ISPRS Journal of Photogrammetry and Remote Sensing 160 (2020) 244–259

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Robust point cloud registration based on topological graph and Cauchy weighted l_q-norm

Jiayuan Li[a,*], Pengcheng Zhao[a,*], Qingwu Hu[a], Mingyao Ai[a,b]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
[b] State Key Laboratory of Information Engineering, Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Keywords:
Point Cloud Registration (PCR)
Coarse-to-fine registration
Feature correspondence
Iterative Closest Point (ICP)
Robust estimation

ABSTRACT

Point Cloud Registration (PCR) is a fundamental and important issue in photogrammetry and computer vision. Its goal is to find rigid transformations that register multiple 3D point sets. This paper proposes a robust and efficient PCR method based on topological graph and Cauchy weighted l_q-norm. Our method does not require initializations and is highly robust to outliers and partial overlaps. It contains two major steps: (1) correspondence-based coarse registration, which is called Weighted l_q Coarse Registration (WCR). In the Wl_qCR method, we represent feature point sets as topological graphs and transform the point matching problem to an edge matching problem. We build a mathematical model for edge correspondence maximization. We also present an edge voting strategy to distinguish potential correct matches from mismatches. Then, we define a concept called edge vector, which has a property that it is invariant to translations. Based on this property, six Degrees of Freedoms (DoF) PCR problem can be simplified into two three DoF subproblems, i.e., rotation estimation and translation estimation. (2) fine registration based on Weighted l_q Iterative Closest Point (Wl_qICP). We propose a new ICP method called Wl_qICP, which is much more robust to partial overlaps compared with traditional ICP. In both rotation estimation and Wl_qICP, we use a new Cauchy weighted l_q-norm ($0 < q < 1$) instead of l_2-norm for object function construction, which has a high degree of robustness. Extensive experiments on both simulated and real data demonstrate the power of the proposed method, i.e., our method is more robust (is able to tolerate up to 99% of outliers) and much faster than compared state-of-the-art methods (Wl_qCR is almost two orders of magnitude faster than RANdom SAmple Consesus (RANSAC) and its variants under 95% of outliers). The source code will be made publicly available in http://www.escience.cn/people/lijiayuan/index.html.

1. Introduction

Point Cloud Registration (PCR) is a fundamental and important issue in photogrammetry, robotics, and computer vision, which has a variety of applications such as Simultaneous Localization and Mapping (SLAM) (Nüchter et al., 2007), 3D modelling (Blais and Levine, 1995; Choi et al., 2015, 2016b), cultural heritage management (Montuori et al., 2014), and landslide monitoring (Ventura et al., 2011), to name a few. Three dimensional laser scanning becomes more and more popular in contemporary surveying and mapping. However, each 3D scan only covers a part of an object scene due to occlusions, it is necessary to capture multiple 3D scans from different viewpoints so that the whole scene can be fully covered. The goal of PCR is to register multiple individual 3D scans with their own coordinate systems into a common reference system, which is usually achieved by sequential pairwise registration.

It is not arguable that Iterative Closest Point (ICP) (Besl and McKay, 1992) algorithm is the de facto standard for the PCR problem, which has been applied in many industrial solutions, such as the on-board software of Zoller + Frohlich[1] scanners and the offline processing software of RIEGL[2] scanners. The basic idea of ICP is to alternate between closest point matching and rigid transformation estimation. However, ICP suffers from a drawback that it requires sufficiently good initial parameters for accurate registration, since ICP finds only a local minimum.

To address this issue, many solutions use external sensors to provide additional information. For example, a RIEGL airborne laser scanner is equipped with a GNSS/INS (Global Navigation Satellite System/Inertial Navigation System) sensor system; point clouds captured by terrestrial laser scanners (for instance, RIEGL VZ400 scanner) usually rely on

* Corresponding authors.
E-mail address: ljy_whu_2012@whu.edu.cn (J. Li).
[1] https://www.zf-laser.com
[2] http://www.riegl.com/products/

https://doi.org/10.1016/j.isprsjprs.2019.12.008
Received 10 September 2019; Received in revised form 10 November 2019; Accepted 9 December 2019
Available online 27 December 2019

artificial markers for coarse registration; VLOAM (visual-LiDAR odometry and mapping) (Zhang and Singh, 2015) uses a panoramic camera and a low-cost MEMS Inertial Measurement Unit (IMU) for Visual-Inertial Odometry (VIO), which enables PCR in GNSS-denied environments. However, GNSS/INS sensors are expensive; artificial marker layout is labor-intensive; methods like VLOAM highly rely on the VIO algorithm.

More well-designed PCR techniques have been proposed to alleviate the dependence on good initializations. These methods usually adapt a coarse-to-fine strategy. First, an initial transformation is obtained by a coarse registration method such as 4-Points Congruent Sets (4PCS) (Aiger et al., 2008), Keypoint-based 4PCS (K4PCS) (Theiler et al., 2014), and Hierarchical Merging based Multiview Registration (HMMR) (Dong et al., 2018), etc. Then, a fine registration approach (e.g., ICP-family) is applied to further refine the initial transformation. Coarse-to-fine technique is very practical in realistic applications. However, 4PCS and K4PCS are sensitive to the proportion of overlapping regions; feature correspondence based coarse registration methods may fail if outlier rates are extremely high. In addition, current ICP-type fine registration methods are not robust enough, which are also sensitive to the proportion of overlapping points.

In this paper, we aim to design a coarse-to-fine registration method that: (1) does not require initialization; (2) has polynomial running time with respect to number of points; (3) is able to deal with cases with extremely high outlier rates (e.g., cases with 99% of outliers). Correspondence-based coarse registration is widely used in practice. However, 3D feature matching is much less accurate compared to its 2D counterparts. As pointed out by (Bustos and Chin, 2017), it is very common to have more than 95% of outliers in initial feature matches. (4) is insensitive to the proportion of overlapping points. To our best knowledge, there is rarely literature that simultaneously addresses all these issues. For instance, ICP requires initialization; Globally Optimal ICP (Go-ICP) (Yang et al., 2015) has exponential running time with respect to number of points; Fast Global Registration (FGR) (Zhou et al., 2016) cannot tolerate extremely high outlier rates; Gaussian Mixture Models (GMMs) (Jian and Vemuri, 2010) is sensitive to the proportion of overlapping points. More specifically, we propose an accurate, robust, and efficient PCR method based on topological graph and Cauchy-weighted l_q-norm cost. It contains two major steps: (1) coarse registration. In this stage, we use topological graphs to represent feature point sets and transform the point matching problem to an edge matching problem. Mismatches can be easily distinguished from potential correct matches since edge correspondences hold an equal-length constraint. Then, we define a concept called edge vector, which has a property that it is invariant to translations. We simplify the six Degree of Freedoms (DoF) PCR problem into two three DoF subproblems, i.e., rotation estimation and translation estimation, based on edge vector. (2) fine registration. We propose a new ICP variant called Weighted l_q ICP (Wl_qICP), which is much more robust to partial overlaps compared with current ICP and its variants. In both rotation estimation and Wl_qICP, we use a new Cauchy-weighted l_q-norm ($0 < q < 1$) for object function construction, which is very robust. The proposed method is validated on both simulated and real datasets. Extensive experimental results show that our method can tolerate more than 99% of outliers, which outperforms the compared state-of-the-arts. The main contributions of this paper are as follows:

- We propose a new Cauchy-weighted l_q-norm ($0 < q < 1$) cost. Compared with traditional l_2-norm cost, the new cost has a high degree of robustness. We use the Augmented Lagrange Function (ALF) and Alternating Direction Method of Multipliers (ADMM) (Boyd et al., 2011) to solve this non-smooth and non-convex object function.
- We transform the point matching problem to an edge matching problem based on topological graph representation. We build a mathematical model for edge correspondence maximization. We then present an edge voting strategy. These strategies can largely improve the inlier rate of initial feature correspondences. Hence, the proposed Weighted l_q Coarse Registration (Wl_qCR) method can tolerate extremely high outlier rates.
- We decompose the six DoF PCR problem into two three DoF subproblems based on edge vector, which largely reduces the parameter space. We also propose a robust rotation estimation algorithm based on weighted l_q-norm.
- We develop a fine registration algorithm called Wl_qICP based on our new cost.

2. Related work

In this section, we will review both coarse registration methods and fine registration methods since the proposed method builds on the coarse-to-fine strategy.

2.1. Coarse registration

As summarized in literature (Dong et al., 2018), a large number of coarse registration methods are based on geometric features (e.g., points, lines, planes, or semantic features). For urban man-made scenes, line and plane features are usually adopted. For example, Stamos and Leordean (2003) presented a line registration method, which used plane-intersection-line correspondences to calculate the transformation parameters. Raposo et al. (2013) developed a plane-based odometry method for indoor mapping, which used plane features as primitives for motion estimation. Similarly, Xiao et al. (2013) proposed a plane-to-plane registration algorithm, which first segmented point clouds into plane primitives, then performed an ICP-like method for transformation estimation. Rabbani et al. (2007) utilized multiple geometric primitives (e.g., planes, cylindrical and spherical surfaces, tori) to register scans captured in industrial environments. Yang et al. (2016) utilized both semantic information and geometrical constraints to match line features derived from pole-like objects and vertical planes. Although these high-level features are very suitable for urban scenes, they work poor in natural scenes such as forest and mountain.

Point feature-based methods have no dependence on scenes and are more popular in realistic applications. These methods typically consist of two major stages, i.e., feature matching and robust transformation estimation. In the first stage, keypoints are first extracted from raw point clouds based on feature detectors. (For example, Local Surface Patches (LSP) (Chen and Bhanu, 2007) detects features with large shape variations measured by shape index. KeyPoint Quality (KPQ) (Mian et al., 2010) and Intrinsic Shape Signatures (ISS) (Zhong, 2009) use eigenvalues of the support region scatter matrix to measure the distinctiveness of points. ISS accepts points whose ratio between two successive eigenvalues is below a given threshold. KPQ uses the ratio between the maximum lengths along the first two principal axes to reject nondistinctive points. MeshDoG (Zaharescu et al., 2009) builds the scale-space on the 3D mesh and applies the Difference-of-Gaussians (DoG) operator to detect features. KeypointNet (Suwajanakorn et al., 2018) presents an end-to-end framework to learn geometrically and semantically consistent keypoints.) Then, each keypoint is described by local feature descriptors, which encode spatial geometric information of a local patch to a compact feature vector so that keypoints can be distinguished from each other. (e.g., spin-image (Johnson and Hebert, 1999) is a data level shape descriptor, which computes 2D histograms of points falling in a cylindrical volume. 3D Shape Context (3DSC) (Frome et al., 2004) describes a feature point based on the 3D histograms of support points. Fast Point Feature Histogram (FPFH) (Rusu et al., 2009) accumulates in a 3D histogram of the three angles between points. Rotational Projection Statistics (RoPS) (Guo et al., 2013) projects points within the support region onto 2D planes and computes their statistical distributions. Signature of Histogram of OrienTations (SHOT) (Salti et al., 2014) introduces a unique and unambiguous local

reference frame based on eigenvalue decomposition. 3DMatch (Zeng et al., 2017), PPFNet (Deng et al., 2018), and 3DSmoothNet (Gojcic et al., 2019) are learned 3D descriptors. 3DMatch learns a local volumetric descriptor based on truncated signed distance field coding; PPFNet presents a N-tuple loss to inject global context into the local descriptor; 3DSmoothNet presents a voxelized smoothed density value representation and develops a Siamese architecture.) Finally, corresponding relationship between two sets of features is established based on feature vector distances and various matching strategies (e.g., nearest neighbor distance ratio (Lowe, 2004) and chi-square test (Zhong, 2009)). In the second stage, a hypothesize-and-verify technique such as RANdom SAmple Consesus (RANSAC) (Fischler and Bolles, 1981) and its variants (e.g., Maximum Likelihood Estimation SAmple Consensus (MLESAC) (Torr and Zisserman, 2000) is a probability-based variant, which maximizes the likelihood instead of the consensus size. BaySAC (Botterill et al., 2009) uses a deterministic strategy that selects the hypothesis set with the highest inlier probability to reduce the number of trails. Optimized BaySAC (Kang et al., 2013) proposes a novel statistical test to estimate the prior probability, which is predictably model-free. Fixed Locally Optimized RANSAC (FLO-RANSAC) (Lebeda et al., 2012) adds a local optimization stage into RANSAC to improve the model accuracy. Universal RANSAC (USAC) (Raguram et al., 2012) incorporates several important practical and computational tricks into a common framework. Normalized Barycentric Coordinate System (NBCS) (Li et al., 2017a) adapts a new coordinate system into a hypothesize-and-verify framework.), a robust optimization algorithm (e.g., FGR (Zhou et al., 2016) introduces a scaled Geman-McClure loss and uses the Black and Rangarajan duality between robust estimation and line processes. L_q estimator (Li et al., 2016a) uses a sparse-inducing norm for robust estimation. Yang and Carlone (Yang and Carlone, 2019) proposed a robust cost based on Truncated Least Squares (TLS). Iterative biconvex optimization (Cai et al., 2018) conducts a deterministic search to refine a given initial solution.), or a filtering method (e.g., guaranteed outlier removal (Bustos and Chin, 2017) uses mixed integer linear programming to remove provably true outliers.) is applied to simultaneously remove outliers and estimate a rigid transformation.

Apart from 3D feature-based methods, 2D features can also be used for the PCR problem. Laser scanners can not only measure distance, but also get the amount of light returning to the scanners. The reflectance information can be projected to 2D gray scale images. First, 2D features are detected and matched from the intensity images; then, these features are projected into 3D space to obtain 3D-to-2D or 3D-to-3D correspondences; finally, a robust transformation estimation algorithm is performed to register these correspondences. For example, Böhm and Becker (2007) used Scale-Invariant Feature Transform (SIFT) (Lowe, 2004) to match features and applied RANSAC for transformation estimation. Barnea and Filin (2008) used the 3D Euclidean distance between keypoint-pairs as match entities. Kang et al. (2009) projected image feature matches to obtain 3D-to-3D correspondences and used a four-point rigid geometric invariance to remove false matches. Similarly, Weinmann et al. (2011) projected SIFT features inside one image to 3D space and obtained a set of 3D-to-2D correspondences. With the rapid progress of feature matching technology, many image feature matching methods can be applied for such a task (e.g., Speeded Up Robust Features (SURF) (Bay et al., 2008), Radiation-variation Insensitive Feature Transform (RIFT) (Li et al., 2018), locality affine-invariant matching (Li et al., 2017b, 2019)). A comprehensive study can be found in literature (Houshiar et al., 2015). However, these methods highly rely on the quality of the projected intensity images. If the density of point clouds is very sparse such as scans captured by Velodyne LiDARs, these methods may not work.

There exist other point-based methods that do not follow the aforementioned framework. Among them, 4PCS (Aiger et al., 2008) is a representative work, which extracts coplanar 4-point correspondence sets based on affine-invariant intersection ratios and performs registration in a hypothesize-and-verify framework. 4PCS has many variants. Super 4PCS (S4PCS) (Mellado et al., 2014) introduces a smart indexing data organization to largely reduce the computational complexity of 4PCS. K4PCS (Theiler et al., 2014) uses the extracted 3D keypoints as the input. Geodesic Distances-based 4PCS (GD-4PCS) (Ge, 2016) extends 4PCS to cope with non-rigid isometric deformations based on geodesic distance. Similar to K4PCS, Semantic-Keypoint-based 4PCS (SK4PCS) (Ge, 2017) extracts semantic keypoints as inputs to a modified 4PCS. Owing to the affine-invariant geometric constraint, these approaches are more robust to noise and uneven point densities. However, they are generally much slower than feature-based methods.

As known, point cloud feature matching is less accurate compared with image feature matching, since point clouds are unorganized and suffer from serious noise, uneven point densities, and texture-less. As a result, feature correspondences matched by local descriptors contain a large number of outliers. Although feature-based methods perform well in most cases, they may fail in some cases with extremely high outlier rates, e.g., cases with 99% of outliers. The reason is that the robust transformation estimation stage cannot tolerate high outlier rates. For example, RANSAC-type methods require a huge number of trails to get an acceptable solution if the outlier rate is higher than 90% (Raguram et al., 2012; Chin and Suter, 2017), which largely decreases their efficiency; M-estimators suffer from a breakdown point of 50% (Rousseeuw and Leroy, 2005).

2.2. Fine registration

Fine registration methods can be roughly categorized into two groups, i.e., local optimal methods and global optimal methods (Yang et al., 2015). Although global optimal methods provide accurate estimations and do not require initializations, their very high computational complexities prevent their usage in realistic applications (Cai et al., 2019). Hence, we focus on local optimal methods in this section, such as ICP-family and probabilistic-based methods.

ICP and its variants: ICP (Besl and McKay, 1992) algorithm can be regarded as a milestone in PCR. As aforementioned, it is a simultaneous correspondence establishment and pose estimation technique. ICP has a number of variants. According to literature (Rusinkiewicz and Levoy, 2001), they can be roughly classified as affecting one of the three subtasks: sampling and matching, error metric, and outlier rejection. (1) sampling and matching. Using subset points instead of the whole point clouds for registration can largely reduce the computational complexity. Some typical strategies include uniform sampling (Turk and Levoy, 1994), random sampling (Masuda et al., 1996), Voxel Grid filtering (Rusu and Cousins, 2011), and octree compression (Schnabel and Klein, 2006). ICP establishes matches based on Euclidean distance nearest neighbor search. Many variants use additional properties or strategies to aid correspondence establishment (e.g., invariant features (Feldmar and Ayache, 1996), surface normal (Serafin and Grisetti, 2015), and Mahalanobis distance (Hansen et al., 2007)). To improve the robustness to partial overlaps, practical ICP-family implementations use a maximum distance threshold to prune the correspondences. (2) error metrics. To improve registration accuracy of the original ICP, many other error metrics are proposed to instead of the point-to-point error distance, such as point-to-line metric in Point-Line ICP (PLICP) (Censi, 2008), point-to-plane metric in (Chen and Medioni, 1992), and plane-to-plane metric in Generalized-ICP (GICP) (Segal et al., 2009). (3) outlier rejection. To cope with this issue, various strategies are used to improve convergence. M-estimators are straightforward, which give large weights to inliers while small weights to outliers (Bergström and Edlund, 2014; Kaneko et al., 2003). Trimmed ICP (TrICP) (Chetverikov et al., 2005) and Anisotropic ICP (A-ICP) (Maier-Hein et al., 2011) use a Least Trimmed Squares (LTS) algorithm for transformation estimation, which only consider a certain number of points with small residuals. However, both M-estimators and LTS have a breakdown point of 0.5, namely, if the overlap between two point clouds is less than 50%, these

methods may not work. Sparse ICP (Bouaziz et al., 2013) uses a sparse cost function instead of the least squares cost. Recently, Yang et al. (Yang et al., 2015) proposed a Go-ICP based on a Branch-and-Bound (BnB) scheme. However, Go-ICP shares the same limitation with global optimal methods. A review of ICP variants can be found in Pomerleau et al. (2013).

Probabilistic-based methods: Normal Distributions Transform (NDT) (Biber and Straßer, 2003) models the distribution of a point set as a density function, which is originally proposed to match 2D laser scans. It first transforms a discrete 2D point set into a piecewise continuous probability density represented by a set of normal distributions. Then, scan matching is defined by maximizing the sum, that the points in another scan score on this density. Takeuchi and Tsubouchi (Takeuchi and Tsubouchi, 2006) extended NDT to 3D. Many extensions have been proposed to improve the calculation of the probability density, such as Segmented Region Growing NDT (SRG-NDT) (Das and Waslander, 2014) and compact 3D-NDT (Stoyanov et al., 2012). GMMs registration (Jian and Vemuri, 2010) uses Gaussian mixture models to represent point clouds and reformulates PCR problem as two Gaussian mixtures alignment problem such that the Kullback-Leibler divergence between these two mixtures is minimized. GMMs-based methods generally employ correlation-based algorithms (e.g., GMMs registration (Jian and Vemuri, 2010) and kernel correlation registration (Tsin and Kanade, 2004)) or Expectation-Maximization (EM) based algorithms (e.g., Coherent Point Drift (CPD) (Myronenko and Song, 2010)) for alignment. As shown in the above-mentioned work, probabilistic-based methods are less sensitive to poor initializations than ICP-family. However, the results of probabilistic-based methods are not as predictable as ICP (Magnusson et al., 2009).

More recently, Cai et al. (2019) proposed a practical optimal method for registration of terrestrial LiDAR scans, which reduces the six DoF PCR problem to a four DoF registration problem based on a highly accurate level compensator. Although this method is fast and accurate, it does not work on other LiDARs that are not equipped with level compensators, such as hand-hold LiDARs, Velodyne LiDARs, and airborne LiDARs.

This paper is organized as follows: Section 3 proposes a correspondence-based coarse registration algorithm (Wl_qCR) and develops a fine registration algorithm based on the proposed cost, called Wl_qICP. Section 4 evaluates the proposed method on both simulated and real data. Finally, a conclusion and discussion section is presented in Section 5.

3. Methodology

This section proposes a robust and efficient PCR method based on topological graph and Cauchy weighted l_q-norm. It is a coarse-to-fine method, which consists of two major stages, i.e., a correspondence-based coarse registration stage (Wl_qCR) and an ICP-like fine registration stage (Wl_qICP).

3.1. Coarse registration based on Wl_qCR

Given a pair of point clouds $(\mathcal{H}_1, \mathcal{H}_2)$, we extract a 3D correspondence set $\mathcal{M} = \{(\boldsymbol{p}_i, \boldsymbol{q}_i) | \boldsymbol{p}_i \in \mathcal{P}; \boldsymbol{q}_i \in \mathcal{Q}; i = 1, 2, ..., n\}$ based on feature detectors and descriptors, where $\boldsymbol{p}_i, \boldsymbol{q}_i \in \mathbb{R}^3$. Note that a correspondence can be also called a match, which is a matching pair. Generally, a correspondence set contains both correct matches (inliers) and false matches (outliers). The goal of point set registration is to estimate an optimal rigid transformation $T \in SE(3)$ that maximizes the number of correct matches in $\mathcal{M}$. The object function is,

$$\underset{T \in SE(3)}{\operatorname{argmax}} \sum_{i=1}^{n} \lfloor C(\boldsymbol{p}_i, T(\boldsymbol{q}_i)) < \varepsilon \rfloor \tag{1}$$

where transformation T is formed by a 3D rotation matrix $\boldsymbol{R}$ and a 3D translation vector $\boldsymbol{t}$; hence, point $\boldsymbol{q}_i$ is transformed to $\tilde{\boldsymbol{q}}_i = \boldsymbol{R}\boldsymbol{q}_i + \boldsymbol{t}$ after applying T;ε is a correct match threshold; and $\lfloor \rfloor$ is an operator that returns 1 if the input predicate is satisfied and 0 otherwise. C is a cost function that measures the similarity between $\boldsymbol{p}_i$ and $T(\boldsymbol{q}_i)$. The most popular cost function is to minimize the sum of l_2-norm distance between the correspondences.

Optimizing this object function faces three major problems: first, 3D correspondence set $\mathcal{M}$ matched by feature descriptors such as FPFH (Rusu et al., 2009) usually contains a large number of outliers. The outlier rate may be higher than 90% due to the differences between point clouds. Such high outlier rates make the optimization very challenging. Second, the rigid transformation T has six DoF. To find an optimal solution of Eq. (1) in a six-dimensional parameter space is difficult. Third, l_2-norm relies on an assumption that observations are free of outliers. Therefore, it is not a robust estimator and is sensitive to outliers, whose solution will be skewed to adjust the residuals. A suitable robust cost function is also very important. To deal with these problems, we first present an edge voting strategy based on topological graphs to improve the inlier rate of $\mathcal{M}$; then, we decompose the six DoF rigid transformation estimation problem into two subproblems (a three DoF rotation problem and a three DoF translation problem) based on graph edge vectors; finally, we introduce a Cauchy weighted l_q-norm cost function instead of l_2-norm for such a task, which has a high degree of robustness.

3.1.1. Topological graph and edge voting

Suppose we have extracted a 3D correspondence set $\mathcal{M}$ from the point cloud pair $(\mathcal{H}_1, \mathcal{H}_2)$. We first construct two attributed graphs $G^{\mathcal{P}}(V^{\mathcal{P}}, E^{\mathcal{P}})$ and $G^{\mathcal{Q}}(V^{\mathcal{Q}}, E^{\mathcal{Q}})$ based on the feature correspondence set $\mathcal{M}$, where $V^{\mathcal{P}}$ ($V^{\mathcal{Q}}$) is a set of nodes and $E^{\mathcal{P}}$ ($E^{\mathcal{Q}}$) is a set of edges. Here, the graph $G^{\mathcal{P}}$ ($G^{\mathcal{Q}}$) is a topological model of point set $\mathcal{P}$ ($\mathcal{Q}$). The nodes $\{v_i^{\mathcal{P}}\}_1^n$ ($\{v_i^{\mathcal{Q}}\}_1^n$) represent point features $\{\boldsymbol{p}_i\}_1^n$ ($\{\boldsymbol{q}_i\}_1^n$) and edges represent pairwise linkages between features (see Fig. 1(a)). Owing to the corresponding relationship between point sets $\mathcal{P}$ and $\mathcal{Q}$, $\{(v_i^{\mathcal{P}} \in V^{\mathcal{P}}, v_i^{\mathcal{Q}} \in V^{\mathcal{Q}})\}_1^n$ is a node correspondence set. It is obviously that if both node correspondences $(v_i^{\mathcal{P}}, v_i^{\mathcal{Q}})$ and $(v_j^{\mathcal{P}}, v_j^{\mathcal{Q}})$ are correct matches, their edge pair $(e^{\mathcal{P}}_{(i-1)\times n+j} \in E^{\mathcal{P}}, e^{\mathcal{Q}}_{(i-1)\times n+j} \in E^{\mathcal{Q}})$ (or can be written as $(e^{\mathcal{P}}_{(i,j)} \in E^{\mathcal{P}}, e^{\mathcal{Q}}_{(i,j)} \in E^{\mathcal{Q}})$) is also a correct match. To make the mathematical deduction simple, we first assume that the correct matches in correspondence set $\mathcal{M}$ are free of noise. Namely, if $(v_i^{\mathcal{P}}, v_i^{\mathcal{Q}})$ is a correct match, i.e., $(\boldsymbol{p}_i, \boldsymbol{q}_i)$ is correct, we have $\boldsymbol{p}_i = \boldsymbol{R}\boldsymbol{q}_i + \boldsymbol{t}$. Hence, if $e^{\mathcal{P}}_{(i,j)}$ and $e^{\mathcal{Q}}_{(i,j)}$ is a correct match, their lengths should be equal, i.e., $\|e^{\mathcal{P}}_{(i,j)}\| = \|e^{\mathcal{Q}}_{(i,j)}\|$, where $\|e^{\mathcal{P}}_{(i,j)}\|$ represents the length of $e^{\mathcal{P}}_{(i,j)}$. If the relationship $\|e^{\mathcal{P}}_{(i,j)}\| = \|e^{\mathcal{Q}}_{(i,j)}\|$ is not hold, there must be at least one false match in $(v_i^{\mathcal{P}}, v_i^{\mathcal{Q}})$ and $(v_j^{\mathcal{P}}, v_j^{\mathcal{Q}})$.

In the above, we assume that there is no noise in the correct matches. This is not true. Correct matches inevitably suffer from noise in real data. If a correspondence node $(v_i^{\mathcal{P}}, v_i^{\mathcal{Q}})$ satisfies $\|\boldsymbol{p}_i - T(\boldsymbol{q}_i)\|_2 < \varepsilon$ (Essentially, a node is equivalent to a feature point), it can be regarded as a correct match. Thus, the equal length constraint is generally not hold. Fortunately, if both $(v_i^{\mathcal{P}}, v_i^{\mathcal{Q}})$ and $(v_j^{\mathcal{P}}, v_j^{\mathcal{Q}})$ are correct matches, the

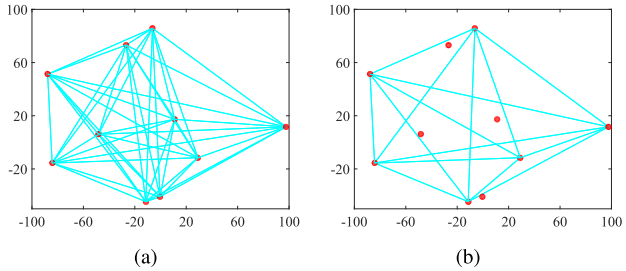


Fig. 1. Point set represented by a topological graph. (a) a complete topological graph, where all points are the nodes of the graph. (b) a compact topological graph after removing potential unreliable edges.

length difference between $\|e^{\mathcal{P}}_{(i,j)}\|$ and $\|e^{\mathcal{Q}}_{(i,j)}\|$ should be small. Based on this observation, we can approximately transform the point matching problem to an edge matching problem, which can be formulated as,

$$\begin{aligned}&\underset{I_E}{\operatorname{argmin}}\, O\left(I_E; E^{\mathcal{P}}, E^{\mathcal{Q}}, \eta\right)\\ &= \underset{I_E}{\operatorname{argmin}}\left[\sum_{i\in I_E}\left(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|\right)^2 + \eta\left(N-K\right)\right]\end{aligned} \tag{2}$$

where $O(I_E; E^{\mathcal{P}}, E^{\mathcal{Q}}, \eta)$ is an object cost function; η is a balance parameter; N is the size of the edge correspondence set; K is the number of inliers; and I_E is a set of indices of inlier edges. In this cost, $\sum_{i\in I_E}(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|)^2$ is a data term, which penalizes edge correspondences with large length differences; the term $(N-K)$ minimizes the number of outliers; and parameter η balances these two terms. Extending Eq. (2) to the whole edge correspondence set, we have,

$$\begin{aligned}&\underset{\boldsymbol{b}}{\operatorname{argmin}}\, O\left(\boldsymbol{b}; E^{\mathcal{P}}, E^{\mathcal{Q}}, \eta\right)\\ &= \underset{\boldsymbol{b}}{\operatorname{argmin}}\left\{\sum_{i=1}^{N} b_i\left(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|\right)^2 + \eta\left(N - \sum_{i=1}^{N} b_i\right)\right\}\\ &= \underset{\boldsymbol{b}}{\operatorname{argmin}}\left\{\sum_{i=1}^{N} b_i\left[\left(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|\right)^2 - \eta\right]\right\}\end{aligned} \tag{3}$$

where $\boldsymbol{b} = \{b_i\}_1^N$ ($b_i \in \{0, 1\}$) is an $N\times 1$ binary vector that assigns a flag for each edge correspondence. Specifically, for a correct edge match, its flag $b_i = 1$; otherwise, $b_i = 0$. If $(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|)^2 > \eta$, it leads to positive cost, which should be penalized; otherwise, if $(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|)^2 \leqslant \eta$, it leads to a non-positive cost, which decreases the total cost and can be regarded as an inlier. Thus, the optimal solution $\boldsymbol{b}$ can be simply obtained by,

$$b_i = \begin{cases} 1 & \left(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|\right)^2 \leqslant \eta \\ 0 & \left(\|e^{\mathcal{P}}_i\| - \|e^{\mathcal{Q}}_i\|\right)^2 > \eta, \end{cases} \quad i = 1, 2, \ldots, N \tag{4}$$

and the inlier edge set is obtained simultaneously,

$$I_E = \{i | b_i = 1, \quad i = 1, 2, \ldots, N\} \tag{5}$$

After obtaining I_E, we can go back to the point matching problem. As mentioned earlier, if both nodes $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $(v^{\mathcal{P}}_j, v^{\mathcal{Q}}_j)$ of an edge correspondence $(e^{\mathcal{P}}_{(i,j)}, e^{\mathcal{Q}}_{(i,j)})$ are correct matches, the edge correspondence is also a correct one. In other words, if $(e^{\mathcal{P}}_{(i,j)}, e^{\mathcal{Q}}_{(i,j)})$ is a correct edge match, $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $(v^{\mathcal{P}}_j, v^{\mathcal{Q}}_j)$ may be correct matches. Hence, we think that $(e^{\mathcal{P}}_{(i,j)}, e^{\mathcal{Q}}_{(i,j)})$ is a support edge pair and votes for $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $(v^{\mathcal{P}}_j, v^{\mathcal{Q}}_j)$, which is called edge voting strategy. Obviously, the higher the number of votes n_i^v of a node correspondence $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$, the better the robustness of $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ is. Fig. 2 gives an example. As shown, $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $\{(v^{\mathcal{P}}_j, v^{\mathcal{Q}}_j)\}_{j=1}^4$ are node correspondences, where nodes $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $\{(v^{\mathcal{P}}_j, v^{\mathcal{Q}}_j)\}_{j=1}^3$ are correct matches while node $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$ is a false one. $\{(e^{\mathcal{P}}_{(i,j)}, e^{\mathcal{Q}}_{(i,j)})\}_{j=1}^4$ are the support edge pairs of $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$. The lengths of these edge pairs are equal, i.e., $\{\|e^{\mathcal{P}}_{(i,j)}\| = \|e^{\mathcal{Q}}_{(i,j)}\|\}_{j=1}^4$ (For simplicity, we assume that there is no noise in the correct matches). Thus, the vote number n_i^v of $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ is 4. Similarly, node $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$ also has four support edge pairs, but only the pair $\{(e^{\mathcal{P}}_{(i,j)}, e^{\mathcal{Q}}_{(i,j)})\}_{j=4}$ has the same length, namely, the vote number n_4^v of node $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$ is 1. Therefore, we can infer that correspondence $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ is more reliable than correspondence $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$.

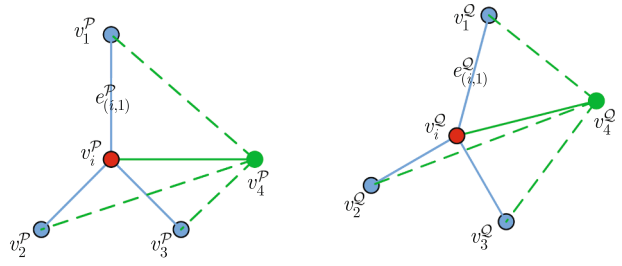


Fig. 2. An example of support edge voting. $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $\{(v^{\mathcal{P}}_j, v^{\mathcal{Q}}_j)\}_{j=1}^4$ are node correspondences, where nodes $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $\{(v^{\mathcal{P}}_j, v^{\mathcal{Q}}_j)\}_{j=1}^3$ are correct matches while node $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$ is a false one. $\{(e^{\mathcal{P}}_{(i,j)}, e^{\mathcal{Q}}_{(i,j)})\}_{j=1}^4$ are the support edge pairs of $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and the n_i^v of $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ is 4. The vote number n_4^v of node $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$ is 1. Thus, correspondence $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ is more reliable than correspondence $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$.

We first build a more compact graph $\widetilde{G}^{\mathcal{P}}(\widetilde{V}^{\mathcal{P}}, \widetilde{E}^{\mathcal{P}})$ based on I_E. Specifically, graph $\widetilde{G}^{\mathcal{P}}$ only preserves the correct edges of $E^{\mathcal{P}}$ and removes false edges and isolated nodes (after removing false edges, some nodes have no connected edges) (see Fig. 1(b)). Then, we calculate the vote number n_i^v for each node $\tilde{v}^{\mathcal{P}}_i \in \widetilde{V}^{\mathcal{P}}$ and sort the nodes $\widetilde{V}^{\mathcal{P}}$ in descending order based on the votes. We regard the first k node correspondences as potential reliable matches. Number k is determined by,

$$\min\left\{k \,\middle|\, \sum_{i=1}^{k} n_i^v \Big/ \sum_{i=1}^{\tilde{n}} n_i^v > \zeta,\ k = 1, 2, \ldots, \tilde{n}\right\} \tag{6}$$

where $\tilde{n}$ is the size of node set $\widetilde{V}^{\mathcal{P}}$; $0 < \zeta < 1$ is a ratio threshold. Finally, we rebuild the graphs $G^{\mathcal{P}}(V^{\mathcal{P}}, E^{\mathcal{P}})$ and $G^{\mathcal{Q}}(V^{\mathcal{Q}}, E^{\mathcal{Q}})$ based on the k potential feature matches. If these potential feature matches are all correct matches, we can easily obtain the registration transformation via least-squares optimization. However, this may not be true. Length of an edge is a scalar feature and is not very robust. Hence, in some cases, false correspondences may also satisfy the equal length constraint. For example, in Fig. 2, correspondences $(v^{\mathcal{P}}_i, v^{\mathcal{Q}}_i)$ and $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$ satisfy the equal length constraint, while $(v^{\mathcal{P}}_4, v^{\mathcal{Q}}_4)$ is a false match. Although edge voting stage can largely improve the robustness of the equal length constraint, it still preserves a portion of outliers. Edge voting stage can improve the inlier rate of initial matches, which is the basis for subsequent steps, i.e., transformation decomposition and transformation estimation.

3.1.2. Edge vector and transformation decomposition

For each edge $e^{\mathcal{P}}_{(i,j)}$ ($j \neq i$), we can define an edge vector as $\vec{\boldsymbol{e}}^{\mathcal{P}}_{(i,j)} = v^{\mathcal{P}}_i - v^{\mathcal{P}}_j = \boldsymbol{p}_i - \boldsymbol{p}_j$. Similarly, the edge vector of $e^{\mathcal{Q}}_{(i,j)}$; can be constructed by $\vec{\boldsymbol{e}}^{\mathcal{Q}}_{(i,j)} = v^{\mathcal{Q}}_i - v^{\mathcal{Q}}_j = \boldsymbol{q}_i - \boldsymbol{q}_j$. Due to the corresponding relationship between $\mathcal{P}$ and $\mathcal{Q}$, $(\vec{\boldsymbol{e}}^{\mathcal{P}}_{(i,j)}, \vec{\boldsymbol{e}}^{\mathcal{Q}}_{(i,j)})$ is an edge vector correspondence. Edge vector has a very important property that it is a translation invariant, which can be used for parameter space decomposition. We assume that both correspondences $(\boldsymbol{p}_i, \boldsymbol{q}_i)$ and $(\boldsymbol{p}_j, \boldsymbol{q}_j)$ are exactly matched and there is no noise in them. Thus, $(\boldsymbol{p}_i, \boldsymbol{q}_i)$ and $(\boldsymbol{p}_j, \boldsymbol{q}_j)$ strictly satisfy the rigid body transformation,

$$\left.\begin{aligned}\boldsymbol{p}_i &= \boldsymbol{R}\boldsymbol{q}_i + \boldsymbol{t}\\ \boldsymbol{p}_j &= \boldsymbol{R}\boldsymbol{q}_j + \boldsymbol{t}\end{aligned}\right\} \tag{7}$$

Then, the relationship between $\vec{\boldsymbol{e}}^{\mathcal{P}}_{(i,j)}$ and $\vec{\boldsymbol{e}}^{\mathcal{Q}}_{(i,j)}$ can be formulated as,

$$\vec{\boldsymbol{e}}^{\mathcal{P}}_{(i,j)} = \boldsymbol{p}_i - \boldsymbol{p}_j = (\boldsymbol{R}\boldsymbol{q}_i + \boldsymbol{t}) - (\boldsymbol{R}\boldsymbol{q}_j + \boldsymbol{t}) = \boldsymbol{R}(\boldsymbol{q}_i - \boldsymbol{q}_j) = \boldsymbol{R}\vec{\boldsymbol{e}}^{\mathcal{Q}}_{(i,j)} \tag{8}$$

From Eq. (8), we know that edge vectors $\vec{\boldsymbol{e}}^{\mathcal{P}}_{(i,j)}$ and $\vec{\boldsymbol{e}}^{\mathcal{Q}}_{(i,j)}$ can be registered by a three DoF rotation matrix. The rotation matrix $\boldsymbol{R}$ is also the one used to register $\mathcal{P}$ and $\mathcal{Q}$. Thus, edge vector is able to eliminate the translation component of rigid transformation and decompose the six DoF problem into two subproblems. First, we use the edge vector sets to estimate an optimal rotation matrix $\boldsymbol{R}$ between $\mathcal{P}$ and $\mathcal{Q}$; after $\boldsymbol{R}$ is found, we then rotationally register $\mathcal{P}$ and $\mathcal{Q}$, and search the optimal translation vector $\boldsymbol{t}$ based on the idea of maximum consensus. Note that the estimation on edge vector pair set is just the same as the estimation on 3D point correspondence set, because an edge vector can also be treated as a point or a 3D vector.

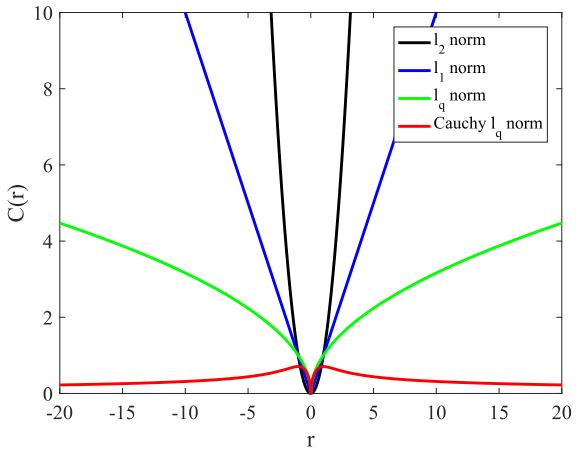


Fig. 3. The cost function curves, where the black, blue, green, and red curves represent the l_2-norm curve, l_1-norm curve, l_q-norm curve ($q = 0.5$), and the proposed Cauchy weighted l_q-norm curve ($q = 0.5$), respectively. (In the figure, r represents residual and $C(r)$ represents cost.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1.3. Cauchy-weighted l_q-norm and model estimation

Motivation: L_2-norm (least-squares) is widely used in photogrammetry, such as exterior orientation and bundle adjustment. However, it is not a robust cost and is very sensitive to outliers. As can be seen in Fig. 3, the curve of l_2-norm cost increases quadratically with residuals, which assigns more emphasis to observations with large residuals (outliers). L_1-norm cost gives equal emphasis to all observations; hence, it is more robust than l_2-norm. However, the influence of outliers towards to the total cost is still very large. The curve of l_q-norm ($q = 0.5$) is much more gradual than l_1-norm. It is able to largely discount the effect of outliers. However, the l_q-norm cost is unbounded. In contrast, the proposed Cauchy weighted l_q-norm cost is a bounded redescending function. As shown, Cauchy weighted l_q-norm gives more emphasis to outliers while less emphasis to observations with small residuals (inliers). It almost completely discounts the influence of outliers and has a high degree of robustness. Compared with RANSAC-type methods, the proposed Cauchy weighted l_q-norm cost has a major advantage. Namely, the proposed cost can obtain optimal solutions while RANSAC-type methods only get approximate solutions. RANSAC-type methods only use a subset instead of all observations for model estimation. For lack of redundant observations, they are sensitive to noise and give no guarantee of the optimality of solutions.

Cauchy weighted l_q-norm cost: Suppose we have obtained a set of edge vector correspondences $\overrightarrow{E}^{\mathcal{P}} = \{\vec{\boldsymbol{e}}_i^{\mathcal{P}}\}_1^N$ and $\overrightarrow{E}^{Q} = \{\vec{\boldsymbol{e}}_i^{Q}\}_1^N$ based on the results of edge voting, where N is the number of edge vectors. If there are no noise and outliers, the transformation between these two sets of observations is,

$$\vec{\boldsymbol{e}}_i^{\mathcal{P}} = \boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} \tag{9}$$

However, as above-mentioned, the results of edge voting still preserve a portion of outliers. Hence, the l_2-norm cost is no longer applicable. Instead, we propose a new cost function based on weighted l_q-norm ($0 < q < 1$), which is still robust even if the outlier rate is high. The new cost function for rotation estimation is,

$$\begin{aligned} &\underset{\boldsymbol{R}}{argmin} \sum_{i=1}^{N} C\left(\vec{\boldsymbol{e}}_i^{\mathcal{P}}, \boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q}\right) \\ &= \underset{\boldsymbol{R}}{argmin} \sum_{i=1}^{N} \left\| w_i\left(\boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} - \vec{\boldsymbol{e}}_i^{\mathcal{P}}\right)\right\|_q^q \end{aligned} \tag{10}$$

where $\mathcal{W} = \{w_i\}_1^N$ is a set of weighting coefficients. Ideally, large weights (close to 1) should be given to inliers while small weights (close to 0) are given to outliers, so that the effect of outliers towards to the cost can be largely discounted. We use the Cauchy function to calculate weights,

$$w_i = \frac{u}{u + r_i^2} \tag{11}$$

where u is a scale factor and r_i is an absolute residual. In Eq. (10), the absolute residual is $r_i = \|\boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} - \vec{\boldsymbol{e}}_i^{\mathcal{P}}\|_2$. Eq. (10) is a non-smooth and non-convex function. For simplicity, we introduce a set of auxiliary variables $\mathcal{S} = \{\boldsymbol{s}_i\}_1^N$ into the function,

$$\underset{\boldsymbol{R}}{\text{argmin}} \sum_{i=1}^{N} \left\|\boldsymbol{s}_i\right\|_q^q \quad s.t. \quad \boldsymbol{\upsilon}_i = w_i\left(\boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} - \vec{\boldsymbol{e}}_i^{\mathcal{P}}\right) - \boldsymbol{s}_i = \boldsymbol{0} \tag{12}$$

This constrained cost can be rewritten as an unconstrained one via Lagrange function and solved by dual ascent method. However, dual ascent has a prerequisite; namely, the function should be strictly convex. Therefore, we adopt the ALF to rewrite the equation as (see Appendix A for details about ALF),

$$\begin{aligned} \mathcal{L}_\rho\left(\boldsymbol{R}, \mathcal{W}, \mathcal{S}, \Lambda\right) &= \sum_{i=1}^{N}\left(\left\|\boldsymbol{s}_i\right\|_q^q + \boldsymbol{\lambda}_i^T\boldsymbol{\upsilon}_i + \frac{\rho}{2}\left\|\boldsymbol{\upsilon}_i\right\|_2^2\right) \\ &= \sum_{i=1}^{N}\left(\left\|\boldsymbol{s}_i\right\|_q^q + \frac{\rho}{2}\left\|\frac{\boldsymbol{\lambda}_i}{\rho} + \boldsymbol{\upsilon}_i\right\|_2^2 - \frac{1}{2\rho}\left\|\boldsymbol{\lambda}_i\right\|_2^2\right) \end{aligned} \tag{13}$$

where $\mathcal{L}_\rho(\boldsymbol{R}, \mathcal{W}, \mathcal{S}, \Lambda)$ is the ALF of Eq. (12); $\Lambda = \{\boldsymbol{\lambda}_i\}_1^N$ is a set of dual variables; and $\rho > 0$ is a penalty coefficient.

Optimization via ADMM: There are two sets of variables in the ALF, i.e., rotation $\boldsymbol{R}$ and weighting coefficients $\mathcal{W}$, and auxiliary variables $\mathcal{S}$. Hence, this function can be solved by alternately optimizing between $(\boldsymbol{R}, \mathcal{W})$ and $\mathcal{S}$ via ADMM (see Appendix B for details about ADMM). ADMM decomposes Eq. (13) into two major steps (note that the multiplier variable update step is not listed),

$$\begin{aligned} \textbf{\textit{step1}}: \quad \underset{\mathcal{S}}{\text{argmin}}\, \mathcal{L}_\rho &= \underset{\mathcal{S}}{\text{argmin}} \sum_{i=1}^{N}\left(\begin{array}{c}\left\|\boldsymbol{s}_i\right\|_q^q + \frac{\rho}{2}\left\|\frac{\boldsymbol{\lambda}_i}{\rho} + \right. \\ \left. w_i^{t-1}(\boldsymbol{R}^{t-1}\vec{\boldsymbol{e}}_i^{Q} - \vec{\boldsymbol{e}}_i^{\mathcal{P}}) - \boldsymbol{s}_i\right\|_2^2\end{array}\right) \\ &= \underset{\mathcal{S}}{\text{argmin}} \sum_{i=1}^{N}\left(\left\|\boldsymbol{s}_i\right\|_q^q + \frac{\rho}{2}\left\|\boldsymbol{\alpha}_i - \boldsymbol{s}_i\right\|_2^2\right) \end{aligned} \tag{14}$$

$$\begin{aligned} \textbf{\textit{step2}}: \quad \underset{\boldsymbol{R},\mathcal{W}}{\text{argmin}}\, \mathcal{L}_\rho &= \underset{\boldsymbol{R},\mathcal{W}}{\text{argmin}} \sum_{i=1}^{N}\left(\begin{array}{c}\left\|\boldsymbol{s}_i^{t-1}\right\|_q^q + \frac{\rho}{2}\left\|\frac{\boldsymbol{\lambda}_i}{\rho} + \right. \\ \left. w_i(\boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} - \vec{\boldsymbol{e}}_i^{\mathcal{P}}) - \boldsymbol{s}_i^{t-1}\right\|_2^2\end{array}\right) \\ &= \underset{\boldsymbol{R},\mathcal{W}}{\text{argmin}} \sum_{i=1}^{N}\frac{\rho}{2}\left\|w_i\left[\boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} - \left(\vec{\boldsymbol{e}}_i^{\mathcal{P}} + \frac{\boldsymbol{s}_i^{t-1}}{w_i} - \frac{\boldsymbol{\lambda}_i}{\rho w_i}\right)\right]\right\|_2^2 \\ &\rightarrow \underset{\boldsymbol{R},\mathcal{W}}{\text{argmin}} \sum_{i=1}^{N}\frac{\rho}{2}\left\|w_i\left[\boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} - \left(\vec{\boldsymbol{e}}_i^{\mathcal{P}} + \frac{\boldsymbol{s}_i^{t-1}}{w_i^{t-1}} - \frac{\boldsymbol{\lambda}_i}{\rho w_i^{t-1}}\right)\right]\right\|_2^2 \\ &= \underset{\boldsymbol{R},\mathcal{W}}{\text{argmin}} \sum_{i=1}^{N}\frac{\rho}{2}\left\|w_i\left(\boldsymbol{R}\vec{\boldsymbol{e}}_i^{Q} - \boldsymbol{\beta}_i\right)\right\|_2^2 \end{aligned} \tag{15}$$

where superscript t is an iteration counter. Note that there is an approximation in Eq. (15). We use a known value w_i^{t-1} to approximate w_i such that $\boldsymbol{\beta}_i$ can also be treated as a known value. $\{\boldsymbol{\alpha}_i\}_1^N$ and $\{\boldsymbol{\beta}_i\}_1^N$ are only used for notation compactness,

$$\left.\begin{aligned} \boldsymbol{\alpha}_i &= \frac{\boldsymbol{\lambda}_i}{\rho} + w_i^{t-1}\left(\boldsymbol{R}^{t-1}\vec{\boldsymbol{e}}_i^{Q} - \vec{\boldsymbol{e}}_i^{\mathcal{P}}\right) \\ \boldsymbol{\beta}_i &= \vec{\boldsymbol{e}}_i^{\mathcal{P}} + \frac{\boldsymbol{s}_i^{t-1}}{w_i^{t-1}} - \frac{\boldsymbol{\lambda}_i}{\rho w_i^{t-1}} \end{aligned}\right\} \tag{16}$$

For ***step 1***, $(\boldsymbol{R}, \mathcal{W})$ is known and $\mathcal{S}$ is the variable. The object function is a l_q-norm penalized Least-Squares (l_q LS) problem (Marjanovic and Solo, 2014), which is separable and hence can be efficiently optimized via coordinate-wise optimal minimization methods. Namely, each s_i can be optimized independently. By performing the minimization of the separated Eq. (14) along j-th coordinate, step ***step 1*** is simplified to the scalar version,

$$\underset{s}{\operatorname{argmin}}\left(\left\|s\right\|_q^q + \frac{\rho}{2}\left\|\alpha - s\right\|_2^2\right) = \underset{s}{\operatorname{argmin}}\left(|s|^q + \frac{\rho}{2}(\alpha - s)^2\right) \tag{17}$$

Then, the shrinkage algorithm can be adapted and the optimal solution is given by (Marjanovic and Solo, 2014),

$$s = \begin{cases} 0 & if\ |\alpha| \leqslant \tau + (q/\rho)\tau^{q-1} \\ \operatorname{sgn}(\alpha)\phi & f\ |\alpha| > \tau + (q/\rho)\tau^{q-1} \end{cases} s.\ t.\quad \tau = \left[\frac{2}{\rho}\left(1-q\right)\right]^{\frac{1}{2-q}} \tag{18}$$

where $\operatorname{sgn}(\cdot)$ is the signum function; ϕ is computed by equation $\phi^{t+1} = |\alpha| - \frac{q}{\rho}(\phi^t)^{q-1}$ with initial guess $\phi^0 = (\tau + |\alpha|)/2$.

For ***step 2***, $\mathcal{S}$ is known and $(\boldsymbol{R}, \mathcal{W})$ is the variable. The object function is a Weighted Least-Squares (WLS) problem, which is generally solved by Iteratively Reweighted Least-Squares (IRLS) (Holland and Welsch, 1977). In each iteration of IRLS, we use the orthonormal matrix method (Horn et al., 1988) to solve a close-form solution for rotation. Based on the ADMM technique, the optimal rotation $\boldsymbol{R}$ can be estimated. Then, the optimal translation vector $\boldsymbol{t}$ can be easily obtained.

3.1.4. Time complexity

The proposed Wl$_q$CR method consists of two main stages, i.e., edge voting and transformation estimation. Since edge voting strategy needs to compute the distance between each two feature points, its time complexity is $O(n^2)$, where n is the number of correspondences in $\mathcal{M}$. The main part of transformation estimation is to estimate the rotation $\boldsymbol{R}$. We use ADMM to optimize the rotation. In the ***step 1***, we use a shrinkage algorithm to calculate $\mathcal{S}$. Its time complexity is $O(N)$, where N is the number of edge vectors. Since edge vectors are constructed based on the k potential feature matches obtained by edge voting stage, the size N should be smaller than k^2, i.e., $N < k^2 \ll n^2$. In the ***step 2***, the time complexity of orthonormal matrix method is $O(N)$; thus, the complexity of the IRLS is $O(t_1 N)$, where t_1 is the number of iterations of the IRLS algorithm. ADMM iteratively performs ***step 1*** and ***step 2*** until convergence. If the number of iterations of ADMM is t_2, the time complexity of rotation estimation is $O(t_2(1 + t_1)N)$. The total time complexity of the proposed Wl$_q$CR is about $O(n^2 + t_2(1 + t_1)N)$. Generally, the numbers of iterations t_1 and t_2 are very small compared with N, and $N \ll n^2$. Thus, the time complexity of Wl$_q$CR can be simplified as $O(n^2)$.

3.2. Fine registration based on Wl$_q$ICP

To further refine the result of correspondence-based registration, we propose an ICP variant, i.e., Wl$_q$ICP. Similar to the ICP, it also alternates between correspondence establishment and transformation estimation, which can be formulated as a two-step optimization,

$$optimation\ 1 \quad \underset{\widetilde{\mathcal{M}}}{argmin} \sum_{i=1}^{n_{\mathcal{H}_2}} (\boldsymbol{R}\boldsymbol{h}_i^2 + \boldsymbol{t} - \boldsymbol{h}_i^1)^2 \tag{19}$$

$$optimation\ 2 \quad \underset{\boldsymbol{R},\boldsymbol{t}}{argmin} \sum_{i=1}^{n_{\mathcal{H}_2}} \left\| w_i\left(\boldsymbol{R}\boldsymbol{h}_i^2 + \boldsymbol{t} - \boldsymbol{h}_i^1\right)\right\|_q^q \tag{20}$$

where $\boldsymbol{h}_i^1$ is a point that belongs to point cloud $\mathcal{H}_1$, i.e., $\boldsymbol{h}_i^1 \in \mathcal{H}_1$;$\boldsymbol{h}_i^2 \in \mathcal{H}_2$;$n_{\mathcal{H}_2}$ is the size of point cloud $\mathcal{H}_2$;$\widetilde{\mathcal{M}}$ is a correspondence set.

In *optimation* 1, transformation parameters $\boldsymbol{R}$ and $\boldsymbol{t}$ are known. Hence, for each point $\boldsymbol{h}_i^2 \in \mathcal{H}_2$, we search its nearest point in $\mathcal{H}_1$ as its correspondence. (Note that the result of coarse registration can be used to calculate an overlap ratio and reject potential false correspondences) *optimation* 2 is also a Cauchy weighted l_q-norm cost, which can be optimized via ADMM as the same as in Section 3.1.3. After replacing the observations and transformation model in Eqs. (14)–(16), the optimization for *optimation* 2 is detailed as follows,

$$\begin{aligned}\underset{\mathcal{S}}{\operatorname{argmin}}\ \mathcal{L}_\rho &= \underset{\mathcal{S}}{\operatorname{argmin}} \sum_{i=1}^{n_{\mathcal{H}_2}} \left(\begin{array}{c} \left\|\boldsymbol{s}_i\right\|_q^q + \frac{\rho}{2}\left\|\frac{\lambda_i}{\rho}+ \right. \\ \left. w_i^{t-1}(\boldsymbol{R}^{t-1}\boldsymbol{h}_i^2 + \boldsymbol{t}^{t-1} - \boldsymbol{h}_i^1) - \boldsymbol{s}_i\right\|_2^2 \end{array}\right) \\ &= \underset{\mathcal{S}}{\operatorname{argmin}} \sum_{i=1}^{n_{\mathcal{H}_2}} \left(\left\|\boldsymbol{s}_i\right\|_q^q + \frac{\rho}{2}\left\|\boldsymbol{\alpha}_i - \boldsymbol{s}_i\right\|_2^2 \right)\end{aligned} \tag{21}$$

$$\underset{\boldsymbol{R},\boldsymbol{t},\mathcal{W}}{\operatorname{argmin}}\ \mathcal{L}_\rho = \underset{\boldsymbol{R},\boldsymbol{t},\mathcal{W}}{\operatorname{argmin}} \sum_{i=1}^{n_{\mathcal{H}_2}} \frac{\rho}{2}\left\| w_i\left(\boldsymbol{R}\boldsymbol{h}_i^2 + \boldsymbol{t} - \boldsymbol{\beta}_i\right)\right\|_2^2 \tag{22}$$

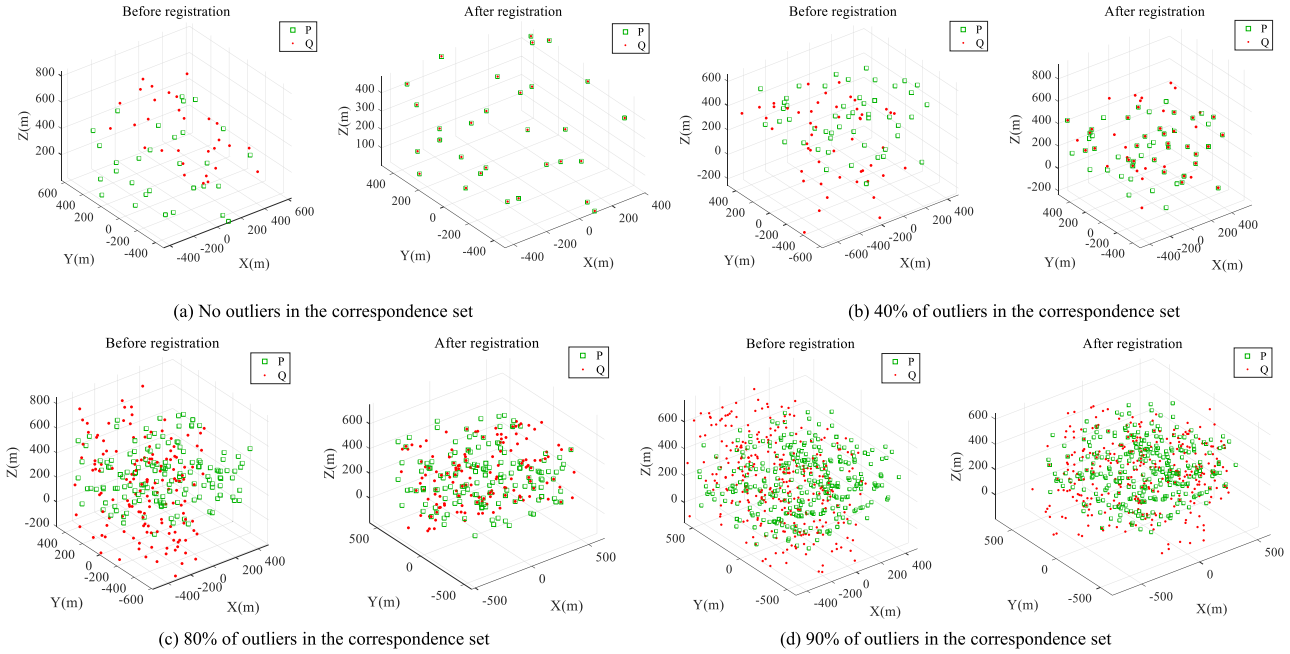
$$\left.\begin{array}{l} \boldsymbol{\alpha}_i = \frac{\lambda_i}{\rho} + w_i^{t-1}\left(\boldsymbol{R}^{t-1}\boldsymbol{h}_i^2 + \boldsymbol{t}^{t-1} - \boldsymbol{h}_i^1\right) \\ \boldsymbol{\beta}_i = \boldsymbol{h}_i^1 + \frac{\boldsymbol{s}_i^{t-1}}{w_i^{t-1}} - \frac{\lambda_i}{\rho w_i^{t-1}} \end{array}\right\} \tag{23}$$

For efficiency, we build a KD-tree for point cloud $\mathcal{H}_1$ in *optimation* 1. Its time complexity is about $O(n_{\mathcal{H}_1}\log n_{\mathcal{H}_1})$, where $n_{\mathcal{H}_1}$ is the size of point cloud $\mathcal{H}_1$. The average time complexity of nearest search in KD-tree is $O(\log n_{\mathcal{H}_1})$. For all $n_{\mathcal{H}_2}$ points in $\mathcal{H}_2$, the total nearest search complexity is about $O(n_{\mathcal{H}_2}\log n_{\mathcal{H}_1})$. Hence, the time complexity of *optimation* 1 is $O((n_{\mathcal{H}_1} + n_{\mathcal{H}_2})\log n_{\mathcal{H}_1})$. As in our previous analysis (Section 3.1.4), the time complexity of *optimation* 2 is $O(t_2(1 + t_1)n_{\mathcal{H}_2})$. Because Wl$_q$ICP iteratively performs *optimation* 1 and *optimation* 2, the total time complexity of Wl$_q$ICP is $O(t_3((n_{\mathcal{H}_1} + n_{\mathcal{H}_2})\log n_{\mathcal{H}_1} + t_2(1 + t_1)n_{\mathcal{H}_2}))$, where t_3 is the number of iterations of Wl$_q$ICP. Suppose the sizes $n_{\mathcal{H}_1}$ and $n_{\mathcal{H}_2}$ are on the same order of magnitude, the total time complexity is simplified as $O(t_3(n_{\mathcal{H}_1}\log n_{\mathcal{H}_1} + t_2 t_1 n_{\mathcal{H}_1}))$.

4. Experiments and evaluations

Here, we comprehensively study the performance of the proposed PCR algorithm on both simulated and real data. First, a parameter study experiment is conducted, which learns a reasonable value for parameter ζ. Then, we evaluate the proposed weighted l_q-norm cost and our correspondence-based coarse registration algorithm based on simulated experiments. Finally, we compare our method with six other state-of-the-art methods on challenging real-world data. Parameter ζ is set to 0.5; parameter q is studied via experiments; $\sqrt{\eta}$ is set to twice of the noise level (or point cloud resolution). We use four metrics for quantitative evaluation, i.e., rotation error $E_{\boldsymbol{R}}$, translation error $E_{\boldsymbol{t}}$, Root-Mean-Square Error (RMSE) (only for simulated experiments), and success rate.

- Rotation error $E_{\boldsymbol{R}} \in [0, \pi)$ (Huynh, 2009), which is computed by,
$$E_{\boldsymbol{R}} = \arccos\frac{tr(\boldsymbol{R}^{gt}(\boldsymbol{R}^e)^T) - 1}{2} \tag{24}$$
where $\boldsymbol{R}^{gt}$ is the ground truth rotation matrix; $\boldsymbol{R}^e$ is the estimated rotation matrix; and $tr(\cdot)$ is the trace of a matrix.
- Translation error $E_{\boldsymbol{t}} = \|\boldsymbol{t}^e - \boldsymbol{t}^{gt}\|_2$, where $\boldsymbol{t}^{gt}$ is the ground truth translation vector and $\boldsymbol{t}^e$ is the estimated one.
- RMSE, whose formula is $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} r_i^2}$, where r_i is the residual error of the i-th correspondence.
- Success rate. In simulated experiments, success rate is the ratio of successful registration times in 1000 tests. The registration is successful only if the RMSE of a test is less than twice of the noise level. In real experiments, we regard a registration whose rotation error is smaller than 2° and translation error is smaller than 1 m as a successful registration.



(a) No outliers in the correspondence set

(b) 40% of outliers in the correspondence set

(c) 80% of outliers in the correspondence set

(d) 90% of outliers in the correspondence set

Fig. 4. Visualization of the simulated 3D correspondence set $\mathcal{M}$. Each subfigure contains two plots, i.e., before registration and after registration. Before registration plot shows two input point sets, where green squares are points in $\mathcal{P}$ and red dots are points in Q. After registration plot gives the registration result of the proposed Wl$_q$CR. As shown, some green squares coincide with the red dots. These correspondences are correct match pairs (inliers) while others are false matches (outliers). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

All the reported running time is calculated on a laptop with an Intel Core i7-8550U @ 1.8 GHz CPU, 8 GB of RAM.

4.1. Parameter ζ study

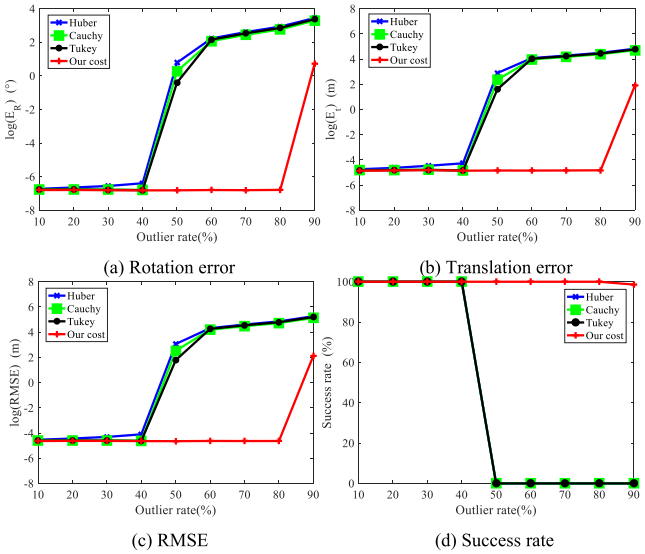
Parameter ζ in Eq. (6) is important to our Wl$_q$CR algorithm. We learn its reasonable value through simulated point set registration experiments. Specifically, a 3D feature point set $Q = \{\boldsymbol{q}_i\}_1^n$ with n points inside a box of [-500 m, 500 m] × [-500 m, 500 m] × [0, 500 m] is first generated. The mean of Q is regard as the ground truth translation $\boldsymbol{t}^{gt}$ and a 3×3 Rodrigues rotation matrix is randomly generated to be the ground truth rotation $\boldsymbol{R}^{gt}$. Thus, the true correspondence set $\mathcal{P}^{gt} = \{\boldsymbol{p}_i^{gt}\}_1^n$ of Q is obtained by $\boldsymbol{p}_i^{gt} = \boldsymbol{R}^{gt}\boldsymbol{q}_i + \boldsymbol{t}^{gt}$. To make the simulation process more realistic, we add Gaussian noise with a zero mean and a 5 cm deviation to $\mathcal{P}^{gt}$, obtaining $\bar{\mathcal{P}}$. We fix the number of correct correspondences to 30, and sequentially increase the outlier rate $n_{outlier}$ ([0, 20%, 40%, 60%, 80%, 85%, 90%, 95%, 99%]). The number n is computed by $n = 30/(1 - n_{outlier})$. Then, we select the last $n - 30$ correspondences from $\bar{\mathcal{P}}$ to add errors (the errors are randomly generated between $[\bar{\mathcal{P}}_{\min}, \bar{\mathcal{P}}_{\max}]$, where $\bar{\mathcal{P}}_{\min}$ and $\bar{\mathcal{P}}_{\max}$ are the smallest and largest values in set $\bar{\mathcal{P}}$, respectively.), obtaining outlier contaminated correspondence set $\mathcal{M} = \{(\boldsymbol{p}_i, \boldsymbol{q}_i) | \boldsymbol{p}_i \in \mathcal{P}; \boldsymbol{q}_i \in Q; i = 1, 2, ..., n\}$. Fig. 4 gives the visualizations of several simulated 3D correspondence sets. We set $\sqrt{\eta}$ to 10 cm (twice of the noise level), $k = 30$ (number of inliers), and estimate ζ based on Eq. (6). The results are summarized in Table 1, where each reported value is the average of 1000 individual tests. As shown, when outlier rate is lower than 80%, inliers hold more than 90% of correct edge correspondences; when outlier rate is 95%, inliers (only 5% of the whole observations) still hold 70% of correct edge correspondences; for extremely high outlier rates, 1% of inliers can hold more than 20% correct edge correspondences. Thus, ζ can be set to 0.7 for easy cases and 0.2 for difficult cases. Actually, $\zeta = 0.2$ is also suitable for easy cases, since a subset of inliers is enough for coarse registration.

Table 1
Study of the parameter ζ.

Outlier rate	0%	20%	40%	60%	80%	85%	90%	95%	99%
ζ	1	0.99	0.98	0.96	0.90	0.87	0.87	0.70	0.23

4.2. Robustness of Cauchy weighted l_q-norm

As aforementioned, the proposed Cauchy weighted l_q-norm cost has a high degree of robustness. To quantitatively evaluate it, we conduct a simulated experiment that is similar to the Section 4.1. The only difference is that the outlier rate $n_{outlier}$ sequentially increases from 10% to 90%. Fig. 4 gives the visualizations of several simulated 3D correspondence sets. We use the generated correspondences as input and



(a) Rotation error

(b) Translation error

(c) RMSE

(d) Success rate

Fig. 5. Comparisons of the proposed cost with classic Huber, Cauchy, and Tukey robust cost on the simulated data. Subfigures (a)~(d) are the rotation error, translation error, RMSE, and success rate results, respectively. For better visualization, the vertical coordinates of subfigures (a)~(c) are plotted in logarithmic..

Table 2
Detailed settings of the compared state-of-the-art algorithms.

Method	Parameters	Implementation	Input
RANSAC	Subset size: 3; stop confidence: 0.99; inlier threshold: 0.1 m; maximum iterations: 1e5.	MATLAB; single thread https://www.peterkovesi.com/matlabfns/index.html#robust	Correspondences
FLO-RANSAC	Subset size: 3; stop confidence: 0.99; inlier threshold: 0.1 m; maximum iterations: 1e5; local optimized subset size: 21; local optimized iterations: 50.	MATLAB; single thread https://zhipengcai.github.io/	Correspondences
FGR	Annealing rate: 1.2; inlier threshold: 0.1 m.	C++; single thread https://github.com/intel-isl/FastGlobalRegistration	Correspondences
S4PCS	Overlap ratio: 0.5; subset size: 5000; registration accuracy: 0.1 m; maximum running time: 1000s.	C++; single thread https://github.com/nmellado/Super4PCS	Downsampled point clouds
K4PCS	Overlap ratio: 0.5; score threshold: 0.001; registration accuracy: 0.1 m; maximum running time: 1000s.	C++; single thread http://pointclouds.org/	ISS keypoints
ICP	maximum correspondence size: 20000; Overlap ratio: 0.5; error metric: point-to-point; maximum iterations: 100; maximum distance threshold: 5 m.	C++; single thread https://www.mathworks.com/help/vision/ref/pcregrigid.html	Downsampled point clouds
Our Wl_qCR	Parameters $q = 0.5$, $\zeta = 0.7$ or 0.2, $\sqrt{\eta} = 0.1m$; maximum iterations of IRLS: 50; maximum ADMM iterations: 50.	MATLAB; single thread http://www.escience.cn/people/lijiayuan/index.html	Correspondences
Our Wl_qICP	maximum correspondence size: 20000; Parameter $q = 0.5$; maximum distance threshold: 5 m; maximum IRLS iterations: 50; maximum ADMM iterations: 50; maximum Wl_qICP iterations: 100.	C++; single thread http://www.escience.cn/people/lijiayuan/index.html	Downsampled point clouds

optimize Eq. (20) to estimate the rigid transformation. We also compare the proposed cost with several classic robust cost functions, including Huber function, Cauchy function, and Tukey function. The rotation error $E_{\boldsymbol{R}}$, translation error $E_{\boldsymbol{t}}$, RMSE, and success rate are reported in Fig. 5. Again, each reported value is the average of 1000 individual tests.

As can be seen, classic robust cost functions including Huber cost, Cauchy cost, and Tukey cost work very well when the outlier rate is lower than 50%. Their results are comparable with our proposed cost. However, they cannot deal with cases with more than 50% of outliers. Once the outlier rate reaches 50%, they are completely ineffective. In contrast, the proposed Cauchy weighted l_q-norm is not very sensitive to outlier rates. It performs perfect when the outlier rate is not higher than 80%. Specifically, the rotation error $E_{\boldsymbol{R}}$ is smaller than 0.002; the translation error $E_{\boldsymbol{t}}$ is smaller than 1 cm; and the success rate is always 100%. Even though the noise level is 5 cm, the RMSE of our cost is still smaller than 1 cm, which indicates that the proposed cost is robust to noise as with the l_2-norm cost. When the outlier rate is higher than 80%, the robustness of the proposed cost slightly decreases. However, our success rate is still higher than 90%. From this experiment, we can infer that the "breakdown point" of our weighted l_q-norm is larger than 0.8, which is much higher than traditional robust estimators, such as M-estimators and LTS. Namely, the proposed WICP is more robust than the TriICP algorithm. Our Wl_qICP is able to deal with cases with 20% of overlaps, while TriICP can only deal with cases with more than 50% of overlaps.

4.3. Evaluation of Wl_qCR on simulated data

The simulated process is the same as Section 4.1. Parameter ζ is set to 0.7. We compare our coarse registration algorithm, i.e., Wl_qCR, with other three correspondence-based methods, including RANSAC (Fischler and Bolles, 1981), FLO-RANSAC (Lebeda et al., 2012), and FGR (Zhou et al., 2016). In RANSAC-type methods, orthonormal matrix algorithm is applied to solve a close-form solution of rigid transformation for each hypothesis. Table 2 summarizes the detailed settings of these algorithms. The average quantitative results are plotted in Fig. 6. As shown, RANSAC and FLO-RANSAC obtain similar results. Both of them perform well when the outlier rate is lower than 95%, i.e., the average $E_{\boldsymbol{R}}$ is about 0.004°, the average $E_{\boldsymbol{t}}$ is about 0.030 m, and the average RMSE is about 0.037 m. However, they often fail to estimate a good solution when the outlier rate is extremely high. For example, the success rate of RANSAC under 99% of outliers is only 7.7%. Although FLO-RANSAC is more robust to high outlier rates than RANSAC, its success rate is only 66.3%. FGR performs much worse than other methods. When the observations are not contaminated by outliers (the outlier rate is 0), FGR achieves similar model fitting accuracy with the proposed Wl_qCR method, which is better than RANSAC-type methods. However, FGR is very sensitive to outliers. Under 20% of outliers, its rotation error is 0.099° and its translation error is 0.68 m. Since our simulated cases are large scale, i.e., a box of [-500 m, 500 m] × [-500 m, 500 m] × [0, 500 m], 0.1° can also lead to large offsets. Its corresponding RMSE is 0.76 m. Moreover, the fitting accuracy of FGR severely decreases as the outlier rate increases. Its $E_{\boldsymbol{R}}$, $E_{\boldsymbol{t}}$, and RMSE under 85% of outliers decrease to 1.60°, 5.876 m, and 6.340 m, respectively. The success rate of FGR is close to 0 when the outlier rate is higher than 20%, since we only regard registrations whose RMSEs are less than 0.1 m as successful registrations. Generally, FGR is not suitable for large-scale or high outlier rate point cloud registration problem, whose results should be refined by other methods such as ICP. In contrast, our Wl_qCR achieves the best performance among these four methods. The proposed Wl_qCR is able to deal with cases with extremely high outlier rates, whose success rate is still 100% under 99% of outliers. The average $E_{\boldsymbol{R}}$, $E_{\boldsymbol{t}}$ and RMSE of all experiments are 0.002°, 0.012 m, and 0.014 m, respectively. The model estimation accuracy of our Wl_qCR is much higher than RANSAC-type methods. The reason is that RANSAC-type methods only use small subsets for model estimation, which is sensitive to noise.

We calculate the running time of each method, which is reported in Table 3. RANSAC and FLO-RANSAC are much slower than the proposed Wl_qCR method, especially in cases with high outlier rates, since the required sampling trials are huge. For example, our Wl_qCR is 500+ times faster than RNASC and 550+ times faster than FLO-RANSAC when the outlier rate is 95%. Even though FGR is implemented in C++, it still cost more running time than our method. In most cases, our Wl_qCR method is almost 40 times faster than FGR.

4.4. Challenging real data experiments

We also evaluate the proposed Wl_qCR and Wl_qICP methods on large scale real data to demonstrate their practicality. We use the challenging ETH LiDAR dataset[3] for evaluation, which contains five categories, i.e., Office, Facade, Courtyard, Arch, and Trees. For efficiency, we apply Voxel Grid downsampling algorithm (Rusu and Cousins, 2011) to decrease the resolution to 0.1 m. For each scan pair ($\mathcal{H}_1$, $\mathcal{H}_2$), we use ISS (Zhong, 2009) detector to extract feature keypoints and use FPFH (Rusu

[3] http://www.prs.igp.ethz.ch/research/completed_projects/automatic_registration_of_point_clouds.html

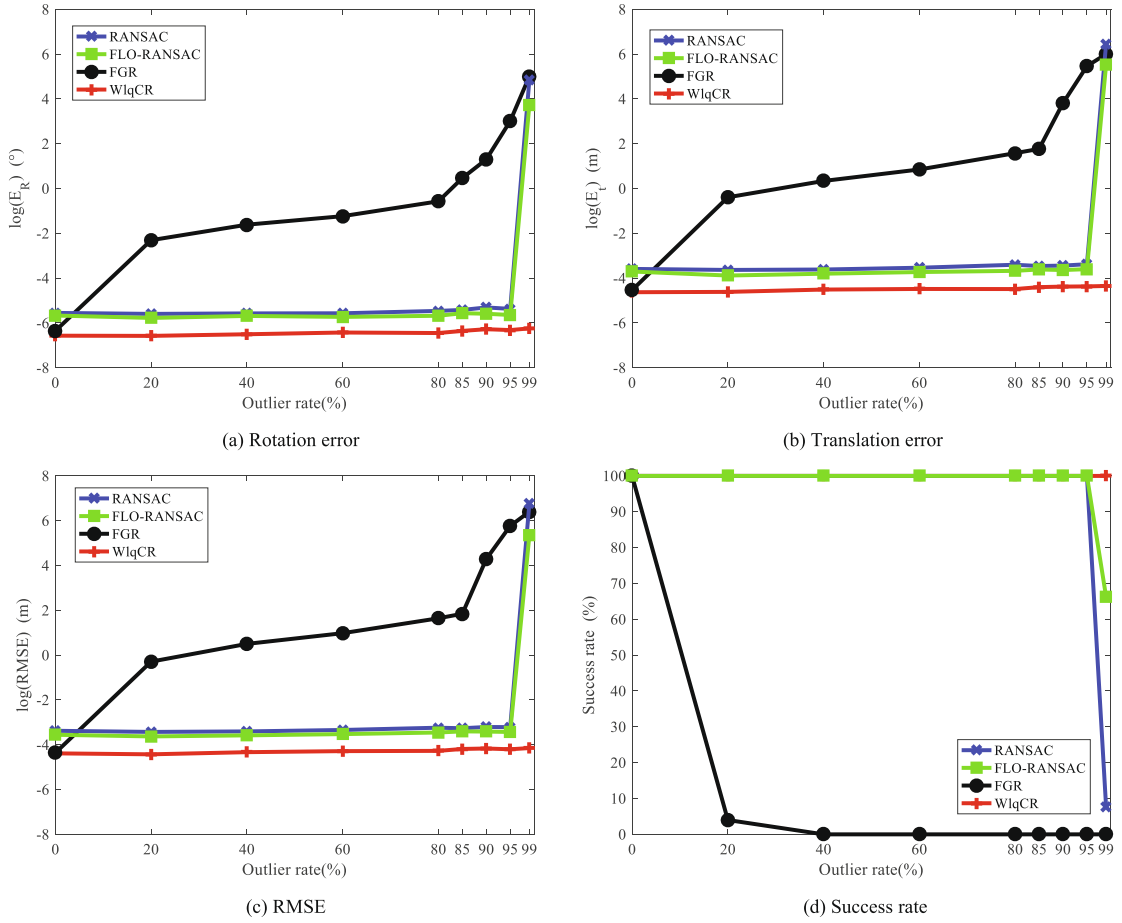


Fig. 6. Comparisons of the proposed Wl$_q$CR with RANSAC, FLO-RANSAC, and FGR on the simulated data. Subfigures (a)~(d) are the rotation error, translation error, RMSE, and success rate results, respectively. The horizontal axis represents the outlier rate of the simulated initial correspondence set, which increases from 0 to 99% ([0, 20%, 40%, 60%, 80%, 85%, 90%, 95%, 99%]). For better visualization, the vertical coordinates of subfigures (a)~(c) are plotted in logarithmic. Smaller values in subfigures (a)~(c) represent higher accuracies, and larger values in subfigure (d) represent better performance.

Table 3
Comparison of running time results (s).

Outlier rate	Methods			
	RANSAC	FLO-RANSAC	FGR	Wl$_q$CR
0%	0.090	0.108	0.083	0.002
20%	0.108	0.127	0.102	0.002
40%	0.108	0.261	0.133	0.002
60%	0.114	0.285	0.213	0.002
80%	0.315	0.334	0.381	0.002
85%	0.726	0.793	0.522	0.002
90%	2.278	2.616	0.763	0.003
95%	19.689	22.191	1.530	0.039
99%	46.818	139.592	8.834	0.505

et al., 2009) descriptor for feature description[4]. Similar to (Cai et al., 2019), initial correspondence set $\mathcal{M}$ is obtained based on K-nearest neighbor search. Specifically, if the descriptor distance between $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ is one of the top-5 nearest to each other, $(\boldsymbol{p}_i, \boldsymbol{q}_i)$ is selected into the initial correspondence set. Table 4 briefly summarizes the information of each LiDAR category, including number of scan pairs N_{SP}, overlap ratios, average number of points $|\mathcal{H}|$, average number of keypoints $|\mathcal{H}^{key}|$, and average number of initial correspondences $|\mathcal{M}|$.

The ground truth transformation T is provided for each scan pair in

[4] Both the implementations of ISS and FPFH are obtained from PCL: http://pointclouds.org/

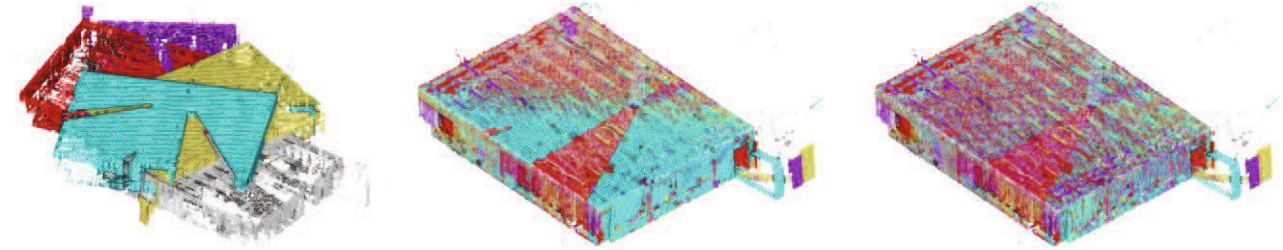
Table 4
Information of the ETH dataset.

Info	Office	Facade	Courtyard	Arch	Trees
N_{SP}	8	21	28	8	10
Overlap	> 80%	60%~70%	40%~70%	30%~40%	≈50%
$\lvert\mathscr{K}\rvert$	10721 K	17343 K	12194 K	27423 K	20248 K
$\lvert\mathcal{H}^{key}\rvert$	1487	2273	4084	5808	10404
$\lvert\mathcal{M}\rvert$	3194	4408	9915	7569	14535

the ETH dataset. If $\|\boldsymbol{p}_i - T(\boldsymbol{q}_i)\|_2 < \varepsilon$, point correspondence $(\boldsymbol{p}_i, \boldsymbol{q}_i)$ is accepted as an inlier. The inlier threshold ε is set to 0.3 m (three times of the point cloud resolution). The correspondence inlier rates of Office, Facade, Courtyard, Arch, and Trees are only 1.77%, 4.91%, 3.86%, 0.67%, and 0.45%, respectively. The lowest outlier rate of these five categories is still higher than 95% and two categories contain more than 99% of outliers. Registration on the ETH dataset is very challenging due to such extremely high outlier rates. The parameters of our Wl$_q$CR method is set to $\zeta = 0.2$ (difficult cases) and $\sqrt{\bar{\eta}} = 0.1m$ (the point cloud resolution). Apart from RANSAC, FLO-RANSAC, and FGR, we add other three state-of-the-arts for comparison, i.e., S4PCS (Mellado et al., 2014), K4PCS (Theiler et al., 2014), and ICP (Besl and McKay, 1992). More details about these algorithms can be found in Table 2.

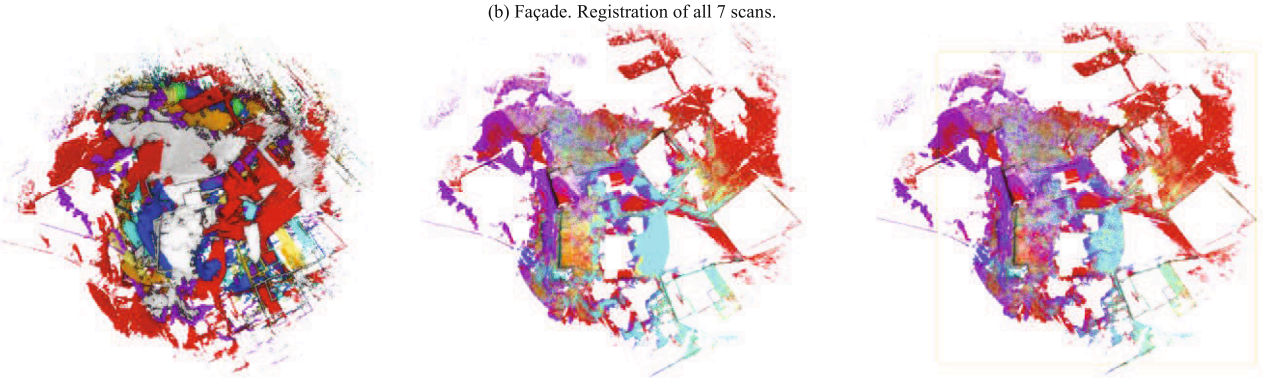
The result of each scan pair is plotted in Fig. 7 (Only half of the scan pairs in Facade and Courtyard categories are displayed). As shown, RANSAC and FLO-RANSAC perform very well on the Office, Facade,

Fig. 7. Comparisons of the proposed Wl_qCR and Wl_qICP with six state-of-the-art methods on the challenging ETH real data. Subfigures (a)~(e) are the results on the Office, Facade, Courtyard, Arch, and Trees, respectively. Left figures plot the rotation error E_R results and right figures plot the translation error E_t results. The horizontal axis represents scan pair IDs. For better visualization, the vertical axis is logarithmic. Smaller values of E_R and E_t represent higher accuracies..



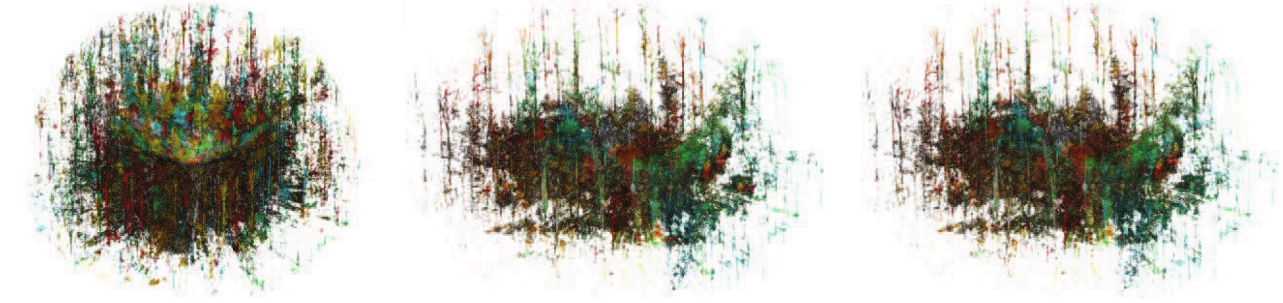
(a) Office. Registration of all 5 scans.

(b) Façade. Registration of all 7 scans.



(c) Courtyard. Registration of all 8 scans.

(d) Arch. Registration of all 5 scans.



(e) Trees. Registration of all 6 scans.

Fig. 8. Complete registration results. Left column is the input LiDAR scans. Each LiDAR scan has its own coordinate system and different scans are represented by different colors. Before registration, point clouds of the input LiDAR scans are messy. Middle column is the results of Wl_qCR and the right one is the results refined by Wl_qICP. As shown, the results after registration are impressing. Since Wl_qCR can obtain good enough results, the results of Wl_qCR and the refined ones look similar. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and Courtyard categories. As demonstrated in Section 4.3, both of them are sensitive to extremely high outlier rate. Thus, they fail to register several scan pairs in Arch and Trees categories, whose rotation errors are even larger than 100°. FGR is not suitable for cases with high outlier rates. Its performance is the worst on the most of scan pairs. For example, its model estimation accuracy is the lowest on the Office and Courtyard categories. S4PCS is comparable with RANSAC-type methods. It works well on the first four categories, which is even better than FLO-RANSAC on the Arch dataset. K4PCS performs less satisfactory than S4PCS. It fails to register several scan pairs of Facade category. K4PCS uses keypoints instead of downsampled point clouds as input, which may be too sparse to construct reliable 4PCS. The proposed Wl$_q$CR obtains the best performance among all coarse registration methods. It gets impressed results in extremely difficult cases, such as Arch and Trees. The registration accuracy of Wl$_q$CR is only slightly lower than the results refined by ICP. In some cases of the Courtyard dataset, ICP refinement even decreases the registration accuracy of our Wl$_q$CR. From Fig. 7, we can see that the proposed Wl$_q$ICP fine registration algorithm has a higher registration accuracy than ICP. The reasons may be two folds: First, the proposed Wl$_q$ICP is robust to low overlap ratios. With the initial transformation provided by our Wl$_q$CR, it does not need to set up the overlap ratio parameters. However, classic ICP is sensitive to overlap ratios. Different values of overlap ratio lead to very different results. Second, the proposed Wl$_q$ICP uses a weighted cost function, which gives large weights to correspondences with small residuals and relatively small weights to noise with relatively large residuals. Therefore, Wl$_q$ICP is more robust to noise than traditional ICP. The complete registration results of each category of the proposed Wl$_q$CR and Wl$_q$CR + Wl$_q$ICP are shown in Fig. 8.

Table 5 and Table 6 report the registration accuracy of each category. As can be seen, FLO-RANSAC is more accurate than RANSAC. The reason is: RANSAC estimates models based on minimum subsets (3-point correspondence subset) while FLO-RANSAC uses much larger subsets to refine the model estimated by RANSAC. Redundant observations are able to improve the robustness to noise. In some categories, S4PCS and K4PCS are less accurate than RANSAC-type methods. Their results can be used as initial results and should be refined by other fine registration methods. The proposed Wl$_q$CR obtains a registration accuracy of 0.21° and 0.038 m on cases with 99.55% of outliers, which is much superior to other coarse registration methods and can be directly used for realistic applications. First, our Wl$_q$CR adapts an equal-length edge constraint and a support edge voting strategy to filter outliers. Second, the cost function used in Wl$_q$CR has a high degree of robustness, i.e., it is still robust even if the outlier rate is 80%~90%. Benefiting from these two aspects, the proposed Wl$_q$CR can tolerate extremely high outlier rates. The proposed Wl$_q$ICP is more accurate than ICP on all the categories. The maximum E_R of Wl$_q$CR + ICP is 0.21° while the maximum E_R of Wl$_q$CR + Wl$_q$ICP is only 0.10°. In terms of E_t, the maximum values of Wl$_q$CR + ICP and Wl$_q$CR + Wl$_q$ICP are 0.082 m and 0.036 m, respectively. Table 7 summarizes the success rate of each method. As reported, the total average success rates of RANSAC, FLO-RANSAC, FGR, S4PCS, K4PCS, Wl$_q$CR, Wl$_q$CR + ICP, and Wl$_q$CR + Wl$_q$ICP are 64.5%, 73.5%, 25.7%, 67.9%, 62.7%, 100%, 98.6%, and 100%, respectively. Our Wl$_q$CR ranks best among coarse registration methods, which achieves 26.5% growth rates compared with the second-best method, i.e., FLO-RANSAC.

Table 8 reports the running time results. Note that the reported value of each method is the optimization time, which does not include the running time of feature extraction. As shown, our Wl$_q$CR algorithm is the fastest. Wl$_q$CR is almost 200 times faster than RANSAC and FLO-RANSAC on the Office. However, if the size of initial correspondence set is large, the proposed Wl$_q$CR becomes much slower. For example, Wl$_q$CR only costs 1.0s on the Facade with 4408 correspondences, whereas it costs 23.4s on the Trees with 14535 correspondences. The reason is that the proposed Wl$_q$CR needs to compute the distance between each two feature points. Thus, the time complexity of our Wl$_q$CR is O(n^2), where n is the number of correspondences. Fortunately, n is much smaller than the size of point clouds $|\mathcal{H}|$. Generally, 10000 correspondences are enough for our Wl$_q$CR to estimate a good solution (if the outlier rate is 99.5%, 10000 correspondences still contain 50 inliers). So, we can set maximum number of correspondences to 10000 for efficiency. On the Courtyard with 9915 correspondences, Wl$_q$CR only takes 8.0s in MATLAB. The proposed Wl$_q$ICP is slower than classic ICP algorithm, because the optimization of weighted l_q-norm is more difficult than traditional l_2-norm. Fortunately, if a good initial guess is provided to Wl$_q$ICP, it is much faster than S4PCS and K4PCS. For example, our Wl$_q$CR + Wl$_q$ICP is almost 10 times faster than S4PCS and K4PCS on the Trees.

Table 5
Rotation error E_R (°).

Method	Office	Facade	Courtyard	Arch	Trees
RANSAC	3.24	0.72	0.30	38.51	31.83
FLO-RANSAC	1.91	0.52	0.12	24.97	21.52
FGR	57.21	2.35	1.42	26.17	39.72
S4PCS	0.74	1.11	0.39	2.66	18.39
K4PCS	0.75	3.96	0.68	5.28	47.88
Wl$_q$CR	0.26	0.32	0.07	0.15	0.21
Wl$_q$CR + ICP	0.21	0.16	0.12	0.07	0.12
Wl$_q$CR + Wl$_q$ICP	0.10	0.07	0.04	0.03	0.05

Table 6
Translation error E_t (m).

Method	Office	Facade	Courtyard	Arch	Trees
RANSAC	0.20	0.16	0.12	7.55	4.34
FLO-RANSAC	0.20	0.10	0.06	4.15	6.07
FGR	2.07	0.67	1.07	8.78	7.71
S4PCS	0.27	0.16	0.23	0.66	3.30
K4PCS	0.10	0.96	0.36	4.67	8.39
Wl$_q$CR	0.025	0.042	0.041	0.043	0.038
Wl$_q$CR + ICP	0.021	0.019	0.082	0.040	0.032
Wl$_q$CR + Wl$_q$ICP	0.009	0.009	0.036	0.014	0.005

Table 7
Success rate (%).

Method	Office	Facade	Courtyard	Arch	Trees
RANSAC	75	100	100	37.5	10
FLO-RANSAC	87.5	100	100	50	30
FGR	25	57.1	46.4	0	0
S4PCS	100	85.7	96.4	37.5	20
K4PCS	100	66.7	89.3	37.5	20
Wl$_q$CR	100	100	100	100	100
Wl$_q$CR + ICP	100	100	92.9	100	100
Wl$_q$CR + Wl$_q$ICP	100	100	100	100	100

Table 8
Running Time (s).

Method	Office	Facade	Courtyard	Arch	Trees
RANSAC	111.2	24.9	101.2	126.2	171.5
FLO-RANSAC	116.9	25.8	104.8	131.4	177.1
FGR	18.9	26.3	42.1	46.6	87.9
S4PCS	1000	1000	686.2	703.3	751.8
K4PCS	12.7	10.8	340.0	674.6	1000
Wl$_q$CR	0.6	1.0	8.0	4.8	23.4
Wl$_q$CR + ICP	1.4	5.2	31.3	54.9	47.4
Wl$_q$CR + Wl$_q$ICP	5.3	6.4	44.5	75.7	78.9

5. Conclusions

In this paper, we propose an accurate, robust, and efficient PCR method in a coarse-to-fine manner. We develop two novel algorithms based on topological graph and Cauchy-weighted l_q-norm cost. Specifically, we first propose a correspondence-based coarse registration approach called Wl_qCR, which can tolerate extremely high outlier rates (higher than 99%). The ultrahigh robustness of Wl_qCR is benefit from two aspects: first, we transform the point matching problem to a high-level edge matching problem and introduce a support edge voting strategy to filter outliers; second, we propose a new cost function, which has a high degree of robustness, i.e., it is still robust even if the outlier rate is 80%~90%. Then, we propose a new ICP variant called Wl_qICP, which is much more robust to partial overlaps compared with current ICP and its variants. In our method, we ignore the scale between point clouds. Thus, it is not suitable for multi-source point cloud registration problem, such as registration of a structure-from-motion point cloud and a LiDAR point cloud. Fortunately, it is easy to extend our method to deal with scale changes. We can calculate a length ratio for each edge correspondence. Then, a length ratio consensus maximization procedure can be used to estimate the optimal scale parameter.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key R&D Program of China (2017YFD0600904), National Natural Science Foundation of China (NSFC) (No. 41901398), China Postdoctoral Science Foundation (No. 2018M640734), and Open Research Fund of Key Laboratory for National Geograophy State Monitoring (National Administration of Surveying, Mapping and Geoinformation) (No. 2018NGCMZD02).

Appendix A. Augmented Lagrange function

Consider an equality constrained problem,

$$\underset{\boldsymbol{x}}{\operatorname{argmin}} f(\boldsymbol{x}) \;\; s.\,t. \;\; g_i(\boldsymbol{x}) = 0 \quad i = 1, 2, \ldots, n \tag{A.1}$$

where $\boldsymbol{x} \in \mathbb{R}^k$. This constrained problem can be converted to an unconstrained problem by the Lagrange function,

$$\mathcal{L}\left(\boldsymbol{x}, \boldsymbol{\lambda}\right) = f(\boldsymbol{x}) + \sum_{i=1}^{n} \lambda_i g_i(\boldsymbol{x}) \tag{A.2}$$

where $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda})$ is a Lagrange function; $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_i, \ldots \lambda_n] \in \mathbb{R}^n$ is a vector of Lagrange multipliers associated with constraints. Eq. (A.2) can be optimized by the method of Lagrange multipliers or dual ascent method. However, both them assume that the objective function is convex. If the cost is not a convex function, they cannot guarantee to obtain the global or even a local solution. Fortunately, augmented Lagrangian methods can yield convergence without assumptions of convexity. The objective of augmented Lagrangian method for Eq. (A.1) is:

$$\mathcal{L}_\rho\left(\boldsymbol{x}, \boldsymbol{\lambda}\right) = f(\boldsymbol{x}) + \sum_{i=1}^{n} \lambda_i g_i(\boldsymbol{x}) + \frac{\rho}{2} \sum_{i=1}^{n} g_i^2(\boldsymbol{x}) \tag{A.3}$$

where $\mathcal{L}_\rho(\boldsymbol{x}, \boldsymbol{\lambda})$ is an augmented Lagrange function (ALF); $\rho > 0$ is a penalty parameter. ALF adds a quadratic penalty to the usual Lagrange function.

Appendix B. Alternating direction method of multipliers

Alternating direction method of multipliers (ADMM) is an extension of the augmented Lagrangian method. It has the superior convergence of the method of multipliers and the decomposability property of the dual ascent method. ADMM is very suitable for solving problems that the variables $\boldsymbol{x}$ can be split into two parts, i.e., it is suitable for problems in the form,

$$\begin{aligned} &\underset{\boldsymbol{x}}{\operatorname{argmin}} f(\boldsymbol{x}) = \underset{\boldsymbol{x}_1, \boldsymbol{x}_2}{\operatorname{argmin}} \left(f_1(\boldsymbol{x}_1) + f_2(\boldsymbol{x}_2)\right) \\ &s.\,t. \;\; g_i(\boldsymbol{x}_1, \boldsymbol{x}_2) = 0 \quad i = 1, 2, \ldots, n \end{aligned} \tag{B.1}$$

where $f(\boldsymbol{x}) = f_1(\boldsymbol{x}_1) + f_2(\boldsymbol{x}_2)$ and $\boldsymbol{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2]$. Using ALF to convert this object to an unconstrained function, we have,

$$\mathcal{L}_\rho\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{\lambda}\right) = f_1(\boldsymbol{x}_1) + f_2(\boldsymbol{x}_2) + \sum_{i=1}^{n} \lambda_i g_i\left(\boldsymbol{x}_1, \boldsymbol{x}_2\right) + \frac{\rho}{2} \sum_{i=1}^{n} g_i^2\left(\boldsymbol{x}_1, \boldsymbol{x}_2\right) \tag{B.2}$$

ADMM decomposes the ALF cost into subproblems, which consists of three-step iterations,

$$\textbf{\textit{step1}}: \boldsymbol{x}_1^{t+1} = \underset{\boldsymbol{x}_1}{\operatorname{argmin}} \mathcal{L}_\rho\left(\boldsymbol{x}_1, \boldsymbol{x}_2^t, \boldsymbol{\lambda}^t\right) \tag{B.3}$$

$$\textbf{\textit{step2}}: \boldsymbol{x}_2^{t+1} = \underset{\boldsymbol{x}_2}{\operatorname{argmin}} \mathcal{L}_\rho\left(\boldsymbol{x}_1^t, \boldsymbol{x}_2, \boldsymbol{\lambda}^t\right) \tag{B.4}$$

step3: $\lambda_i^{t+1} = \lambda_i^t + \rho g_i(\boldsymbol{x}_1^{t+1}, \boldsymbol{x}_2^{t+1}) \quad i = 1, 2, ..., n$ (B.5)

where superscript t is an iteration counter. ***step 1*** is the $\boldsymbol{x}_1$-minimization step, where $\boldsymbol{x}_1$ is the only variable and ($\boldsymbol{x}_2$, *bmlambda*) is known; ***step 2*** is the $\boldsymbol{x}_2$-minimization step, where $\boldsymbol{x}_2$ is the only variable and ($\boldsymbol{x}_1$, *bmlambda*) is known; ***step 3*** is a multiplier variable update step. More details about ALF and ADMM can be found in literature (Boyd et al., 2011).

References

Aiger, D., Mitra, N.J., & Cohen-Or, D. (2008). 4-points congruent sets for robust pairwise surface registration. In ACM transactions on graphics (TOG) (p. 85). Acm volume 27.

Barnea, S., Filin, S., 2008. Keypoint based autonomous registration of terrestrial laser point-clouds. ISPRS J. Photogramm. Remote Sens. 63, 19–35.

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-up robust features (surf). Comput. Vision Image Understand. 110, 346–359.

Bergström, P., Edlund, O., 2014. Robust registration of point sets using iteratively reweighted least squares. Comput. Optim. Appl. 58, 543–561.

Besl, P., McKay, N.D., 1992. A method for registration of 3-d shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14, 239–256.

Biber, P., Straßer, W. 2003. 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) (Cat. No. 03CH37453). IEEE, vol. 3, pp. 2743–2748.

Blais, G., Levine, M.D., 1995. Registering multiview range data to create 3d computer objects. IEEE Trans. Pattern Anal. Mach. Intell. 17, 820–824.

Böhm, J., Becker, S., 2007. Automatic marker-free registration of terrestrial laser scans using reflectance. In: Proceedings of the 8th Conference on Optical 3D Measurement Techniques. Zurich, Switzerland, pp. 9–12.

Botterill, T., Mills, S., Green, R.D. 2009. New conditional sampling strategies for speeded-up ransac. In: BMVC. Citeseer, pp. 1–11.

Bouaziz, S., Tagliasacchi, A., Pauly, M. 2013. Sparse iterative closest point. In Computer Graphics Forum. Wiley Online Library, vol. 32, pp. 113–123.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. 3, 1–122.

Bustos, Á.P., Chin, T.-J., 2017. Guaranteed outlier removal for point cloud registration with correspondences. IEEE Trans. Pattern Anal. Mach. Intell. 40, 2868–2882.

Cai, Z., Chin, T.-J., Bustos, A.P., Schindler, K., 2019. Practical optimal registration of terrestrial lidar scan pairs. ISPRS J. Photogramm. Remote Sens. 147, 118–131.

Cai, Z., Chin, T.-J., Le, H., Suter, D., 2018. Deterministic consensus maximization with biconvex programming. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 685–700.

Censi, A. 2008. An icp variant using a point-to-line metric. In: 2008 IEEE International Conference on Robotics and Automation. IEEE, pp. 19–25.

Chen, H., Bhanu, B., 2007. 3d free-form object recognition in range images using local surface patches. Pattern Recogn. Lett. 28, 1252–1262.

Chen, Y., Medioni, G., 1992. Object modelling by registration of multiple range images. Image Vision Comput. 10, 145–155.

Chetverikov, D., Stepanov, D., Krsek, P., 2005. Robust euclidean alignment of 3d point sets: the trimmed iterative closest point algorithm. Image Vision Comput. 23, 299–309.

Chin, T.-J., Suter, D., 2017. The maximum consensus problem: recent algorithmic advances. Synthesis Lect. Comput. Vision 7, 1–194.

Choi, S., Zhou, Q.-Y., Koltun, V., 2015. Robust reconstruction of indoor scenes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5556–5565.

Das, A., Waslander, S.L., 2014. Scan registration using segmented region growing ndt. Int. J. Robot. Res. 33, 1645–1663.

Deng, H., Birdal, T., Ilic, S., 2018. Ppfnet: Global context aware local features for robust 3d point matching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 195–205.

Dong, Z., Yang, B., Liang, F., Huang, R., Scherer, S., 2018. Hierarchical registration of unordered tls point clouds based on binary shape context descriptor. ISPRS J. Photogramm. Remote Sens. 144, 61–79.

Feldmar, J., Ayache, N., 1996. Rigid, affine and locally affine registration of free-form surfaces. Int. J. Comput. Vision 18, 99–119.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24, 381–395.

Frome, A., Huber, D., Kolluri, R., Bülow, T., Malik, J., 2004. Recognizing objects in range data using regional point descriptors. In: European Conference on Computer Vision. Springer, pp. 224–237.

Ge, X., 2016. Non-rigid registration of 3d point clouds under isometric deformation. ISPRS J. Photogramm. Remote Sens. 121, 192–202.

Ge, X., 2017. Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets. ISPRS J. Photogramm. Remote Sens. 130, 344–357.

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

Guo, Y., Sohel, F., Bennamoun, M., Lu, M., Wan, J., 2013. Rotational projection statistics for 3d local surface description and object recognition. Int. J. Comput. Vision 105, 63–86.

Hansen, M.F., Blas, M.R., Larsen, R. 2007. Mahalanobis distance based iterative closest point. In: SPIE Medical Imaging: Image Processing. SPIE.

Holland, P.W., Welsch, R.E., 1977. Robust regression using iteratively reweighted least-squares. Commun. Stat.-Theory Methods 6, 813–827.

Horn, B.K.P., Hilden, H.M., Negahdaripour, S., 1988. Closed-form solution of absolute orientation using orthonormal matrices. J. Opt. Soc. Am. A 5, 1127–1135.

Houshiar, H., Elseberg, J., Borrmann, D., Nüchter, A., 2015. A study of projections for key point based registration of panoramic terrestrial 3d laser scan. Geo-spatial Inf. Sci. 18, 11–31.

Huynh, D.Q., 2009. Metrics for 3d rotations: comparison and analysis. J. Math. Imaging Vision 35, 155–164.

Jian, B., Vemuri, B.C., 2010. Robust point set registration using gaussian mixture models. IEEE Trans. Pattern Anal. Mach. Intell. 33, 1633–1645.

Johnson, A.E., Hebert, M., 1999. Using spin images for efficient object recognition in cluttered 3d scenes. IEEE Trans. Pattern Anal. Mach. Intell. 21, 433–449.

Kaneko, S., Kondo, T., Miyamoto, A., 2003. Robust matching of 3d contours using iterative closest point algorithm improved by m-estimation. Pattern Recogn. 36, 2041–2047.

Kang, Z., Li, J., Zhang, L., Zhao, Q., Zlatanova, S., 2009. Automatic registration of terrestrial laser scanning point clouds using panoramic reflectance images. Sensors 9, 2621–2646.

Kang, Z., Zhang, L., Wang, B., Li, Z., Jia, F., 2013. An optimized baysac algorithm for efficient fitting of primitives in point clouds. IEEE Geosci. Remote Sens. Lett. 11, 1096–1100.

Lebeda, K., Matas, J., Chum, O., 2012. Fixing the locally optimized ransac–full experimental evaluation. In: British machine vision conference. Citeseer, pp. 1–11.

Li, J., Hu, Q., Ai, M., 2016a. Robust feature matching for remote sensing image registration based on l_ q-estimator. IEEE Geosci. Remote Sens. Lett. 13, 1989–1993.

Li, J., Hu, Q., Ai, M., 2017a. Robust feature matching for geospatial images via an affine-invariant coordinate system. Photogram. Rec. 32, 317–331.

Li, J., Hu, Q., Ai, M. 2018. Rift: Multi-modal image matching based on radiation-invariant feature transform. arXiv preprint arXiv:1804.09493.

Li, J., Hu, Q., Ai, M., 2019. Lam: Locality affine-invariant feature matching. ISPRS J. Photogramm. Remote Sens. 154, 28–40.

Li, J., Hu, Q., Ai, M., Zhong, R., 2017b. Robust feature matching via support-line voting and affine-invariant ratios. ISPRS J. Photogramm. Remote Sens. 132, 61–76.

Li, J., Zhong, R., Hu, Q., Ai, M., 2016b. Feature-based laser scan matching and its application for indoor mapping. Sensors 16, 1265.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60, 91–110.

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

Maier-Hein, L., Franz, A.M., Dos Santos, T.R., Schmidt, M., Fangerau, M., Meinzer, H.-P., Fitzpatrick, J.M., 2011. Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error. IEEE Trans. Pattern Anal. Mach. Intell. 34, 1520–1532.

Marjanovic, G., Solo, V., 2014. l_q sparsity penalized linear regression with cyclic descent. IEEE Trans. Signal Process. 62, 1464–1475.

Masuda, T., Sakaue, K., Yokoya, N. (1996). Registration and integration of multiple range images for 3-d model construction. In: Proceedings of 13th international conference on pattern recognition. IEEE. vol. 1, pp. 879–883.

Mellado, N., Aiger, D., Mitra, N.J. 2014. Super 4pcs fast global pointcloud registration via smart indexing. In: Computer Graphics Forum. Wiley Online Library, vol. 33, pp. 205–215.

Mian, A., Bennamoun, M., Owens, R., 2010. On the repeatability and quality of keypoints for local feature-based 3d object retrieval from cluttered scenes. Int. J. Comput. Vision 89, 348–361.

Montuori, A., Luzi, G., Stramondo, S., Casula, G., Bignami, C., Bonali, E., Bianchi, M.G., Crosetto, M. 2014. Combined use of ground-based systems for cultural heritage conservation monitoring. In: 2014 IEEE Geoscience and Remote Sensing Symposium. IEEE, pp. 4086–4089.

Myronenko, A., Song, X., 2010. Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. 32, 2262–2275.

Nüchter, A., Lingemann, K., Hertzberg, J., Surmann, H., 2007. 6d slam3d mapping outdoor environments. J. Field Robot. 24, 699–722.

Pomerleau, F., Colas, F., Siegwart, R., Magnenat, S., 2013. Comparing icp variants on real-world data sets. Autonomous Robots 34, 133–148.

Rabbani, T., Dijkman, S., van den Heuvel, F., Vosselman, G., 2007. An integrated approach for modelling and global registration of point clouds. ISPRS J. Photogramm. Remote Sens. 61, 355–370.

Raguram, R., Chum, O., Pollefeys, M., Matas, J., Frahm, J.-M., 2012. Usac: a universal framework for random sample consensus. IEEE Trans. Pattern Anal. Mach. Intell. 35, 2022–2038.

Raposo, C., Lourenço, M., Antunes, M., Barreto, J.P. 2013. Plane-based odometry using an rgb-d camera. In: BMVC.

Rousseeuw, P.J., Leroy, A.M., 2005. Robust regression and outlier detection, vol. 589 John wiley & sons.

Rusinkiewicz, S., Levoy, M. 2001. Efficient variants of the icp algorithm. In: 3dim. vol. 1, pp. 145–152.
Rusu, R.B., Blodow, N., Beetz, M., 2009. Fast point feature histograms (fpfh) for 3d registration. In: 2009 IEEE International Conference on Robotics and Automation. IEEE, pp. 3212–3217.
Rusu, R.B., Cousins, S., 2011. 3d is here: Point cloud library (pcl). In: In 2011 IEEE International Conference on Robotics and Automation. IEEE, pp. 1–4.
Salti, S., Tombari, F., Di Stefano, L., 2014. Shot: Unique signatures of histograms for surface and texture description. Comput. Vis. Image Underst. 125, 251–264.
Schnabel, R., Klein, R., 2006. Octree-based point-cloud compression. Spbg 6, 111–120.
Segal, A., Haehnel, D., & Thrun, S. (2009). Generalized-icp. In Robotics: science and systems (p. 435). Seattle, WA volume 2.
Serafin, J., Grisetti, G., 2015. Nicp: Dense normal based point cloud registration. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 742–749.
Stamos, I., Leordeanu, M. 2003. Automated feature-based range registration of urban scenes of large scale. In: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings. IEEE. vol. 2, pp. II–Ii.
Stoyanov, T., Magnusson, M., Andreasson, H., Lilienthal, A.J., 2012. Fast and accurate scan registration through minimization of the distance between compact 3d ndt representations. Int. J. Robot. Res. 31, 1377–1393.
Suwajanakorn, S., Snavely, N., Tompson, J.J., Norouzi, M. (2018). Discovery of latent 3d keypoints via end-to-end geometric reasoning. In: Advances in Neural Information Processing Systems, pp. 2059–2070.
Takeuchi, E., Tsubouchi, T., 2006. A 3-d scan matching using improved 3-d normal distributions transform for mobile robotic mapping. In: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, pp. 3068–3073.
Theiler, P.W., Wegner, J.D., Schindler, K., 2014. Keypoint-based 4-points congruent sets–automated marker-less registration of laser scans. ISPRS J. Photogramm. Remote Sens. 96, 149–163.
Torr, P.H., Zisserman, A., 2000. Mlesac: A new robust estimator with application to estimating image geometry. Comput. Vision Image Understand. 78, 138–156.
Tsin, Y., Kanade, T., 2004. A correlation-based approach to robust point set registration. In: European Conference on Computer Vision. Springer, pp. 558–569.
Turk, G., Levoy, M., 1994. Zippered polygon meshes from range images. In: Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques. ACM, pp. 311–318.
Ventura, G., Vilardo, G., Terranova, C., Sessa, E.B., 2011. Tracking and evolution of complex active landslides by multi-temporal airborne lidar data: the montaguto landslide (southern italy). Remote Sens. Environ. 115, 3237–3248.
Weinmann, M., Weinmann, M., Hinz, S., Jutzi, B., 2011. Fast and automatic image-based registration of tls data. ISPRS J. Photogramm. Remote Sens. 66, S62–S70.
Xiao, J., Adler, B., Zhang, J., Zhang, H., 2013. Planar segment based three-dimensional point cloud registration in outdoor environments. J. Field Robot. 30, 552–582.
Yang, B., Dong, Z., Liang, F., Liu, Y., 2016. Automatic registration of large-scale urban scene point clouds based on semantic feature points. ISPRS J. Photogramm. Remote Sens. 113, 43–58.
Yang, H., Carlone, L. 2019. A polynomial-time solution for robust registration with extreme outlier rates. arXiv preprint arXiv:1903.08588.
Yang, J., Li, H., Campbell, D., Jia, Y., 2015. Go-icp: A globally optimal solution to 3d icp point-set registration. IEEE Trans. Pattern Anal. Mach. Intell. 38, 2241–2254.
Zaharescu, A., Boyer, E., Varanasi, K., Horaud, R., 2009. Surface feature detection and description with applications to mesh matching. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 373–380.
Zeng, A., Song, S., Nießner, M., Fisher, M., Xiao, J., Funkhouser, T., 2017. 3dmatch: Learning local geometric descriptors from rgb-d reconstructions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1802–1811.
Zhang, J., Singh, S., 2015. Visual-lidar odometry and mapping: Low-drift, robust, and fast. In: 2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 2174–2181.
Zhong, Y., 2009. Intrinsic shape signatures: a shape descriptor for 3d object recognition. In: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops. IEEE, pp. 689–696.
Zhou, Q.-Y., Park, J., Koltun, V., 2016. Fast global registration. In: European Conference on Computer Vision. Springer, pp. 766–782.