{R}\vec{\boldsymbol{x}}_{ij} + \vec{\boldsymbol{o}}_{ij} + \vec{\boldsymbol{\epsilon}}_{ij} \tag{9}$$

where $\vec{\boldsymbol{o}}_{ij} = \boldsymbol{o}_i - \boldsymbol{o}_j$ and $\vec{\boldsymbol{\epsilon}}_{ij} = \boldsymbol{\epsilon}_i - \boldsymbol{\epsilon}_j$. Obviously, the length of $\vec{\boldsymbol{x}}_{ij}$ is an invariant to rotations. Then, the following constraint is deduced based on the triangle inequality,

$$|\| \vec{\boldsymbol{y}}_{ij} \| - \| \vec{\boldsymbol{x}}_{ij} \|| \leq \| \vec{\boldsymbol{\epsilon}}_{ij} + \vec{\boldsymbol{o}}_{ij} \| \tag{10}$$

Case 1: If both $(\boldsymbol{x}_i,\ \boldsymbol{y}_i)$ and $(\boldsymbol{x}_j,\ \boldsymbol{y}_j)$ are inliers, then, $\vec{\boldsymbol{o}}_{ij} = \mathbf{0}$, equation (10) becomes

$$|\| \vec{\boldsymbol{y}}_{ij} \| - \| \vec{\boldsymbol{x}}_{ij} \|| \leq \| \vec{\boldsymbol{\epsilon}}_{ij} \| < 2\varepsilon \tag{GCC}$$

which is called geometric consistency constraint (GCC). All inlier pairs must obey the GCC.

Case 2: If at least one of $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ and $(\boldsymbol{x}_j, \boldsymbol{y}_j)$ is an outlier, i.e., $(\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathcal{O}^*$ or/and $(\boldsymbol{x}_j, \boldsymbol{y}_j) \in \mathcal{O}^*$, then, the length residual $d_{ij} = \| \vec{\boldsymbol{y}}_{ij} \| - \| \vec{\boldsymbol{x}}_{ij} \|$ is an arbitrary since $(\vec{\boldsymbol{x}}_{ij}, \vec{\boldsymbol{y}}_{ij})$ is an outlier. Actually, the length residuals d_{ij} is a random variable. Without loss of generality, we assume that $|d_{ij}|$ are identically distributed within $[0, D]$, where D is the range of $|d_{ij}|$. The probability density function PDF of $|d_{ij}|$ is,

$$PDF(|d_{ij}|) = \frac{1}{D}, \quad 0 \leq |d_{ij}| \leq D \tag{11}$$

Hence, the probability that $(\vec{\boldsymbol{x}}_{ij}, \vec{\boldsymbol{y}}_{ij})$ satisfies the constraint (GCC) is,

$$P = PDF(0 \leq |d_{ij}| < 2\varepsilon) = \frac{2\varepsilon}{D}, \quad \varepsilon \ll D \tag{12}$$

As analyzed, all inlier line correspondences satisfy the (GCC) while most of outliers conflict with it. Note that we take the identical distribution as an example for computational convenience. The same conclusion can also be drawn under normal distribution assumption. Thus, the (GCC) can provide a reasonable condition $\mathbb{C}$.

Voting. Now, let's see the subproblem $\mathcal{P}_k$. Correspondence $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ can form $N-1$ line correspondences with $\mathcal{H}\backslash(\boldsymbol{x}_k, \boldsymbol{y}_k)$. If $(\vec{\boldsymbol{x}}_{ki}, \vec{\boldsymbol{y}}_{ki})$ formed by $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ and $(\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathcal{H}\backslash(\boldsymbol{x}_k, \boldsymbol{y}_k)$ satisfies the constraint (GCC), then, we regard that $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ gets one vote from $(\boldsymbol{x}_i, \boldsymbol{y}_i)$. If $(\boldsymbol{x}_k, \boldsymbol{y}_k) \in \mathcal{I}^*$ is a true inlier, the constructed line correspondences consist of $(1-\eta)N - 1$ inliers and ηN outliers: (1) it obtains $(1-\eta)N - 1 = |\mathcal{I}^*| - 1$ votes from inlier set $\mathcal{I}^*$ based on the analysis in **Case 1**, where η is the outlier rate; (2) it gets $\eta PN = \frac{2\varepsilon\eta N}{D}$ votes from $\mathcal{O}^*$ according to the analysis in **Case 2**. Then, the total number of votes is $v_k = |\mathcal{I}^*| + \eta PN - 1$. Since $p_k = |\mathcal{I}^*|$ and $\eta PN \geq 0$, we must have $v_k + 1 \geq p_k = |\mathcal{I}^*| > l$. Thus, $v_k + 1$ can be an upper bound of p_k, and it is guaranteed that any inlier will not be rejected by **Theorem 1**.

There are an infinite number of upper bounds for p_k, but only tight ones can be effectively applied for true outlier removal. For example, if we have an upper bound u_k and it is larger than l even for $(\boldsymbol{x}_k, \boldsymbol{y}_k) \in \mathcal{O}^*$, this bound is not a tight one and **Theorem 1** cannot be applied. To check the tightness of bound $v_k + 1$, let's see the case that $(\boldsymbol{x}_k, \boldsymbol{y}_k) \in \mathcal{O}^*$ is a true outlier, where all $N-1$ line correspondences are outliers. Thus, the total number of votes is $v_k = P(N-1) = \frac{2\varepsilon(N-1)}{D}$. If we want to reject most of the true outliers, $v_k + 1$ should be smaller than l in most cases (Please note that it does not need to guarantee that $v_k + 1 < l$ is always hold, since the theory of GORE is only to reject a portion of true outliers, not all). We first calculate the ratio λ between $v_k + 1$ and $|\mathcal{I}^*|$,

$$\lambda = \frac{v_k + 1}{|\mathcal{I}^*|} = \frac{PN + (1-P)}{(1-\eta)N} \approx \frac{P}{1-\eta} \tag{13}$$

Empirically, the lower bound l obtained by RANSAC is generally larger than 80% of $|\mathcal{I}^*|$, i.e., $\frac{l}{|\mathcal{I}^*|} > 80\%$.

As known, the resolutions of 3D point clouds captured by structured light and high-precision LiDARs (e.g., Faro and Riegl LiDARs) are centimetres, millimetres, or even sub-millimetres. The noise bound ε is generally set to be $2 \sim 3$ times of the resolution. For these high-precision sensors, the probability $\frac{2\varepsilon}{D}$ is smaller than $\frac{1}{1000}$, i.e., $\lambda < 10\%$ when the outlier rate is 99%. Even for the sparse KITTI dataset, in which ε is generally set to $0.5m$ and $D \approx 200m$ is about the twice of the measure range, the probability $P \approx 0.005$ and the ratio $\lambda \approx 50\%$ at an outlier rate of 99%. Thus, $v_k + 1$ is generally a tight bound and most true outliers can be removed.

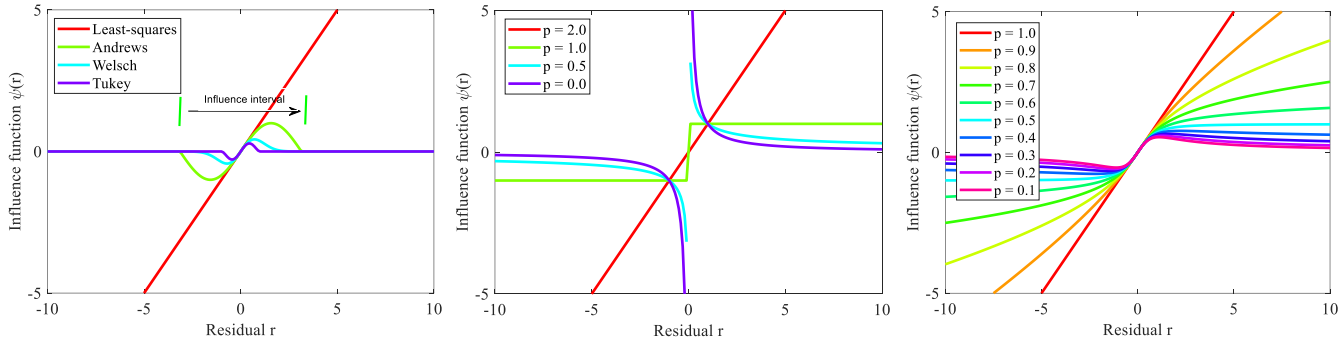
Even in extreme cases, where $\lambda > 80\%$ due to unsatisfied identity distribution assumption of $|d_{ij}|$, high noise levels, and extreme outlier ratios, we can still perform a second order GCC voting on the concensus set of subproblem $\mathcal{P}_k$ to obtain a tight upper bound. We will further compare its tightness with the original one of GORE in the experimental section.

3.3 Our l_p-like Robust Estimator

QGORE is an outlier pruning algorithm. Hence, a robust model fitting algorithm should be applied for accurate registration. The BnB [83] algorithm is generally adopted to obtain the globally optimal solution. However, BnB is still very slow after pruning. Therefore, this paper also proposes a local optimal method (adaptive l_p estimator) for choice, which is several orders of magnitude faster than BnB and more accurate than RANSAC-type methods.

Least-squares versus M-estimators. Influence function $\psi(r) = \frac{\partial\rho(r)}{\partial r}$ can intuitively reflect the impact that an observation has on the energy cost. Fig. 3(a) plots the influence functions of least-squares and M-estimators (Andrews, Welsch, and Tukey estimators). As shown, least-squares has a linear relationship with the residual. Large residuals have more impact on the solution. Thus, least-squares is not a robust cost and very sensitive to outliers. In contrast, the influence functions of M-estimators are bounded redescending functions. Observations with large residuals almost have no impact on the energy. Hence, M-estimators have good robustness. However, the disadvantage of traditional M-estimators is that their influence interval is small. As illustrated in Fig. 3(a), influence interval is the residual range that has a large impact on the energy cost. For example, only observations with residuals $|r| < \pi$ will participate in the optimization of Andrews estimator. In the cases with high outlier rates, initial models may greatly deviate from the ground truth, resulting in large residuals for inlier observations. These inlier observations will be ignored by the optimization at the beginning, which makes the optimization diverge. Least-squares, on the other hand, has the characteristic of a large impact interval, i.e., all observations have an impact on the optimization. Therefore,

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2023.3262780



(a) Least-squares versus M-estimators (b) l_p (Least-power) estimator (c) Our improved l_p cost

Fig. 3. Influence functions. Least-squares is a linear function; M-estimators are bounded redescending functions; both l_p estimator and our improved l_p cost are a family of cost functions with a robustness-control parameter.

an ideal robust cost should have the advantages of both the least-squares and M-estimators: (1) the influence interval decreases from wide to narrow along with iterations; (2) the influence function gradually changes from an unbounded function to a bounded redescending function (robustness increases from low to high).

***l_p* (Least-power) estimator [84].** The influence function of l_p cost is $\psi(r) = sgn(r)|r|^{p-1}$, which is essentially a family of cost functions. Specifically, it is a least-squares with $p = 2$, a l_1 cost with $p = 1$, and a Schatten p-norm cost with $0 < p < 1$. It has no robustness when $p \geq 2$. The l_1 cost has a certain degree of robustness since it is a bounded function (Fig. 3(b)). However, large residuals have the same influence as small ones. Schatten p-norm gets better robustness than l_1 as its curve is descending (Fig. 3(b)). As shown, the robustness of l_p estimator increases as the value of p decreases. Therefore, if we decrease p along with iterations with an initial value of 2, the influence interval gradually decreases and the robustness increases, which perfectly satisfies the characteristics of an ideal robust cost summarized in the above. However, many difficulties are encountered when $p < 1$ due to the troublesomeness arose from zero residuals. Moreover, inliers with residuals around 0 have too much influence on the energy compared to inliers with relatively large residuals, which will bias the solution.

Our improved *l_p* cost. First, we use r^2 to eliminate the absolute value in l_p cost and add a constant to avoid division by zero. Then, our l_p-like ψ-type robust cost is,

$$\psi(r) = r(1 + r^2)^{p-1} \tag{14}$$

The same as l_p cost, it is also a family of functions. When $p = 1$, it is an ordinary least-squares; when $p = 0.5$, it is a l_1-l_2 cost; it becomes a Cauchy function if $p = 0$; and it yields a Geman-McClure estimate when $p = -1$ (Note that our l_p-like adaptive robust cost is in ψ-type instead of ρ-type). As shown in Fig. 3(c), the cost function can be gradually changed from the least-squares to M-estimators by controlling the parameter p, which fills the robustness and influence interval gaps between least-squares and M-like costs. Compared with the Adaptive loss consisted of four pieces [85], our proposed l_p-like cost does not suffer from any singularity, resulting in a more concise and elegant robust estimate.

Algorithm 1: Our proposed QGORE algorithm

Input: A correspondence set $\mathcal{H}$ with LRFs.
Output: Subset $\mathcal{H}'$ with much higher inlier rate.

```
// ε: three times of noise level
```
Initialization: $\varepsilon = 3 * \sigma$; $\mathcal{H}' = \mathcal{H}$.

```
// Use one-point RANSAC to seek l;
// lo: local optimization;
// GC: graph-cut;
```
$l = 1\text{-}ptRANSAC(\mathcal{H}, m = 1, c = 0.99, lo = GC)$;

```
// Seek u_k for each P_k based on GCC;
// k represents (x_k, y_k), u_k = v_k + 1;
```
for *each* $k \in \mathcal{H}$ **do**
```
    // geometric consistency voting
```
    $v_k = GCC\_voting(\mathcal{H} \backslash k, 2\varepsilon)$;
    if $v_k + 1 < l$ **then** $k \in \mathcal{O}^*$, $\mathcal{H}' = \mathcal{H}' \backslash k$;
end
return $\mathcal{H}'$

IRLS Optimization. We reformulate our cost as a Black-Rangarajan duality problem based on Eq. (4), where the penalty function Ψ of our ψ is,

$$\Phi(w_i) = w_i - \frac{p-1}{p} w_i^{\frac{p}{p-1}} \tag{15}$$

We use an adaptive IRLS strategy to optimize problem (4). There are three main steps at each inner iteration t of the adaptive IRLS:

1) **Variable update**: optimize $\boldsymbol{\delta}^{(t)} = (\mathbf{R}^{(t)}, \boldsymbol{t}^{(t)})$ with fixed weights $w_i^{(t-1)}$ in Eq. (4),

$$\boldsymbol{\delta}^{(t)} = \arg\min_{\boldsymbol{\delta}} \sum_{i=1}^{n} w_i^{(t-1)} r_i^2 \tag{16}$$

where the second term in (4) is a constant with a given $w_i^{(t-1)}$ and is dropped. This is a simple WLS problem, which can be solved globally using non-minimal solvers.

TABLE 1
Comparison of the tightness of upper bounds.

Method	$\|\mathcal{H}\|$	l	$\|\mathcal{I}^*\|$	$\|\mathcal{H}'\|$	Time (s)
GORE	10000	88.42	96.96	100.59	43.47
Our QGORE				101.10	2.72

2) **Weight update**: optimize weights $\boldsymbol{w}^{(t)} = \{w_i^{(t)}\}_1^N$ based on Eq. (4) with fixed $\boldsymbol{\delta}^{(t)}$,

$$\boldsymbol{w}^{(t)} = \arg\min_{w_i \in [0,1]} \sum_{i=1}^{n} w_i \left(r_i^{(t)}\right)^2 + \Psi(w_i) \tag{17}$$

where $r_i^{(t)}$ is a constant with a given $\boldsymbol{\delta}^{(t)}$. Essentially, the solution of (17) is equivalent to,

$$w_i^{(t)} = \left(1 + \left(r_i^{(t)}\right)^2\right)^{p-1} \tag{18}$$

3) p **update**: change the weight/cost function via decreasing parameter p by a step-size τ, i.e., $p^{(t)} = p^{(t-1)} - \tau$.

Our adaptive IRLS is similar to the traditional IRLS. The only difference is that the adaptive IRLS has one more step of p update, so that the robustness of cost function can be gradually increased.

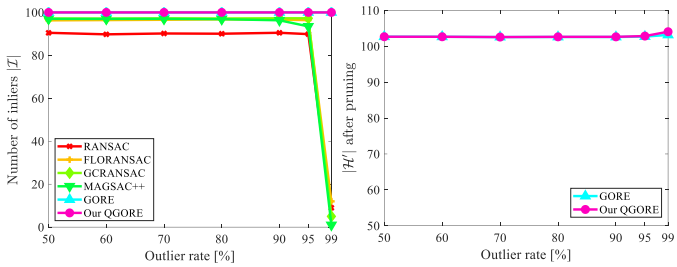
3.4 Main Algorithm and Complexity

Algorithm 1 summarizes our proposed QGORE method. Given a set of "Rotation correspondences" $\mathcal{H}$, QGORE first seeks a reasonable bound l for problem (2) by using one-point RANSAC and a tight upper bound u_k for each sub-problem $\mathcal{P}_k$ based on geometric consistency voting. True outliers $O \subseteq \mathcal{O}^*$ are rejected by the bound confliction theorem, obtaining a subset $\mathcal{H}'$ with a much higher inlier rate. Then, our adaptive l_p-like estimator (abbreviated as Ada_l_p) is adapted on the result of QGORE to estimate the final registration transformation.

In our whole QGORE+Ada_l_p pipeline, one-point RANSAC costs $O(T_1 N)$ time, where T_1 is the number of iterations of RANSAC; The "for" loop (Line 3∼6) costs $O(N^2)$ time since it needs to calculate pairwise distances between correspondences; l_p-like estimator consumes $O(T_2 N)$, where T_2 is the number of iterations of IRLS. As aforementioned, $T_1 = 459$ even if the outlier rate reaches 99%. It is generally smaller than N. The maximum number of iterations of IRLS is set to 50, i.e., $T_2 \leq 50$. Thus, our total time complexity is approximate $O(N^2)$.

4 Experiments and Evaluations

In this section, we first compare the tightness of the upper bound of QGORE with the one in GORE on simulated data; then, we qualitatively and quantitatively evaluate our registration accuracy on both simulated and real datasets; finally, we develop a 3D laser odometry based on our QGORE+Ada_l_p pipeline and show the potentials of our method in the simultaneous localization and mapping (SLAM) task. All the following experiments are performed on a laptop with a i7-8550U @ 1.8GHz CPU, and 8 GB of RAM.



(a) Number of inliers after pruning (b) Subset size $|\mathcal{H}'|$ after pruning

Fig. 4. Comparison results of outlier pruning.

4.1 Tightness of Upper Bound

We first randomly generate N 3D points $\mathcal{X} = \{\boldsymbol{x}_i\}_1^N$ by using a normal distribution generator $\mathcal{N}(0, 100^2)$. The true correspondence set $\mathcal{Y}^t$ of $\mathcal{X}$ is obtained via transformation $\boldsymbol{y}_i^t = \mathbf{R}^t \boldsymbol{x}_i + \boldsymbol{t}^t$, where $\mathbf{R}^t$ represents the ground truth rotation generated within $[-\pi/2,\ \pi/2]$ and $\boldsymbol{t}^t$ is the ground truth translation generated within [-100, 100]. Given an outlier rate η (Note that our method does not require η as an input. It is just to show how difficult situations our method can cope with.), we randomly select $\lceil \eta N \rceil$ points from $\mathcal{Y}^t$ to add gross errors generated by $\mathcal{N}(0, (100\sigma)^2)$ and the other points are added Gaussian noise $\mathcal{N}(0, \sigma^2)$, where $\sigma = \frac{\varepsilon}{3}$ is the standard deviation of noise, obtaining contaminated correspondences $\mathcal{Y}$ and $\mathcal{H} = \{\mathcal{X}, \mathcal{Y}\}$. In this experiment, we set $N = 10^4$, $\eta = 99\%$, and $\varepsilon = 0.6$. We test GORE and our QGORE for 100 independent runs. To fairly compare the tightness of upper bound, we fix the lower bound of QGORE to be the same as that of GORE. In each run, we record the following measures: lower bound l, consensus size of global solution $|\mathcal{I}^*|$, the number of remained correspondences $|\mathcal{H}'|$, and running time.

The results are reported in Table 1. As shown, both GORE and QGORE can reject more than 99.9% true outliers, i.e., $\frac{|\mathcal{O}|}{|\mathcal{O}^*|} > 99.9\%$. Our QGORE achieves comparable tightness with GORE while being 15 times faster.

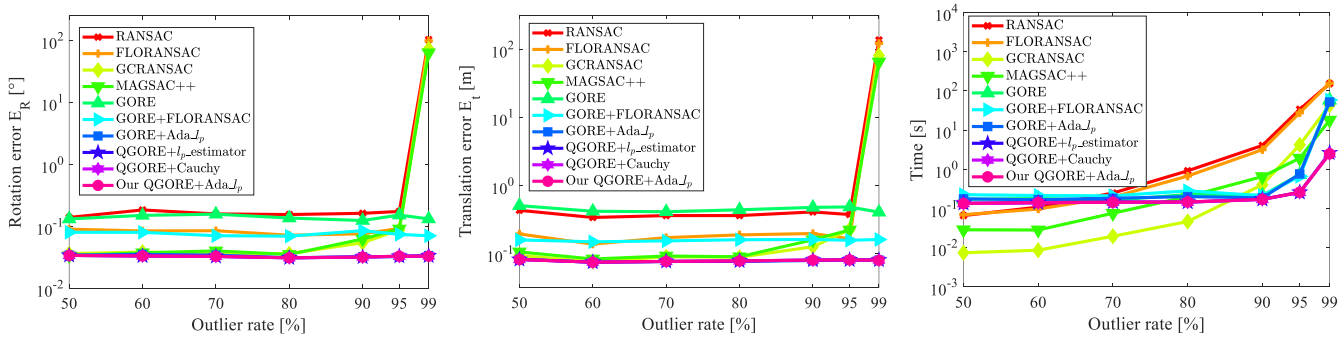
4.2 Registration Performance Evaluation

Here, we comprehensively evaluate the registration performance of our method on both simulated and real datasets. We compare our QGORE/QGORE+Ada_l_p with 11 algorithms/pipelines, including RANSAC [18], FLORANSAC [35], GCRANSAC [36], MAGSAC++ [86], GORE [1], 4DOFReg [87], GORE+FLORANSAC, GORE+Ada_l_p, 1-ptRS+Ada_l_p, QGORE+l_p_estimator, and QGORE+Cauchy, where Cauchy is the Cauchy M-estimator. For a fair comparison, we apply our 1-ptRS as a "preprocessor" for GORE as long as 1-ptRS has operating conditions. Settings for each compared method are summarized in Table 2. We record rotation error $E_{\mathbf{R}}$, translation error $E_{\boldsymbol{t}}$, and running time metrics for quantitative evaluations. The computational formulas of $E_{\mathbf{R}}$ and $E_{\boldsymbol{t}}$ are as follows,

$$\begin{cases} E_{\boldsymbol{t}} = \| \boldsymbol{t}^t - \boldsymbol{t}^e \| \\ E_{\mathbf{R}} = \arccos \frac{tr(\mathbf{R}^t (\mathbf{R}^e)^T) - 1}{2} \end{cases} \tag{19}$$

TABLE 2
Detailed settings of the compared algorithms (MNI represents maximum number of iterations).

Method	Parameters	Implementations
RANSAC	$m = 3$; confidence: 0.99; MNI: 10^5.	MATLAB code; single thread https://www.peterkovesi.com/matlabfns/index.html
FLORANSAC	$m = 3$; confidence: 0.99; LO size: 21; LO iterations: 50; MNI: 10^5.	MATLAB code; single thread https://zhipengcai.github.io
GCRANSAC	$m = 3$; confidence: 0.99; spatial weight: 0.14; number of neighborhoods: 20; LO iterations: 50; MNI: 10^5.	C++ code; single thread https://github.com/danini/graph-cut-ransac
MAGSAC++	$m = 3$; confidence: 0.99; maximum threshold: 50; MNI: 10^5.	C++ code; single thread https://github.com/danini/magsac
GORE	Lower bound: 0; repeat?: true.	C++ code; single thread https://cs.adelaide.edu.au/ aparra/project/gore
4DOFReg	Lower bound: 0; gap: 0; enforceOneToOne?: false.	C++ code; single thread https://github.com/ZhipengCai/
l_p_estimator	$p^{(0)} = 2$; $\tau = 0.1$; MNI: 100; convergence threshold: e^{-4}.	MATLAB code; single thread https://ljy-rs.github.io/web
Cauchy	Tuning: $c = 2.385$; MNI: 100; convergence threshold: e^{-4}.	MATLAB code; single thread https://ljy-rs.github.io/web
QGORE	$m = 1$; confidence: 0.99; lo = GC; MNI: 1000.	MATLAB code; single thread https://ljy-rs.github.io/web
Ada_l_p	$p^{(0)} = 1$; $\tau = 0.1$; MNI: 100; convergence threshold: e^{-4}.	MATLAB code; single thread https://ljy-rs.github.io/web



(a) $E_{\mathbf{R}}$ (b) $E_{\boldsymbol{t}}$ (c) running time

Fig. 5. Registration results on simulated data. Subfigures (a), (b), and (c) plot the angular error $E_{\mathbf{R}}$, translation error $E_{\boldsymbol{t}}$, and running time, respectively.

where $\boldsymbol{t}^e$ and $\mathbf{R}^e$ are estimated rigid model parameters; and $tr(\mathbf{X})$ represents the trace of $\mathbf{X}$.

4.2.1 Simulations

The simulation process is similar to Section 4.1. Differently, we fix the number of inliers ($\lceil(1-\eta)N\rceil = 100$) instead of N and increase the standard deviation of noise $\sigma = 0.5$. Because it is difficult to simulate the LRFs of sparse keypoints, 1-ptRS can not be applied. we use our previous work [3] for lower bound calculation in this simulation experiment. The 4DOFReg is not included for comparison since it is not suitable for 6-DoF registration simulated in this section.

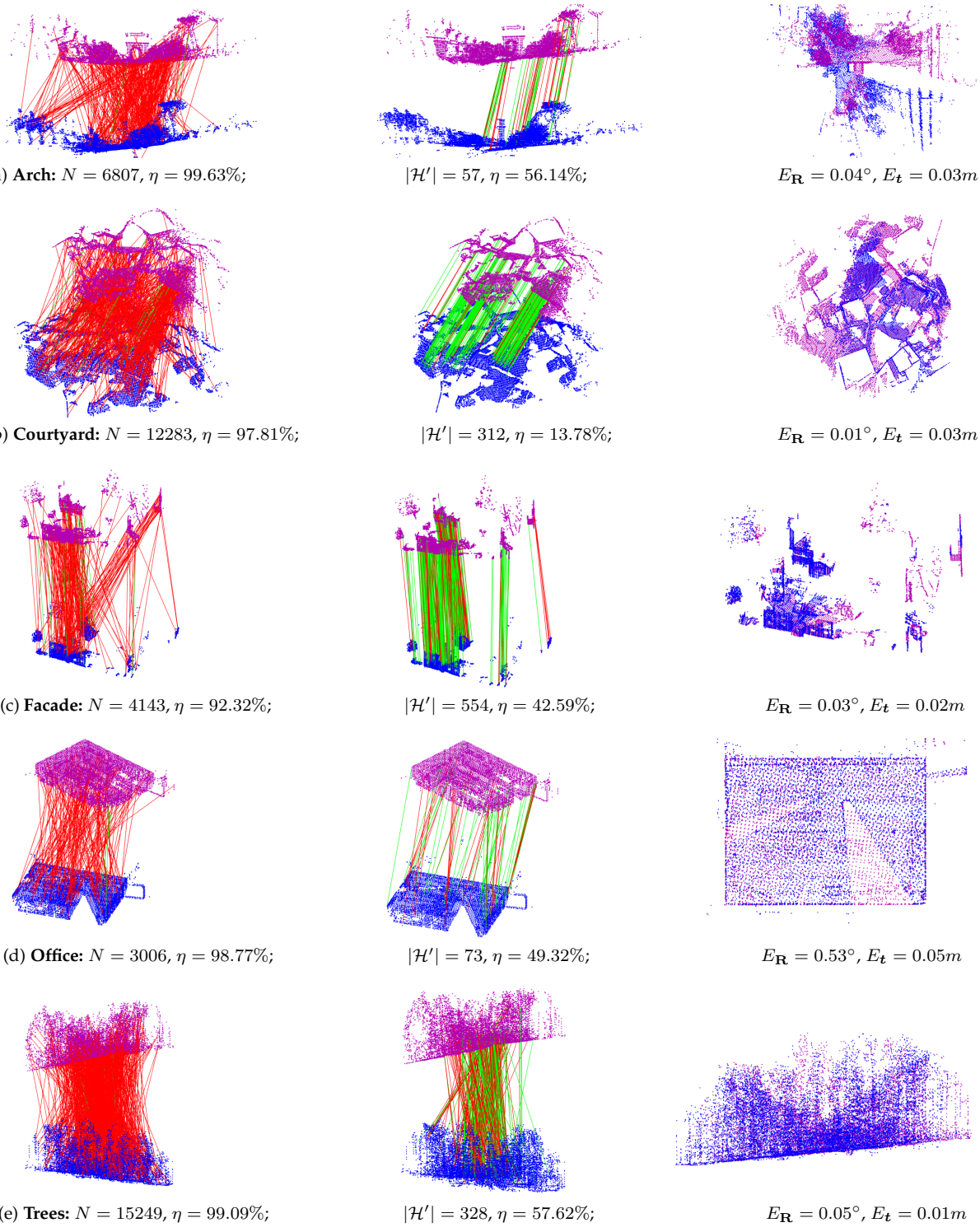
We first compare the outlier pruning performance of our QGORE with that of the RANSAC-type methods and GORE. From Fig. 4(a), we can see that only GORE and our QGORE preserve all inliers after pruning. RANSAC-type methods compromise the global optimality of the solution, since some true inliers are rejected by them. Actually, RANSAC-type methods can only obtain approximate solutions [19]. As shown in Fig. 4(b), the subset size $|\mathcal{H}'|$ of QGORE is almost the same as the one of GORE.

Figure 5 plots the comparison results. As shown, RANSAC-type methods cannot tolerate with 99% of outliers with 10^5 iterations. RANSAC and FLORANSAC only obtain approximate solutions, so their registration accuracies are worse than our QGORE+Ada_l_p. GCRANSAC and MAGSAC++ use graph cut or M-estimation for further model refinement. Hence, they are comparable to our pipeline at low outlier rates. Once the outlier rate reaches 90%, their registration accuracies largely decrease. GORE-type pipelines are still robust even at an outlier rate of 99%. Comparison of GORE+FLORANSAC and GORE+Ada_l_p, we can infer that the proposed adaptive l_p estimator has better fitting accuracy than RANSAC variants. Since the outlier rate after QGORE pruning is very low ($< 1\%$ as shown in Section 4.1), traditional l_p estimator and Cauchy

TABLE 3
Detailed information of the ETH dataset

Info	Arch	Courtyard	Facade	Office	Trees
Overlap	30%∼40%	40%∼70%	60%∼70%	>80%	≈50%
N_p	8	28	21	8	10
N_k	5808	4084	2273	1487	10404
N_c	7569	9915	4408	3194	14535
η	99.33%	96.14%	95.09%	98.23%	99.55%

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2023.3262780



(a) **Arch:** $N = 6807$, $\eta = 99.63\%$; $|\mathcal{H}'| = 57$, $\eta = 56.14\%$; $E_{\mathbf{R}} = 0.04°$, $E_{\boldsymbol{t}} = 0.03m$

(b) **Courtyard:** $N = 12283$, $\eta = 97.81\%$; $|\mathcal{H}'| = 312$, $\eta = 13.78\%$; $E_{\mathbf{R}} = 0.01°$, $E_{\boldsymbol{t}} = 0.03m$

(c) **Facade:** $N = 4143$, $\eta = 92.32\%$; $|\mathcal{H}'| = 554$, $\eta = 42.59\%$; $E_{\mathbf{R}} = 0.03°$, $E_{\boldsymbol{t}} = 0.02m$

(d) **Office:** $N = 3006$, $\eta = 98.77\%$; $|\mathcal{H}'| = 73$, $\eta = 49.32\%$; $E_{\mathbf{R}} = 0.53°$, $E_{\boldsymbol{t}} = 0.05m$

(e) **Trees:** $N = 15249$, $\eta = 99.09\%$; $|\mathcal{H}'| = 328$, $\eta = 57.62\%$; $E_{\mathbf{R}} = 0.05°$, $E_{\boldsymbol{t}} = 0.01m$

Fig. 6. Our qualitative results on the ETH dataset. Left column: initial correspondence set $\mathcal{H}$, where green lines and red lines represent inliers and outliers, respectively. Middle column: results of our QGORE pruning. Right column: final registration results of our QGORE+Ada_l_p pipeline.

TABLE 4
Average angular error ($E_{\mathbf{R}} \downarrow$), translation error ($E_{\boldsymbol{t}} \downarrow$), and running time ($\downarrow$) results on the ETH dataset. **Bold** fonts denote the first place.

Method	Arch			Courtyard			Facade			Office			Trees		
	$E_{\mathbf{R}}$ (°)	$E_{\boldsymbol{t}}$ (m)	time (s)	$E_{\mathbf{R}}$ (°)	$E_{\boldsymbol{t}}$ (m)	time (s)	$E_{\mathbf{R}}$ (°)	$E_{\boldsymbol{t}}$ (m)	time (s)	$E_{\mathbf{R}}$ (°)	$E_{\boldsymbol{t}}$ (m)	time (s)	$E_{\mathbf{R}}$ (°)	$E_{\boldsymbol{t}}$ (m)	time (s)
RANSAC	28.53	7.03	178.81	0.20	0.08	108.60	0.46	0.09	62.85	1.65	0.13	99.81	68.02	9.45	166.32
FLORANSAC	34.89	6.88	181.32	0.11	0.05	108.38	0.28	0.06	55.64	0.83	0.07	103.08	68.27	9.34	168.28
GCRANSAC	60.15	7.42	32.94	**0.04**	0.04	30.75	0.12	**0.02**	4.72	0.43	0.04	11.82	56.61	6.78	55.15
MAGSAC++	3.03	1.32	8.85	0.41	0.24	9.44	1.16	0.23	2.40	2.43	0.30	3.24	4.02	1.58	20.08
GORE	0.53	0.12	26.59	0.13	0.07	24.81	0.32	0.08	51.32	0.51	0.06	13.44	0.24	0.09	271.31
4DOFReg	**0.17**	0.14	3.17	0.06	0.05	3.98	0.15	0.07	1.29	1.46	0.13	0.80	0.23	0.13	11.04
GORE+FLORANSAC	0.31	0.09	34.89	**0.04**	0.04	25.06	0.25	0.05	51.38	0.43	0.05	13.97	0.22	0.05	279.02
GORE+Ada_l_p	**0.17**	0.05	26.62	**0.04**	**0.03**	24.83	**0.07**	**0.02**	51.33	0.29	**0.02**	13.46	0.17	0.05	271.35
1-ptRS+Ada_l_p	0.38	0.05	**0.92**	**0.04**	0.04	**1.49**	0.12	0.03	**0.25**	0.37	0.04	**0.22**	0.26	0.07	**2.47**
QGORE+l_p_estimator	1.58	0.96	3.71	**0.04**	0.04	7.21	0.39	0.09	1.75	5.14	0.38	0.80	0.53	0.09	10.51
QGORE+Cauchy	2.49	1.43	3.90	**0.04**	**0.03**	7.17	0.19	0.04	1.69	5.84	0.59	0.77	0.34	0.10	10.93
QGORE+Ada_l_p	0.18	**0.04**	3.95	**0.04**	**0.03**	6.74	**0.07**	**0.02**	1.88	**0.28**	**0.02**	0.84	**0.13**	**0.04**	10.66

TABLE 5
The lower bound l and subset size $|\mathcal{H}'|$ of GORE and QGORE.

Data	$\lvert\mathcal{I}^*\rvert$	1-ptRS+Ada_l_p	GORE		QGORE	
		$\lvert\mathcal{I}\rvert$	l	$\lvert\mathcal{H}'\rvert$	l	$\lvert\mathcal{H}'\rvert$
Arch	55	46	54	461	47	327
Courtyard	338	326	332	448	326	402
Facade	216	184	212	392	186	505
Office	44	40	42	137	40	162
Trees	75	64	68	352	65	224

M-estimator are comparable to our adaptive l_p estimator in this experiment.

From Fig. 5(c), we can see that GORE+Ada_l_p and GORE have similar running times, which indicated that GORE-type methods can effectively remove most of the outliers and greatly improve the efficiency of traditional methods. Moreover, our QGORE+Ada_l_p (implemented in MATLAB) is 15+ times faster than GORE+Ada_l_p (GORE is implemented in C++) at an outlier rate of 99%. Compared with RANSAC-type methods, our advantage is very obvious in cases with high outlier rates. For example, QGORE+Ada_l_p is about 10 times faster than GCRANSAC and MAGSAC++ (both of them are implemented in C++) at an outlier rate of 99%.

4.2.2 Real LiDAR experiment

Dataset: We use the challenging large-scale ETH LiDAR dataset[1] with ground truth for a real experiment, which consists of five scan categories:

- **Arch:** acquired around a Roman arch. The scans suffer from low overlap and moving objects.
- **Courtyard:** acquired in an ancient tomb courtyard. The scans lack of very low vertical variations.
- **Facade:** an urban scene with typical structural features. The scans contain many moving objects.
- **Office:** a representative indoor dataset of an office room, which suffers from self-similar structures.
- **Trees:** a forest scene with a large amount of underwood, which suffers from repeated texture.

1. http://www.prs.igp.ethz.ch/research/completed_projects/automatic_registration_of_point_clouds.html

To reduce the size of LiDAR scans, we downsample the resolution to 0.1m and set $\varepsilon = 0.3$. We use ISS [21] and FPFH [27] to extract features, and establish initial correspondence set $\mathcal{H}$ based on the top-5 nearest search. The LRF of each feature in $\mathcal{H}$ is also estimated via the method described in [28]. The detailed information is summarized in Table 3, including overlap ratio, number of scan pairs N_p, number of keypoints N_k, number of correspondences N_c, and outlier ratio η.

Qualitative Result: Figure 6 shows our qualitative results, where one LiDAR pair is selected from each category for evaluation. Our QGORE reduces the average number of correspondences from $|\mathcal{H}| = 8298$ to $|\mathcal{H}'| = 265$, i.e., more than 95% of correspondences are removed. Importantly, these rejected correspondences are guaranteed to be true outliers. Thus, it does not compromise the optimality of the solution. After pruning, the average outlier rate is reduced from 97.52% to 43.89%, which greatly simplifies the PCR problem.

Arch suffers from low overlaps; Courtyard lacks of vertical variations; Facade contains many moving objects; Office contains self-similar structures; and Trees is full of repeated texture. Although registration on these scan pairs is very challenging, our QGORE+Ada_l_p pipeline still yields very good results, i.e., the worst angular error is smaller than 1°.

Quantitative Comparison: Table 4 summarizes the average quantitative results. As shown, RANSAC-type algorithms perform well on cases with relatively low outlier rates (e.g., Courtyard and Facade), while getting poor results on Arch and Trees whose outlier rates are higher than 99%. This is consistent with the simulation results, in which RANSAC-type methods break down at 99% of outliers. 1-ptRS+Ada_l_p is less accurate than our QGORE+Ada_l_p, since 1-ptRS does not preserve the global solution. GORE-type methods (GORE, 4DOFReg, GORE+FLORANSAC, GORE+Ada_l_p, and our QGORE+Ada_l_p) achieve surprising robustness. They can successfully register all scan pairs. GORE itself only finds a suboptimal solution (by-product), which results in low registration accuracy. Fortunately, it only takes a few seconds to refine the results by FLORANSAC or l_p-like estimator after GORE pruning. Comparing GORE+FLORANSAC with GORE+Ada_l_p, our l_p-like estimator achieves much better registration accuracy than

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2023.3262780

TABLE 6
Comparison against GORE with a global optimiser BnB.

Data	$\|\mathcal{H}\|$	η (%)	GORE			QGORE			GORE+BnB		QGORE+BnB	
			l	$\|\mathcal{H}'\|$	time (s)	l	$\|\mathcal{H}'\|$	time (s)	$\|\mathcal{I}^*\|$	time (s)	$\|\mathcal{I}^*\|$	time (s)
Bimba	535	94.02	29	127	4.49	28	116	0.53	32	278.69	32	158.80
Children	530	95.09	23	124	2.33	20	104	0.43	26	385.47	26	131.95
Dragon	645	93.95	36	103	4.21	36	112	0.65	39	164.18	39	219.42
Angle	668	94.61	32	150	3.26	32	102	0.57	36	3547.06	36	265.31
Bunny	555	93.87	29	78	1.93	29	83	0.36	34	67.31	34	78.18

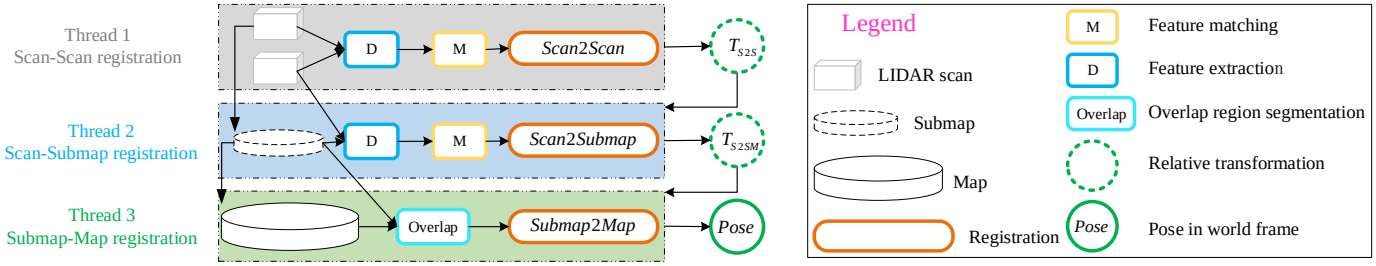


Fig. 7. The framework of the proposed laser odometry, which contains three threads, i.e., scan-to-scan, scan-to-submap, and submap-to-map registration threads.

FLORANSAC. GORE+Ada_l_p and QGORE +Ada_l_p obtain the best registration performance, whose angular errors are smaller than 0.5°and translation errors are better than 0.1m. The reasons are twofold: First, GORE-type methods preserve the global solution in $\mathcal{H}'$. Second, our l_p-like estimator has high robustness and is not easy to fall into local minima. Traditional l_p estimator and Cauchy M-estimator perform much worse than our Ada_l_p. For example, QGORE+l_p-estimator and QGORE+Cauchy get the worst performance on the Office category among 12 compared pipelines. In this experiment, the outlier rate after QGORE pruning is still very high compared with the one in Section 4.2.1. Traditional l_p estimator and Cauchy M-estimator become unreliable at such high outlier rates.

Another advantage of our QGORE+Ada_l_p is the high computational efficiency. 1-ptRS+Ada_l_p is the fastest; however, it breaks the global solution. 4DOFReg ranks the second, since it only solves a 4-DoF registration problem (1-DoF rotation + 3-DoF translation), which largely reduces the parameter space. However, this also greatly limits its application scenarios. Our QGORE+Ada_l_p is 15+ times faster than GORE+Ada_l_p on the whole dataset. This difference in computational efficiency is even more pronounced in difficult cases, which is determined by the possible exponential complexity of GORE. For example, in Trees with $\eta = 99.55\%$, GORE+Ada_l_p is 25+ times slower than our QGORE+Ada_l_p.

Table 5 compares the outlier pruning performance of GORE and QGORE on the ETH dataset, which reports the average size of optimal true inlier set $|\mathcal{I}^*|$, lower bound l, and the size of remaining correspondences $|\mathcal{H}'|$ on each category. As can be seen, our $|\mathcal{H}'|$ is much smaller than the one of GORE except for Facade. The average $|\mathcal{H}'|$ of QGORE on the whole dataset is 374, while GORE remains $|\mathcal{H}'| = 388$ correspondences. Our weakness is that the tightness of lower bound is not as good as GORE due to the noise of LRF estimation. This is the reason why our QGORE remains more correspondences than GORE on the Facade dataset.

4.3 Comparison against GORE with a BnB optimiser

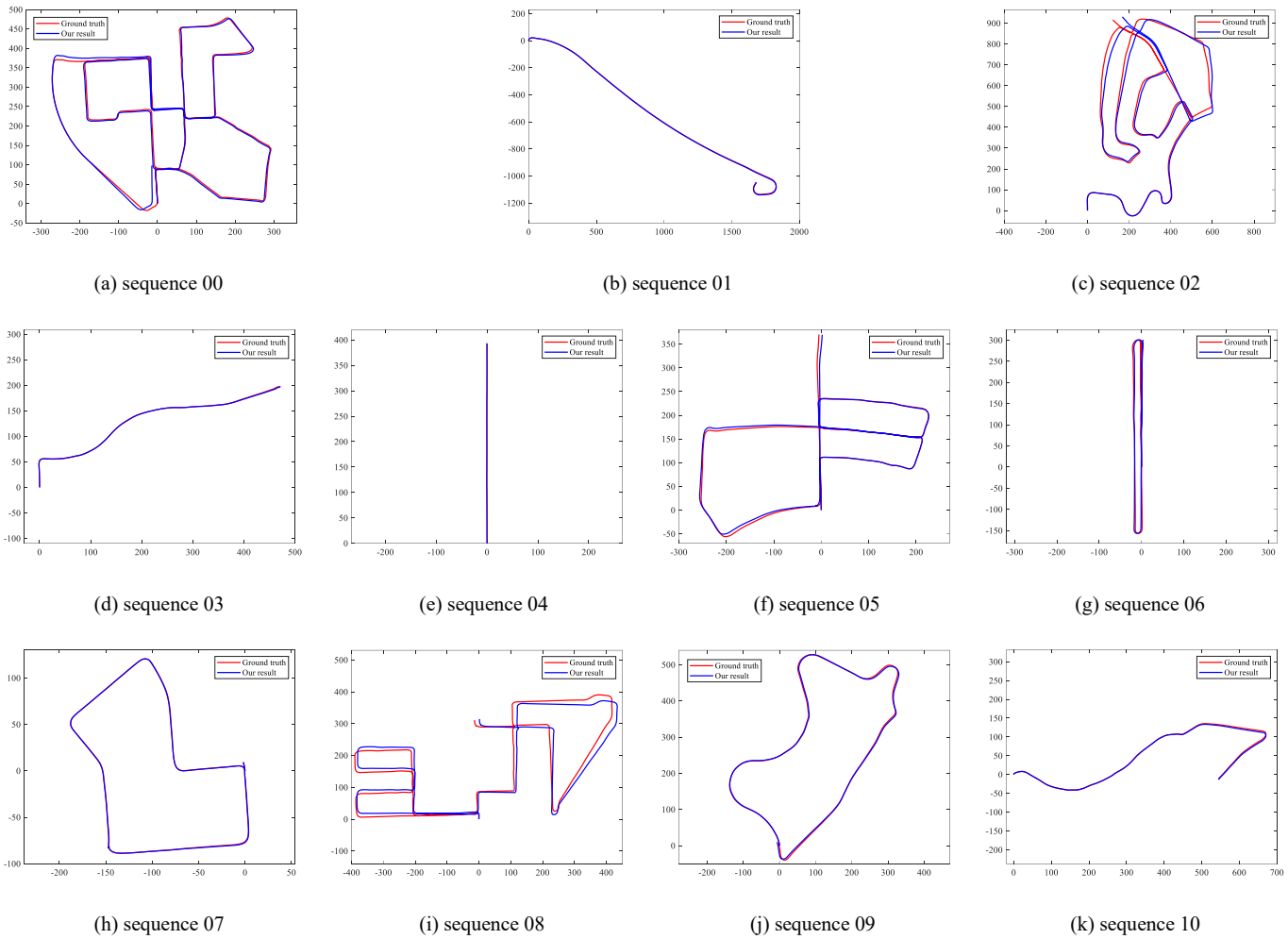
As aforementioned, our QGORE outlier pruning algorithm does not harm the globally optimal solution. To show its effectiveness, we compare QGORE against GORE with a global BnB optimiser [83] here. Since GORE+BnB is very slow on the ETH dataset (the first registration instance of Arch takes more than 12 hours on our laptop), we use an easier dataset, i.e., the FGR dataset [9], for evaluation. The FGR dataset contains five point cloud models with partial overlaps, including Bimba, Children, Chinese Dragon, Angle, and Bunny. Each model category has five point cloud pairs. For each registration instance, we record the following measures: lower bound l, number of remaining correspondences after pruning $|\mathcal{H}'|$, consensus size of global solution $|\mathcal{I}^*|$, and running time. The results are summarized in Table 6.

As shown, QGORE+BnB yields the same global solution $|\mathcal{I}^*|$ as GORE+BnB, while being almost 5 times faster than GORE+BnB. This benefits from the good outlier pruning performance of our QGORE, i.e., the number of remaining correspondences after QGORE is smaller than the one of GORE (116 versus 103 in average). Moreover, our QGORE is 6 times faster than GORE on this dataset.

4.4 SLAM Experiment on KITTI

To show the potential of our QGORE in the SLAM task, we develop a 3D laser-based odometry. As shown in Figure 7, there are three threads in our framework. In the first thread, a scan-to-scan registration is performed to estimate a relative rigid transformation T_{S2S} between two consecutive laser frames. Thread 2 uses the transformation T_{S2S} to guide

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2023.3262780



(a) sequence 00 (b) sequence 01 (c) sequence 02

(d) sequence 03 (e) sequence 04 (f) sequence 05 (g) sequence 06

(h) sequence 07 (i) sequence 08 (j) sequence 09 (k) sequence 10

Fig. 8. Qualitative evaluation on the KITTI dataset, where blue lines correspond to our estimated trajectories and red ones are the ground truth.

TABLE 7
Quantitative evaluation (metric: ATE(%)↓) on KITTI Dataset. **Bold** fonts denote the first place.

Method	KITTI sequence ID											
	00	01	02	03	04	05	06	07	08	09	10	Mean
LOAM	0.78	1.43	0.92	0.86	0.71	0.57	0.65	0.63	1.12	0.77	0.79	0.84
IMLS-SLAM	0.50	0.82	**0.53**	0.68	0.33	0.32	0.33	0.33	**0.80**	0.55	0.53	0.52
MC2SLAM	0.51	0.79	0.54	0.65	0.44	0.27	0.31	0.34	0.84	0.46	0.52	0.52
SUMA++	0.64	1.60	1.00	0.67	0.37	0.40	0.46	0.34	1.10	0.47	0.66	0.70
LO-Net	0.78	1.42	1.01	0.73	0.56	0.62	0.55	0.56	1.08	0.77	0.92	0.83
MULLS	0.51	0.62	0.54	**0.61**	0.35	0.28	0.24	**0.29**	**0.80**	0.48	0.63	0.49
Ours	**0.49**	**0.61**	0.55	0.64	**0.28**	**0.23**	**0.19**	0.33	0.82	**0.41**	**0.48**	**0.46**

the feature matching stage since it provides accurate motion information. In this thread, every K consecutive laser scans are regarded as a group. A scan-to-submap registration is performed to merge laser scans into submaps and estimate the relative transformation T_{S2SM} between consecutive submaps. The transformation T_{S2SM} is then passed to the thread 3 to segment overlapping regions. Finally, a submap-to-map registration is performed to recover the final poses of laser scans and the 3D scene map. The scan-to-scan and scan-to-submap registrations are based on our QGORE algorithm, which takes 3D feature correspondences as inputs. The submap-to-map registration takes T_{S2SM} as the initialization and uses ICP for fine registration.

The proposed odometry is evaluated on the KITTI training dataset[2], which contains 11 laser sequences with a total of 23,201 scans captured by a Velodyne HDL-64E LiDAR. KITTI covers three typical scenarios, including country scenery, unban scenery, and highway. Each sequence is associated with a ground truth pose file measured by a high precision GPS-INS equipment. The same as in [88], [89], we use the average translation error (ATE) as the evaluation

2. http://www.cvlibs.net/datasets/kitti/eval_odometry.php

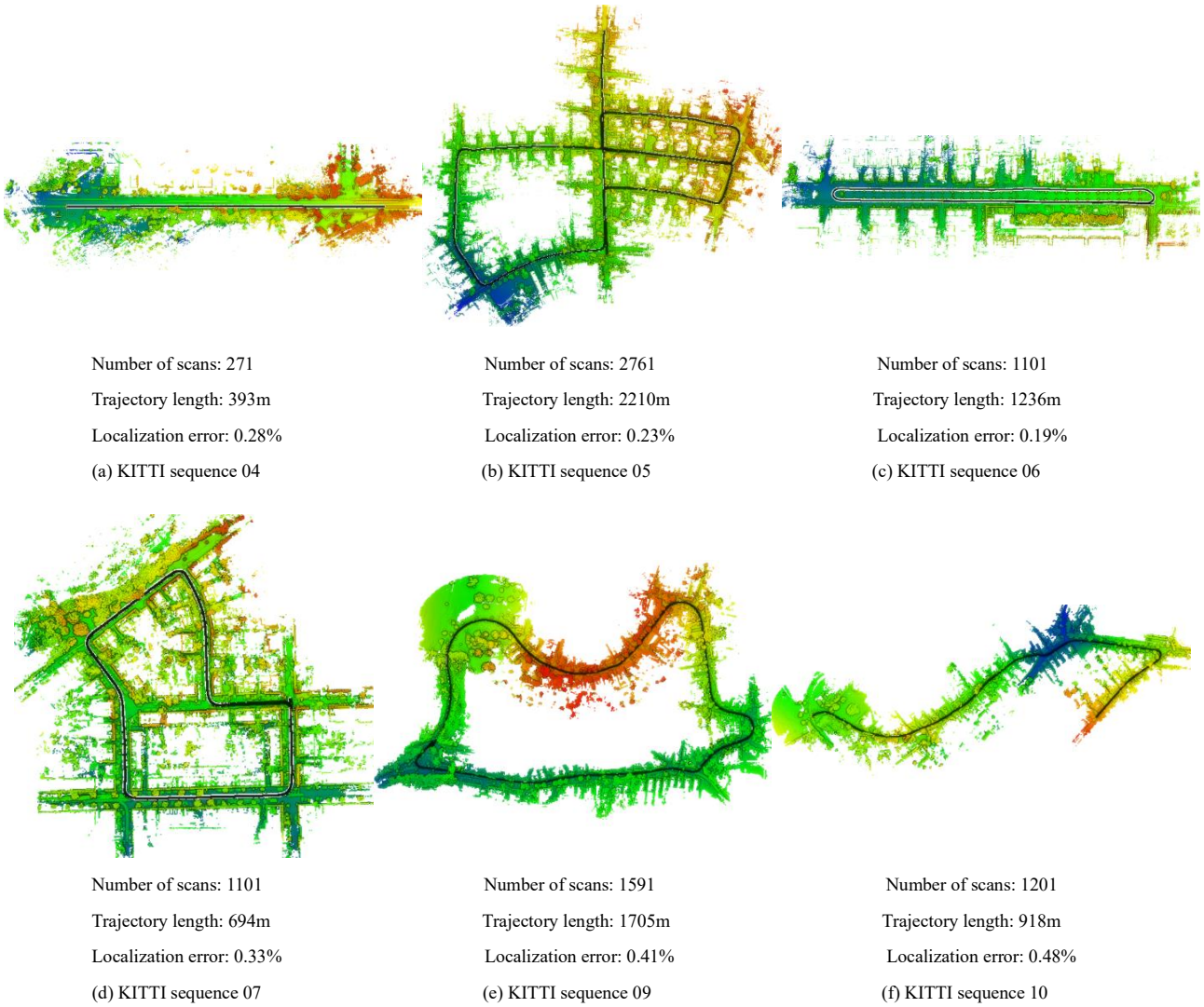


Fig. 9. Our reconstructed 3D scene maps on the KITTI dataset. The color of point clouds is rendered by height. The black line inside each subfigure is our estimated trajectory of the vehicle. Sequence 04, 06, and 09 have been rotated for better visualization.

metric. Our method is compared with six state-of-the-art laser odometry methods, i.e., lidar odometry and mapping (LOAM) [88], implicit moving least squares SLAM (IMLS-SLAM) [89], motion compensation SLAM (MC2SLAM) [90], semantic surfel-mapping (SUMA++) [91], lidar odometry net (LO-Net) [92], and multi-metric linear least square (MULLS) [93]. Figure 8 provides our estimated trajectories. As shown, our method obtains globally consistent results, which are very close to the ground truth. The quantitative results are reported in Table 7, where the results of LOAM, IMLS-SLAM, MC2SLAM, SUMA++, LO-Net, and MULLS are obtained from their original papers. Our method ranks the best on 7 sequences, the second on 3 sequences, and the fourth on one sequence. We achieve the best ATE on these 11 sequences, i.e., 0.46%. Figure 9 gives our reconstructed 3D scene maps. The results are impressive and there is no obvious ghost phenomenon. KITTI also provides a test dataset for online evaluations, which contains 11 laser sequences without open ground truth. In the leaderboard, our proposed method named FBLO currently ranks the sixth among all pure laser SLAM algorithms. It achieves an ATE accuracy of 0.62% and an average rotation error (ARE) of 0.0014 °/m.

4.5 Limitations

The limitations of our method are as follows:

- LRF estimation is relatively sensitive to sparse point density. Hence, if point sets are too sparse, our one-point RANSAC algorithm may not work well.
- Although QGORE reduces the efficiency from exponential to quadratic, it is still far from real-time operation and difficult to be applied in SLAM.
- QGORE is not suitable for photogrammetric point cloud pairs, which suffer from scale changes.

5 Conclusion

In this paper, we propose the first quadratic-time guaranteed outlier removal (QGORE) method for point cloud registration. QGORE introduces the ”rotation correspondence” to establish a one-point RANSAC for lower bound

estimation and proposes the geometric consistency voting for tight upper bound seeking. We also present a l_p-like robust estimator for transformation fitting and provide a QGORE+Ada_l_p pipeline. Extensive experiments demonstrate that our QGORE (QGORE+Ada_l_p) is superior to state-of-the-arts: compared with RANSAC-type methods, QGORE+Ada_l_p achieves better robustness and registration accuracy; compared with GORE, QGORE decreases the computational complexity from exponential to quadratic; compared with our previous work [3], QGORE can guarantee the optimality of the solution.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 42030102 and 42271444, and the Science and Technology Major Project of Hubei Province under Grant 2021AAA010.

REFERENCES

[1] Á. P. Bustos and T.-J. Chin, "Guaranteed outlier removal for point cloud registration with correspondences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 2868–2882, 2017.

[2] X. Huang, G. Mei, J. Zhang, and R. Abbas, "A comprehensive survey on point cloud registration," *arXiv preprint arXiv:2103.02690*, 2021.

[3] J. Li, "A practical o(n2) outlier removal method for correspondence-based point cloud registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 3926–3939, 2022.

[4] P. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[5] S. Bouaziz, A. Tagliasacchi, and M. Pauly, "Sparse iterative closest point," in *Computer graphics forum*, vol. 32, no. 5. Wiley Online Library, 2013, pp. 113–123.

[6] S. Rusinkiewicz, "A symmetric objective function for icp," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, pp. 1–7, 2019.

[7] J. Zhang, Y. Yao, and B. Deng, "Fast and robust iterative closest point," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[8] J. Li, Q. Hu, Y. Zhang, and M. Ai, "Robust symmetric iterative closest point," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 185, pp. 219–231, 2022.

[9] Q.-Y. Zhou, J. Park, and V. Koltun, "Fast global registration," in *European conference on computer vision*. Springer, 2016, pp. 766–782.

[10] H. Yang, J. Shi, and L. Carlone, "Teaser: Fast and certifiable point cloud registration," *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 314–333, 2020.

[11] J. Li, Q. Hu, and M. Ai, "Gesac: Robust graph enhanced sample consensus for point cloud registration," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 167, pp. 363–374, 2020.

[12] C. Choy, W. Dong, and V. Koltun, "Deep global registration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2514–2523.

[13] J. Li, Q. Hu, and M. Ai, "Point cloud registration based on one-point ransac and scale-annealing biweight estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9716–9729, 2021.

[14] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser, "3dmatch: Learning local geometric descriptors from rgb-d reconstructions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1802–1811.

[15] Z. Gojcic, C. Zhou, J. D. Wegner, and A. Wieser, "The perfect match: 3d point cloud matching with smoothed densities," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5545–5554.

[16] S. Ao, Q. Hu, B. Yang, A. Markham, and Y. 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*, 2021, pp. 11 753–11 762.

[17] H. Wang, Y. Liu, Z. Dong, W. Wang, and B. Yang, "You only hypothesize once: Point cloud registration with rotation-equivariant descriptors," *arXiv preprint arXiv:2109.00182*, 2021.

[18] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[19] T.-J. Chin and D. Suter, "The maximum consensus problem: recent algorithmic advances," *Synthesis Lectures on Computer Vision*, vol. 7, no. 2, pp. 1–194, 2017.

[20] R. Horst and H. Tuy, *Global optimization: Deterministic approaches*. Springer Science & Business Media, 2013.

[21] Y. Zhong, "Intrinsic shape signatures: A shape descriptor for 3d object recognition," in *IEEE International Conference on Computer Vision Workshops*, 2009, pp. 689–696.

[22] S. Suwajanakorn, N. Snavely, J. J. Tompson, and M. Norouzi, "Discovery of latent 3d keypoints via end-to-end geometric reasoning," in *Advances in Neural Information Processing Systems*, 2018, pp. 2059–2070.

[23] J. Li and G. H. Lee, "Usip: Unsupervised stable interest point detection from 3d point clouds," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 361–370.

[24] X. Bai, Z. Luo, L. Zhou, H. Fu, L. Quan, and C.-L. 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*, 2020, pp. 6359–6367.

[25] F. Lu, G. Chen, Y. Liu, Z. Qu, and A. Knoll, "Rskdd-net: Random sample-based keypoint detector and descriptor," *Advances in Neural Information Processing Systems*, vol. 33, pp. 21 297–21 308, 2020.

[26] Z. Luo, W. Xue, J. Chae, and G. Fu, "Skp: Semantic 3d keypoint detection for category-level robotic manipulation," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5437–5444, 2022.

[27] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (fpfh) for 3d registration," in *IEEE International Conference on Robotics and Automation*, 2009, pp. 3212–3217.

[28] Y. Guo, F. Sohel, M. Bennamoun, M. Lu, and J. Wan, "Rotational projection statistics for 3d local surface description and object recognition," *International journal of computer vision*, vol. 105, no. 1, pp. 63–86, 2013.

[29] C. Choy, J. Park, and V. Koltun, "Fully convolutional geometric features," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 8958–8966.

[30] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[31] P. H. Torr, S. J. Nasuto, and J. M. Bishop, "Napsac: High noise, high dimensional robust estimation-its in the bag," in *British Machine Vision Conference (BMVC)*, 2002.

[32] O. Chum and J. Matas, "Matching with prosac-progressive sample consensus," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1. IEEE, 2005, pp. 220–226.

[33] D. Barath, M. Ivashechkin, and J. Matas, "Progressive napsac: sampling from gradually growing neighborhoods," *arXiv preprint arXiv:1906.02295*, 2019.

[34] O. Chum, J. Matas, and J. Kittler, "Locally optimized ransac," in *Joint Pattern Recognition Symposium*. Springer, 2003, pp. 236–243.

[35] K. Lebeda, J. Matas, and O. Chum, "Fixing the locally optimized ransac–full experimental evaluation," in *British Machine Vision Conference*. Citeseer, 2012, pp. 1–11.

[36] D. Barath and J. Matas, "Graph-cut ransac," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6733–6741.

[37] J. Matas and O. Chum, "Randomized ransac with td, d test," *Image and vision computing*, vol. 22, no. 10, pp. 837–842, 2004.

[38] J. Matas and O. Chum, "Randomized ransac with sequential probability ratio test," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 2. IEEE, 2005, pp. 1727–1732.

[39] O. Chum and J. Matas, "Optimal randomized ransac," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 8, pp. 1472–1482, 2008.

[40] L. Moisan, P. Moulon, and P. Monasse, "Automatic homographic registration of a pair of images, with a contrario elimination of outliers," *Image Processing On Line*, vol. 2, pp. 56–73, 2012.

[41] D. Barath, J. Matas, and J. Noskova, "Magsac: marginalizing sample consensus," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10 197–10 205.

[42] D. Barath, J. Noskova, and J. Matas, "Marginalizing sample consensus," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[43] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J.-M. Frahm, "Usac: A universal framework for random sample consensus," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 2022–2038, 2012.

[44] M. Ivashechkin, D. Barath, and J. Matas, "Usacv20: robust essential, fundamental and homography matrix estimation," *arXiv preprint arXiv:2104.05044*, 2021.

[45] M. Ivashechkin, D. Barath, and J. Matas, "Vsac: Efficient and accurate estimator for h and f," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 15 243–15 252.

[46] J. Li, Q. Hu, and M. Ai, "Rift: Multi-modal image matching based on radiation-variation insensitive feature transform," *IEEE Transactions on Image Processing*, vol. 29, pp. 3296–3310, 2020.

[47] J. Li, W. Xu, P. Shi, Y. Zhang, and Q. Hu, "Lnift: Locally normalized image for rotation invariant multimodal feature matching," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022, doi:10.1109/TGRS.2022.3165940.

[48] P. J. Huber, *Robust statistics*. Springer, 1981.

[49] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *International journal of computer vision*, vol. 19, no. 1, pp. 57–91, 1996.

[50] H. Yang, P. Antonante, V. Tzoumas, and L. Carlone, "Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1127–1134, 2020.

[51] J. Li, Q. Hu, and M. Ai, "Robust geometric model estimation based on scaled welsch q-norm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 8, pp. 5908–5921, 2020.

[52] Y. Liu, C. Wang, Z. Song, and M. Wang, "Efficient global point cloud registration by matching rotation invariant features through translation search," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 448–463.

[53] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 1–4.

[54] R. Schnabel and R. Klein, "Octree-based point-cloud compression." *Spbg*, vol. 6, pp. 111–120, 2006.

[55] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image and Vision Computing*, vol. 10, pp. 145–155, 1992.

[56] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in *Robotics: Science and Systems*, vol. 2, no. 4. Seattle, WA, 2009, p. 435.

[57] D. Chetverikov, D. Stepanov, and P. Krsek, "Robust euclidean alignment of 3d point sets: the trimmed iterative closest point algorithm," *Image and Vision Computing*, vol. 23, no. 3, pp. 299–309, 2005.

[58] L. Maier-Hein, A. M. Franz, T. R. Dos Santos, M. Schmidt, M. Fangerau, H.-P. Meinzer, and J. M. Fitzpatrick, "Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1520–1532, 2011.

[59] S. Rusinkiewicz and M. Levoy, "Efficient variants of the icp algorithm," in *Proceedings third international conference on 3-D digital imaging and modeling*. IEEE, 2001, pp. 145–152.

[60] A. L. Pavlov, G. W. Ovchinnikov, D. Y. Derbyshev, D. Tsetserukou, and I. V. Oseledets, "Aa-icp: Iterative closest point with anderson acceleration," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 3407–3412.

[61] J. Yang, H. Li, D. Campbell, and Y. Jia, "Go-icp: A globally optimal solution to 3d icp point-set registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 11, pp. 2241–2254, 2015.

[62] P. Biber and W. Straßer, "The normal distributions transform: A new approach to laser scan matching," in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, vol. 3. IEEE, 2003, pp. 2743–2748.

[63] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[64] O. Hirose, "A bayesian formulation of coherent point drift," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 7, pp. 2269–2286, 2020.

[65] B. Jian and B. C. Vemuri, "Robust point set registration using gaussian mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1633–1645, 2010.

[66] J. Fan, J. Yang, D. Ai, L. Xia, Y. Zhao, X. Gao, and Y. Wang, "Convex hull indexed gaussian mixture model (ch-gmm) for 3d point set registration," *Pattern Recognition*, vol. 59, pp. 126–141, 2016.

[67] M. Magnusson, A. Nuchter, C. Lorken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3d registration reliability and speed-a comparison of icp and ndt," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 3907–3912.

[68] H. Deng, T. Birdal, and S. Ilic, "3d local features for direct pairwise registration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3244–3253.

[69] W. Lu, G. Wan, Y. Zhou, X. Fu, P. Yuan, and S. Song, "Deepvcp: An end-to-end deep neural network for point cloud registration," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 12–21.

[70] G. D. Pais, S. Ramalingam, V. M. Govindu, J. C. Nascimento, R. Chellappa, and P. Miraldo, "3dregnet: A deep neural network for 3d point registration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 7193–7203.

[71] Y. Aoki, H. Goforth, R. A. Srivatsan, and S. Lucey, "Pointnetlk: Robust & efficient point cloud registration using pointnet," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7163–7172.

[72] W. Yuan, B. Eckart, K. Kim, V. Jampani, D. Fox, and J. Kautz, "Deepgmr: Learning latent gaussian mixture models for registration," in *European conference on computer vision*. Springer, 2020, pp. 733–750.

[73] X. Bai, Z. Luo, L. Zhou, H. Chen, L. Li, Z. Hu, H. Fu, and C.-L. Tai, "Pointdsc: Robust point cloud registration using deep spatial consistency," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 15 859–15 869.

[74] F. Lu, G. Chen, Y. Liu, L. Zhang, S. Qu, S. Liu, and R. Gu, "Hregnet: A hierarchical network for large-scale outdoor lidar point cloud registration," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 16 014–16 023.

[75] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Communications in Statistics - Theory and Methods*, vol. 6, no. 9, pp. 813–827, 1977.

[76] C. Raposo and J. P. Barreto, "Theory and practice of structure-from-motion using affine correspondences," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 5470–5478.

[77] D. Barath and L. Hajder, "Efficient recovery of essential matrix from two affine correspondences," *IEEE Transactions on Image Processing*, vol. 27, no. 11, pp. 5328–5337, 2018.

[78] B. Guan, J. Zhao, Z. Li, F. Sun, and F. Fraundorfer, "Minimal solutions for relative pose with a single affine correspondence," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 1929–1938.

[79] D. Barath, M. Polic, W. Förstner, T. Sattler, T. Pajdla, and Z. Kukelova, "Making affine correspondences work in camera geometry computation," in *European Conference on Computer Vision*. Springer, 2020, pp. 723–740.

[80] H. Chen and B. Bhanu, "3d free-form object recognition in range images using local surface patches," *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1252–1262, 2007.

[81] J. Yang, Y. Xiao, Z. Cao, and W. Yang, "Ranking 3d feature correspondences via consistency voting," *Pattern Recognition Letters*, vol. 117, pp. 1–8, 2019.

[82] J. Yang, K. Xian, P. Wang, and Y. Zhang, "A performance evaluation of correspondence grouping methods for 3d rigid data matching," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 6, pp. 1859–1874, 2019.

[83] A. Parra Bustos, T.-J. Chin, and D. Suter, "Fast rotation search with stereographic projections for 3d registration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 3930–3937.

[84] W. J. Rey, *Introduction to robust and quasi-robust statistical methods*. Springer Science & Business Media, 2012.

[85] J. T. Barron, "A general and adaptive robust loss function," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4331–4339.
[86] D. Barath, J. Noskova, M. Ivashechkin, and J. Matas, "Magsac++, a fast, reliable and accurate robust estimator," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 1304–1312.
[87] Z. Cai, T.-J. Chin, A. P. Bustos, and K. Schindler, "Practical optimal registration of terrestrial lidar scan pairs," *ISPRS journal of photogrammetry and remote sensing*, vol. 147, pp. 118–131, 2019.
[88] J. Zhang and S. Singh, "Loam: Lidar odometry and mapping in real-time." in *Robotics: Science and Systems*, vol. 2, no. 9, 2014.
[89] J.-E. Deschaud, "Imls-slam: scan-to-model matching based on 3d data," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2480–2485.
[90] F. Neuhaus, T. Koß, R. Kohnen, and D. Paulus, "Mc2slam: Real-time inertial lidar odometry using two-scan motion compensation," in *German Conference on Pattern Recognition*. Springer, 2018, pp. 60–72.
[91] X. Chen, A. Milioto, E. Palazzolo, P. Giguere, J. Behley, and C. Stachniss, "Suma++: Efficient lidar-based semantic slam," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 4530–4537.
[92] Q. Li, S. Chen, C. Wang, X. Li, C. Wen, M. Cheng, and J. Li, "Lo-net: Deep real-time lidar odometry," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8473–8482.
[93] Y. Pan, P. Xiao, Y. He, Z. Shao, and Z. Li, "Mulls: Versatile lidar slam via multi-metric linear least square," *arXiv preprint arXiv:2102.03771*, 2021.

Jiayuan Li received the B.Eng., M.Eng., and Ph.D. degrees from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China. He is currently an Associate Researcher with Wuhan University. He has authored more than 40 peer-reviewed articles in international journals. His research is mainly focused on SLAM, image matching, and point cloud registration. He was awarded the Best Youth Author Award by ISPRS in 2021 and the Talbert Abrams Award by ASPRS in 2018.

Pengcheng Shi received the B.S. degree in remote sensing science and technology from Liaoning Technical University, China, in 2018, and the M.S. degree in surveying and mapping engineering from Tongji University, China, in 2021. Currently, he is pursuing the Ph.D. degree with the School of Computer Science, Wuhan University. His research interests include simultaneous localization and mapping (SLAM) and point cloud registration.

Qingwu Hu received the B.Eng. and M.Eng. degrees in photogrammetry and remote sensing from the Wuhan Technical University of Surveying and Mapping, Wuhan, China, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, in 2007. He has authored more than 60 peerreviewed articles in international journals. His research interests include methods, techniques and applications of remote sensing, GIS and GPS integration, and photogrammetry.

Yongjun Zhang received the B.S. degree in geodesy, the M.S. degree in geodesy and surveying engineering, and the Ph.D. degree in geodesy and photography from Wuhan University, Wuhan, China, in 1997, 2000, and 2002, respectively. He is currently the Dean of the School of Remote Sensing and Information Engineering, Wuhan University. He has published more than 180 research articles and one book. His research interests include aerospace and low-attitude photogrammetry, image matching, combined block adjustment with multisource data sets, object information extraction and modeling with artificial intelligence, integration of LiDAR point clouds and images, and 3D city model reconstruction. He is the Co-Editor-in-Chief of The Photogrammetric Record.