Point Cloud Registration Based on One-Point RANSAC and Scale-Annealing Biweight Estimation

Jiayuan Li, Qingwu Hu, and Mingyao Ai

***Abstract*— Point cloud registration (PCR) is an important task in photogrammetry and remote sensing, whose goal is to seek a seven-parameter similarity transformation to register a pair of point clouds. Traditional iterative closest point (ICP) variants highly rely on the initial parameters, and most of them cannot deal with cross-source (multisource) point clouds with scale changes. In this article, we propose a flexible correspondence-based PCR method, which is initial-guess free, fast, and robust. We first decompose the full seven-parameter registration problem into three subproblems, i.e., scale, rotation, and translation estimations, based on line vectors. Then, we propose a one-point random sample consensus (RANSAC) algorithm to estimate the scale and translation parameters. For the rotation estimation, we introduce a graduated optimization strategy into Tukey's biweight function and propose a scale-annealing biweight estimator. We evaluate the proposed method on both same-source and cross-source data. Results show that the proposed method is robust against over 99% outliers and is one to two orders of magnitude faster than its competitors. The source code of our method will be made public.**

***Index Terms*— Biweight estimator, correspondence, cross-source (multisource), point cloud registration (PCR), random sample consensus (RANSAC).**

I. Introduction

THE 3-D measurement, such as light detection and ranging (LiDAR) scanning, is a basic and important technique, which has been widely applied in the fields of remote sensing, photogrammetry, and computer vision, such as 3-D city modeling [1], digital elevation model (DEM) generation [2], simultaneous localization and mapping (SLAM) [3], indoor mapping [4], and archeology [5], to name a few. Unfortunately, a single point cloud only covers a part of the scene due to viewpoint occlusions and limited filed-of-view (FoV). For example, the maximum measuring range of RIGEL VZ-400[1] scanner is 600 m, and its vertical FoV is 100°. To cope with this problem, multiple 3-D scans recorded from different viewpoints are required to cover the whole scene. Similar to image-based 3-D reconstruction, these scans form a scan sequence where each adjacent point cloud pair has overlaps. Point cloud registration (PCR) can merge these individual scans into a 3-D panorama, whose basic idea is to seek optimal similarity transformations to convert point clouds with local coordinate systems into the same reference system [6].

Iterative closest point (ICP) [7] is probably the most popular PCR method, which has become a de-facto standard for engineering solutions. However, the performance of ICP highly relies on the initializations for rigid transformations since it establishes correspondences based on the nearest distance in each iteration and only converges to a local minimum. In practice, engineering solutions require additional information to improve the robustness of ICP. For example, RIEGL airborne LiDAR equips a GNSS/INS navigation system to provide real-time pose for accurate point cloud fusion; ground laser LiDARs, such as Z+F IMAGER[2] scanner, use high-reflection targets as control points to estimate initial solutions. However, GNSS/INS systems are expensive, and the control point layout is labor-intensive. Moreover, ICP cannot deal with cross-source point clouds with scale changes, such as multiview stereo (MVS) and LiDAR point cloud pairs [8], [9]. The MVS-LiDAR registration can reconstruct more realistic and accurate 3-D scenes. LiDAR provides accurate geometric information, while MVS provides rich spectral information. Both geometric information and spectral information are important for scene interpretation. For example, LiDAR-MVS fusion can improve the accuracies of object detection, classification, semantic segmentation, and so on.

Due to the development of 3-D keypoint technology, correspondence-based PCR methods have received more and more attention because they are initialization-free, cheap, and flexible. Unlike ICP, these methods preestablish correspondences by a 3-D feature matching technique, such as fast point feature histogram (FPFH) [10] algorithm. Then, a geometric transformation (seven-parameter similarity model) is fitted via robust estimation. Perhaps, random sample consensus (RANSAC) [11] is the most widely used method for robust model fitting. Unfortunately, 3-D keypoint matching [12] is

Manuscript received June 26, 2020; revised October 30, 2020; accepted December 9, 2020. Date of publication January 8, 2021; date of current version October 26, 2021. This work was supported in part by the Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, under Grant 18E02 and in part by the State Key Laboratory of Rail Transit Engineering Informatization (FSDI) under Grant SKLK19-06. *(Corresponding author: Qingwu Hu.)*

The authors are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China (e-mail: ljy_whu_2012@whu.edu.cn; huqw@whu.edu.cn).

Digital Object Identifier 10.1109/TGRS.2020.3045456

[1]http://www.riegl.com/products/

[2]https://www.zf-laser.com

much more challenging than 2-D feature matching, such as scale-invariant feature transform (SIFT) [13], speeded up robust features (SURFs) [14], and radiation-variation insensitive feature transform (RIFT) [15], because of the problems of texture-less, noise, and density variations, which results in extremely high outlier rates in putative correspondences. As pointed out in [16], the outlier rates of initial correspondences are often higher than 95%. At such high outlier rates, RANSAC and its variants require a huge number of subset sampling trials to obtain a satisfactory result, which dramatically decreases their efficiency. For instance, to obtain an outlier-free minimal size subset with 0.99 confidence for seven-parameter registration at an outlier rate of 99%, RANSAC theoretically requires 4 605 168 trials (please see Section III-B for details), which is too huge to be practical.

In this article, we aim to design a PCR method with the following properties: 1) initialization-free; 2) high efficiency; 3) high robustness (robust against over 99% outliers); and 4) seven-parameter registration (suitable for cross-source point clouds with scale changes). We extend the concept of line vector to cope with scale changes and use it to decompose the full seven-parameter registration problem into three subproblems, which largely reduces the parameter space. This decomposition lays a foundation for the subsequent model estimation, which makes the seven degree-of-freedom (DoF) registration problem become much easier. We propose a one-point RANSAC for scale and translation estimations, which dramatically reduces the computational complexity. Our one-point RANSAC does not require any prior information since the transformation decomposition does not rely on other sensor information or assumptions. To our best knowledge, one-point RANSAC has never been used in PCR before. We also introduce a graduated optimization strategy into the biweight function to solve the rotation problem, which is fast and robust. These improvements guarantee the high robustness and high efficiency of the proposed method. The proposed method is a correspondence-based method, which does not require initial parameters to establish correspondences, such as ICP-type methods. The estimations of scale, rotation, and translation also do not rely on initial solutions. Thus, it is an initialization-free method, unlike ICP. We conduct two real experiments, i.e., same-source registration and cross-source registration. Both experiments show that the proposed method is much superior to other compared baselines and state of the arts, especially on cases with extremely high outlier rates.[3]

The remainder of this article is organized as follows. Section II briefly reviews some related techniques. Section III describes the details of the proposed PCR method. The experimental setup, data sets, performance metrics, and result discussions are provided in Section IV. Section V presents the conclusions.

II. Related Work

In this section, we briefly review some techniques that are highly related to PCR, including 3-D keypoint matching, correspondence-based registration, and point-based registration methods. More comprehensive surveys can be found in [17] and [18].

A. 3-D Keypoint Matching

The same as image feature matching, 3-D keypoint matching also contains three main steps: detection, description, and matching.

1) Keypoint Detectors: Detectors usually analyze the distribution properties of local surfaces to identify distinctive points (keypoints), which can be grouped into two categories, i.e., handcrafted detectors and learning-based ones. Handcrafted detectors are mostly driven by mathematical analysis. The representative methods are local surface patches (LSP) [19], intrinsic shape signatures (ISSs) [20], Mesh-DoG [21], and so on. Differently, learning-based ones are driven by data, e.g., KeypointNet [22], unsupervised stable interest point (USIP) [23], and salient keypoint detection (SKD) [24].

2) Keypoint Description: Descriptors encode local patches of keypoints into feature vectors so that matching patches and nonmatching patches can be easily distinguished. Unique shape context (USC) [25], binary shape context (BSC) [26], [27], FPFH [10], and rotational projection statistics (RoPS) [28] are representative handcrafted descriptors. Recently, learning-based descriptors (such as 3DMatch [29], PPFNet [30], and 3DSmoothNet [31]) become popular, which has shown potentials compared with their handcrafted counterparts.

3) Feature Matching: Establish one-to-one corresponding relationship based on matching scores between feature vectors of keypoints, e.g., the chi-square test method [20] and nearest neighbor distance ratio strategy [13].

B. Correspondence-Based Registration

1) RANSAC-Family: RANSAC is the most widely used robust fitting method in computer vision, whose basic idea relies on a hypothesize-and-verify technique. RANSAC alternately performs minimal subset sampling and model fitting until the stopping criterion is reached. It has many variants [32]–[41]. For example, maximum likelihood estimation sample consensus (MLESAC) [32] generalizes RANSAC based on a probabilistic interpretation. Group sample consensus (GroupSAC) [34] uses a prior that inliers are more similar to each other to improve the sampling stage. Locally optimized RANSAC (LORANSAC) [33] and fixed LORANSAC (FLORANSAC) [36] do not agree with the assumption that a model fitted by an outlier-free subset is consistent with all inliers. They locally refine the estimated model based on the current inlier set. Chum and Matas [35] proposed a model verification strategy called modified sequential probability ratio (SPR) test to reduce the time complexity. Universal RANSAC (USAC) [37] incorporates many practical tricks, such as sampling, model checking, model verification, and model refinement into a common framework. Marginalizing sample consensus (MAGSAC++) [39] proposes a scoring function that the inlier threshold parameter is not required

[3]https://www.researchgate.net/profile/Jiayuan_Li2

and fits the model by iteratively reweighted least-squares (IRLS) [42]. Differentiable sample consensus (DSAC) [40] uses a probabilistic-based hypothesis selection to learn good samples in an end-to-end manner. Neural-guided RANSAC (NG-RANSAC) [41] uses prior information to guide hypothesis sampling and combines this strategy with DSAC to build neural networks. As mentioned earlier, the major limitation of these methods is that they require a huge number of iterations to seek a satisfactory solution under high outlier rates.

2) M-Like Estimators: M-like estimators penalize outliers by giving them very small weights based on robust cost functions [43]–[50]. Unfortunately, traditional M-estimators and S-estimators are sensitive to outlier rates. They cannot tolerate more than 50% outliers. Recently, several robust methods that are robust to high outlier rates are proposed. For example, Zhou *et al.* [47] introduced the Geman-McClure estimator loss for PCR and optimized this robust cost based on line processes. Similar to our method, truncated least-squares estimation and semidefinite relaxation (TEASER++) first decouples the registration problem into subproblems. Differently, it uses a truncated least-squares (TLS) cost for each subproblem estimation. The complexity of scale estimation in TEASER++ is $O(N^4)$, which is too slow to be practical.

C. Point-Based Registration

1) ICP-Family: ICP [7] is a milestone for PCR, which simultaneously establishes correspondences and estimates relative pose between point clouds. Specifically, ICP first finds point correspondences based on the criterion of nearest distance; then, a rigid transformation is fitted by using these correspondences. ICP alternates these two stages until converging to a local minimum. Many variants are proposed to improve one or more subtasks (sampling, error metric, outlier rejection, and optimality) of ICP. The sizes of point clouds are huge, an appropriate sampling, such as voxel-grid filtering [51] and octree compression [52], can largely improve the efficiency. More accurate error metrics, such as point-to-line [53] metric for 2-D point clouds and point-to-plane [54] and plane-to-plane [55] metrics for 3-D point clouds, are introduced into ICP. Original ICP is sensitive to partial overlaps, which leads to outliers in the correspondence establishment stage. To improve the robustness to partial overlaps, many methods introduce a robust estimation technique in the pose estimation stage. For example, least trimmed squares (LTS) robust estimator is used in Trimmed ICP [56] and Anisotropic ICP [57]; a sparse cost function is introduced by Sparse ICP [58]; and weighted l_q-norm ICP [6] adapts a weighted l_q estimator with high robustness. Yang *et al.* [59] proposed a global optimal ICP (Go-ICP), which uses a branch-and-bound (BnB) algorithm to globally search the rotation and translation. As aforementioned, ICP-family highly relies on the initializations and is not suitable for cross-source data with scale changes.

2) Other Methods: A representative work is four-point congruent sets (4PCS) [60], which uses affine-invariant intersection ratios to find coplanar four-point correspondence bases and estimates an optimal rigid model based on these selected bases. Various variants of 4PCS have been presented. For example, Super 4PCS (S4PCS) [61] introduces a smart index to organize data and converts the original 4PCS as an instance problem to significantly reduce its computational complexity. Ge [62] proposed a nonrigid variant to cope with isometric deformations. Keypoint 4PCS (K4PCS) [63] extracts 3-D DoG and Harris keypoints as the input of 4PCS. However, their computational complexities are still very high at high outlier rates.

III. Methodology

A. Problem Formulation

Given a set of 3-D correspondences $M = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_1^N$ without outliers, where $\boldsymbol{x}_i,\ \boldsymbol{y}_i \in \mathbb{R}^3$, such that the geometric relationship between $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ can be exactly modeled by

$$\boldsymbol{y}_i = s\mathbf{R}\boldsymbol{x}_i + \boldsymbol{t} + \boldsymbol{n}_i \tag{1}$$

where $s > 0$ is a scale factor, $\mathbf{R} \in SO(3)$ is an orthogonal rotation matrix; $\boldsymbol{t} \in \mathbb{R}^3$ is a translation vector; and $\boldsymbol{n}_i$ represents measurement noise, which can be assumed that it is unknown but bounded. However, there must be outliers in the correspondence set M obtained by a 3-D keypoint matching technique. Correspondence-based PCR should be formulated as a robust estimation problem

$$\underset{s,\boldsymbol{R},\boldsymbol{t}}{\text{minimize}} \sum_{i=1}^{N} \rho\big(\parallel \boldsymbol{y}_i - (s\mathbf{R}\boldsymbol{x}_i + \boldsymbol{t}) \parallel\big) \tag{2}$$

where $\rho(\cdot)$ is a robust function and $\|\cdot\|$ is the l_2-norm.

B. Transformation Decomposition

1) Motivation: Probably, RANSAC is the most widely used technique for solving problem (2). It is reliable and efficient in cases with low outlier rates. However, its computational complexity increases exponentially as outlier rate increases. The minimal number of trials N_T required by RANSAC to generate at least one good subset with purely of inliers is computed as follows:

$$N_T = \left\lceil \frac{\log(1-p)}{\log\big(1-(1-p)^m\big)} \right\rceil \tag{3}$$

where p is the confidence of good subsets, which is generally set to 0.99; m is the size of minimal subsets ($m = 3$ for seven-parameter registration); and $\lceil\rceil$ is a ceiling function. At an outlier rate of 99%, RANSAC requires at least 4 605 168 trials to get a probably correct solution for registration problem, which is too large to be practical. Fortunately, if we can decompose the full seven-parameter registration problem into scale, rotation, and translation estimation subproblems, the values of m become small, i.e., $m = 1$ for scale and translation estimations and $m = 2$ for rotation estimation. Then, at the same outlier rate, RANSAC only requires 459 trials and 46 050 trials for $m = 1$ and $m = 2$, which are the 1/10 000 and 1/100 of the original value, respectively. In our method, we use a scale-annealing biweight estimator instead of two-point RANSAC for rotation estimation, which further speeds up the optimization.

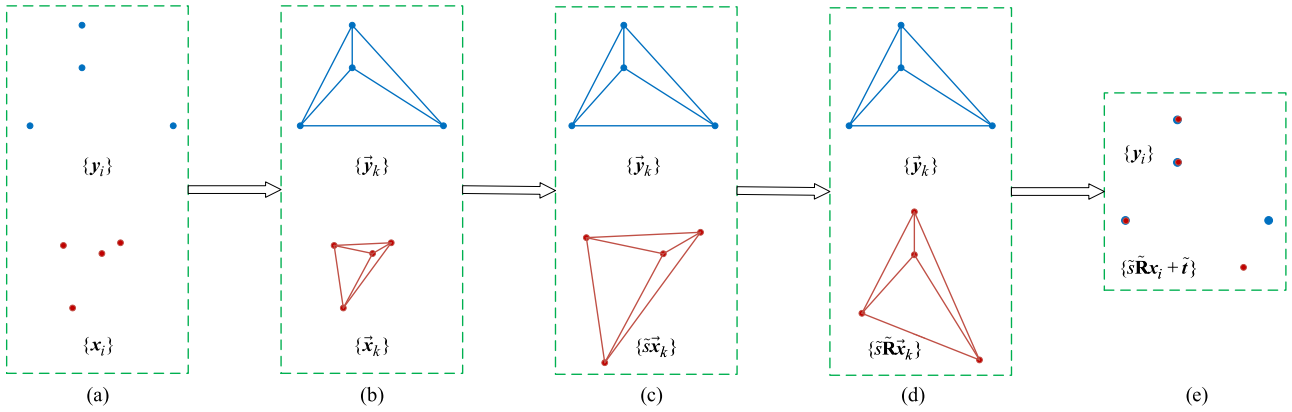

Fig. 1. Illustrating the framework of our proposed seven-parameter PCR method. (a) Given a set of 3-D correspondences $\{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_1^N$, (b) we first construct a set of line vector correspondences $\{(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{y}}_k)\}_1^K$, which are used for scale estimation based on model (7). (c) Scale difference is eliminated according to $\tilde{s}$. (d) Then, the rotation is estimated based on model (4). (e) Finally, a translation estimation step is performed to register these two point sets.

2) Line Vector: In our previous work [6], we present an edge line vector to decompose a six DoF rigid transformation into two 3-DoF subproblems. In this article, we extend this line vector to cope with scale differences and use it to decompose the full seven-parameter registration problem.

Specifically, given a pair of point correspondences $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ and $(\boldsymbol{x}_j, \boldsymbol{y}_j)$, we can construct a line vector correspondence $(\vec{\boldsymbol{x}}_{ij} = \boldsymbol{x}_i - \boldsymbol{x}_j, \vec{\boldsymbol{y}}_{ij} = \boldsymbol{y}_i - \boldsymbol{y}_j)$. If both $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ and $(\boldsymbol{x}_j, \boldsymbol{y}_j)$ are inliers, we have

$$\begin{aligned}\vec{\boldsymbol{y}}_{ij} &= s\mathbf{R}(\boldsymbol{x}_i - \boldsymbol{x}_j) + (\boldsymbol{n}_i - \boldsymbol{n}_j) \\ &= s\mathbf{R}\vec{\boldsymbol{x}}_{ij} + \vec{\boldsymbol{n}}_{ij}\end{aligned} \tag{4}$$

where $\vec{\boldsymbol{n}}_{ij} = \boldsymbol{n}_i - \boldsymbol{n}_j$ is a noise vector. If the noise bound is τ, then $\|\vec{\boldsymbol{n}}_{ij}\| \leq 2\tau$. It is clear that the translation term is eliminated in (4). Thus, the line vector is invariant to translation and only related to scale and rotation.

Furthermore, the length of a line vector is invariant to both rotation and translation. Thus, we take the l_2-norm operation on (4) and obtain

$$\|\vec{\boldsymbol{y}}_{ij}\| = \|s\mathbf{R}\vec{\boldsymbol{x}}_{ij} + \vec{\boldsymbol{n}}_{ij}\|. \tag{5}$$

According to the triangle inequality, the above equation can be reformulated as

$$-\|\vec{\boldsymbol{n}}_{ij}\| \leq \|\vec{\boldsymbol{y}}_{ij}\| - \|s\mathbf{R}\vec{\boldsymbol{x}}_{ij}\| \leq \|\vec{\boldsymbol{n}}_{ij}\|. \tag{6}$$

To eliminate the rotation term, (6) is divided by $\|\vec{\boldsymbol{x}}_{ij}\|$, which leads to

$$|s_{ij} - s| \leq \tau_{ij} \tag{7}$$

where $s_{ij} = (\|\vec{\boldsymbol{y}}_{ij}\|)/(\|\vec{\boldsymbol{x}}_{ij}\|)$ is the scale difference between lines $\vec{\boldsymbol{x}}_{ij}$ and $\vec{\boldsymbol{y}}_{ij}$, and $\tau_{ij} = (2\tau)/(\|\vec{\boldsymbol{x}}_{ij}\|)$. Equation (7) is a scalar function, where the scale s is the only unknown.

Therefore, the full seven-parameter registration problem is divided into three subproblems, as illustrated in Fig. 1. First, we estimate the optimal scale parameter $\tilde{s}$ based on the model (7). Then, the scale difference is eliminated according to $\tilde{s}$, and the optimal rotation $\tilde{\mathbf{R}}$ is estimated based on the model (4). Finally, we put $\tilde{s}$ and $\tilde{\mathbf{R}}$ into problem (1) and seek the optimal translation $\tilde{\boldsymbol{t}}$. Note that the operations on line vectors are the same as the ones on points since a line vector can also be treated as a 3-D-coordinate vector.

C. Scale Estimation

In fact, robust scale estimation based on model (7) is equivalent to a maximum consensus problem, which seeks a scalar scale that maximizes the number of inliers

$$\begin{aligned}&\max_{s, I^s \subseteq H} |I^s| \\ &\text{s.t.} \quad \frac{|s_k - s|}{\tau_k} \leq 1 \quad \forall k \in I^s\end{aligned} \tag{8}$$

where, for simplicity, we use subscript k instead of ij, $\mathcal{H} = \{1, 2, \ldots, K\}$ is an index set of the scale observations $\{s_k\}_1^K$, $\{\tau_k\}_1^K$ is a set of inlier thresholds, subset I^s is a consensus set that consists of inliers, and $|I^s|$ denotes the number of inliers (or called the size of inlier set). The optimal scale $\tilde{s}$ corresponds to the largest inlier consensus set $\tilde{I}^s$. Hereafter, the indices and their corresponding data are treated as equivalent.

This problem can be solved by a straightforward enumeration method. Specifically, for each s_k, we regard it as the solution of the above problem and calculate its corresponding consensus set. Then, we find the consensus set $\tilde{I}^s$ with the largest size and use $\tilde{I}^s$ to estimate an optimal scale $\tilde{s}$ based on the least-squares cost

$$\tilde{s} = \arg\min_s \sum_{k \in \tilde{I}^s} \left(\frac{s_k - s}{\tau_k}\right)^2. \tag{9}$$

The solution is

$$\tilde{s} = \left(\sum_{k \in \tilde{I}^s} \frac{1}{\tau_k^2}\right)^{-1} \sum_{k \in \tilde{I}^s} \frac{s_k}{\tau_k^2}. \tag{10}$$

This enumeration method suffers from high computational complexity, which is not acceptable in practice. For a correspondence set with N points, we can construct $K = (N(N-1))/(2)$ line vectors. Hence, the computational complexity of the enumeration method is $O(K^2) = O(N^4)$. To reduce complexity, we present a one-point RANSAC

Algorithm 1 One-Point RANSAC for Scale Estimation

Input: $\{s_k\}_1^K$ and $\{\tau_k\}_1^K$.
Initialize: $p = 0.99$, $i =: 0$, $N_T := 10^5$, and $|\tilde{I}^s| := 0$.
while $i \leq N_T$ **do**
  Randomly pick a scale s_i from $\{s_k\}_1^K$.
  Find a consensus set I_i^s according to (8).
  if $|I_i^s| > |\tilde{I}^s|$ **then**
    Estimate a refined scale s_i' according to (10).
    Find a refined consensus set $I_i'^s$.
    $|\tilde{I}^s| := |I_i^s|$, $\tilde{I}^s := I_i^s$, $\tilde{s} := s_i'$.
    Update N_T according to (3).
  end if
  $i := i + 1$;
end while
Output: the optimal scale $\tilde{s}$ and the inlier set $\tilde{I}^s$.

algorithm for robust scale estimation. In fact, one-point RANSAC means that the size of minimal subsets is $m = 1$, which had been used in pose estimation [64] and visual odometry [65]–[67]. These methods highly rely on prior information to reduce the DoF of parameter space, which limits their applications. For example, Scaramuzza [67] proposed a one-point RANSAC algorithm for motion estimation. However, it is only suitable for cases that the camera is installed on a nonholonomic wheeled vehicle. Civera *et al.* [66] used the prior probabilistic information from the extended Kalman filter (EKF) to reduce the DoF. Lee *et al.* [64] used a ground object assumption and a 2-D object bounding box as additional observations. Differently, our method does not require any prior information since the transformation decomposition does not rely on other sensor information or assumptions. Moreover, to the best of our knowledge, one-point RANSAC has never been used in PCR before.

We also introduce a local optimization strategy in the one-point RANSAC to refine the solution. First, we randomly sample a scale s_i from $\{s_k\}_1^K$ and find a consensus set I_i^s. Then, I_i^s is used to estimate a refined scale s_i' based on (10), and a refined consensus set $I_i'^s$ is calculated according to s_i'. These two steps are alternated until the stopping criterion is reached, and the scale with the largest consensus set is accepted as the optimal solution $\tilde{s}$. The details of the one-point RANSAC are summarized in Algorithm 1.

D. Rotation Estimation

In the scale estimation step, a large portion of outlying line vector correspondences can be filtered. Namely, the inlier set $\tilde{I}^s$ can be used for rotation estimation, which has a much lower outlier rate (generally lower than 80%). Based on this observation, we propose a generalized M-estimation called scale-annealing biweight estimator. It can tolerate 80% $\sim$ 90% outliers and is much faster than a two-point RANSAC algorithm. The objective function of the proposed estimator is

$$\underset{\mathbf{R}}{\text{minimize}} \sum_{(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{y}}_k) \in \tilde{I}^s} \rho\big(\| \vec{\boldsymbol{y}}_k - \tilde{s}\mathbf{R}\vec{\boldsymbol{x}}_k \|\big). \tag{11}$$

Algorithm 2 Scale-Annealing Biweight for Rotation Estimation

Input: $\tilde{s}$ and $\tilde{I}^s$.
Initialize: $u := 1000$, $\alpha = 1.3$, and $\{w(r_k)\} := 1$.
while not converged **do**
  Formulate a WLS problem according to (13).
  Estimate a solution $\mathbf{R}$ via the SVD.
  Calculate residuals $\{r_k\}$ for $\tilde{I}^s$.
  Update weights $\{w(r_k)\}$ according to (14).
  Anneal the scale by $u := \frac{u}{\alpha}$.
end while
Compute $\tilde{\mathbf{R}}$ via the SVD with the newest weights.
Output: the optimal rotation $\tilde{\mathbf{R}}$.

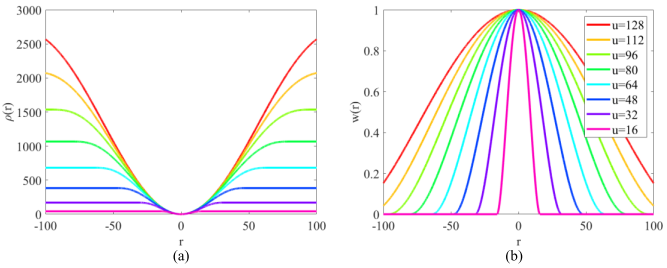

Fig. 2. Visualization of the proposed scale-annealing biweight cost. (a) Its weight function. (b) With different scale parameters.

If $(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{y}}_k)$ is an inlier, then $r_k = \| \vec{\boldsymbol{n}}_k \| \leq 2\tau$, where $r_k = \| \vec{\boldsymbol{y}}_k - \tilde{s}\mathbf{R}\vec{\boldsymbol{x}}_k \|$ is the residual of $(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{y}}_k)$; otherwise, the residual can be $\gg 2\tau$. It means that the energy is dominated by outliers. The robust function $\rho(r_k)$ reflects the influence of r_k toward the total optimization energy, which is critical to robust estimation problems. A good robust function should have a property that minimizes the distances between correspondences while disabling spurious data. In our method, we adapt a scaled biweight function

$$\rho(r_k, u) = \begin{cases} \dfrac{u^2}{6}\left(1 - \left(1 - \dfrac{r_k^2}{u^2}\right)^3\right), & |r_k| \leq u \\ \dfrac{u^2}{6}, & |r_k| > u \end{cases} \tag{12}$$

where u is a scale parameter, which controls the shape of function ρ. Fig. 2(a) visualizes the biweight functions with different scales. As shown, the larger the scale u, the greater the impact of outliers on the total energy.

Instead of minimizing the function ρ directly, it is simpler to solve the following weighted least-squares (WLS) problem:

$$\underset{\mathbf{R}}{\text{minimize}} \sum_{(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{y}}_k) \in \tilde{I}^s} w(r_k) \| \vec{\boldsymbol{y}}_k - \tilde{s}\mathbf{R}\vec{\boldsymbol{x}}_k \|^2 \tag{13}$$

where $w(r_k) = (\partial\rho)/(\partial r_k)/r_k$ is a weight. $(\partial\rho)/(\partial r_k)$ is the derivative of ρ with respect to the residual r_k. The weight function of robust cost $\rho(r_k, u)$ is

$$w(r_k, u) = \begin{cases} \left(1 - \dfrac{r_k^2}{u^2}\right)^2, & |r_k| \leq u \\ 0, & |r_k| > u. \end{cases} \tag{14}$$

The weight functions with different scales are displayed in Fig. 2(b). As shown, the larger the scale, the smoother

TABLE I
DETAILED SETTINGS OF THE COMPARED ALGORITHMS (MNI REPRESENTS MAXIMUM NUMBER OF ITERATIONS)

Method	Parameters	Implementations	Input	Complexity
FLORANSAC	$m = 3$; $p = 0.99$; LO size: 21; LO iterations: 50; $N_T = 10^5$.	MATLAB code; single thread https://zhipengcai.github.io/	Correspondences	$O(N_T N)$
K4PCS	Inlier ratio: 0.5; score bound: 10^{-3}; maximum time: 10^3s.	C++ code; single thread https://pointclouds.org/	ISS keypoints	$O(N_{kp}^2 + L)$
FMP+BnB	degrees of freedom: 4; Lower bound: 0; repeat?: true.	C++ code; single thread https://github.com/ZhipengCai/Demo—Practical-optimal-registration-of-terrestrial-LiDAR-scan-pairs	Correspondences	$O(log(N)N^2)$
TEASER++	GNC factor: 1.4; max clique?: false; kcore heuristic threshold: 0.5; MNI: 100	C++ code; single thread https://github.com/MIT-SPARK/TEASER-plusplus	Correspondences	$O(N^4)$
Ours	$m = 1$; $p = 0.99$; $N_T = 10^5$; $\alpha = 1.3$.	MATLAB code; single thread http://www.escience.cn/people/lijiayuan/index.html	Correspondences	$O(N_T^s N^2)$

the function curve. Thus, a weight function with a large u allows more observations to take part in the optimization stage, which largely avoids that inliers are disabled by an inaccurate solution of $\mathbf{R}$. As u decreases, the estimated $\mathbf{R}$ becomes more and more precise. Then, large weights are assigned to data with small residuals, while small weights (close to 0) are given to outliers, which disables outliers in the optimization.

Based on this property, we use a graduated optimization strategy instead of fixing the scale in the IRLS. Specifically, the initial scale u is assigned a large value, e.g., $u = 1000$. Then, the scale is decreased by a step-size α, i.e., $u := (u/\alpha)$, along with iterations. In each iteration, we use a singular value decomposition (SVD) algorithm to solve a solution for $\mathbf{R}$. This graduated optimization strategy can largely avoid that the solver gets stuck in local minima compared with fixed-scale IRLS. The proposed scale-annealing biweight estimator for robust rotation estimation is summarized in Algorithm 2.

E. Translation Estimation

We can obtain an inlier line vector set In based on the optimal scale $\tilde{s}$ and rotation $\tilde{\mathbf{R}}$

$$\text{In} = \left\{ (\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{y}}_k) | \parallel \vec{\boldsymbol{y}}_k - \tilde{s}\tilde{\mathbf{R}}\vec{\boldsymbol{x}}_k \parallel \leq 2\tau;\ k = 1, \ldots, K' \right\}. \quad (15)$$

Since a line vector corresponds to two points in set M, we project the line vector set In to original point space and count the frequency of each point. The higher the frequency of a point, the larger the probability that it is an inlier. Thus, we sort these points according to the frequency and select the best 70% point correspondences M' as potential inliers to estimate the translation. Compared with M, M' has a much lower outlier rate, which speeds up the following optimization.

Essentially, robust translation estimation has the same cost function with scale estimation. It is also a maximum consensus problem

$$\begin{aligned} &\max_{\boldsymbol{t}, I^t \subseteq M'} |I^t| \\ &\text{s.t.} \ \parallel \boldsymbol{t}_i - \boldsymbol{t} \parallel \leq 2\tau \quad \forall i \in I^t \end{aligned} \quad (16)$$

where $\boldsymbol{t}_i = (\boldsymbol{y}_i - \tilde{s}\tilde{\mathbf{R}}\boldsymbol{x}_i)$. We also use the one-point RANSAC to optimize this problem. The process is almost the same as in Algorithm 1. The least-squares solution of $\boldsymbol{t}$ is just the mean of the consensus set.

F. Computational Complexity

In our method, line vector construction consumes $O(N^2)$ time. The complexity of scale estimation is $O(N_T^s N^2)$, where N_T^s is the minimal number of trials required by RANSAC in scale estimation. Rotation estimation costs $O(|\tilde{I}^s|)$ time ($|\tilde{I}^s| < N^2$). The complexity of translation estimation is smaller than $O(N_T^t N)$, where N_T^t is the minimal number of trials required by RANSAC in translation estimation. Thus, the scale estimation is the most time-consuming step, and the total computational complexity of our method can be simplified as $O(N_T^s N^2)$. If the correspondence set is large, then $N^2 \gg N_T^s$, and the complexity becomes $O(N^2)$.

G. Multiview Scans Registration

Directly using the proposed method to register multiple scans is difficult because of the RANSAC framework. There are two ways to achieve the purpose. First, we can perform pairwise registration to obtain initial poses. These poses are used to initialize a multiple-scans joint registration method, such as in [68]. Second, we can use pairwise registration to construct a pose-graph and optimize this pose-graph based on an SLAM back-end toolbox, such as g^2o [69].

IV. EXPERIMENTS

This section comprehensively evaluates the proposed method on real experiments. First, we evaluate the proposed method on same-source LIDAR point clouds, which do not suffer from scale changes (in this experiment, scale estimation is disabled). Second, our method is assessed on cross-source point clouds with unknown scales, i.e., MVS and structured light point-cloud pairs. Our method is compared with several baselines and state of the arts, including FLORANSAC [36], K4PCS [63], fast match pruning with BnB (FMP+BnB) [70], and TEASER++ [49]. The information of parameters, implementation details, inputs, and time complexity of each compared method is summarized in Table I, where L denotes the number of candidate congruent four-points and N_{kp} is the number of keypoints. $N_T \gg N$ at high outlier rates.

We use the ISS algorithm to detect keypoints and FPFH algorithm for description. A correspondence $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ is selected into the initial correspondence set M only if $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$

TABLE II
DETAILED INFORMATION OF DATA SETS (K HERE MEANS THOUSANDS)

Info	Andreashaus (Mean±Std)	Bremen_city (Mean±Std)	Campus (Mean±Std)
N_{pr}	14	12	73
Noise τ	0.02m	0.1m	0.1m
N_{pt}	3122K±1014K	1295K±523K	1232K±301K
N_{kp}	5919±1433	6335±2549	6988±1021
N	8644±2827	8670±2938	9865±1162
Overlap	45%±18%	49%±13%	75%±7%
$r_{outlier}$	99.39%±0.62%	98.74%±0.47%	98.98%±0.52%

are one of the top-k best matches ($k = 5$ in our experiments) to each other. We regard the average point cloud resolution as the noise bound τ. We use common metrics for quantitative assessment, i.e., translation error E_t and rotation error E_R

$$\begin{cases} E_t = \| \boldsymbol{t}^t - \boldsymbol{t}^e \| \\ E_R = \arccos \dfrac{\mathrm{tr}\left(\boldsymbol{R}^t (\boldsymbol{R}^e)^{\mathrm{T}}\right) - 1}{2} \end{cases} \tag{17}$$

where superscript t and superscript e represent the ground-truth (GT) one (scale, rotation, or translation) and the estimated one, respectively; tr(·) is the trace of a matrix. For the second experiment, the scale difference $E_s = |s^e - s^t|$ is also calculated for evaluation. All the results are calculated on a laptop with a single CPU Core i7-8550U at 1.8 GHz and 8 GB of RAM.

A. Same-Source Registration

1) Data Set: We use three challenging real data sets for evaluation, i.e., Andreashaus, Bremen_city, and Campus data sets.[4] The Andreashaus is a middle-scale indoor data set that was collected at 15 different poses inside a residential house. The other two are large-scale outdoor data sets, where the Bremen_city with 13 LIDAR scans was taken in the city center of Bremen and the Campus with 74 LIDAR scans was taken on the campus of the Jacobs University Bremen. All these data sets were recorded using a Riegl VZ-400 LIDAR scanner. Since the sizes of original LIDAR scans are huge (more than tens of millions of points in each scan), the voxel-grid filter is applied to downsample the resolutions of these point clouds. The detailed information about these three LiDAR data sets is summarized in Table II, including the number of scan pairs N_{pr}, point cloud resolution (it is regarded as the noise bound), the average number of points N_{pt}, the average number of keypoints N_{kp}, the average number of feature correspondences N, average overlapping ratio, and average outlier ratio r_{outlier} in the initial correspondences M. Each scan pair of these data sets is assigned an approximate GT rigid transformation $(\boldsymbol{R}^t, \boldsymbol{t}^t)$, which is obtained by a coarse-to-fine registration strategy. Specifically, scans are first manually registered by using reflective markers. Then, the 3-D Toolkit[5] (a graph-based pose optimization algorithm) is used to refine the results of manual registration and provide accurate registration parameters. Thus, if a correspondence $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ satisfies $\|\boldsymbol{y}_i^t - (\boldsymbol{R}^t \boldsymbol{x}_i + \boldsymbol{t}^t)\| \leq 2\tau$, it is an inlier; otherwise, it is an outlier. As shown in Table II, the lowest r_{outlier} is still higher than 98%, which makes the correspondence-based or keypoint-based registration very difficult.

2) Qualitative Assessment: One scan pair is selected from each data set for qualitative evaluation. The pair from Andreashaus and the one from Bremen_city have very low overlapping ratios, which are only 25.72% and 33.18%, respectively. The overlap of the pair from Campus is relatively large, i.e., 71.52%. All these pairs are partially overlapping. Therefore, they are challenging for K4PCS. The outlier rates of these three pairs are 99.83%, 99.51%, and 99.56%, respectively. Such high outlier rates bring great challenges to current correspondence-based registration methods, such as RANSAC-type methods and TEASER++. The comparison results are displayed in Fig. 3.

As shown, none of these four compared methods obtain satisfactory results on all the three scan pairs. FLORANSAC and TEASER++ cannot handle cases with extremely high outlier rates, e.g., $>$ 99%. Since the outlier rates of initial correspondences are extremely high, they are also extremely high in the keypoints. K4PCS is sensitive to the outlier rates of keypoints. FMP+BnB performs better than the aforementioned methods. However, FMP+BnB only solves a 4-DoF registration problem. Its registration accuracy is not very high since it relies on an accurate level compensator. As a result, it is not suitable for the registration of point clouds captured by handheld laser scanners. In contrast, the proposed method achieves good alignments on all these scan pairs. Our results are very close to the GT, i.e., the rotation accuracy is better than 0.3° and the translation errors are smaller than 0.05 m. Our method is able to tolerate more than 99% outliers, which is more robust than FLORANSAC and TEASER++.

In fact, each data set is a LIDAR scan sequence, which can be sequentially registered to reconstruct the 3-D scene. Thus, we first estimate a rigid transformation for each of the two adjacent scans. Then, we regard the first scan as the base and transform other scans into the basic coordinate system based on the estimated transformations. This process is called laser odometry. The Campus data set consists of 74 LIDAR scans. To reduce error accumulation in Campus, we treat ten scans as a group and use our laser odometry to get a submap. Then, we use the proposed method to register consecutive submaps to obtain the final 3-D scene map. The reconstructed results are displayed in Fig. 4. As shown, our results are impressive. There is almost no "ghosting" in all 3-D maps. We calculate the mean absolute error (MAE) between our estimated trajectory and the one of the 3-D Toolkit for evaluation. The MAEs of the Andreashaus, Bremen_city, and Campus are 0.05, 0.12, and 0.18 m, respectively. Our method can be a good scan matcher of the front-end of a 3-D laser SLAM.

3) Quantitative Assessment: Fig. 5 reports ten individual quantitative results of each data set. FLORANSAC performs better on the Bremen_city and Campus than on the Andreashaus since the outlier rate of Andreashaus is higher. FLORANSAC generally cannot seek a satisfactory solution

[4] http://kos.informatik.uni-osnabrueck.de/3Dscans/
[5] http://slam6d.sourceforge.net/

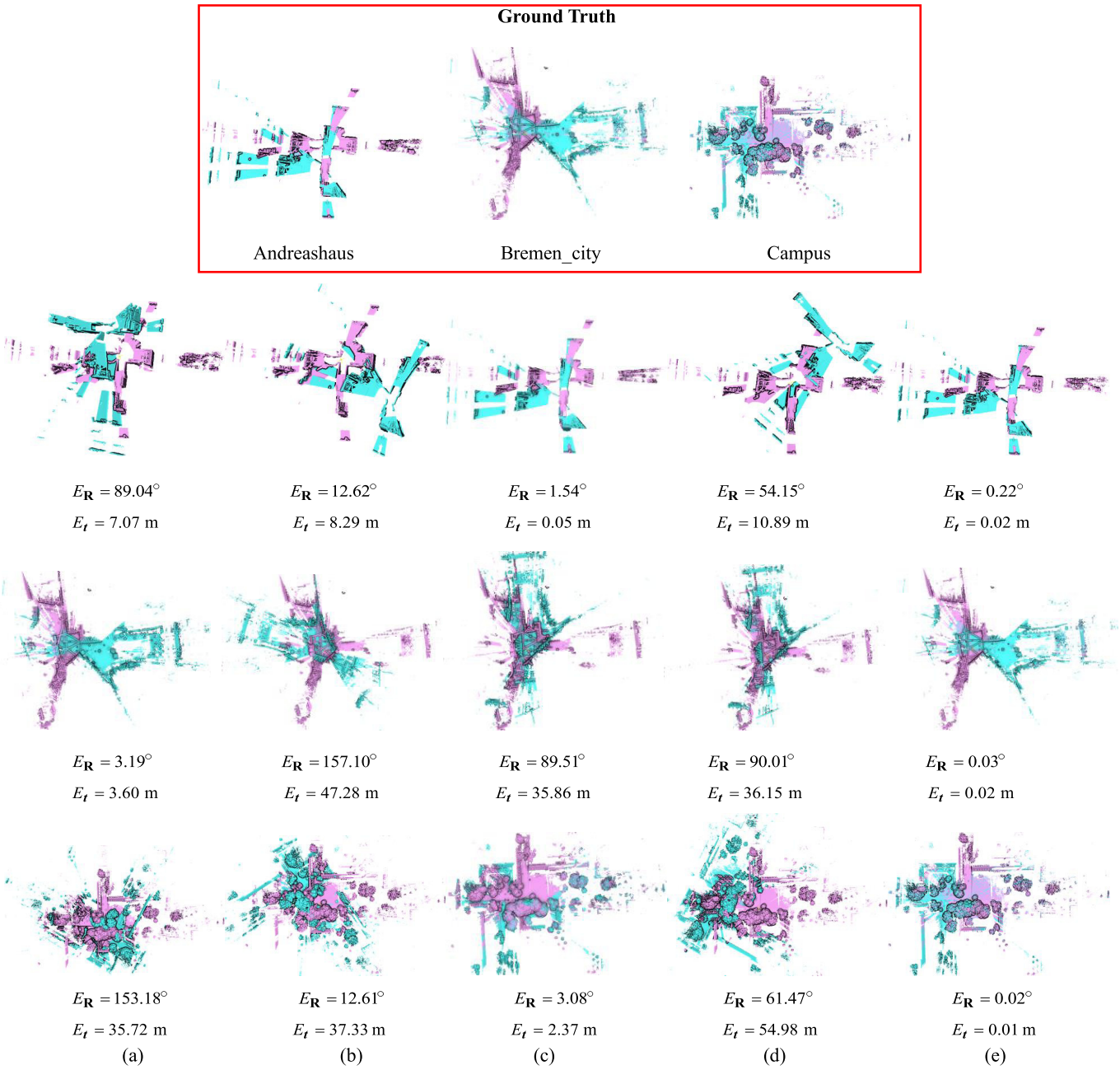

Fig. 3. Qualitative results on three selected scan pairs. The first ∼ fifth columns are the results of (a) FLORANSAC, (b) K4PCS, (c) FMP+BnB, (d) TEASER++, and (e) our method, respectively, where magenta point clouds are target point clouds and cyan ones are source point clouds. The registration accuracy ($E_{\mathbf{R}}$ and E_t) of each result is given, and the GT registrations are provided in the red box for reference.

within 10^5 iterations at an outlier rate above 99%. For example, it completely fails on the seventh and eighth scan pairs of Andreashaus (the rotation errors of FLORANSAC are larger than 100°). K4PCS is sensitive to overlapping ratios. It achieves better performance on the Campus than on the other two data sets since the overlap of Campus (≈75%) is much higher than the others (< 25%). TEASER++ performs worse than FLORANSAC and K4PCS. The reason may be that TLSs used in TEASER++ become unreliable at high outlier rates. The proposed method obtains the best performance, which gets good alignments on all these scan pairs. Our method decomposes the full registration problem into subproblems. As a result, our one-point RANSAC only requires a small number of trials to obtain satisfactory solutions. We also introduce a graduated optimization strategy to avoid local minima in the rotation estimation. Both strategies guarantee the high robustness of the proposed method, i.e., it is suitable for cases with extremely high outlier rates (> 99%).

Table III summarizes the average quantitative results of each compared method. It is clear that the proposed method achieves the best registration accuracy, which is far better than others. Its average rotation error is smaller than 0.25°, and its translation accuracy is better than 0.02 m. Our method can be directly used in practical applications without any fine

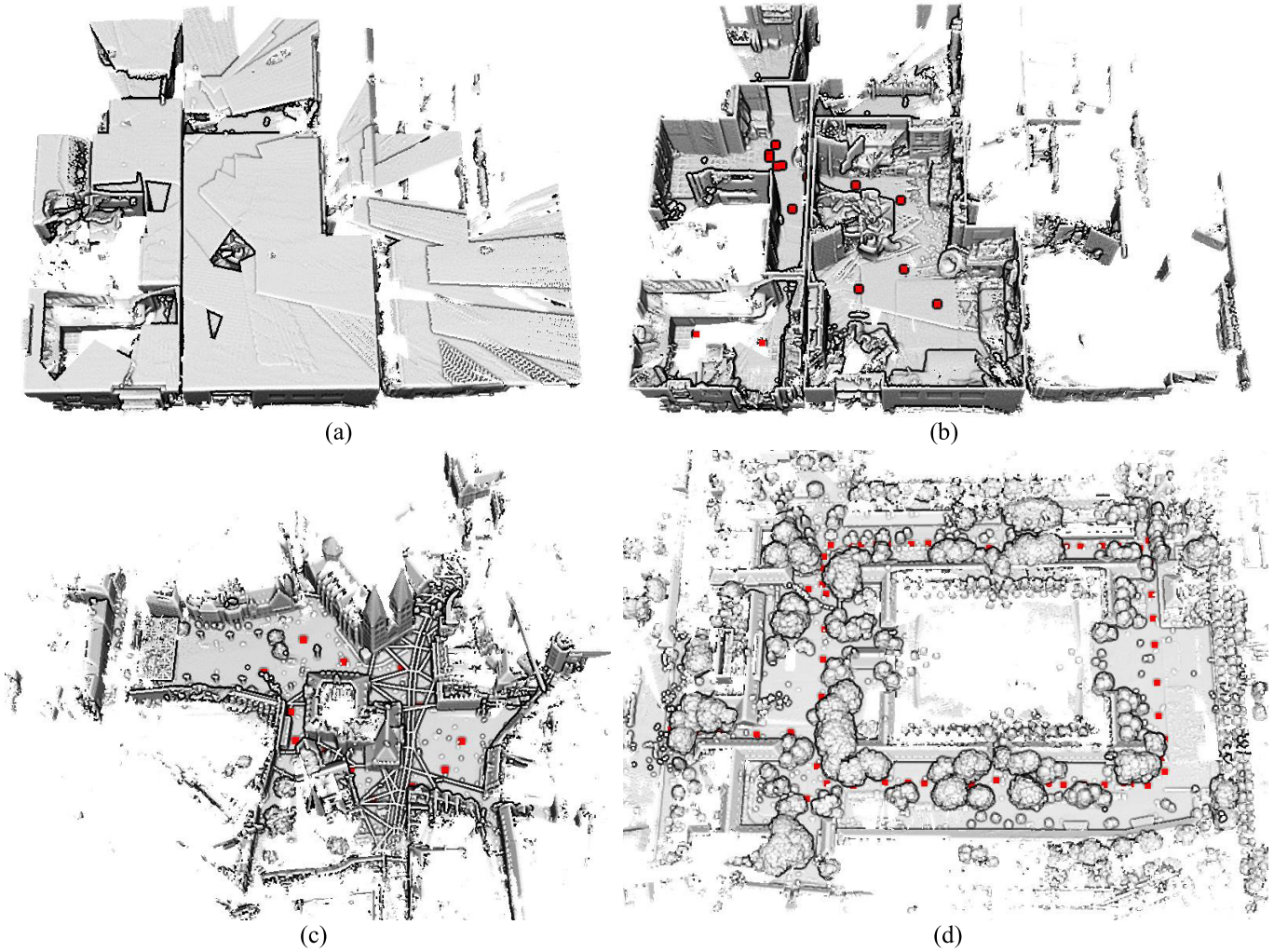

Fig. 4. Scene reconstruction result for each data set. (a) Andreashaus is reconstructed by 15 LIDAR scans. (b) Andreashaus reconstruction after roof removal. (c) Bremen_city is reconstructed by 13 LIDAR scans. (d) Campus is reconstructed by 74 LIDAR scans. Red dots in the figure represent locations estimated by our method where LIDAR scans were recorded.

TABLE III
QUANTITATIVE REGISTRATION RESULTS ON SAME-SOURCE DATA

Method	Andreashaus (Mean±Std)			Bremen_city (Mean±Std)			Campus (Mean±Std)		
	$E_{\mathbf{R}}$ (°)	E_t (m)	time (s)	$E_{\mathbf{R}}$ (°)	E_t (m)	time (s)	$E_{\mathbf{R}}$ (°)	E_t (m)	time (s)
FLORANSAC	47.36±66.44	1.66±2.64	186.05±25.02	15.38±51.53	2.79±8.27	179.82±14.35	12.75±42.30	4.33±11.13	194.58±44.66
K4PCS	26.80±46.72	2.02±2.53	112.46±74.01	71.71±61.91	27.94±21.56	625.30±364.52	10.46±36.79	7.11±23.22	819.56±279.37
FMP+BnB	1.33±0.24	0.05±0.06	**2.49**±0.71	9.27±25.29	4.69±9.93	**2.16**±0.65	1.15±0.86	0.45±0.55	**2.21**±0.89
TEASER++	59.77±68.70	4.43±4.40	8.11±2.74	19.25±42.61	5.78±11.22	8.69±2.46	28.88±52.86	30.28±46.32	9.46±4.93
Ours	**0.24**±0.24	**0.01**±0.01	9.30±2.54	**0.05**±0.04	**0.02**±0.01	9.65±2.44	**0.03**±0.03	**0.01**±0.01	10.44±1.22

registration step, such as ICP refinement. The success rates are reported in Table IV. Under a stringent criterion ($E_R < 1°$ and $E_t < 0.2$ m), the success rate of our method is 100%, which gains a growth rate of 48.48% compared with the second best method, i.e., FLORANSAC. If we relax the criterion to $E_R < 5°$ and $E_t < 2$ m, the success rate of our method is still 16.16% higher than the one of FMP+BnB.

Another major advantage of the proposed method is its high efficiency. Table III also reports the running time of each method. Note that the time of point cloud reading and feature extraction is not included. FMP+BnB is the fastest because it is only a 4-DoF registration solution, while others solve the 6-DoF registration problem. The parameter space of 4-DoF registration is much smaller than the one of 6-DoF registration. Our method consumes about 10 s to register a correspondence set with 9000 points and an outlier rate of 99%, which is much faster than FLORANSAC and K4PCS. For example, our method is 18+ times faster than FLORANSAC and 60+ times

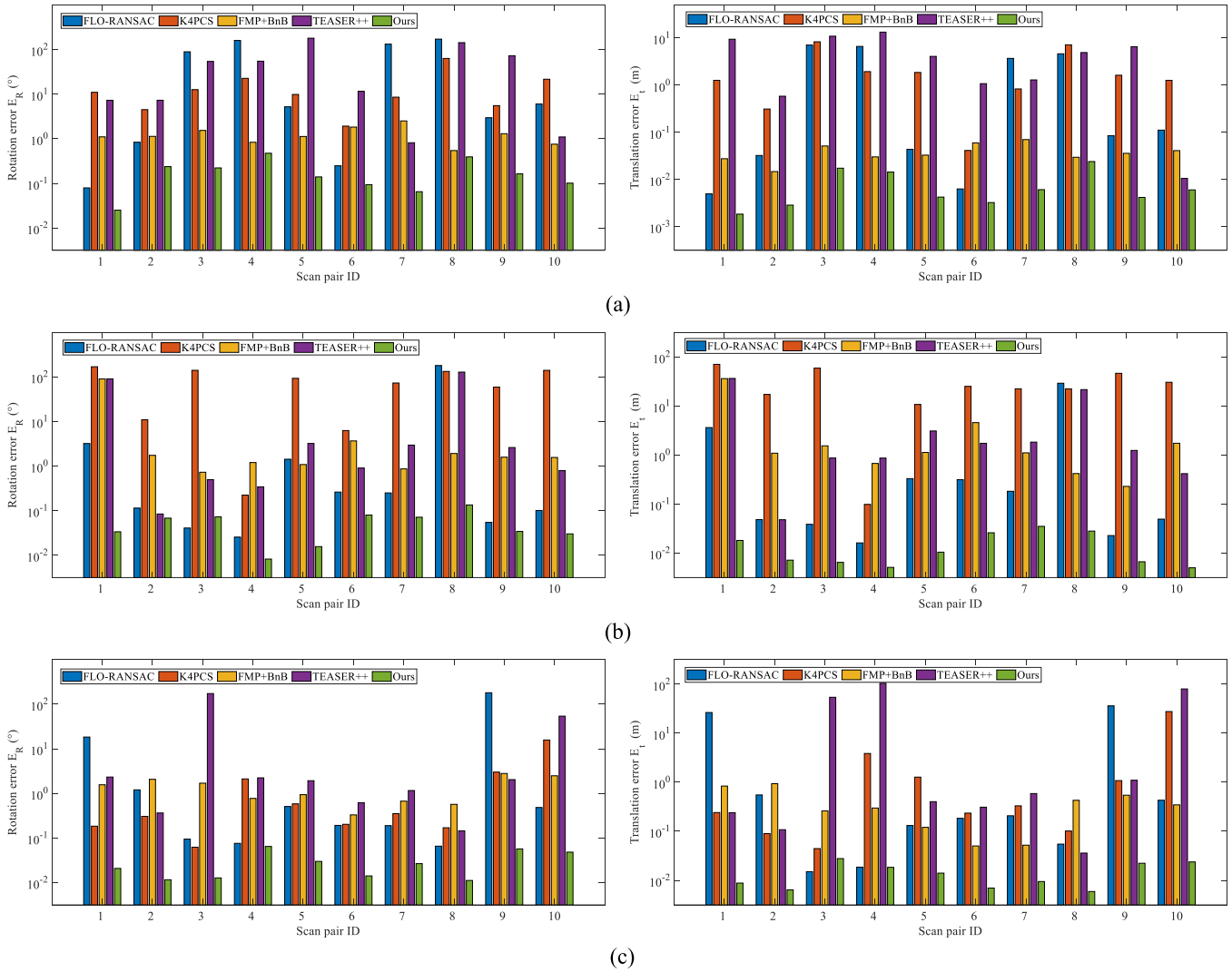

Fig. 5. Individual evaluations on the same-source LIDAR data sets. (a)–(c) Results on the Andreashaus, Bremen_city, and Campus, respectively. (Left column) Rotation errors $E_{\boldsymbol{R}}$. (Right column) Translation errors $E_{\boldsymbol{t}}$. For better visualization, only the first ten scan pairs for each data set are displayed.

TABLE IV
TOTAL SUCCESS RATES ON THE SAME-SOURCE DATA

Method	Success rate (%)	
	$E_{\mathbf{R}}<0.5^\circ$, $E_{\boldsymbol{t}}<0.2$m	$E_{\mathbf{R}}<5^\circ$, $E_{\boldsymbol{t}}<2$m
FLORANSAC	51.52	74.75
K4PCS	25.25	54.55
FMP+BnB	32.32	83.84
TEASER++	16.16	35.35
Ours	**100.00**	**100.00**

faster than K4PCS on the Bremen_city data set. In fact, the complexity of the proposed method $[O(N_T^s N^2)]$ is close to the one of FPM+BnB $[O(\log(N)N^2)]$. FPM+BnB is four times faster than our method in the table because of the C++ implementations. TEASER++ is slightly faster than our method, which benefits from the C++ implementation. In our method, all three subproblems are solved by iteration-based methods. Thus, the time efficiency can be further improved by a programming language that is more suitable for iterative calculation, such as C++. Even considering the time required for data reading and feature extraction, our method can still register a point cloud pair in less than 1 min. For instance, the time of preprocessing (data reading and feature extraction) on the Andreashaus data set is 22.81±7.18 s, and it costs 32.95±11.16 s on the Bremen_city data set.

4) With/Without Scale Estimation: In the above, we disabled the scale estimation since the scale is known, i.e., 1. To evaluate the performance of our full seven-parameter registration method on the same-source data set, we perform a simple experiment that enables the scale estimation on the Bremen_city data set. The rotation and translation errors are 0.053°±0.037° and 0.034 m±0.035 m, respectively, which are comparable to the ones without scale estimation. The major difference is that scale estimation is much slower (running time is 148.29±69.86 s). As analyzed, scale estimation is the most time-consuming part of our method.

B. Cross-Source Registration

In this section, we consider scale differences between point clouds and evaluate the proposed method on a cross-source point cloud data set. The proposed method is compared with FLORANSAC, which ranks second under the stringent criterion in the same-source registration experiment. K4PCS and

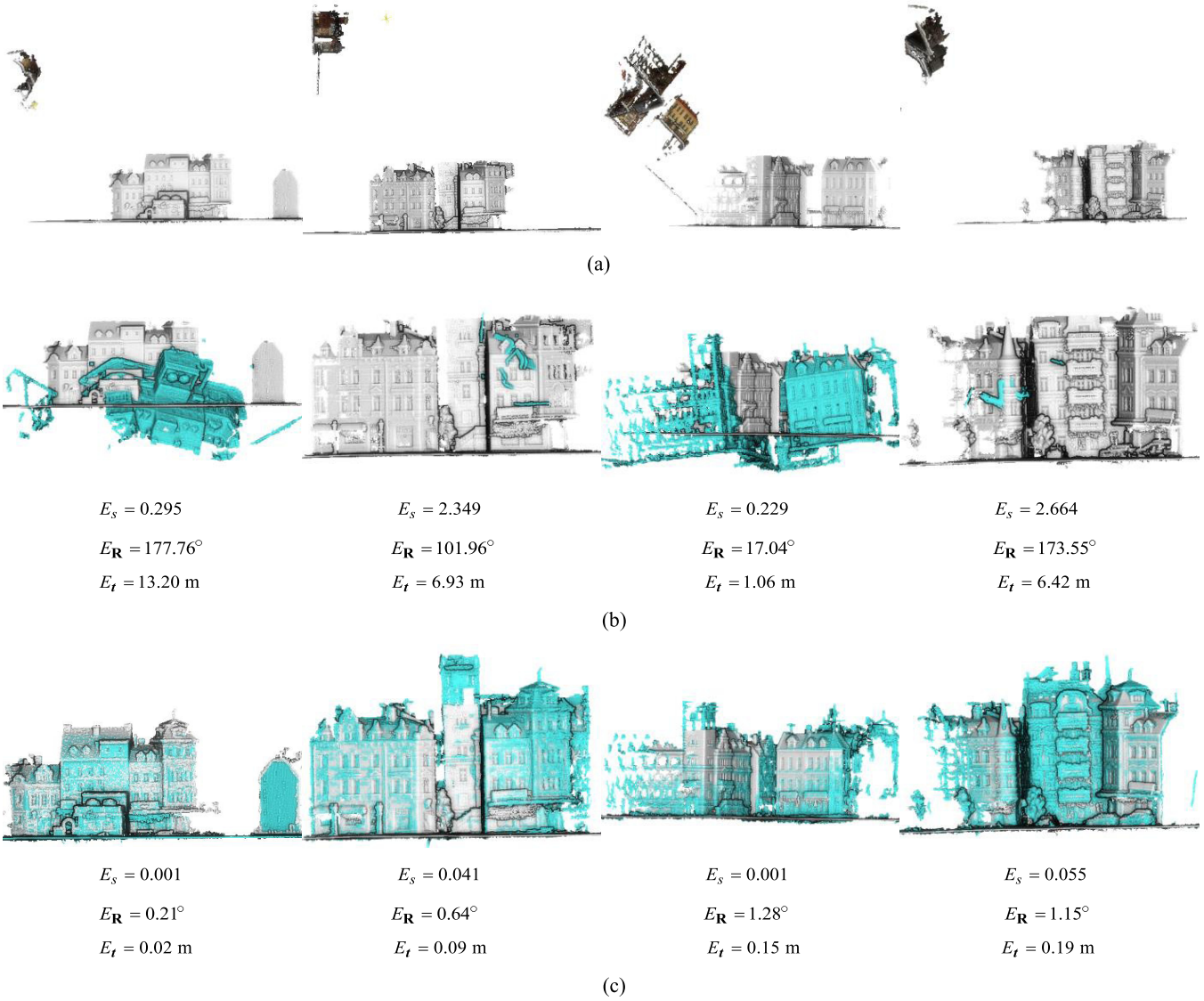

Fig. 6. Qualitative results on the cross-source data set. (a) Input MVS and structured light point cloud pairs. (b) Results of FLORANSAC. (c) Results of our method. The gray point clouds are captured by a structured light sensor, and the cyan ones are generated by an MVS technique. The scale, rotation, and translation errors are provided below each result.

FMP+BnB cannot deal with point clouds with scale changes. The computational complexity of TEASER++ is $O(N^4)$, which is too slow to be practical. For example, TEASER++ requires more than 10^{14} iterations for scale estimation with 5000 correspondences, which takes more than one day in our computational environment. Thus, these methods are excluded from the comparison.

1) Data Set: We collect 19 scenes from the MVS data set[6] for assessment. Each scene of this data set consists of a structured light point cloud, an image sequence, and a calibration file. We use the visualSFM software[7] and patch-based multiview stereo software (PMVS) [71] to generate an MVS point cloud for each image sequence. For each scene, we manually pick four correspondences and estimate an initial seven-parameter similarity transformation. Then, the ICP algorithm is used to refine the initial parameters (rotation and translation) and estimate an accurate transformation, which is regarded as the GT. To reduce the sizes of point clouds, we downsample their resolutions to 0.05 m. The detailed information of this data set is summarized as follows: $N_{pr} = 19$, $N_{pt} = 5.42 \times 10^5$, $N = 3391$, $r_{\text{outlier}} = 98.86\%$, overlap $> 80\%$, resolution $= 0.05$ m.

2) Qualitative Assessment: Four difficult scenes are selected for qualitative evaluation [see Fig. 5(a)]. Although the overlaps of these point cloud pairs are large, the outlier rates in the initial correspondence sets are still very high. Their outlier rates are 99.16%, 99.47%, 99.33%, and 99.48%, respectively. The reasons may be twofold: first, MVS methods cannot generate points in textureless images, which results in holes in the point clouds. Then, many keypoints in the structured light point cloud do not have corresponding points in the MVS point cloud. Second, the MVS point cloud suffers from geometric distortions. Error accumulations cannot be eliminated completely even with a bundle adjustment step. High outlier rates and large scale changes make registration on these pairs challenging. Fig. 6 shows the results of the proposed method

[6]http://roboimagedata.compute.dtu.dk/
[7]http://ccwu.me/vsfm/

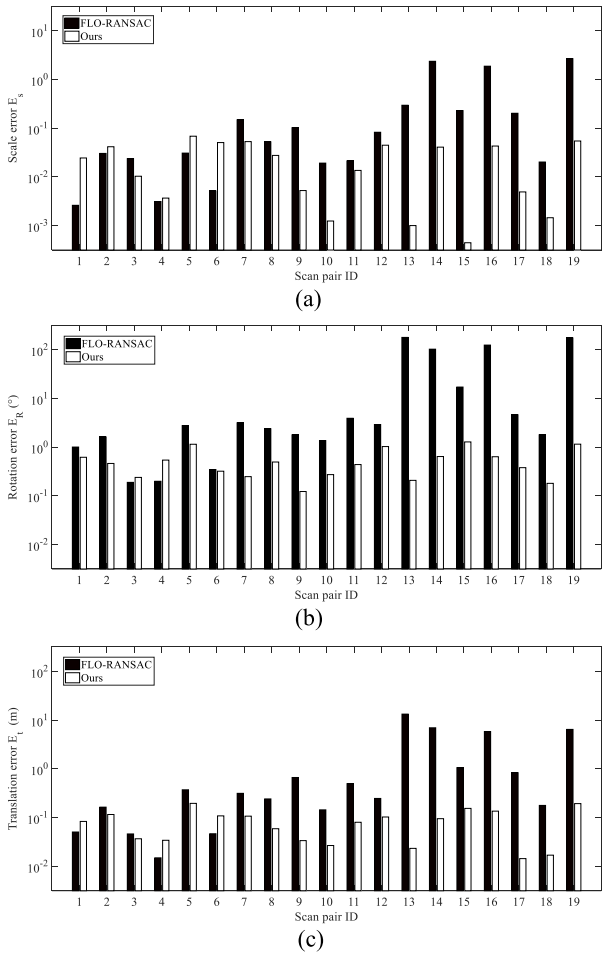

Fig. 7. Individual evaluations on the cross-source data. (a)–(c) Results of scale (E_s), rotation ($E_{\mathbf{R}}$), and translation errors (E_t), respectively.

and FLORANSAC. As can be seen, our method achieves good results, while FLORANSAC fails to register any point cloud pair. As aforementioned, the proposed method is more robust than FLORANSAC and suitable for cases with extremely high outlier rates. Generally, the registration accuracy of our method on cross-source data is slightly lower than the one on same-source data. The reasons are as follows. First, MVS point clouds and structured light point clouds cannot be perfectly registered due to geometric distortions. Second, same-source point clouds have an exact scale of one. However, the scales between cross-source point clouds are estimated, which are only optimal approximations. The scale errors will decrease the accuracy of rotation and translation estimations.

3) Quantitative Assessment: The individual quantitative results are plotted in Fig. 7. As shown, our method performs much better than FLORANSAC on most of the scenes. For example, FLORANSAC gets rotation errors that are even larger than 100° on the last several scenes. In contrast, the maximum rotation error and maximum translation error of our method are 1.28° and 0.19 m, respectively.

The quantitative results are reported in Table V. Our method achieves much higher registration accuracy than FLORANSAC. Due to the scale errors, the rotation and translation

TABLE V
DETAILED INFORMATION OF DATA SETS (K HERE MEANS THOUSANDS)

Method	E_s (Mean±Std)	$E_{\mathbf{R}}$ (°) (Mean±Std)	$E_{\boldsymbol{t}}$ (m) (Mean±Std)
FLORANSAC	0.429±0.844	32.74±61.20	1.96±3.55
Ours	**0.026**±0.022	**0.55**±0.36	**0.09**±0.05

TABLE VI
TOTAL SUCCESS RATES ON THESE THREE DATA SETS

Method	Success rate (%)	
	$E_s < 0.05$, $E_{\mathbf{R}} < 1°$, $E_{\boldsymbol{t}} < 0.2$m	$E_s < 0.1$, $E_{\mathbf{R}} < 5°$, $E_{\boldsymbol{t}} < 1$m
FLORANSAC	15.79	57.89
Ours	**68.42**	**100.00**

accuracies of the proposed method are less impressing than the ones in the above section. Fortunately, under a stringent criterion ($E_s < 0.05$, $E_{\boldsymbol{R}} < 1°$, and $E_{\boldsymbol{t}} < 0.2$ m), our method still obtains a success rate of 68.42%, which is 52.63% higher than the one of FLORANSAC. Under a relaxed criterion ($E_s < 0.1$, $E_{\boldsymbol{R}} < 5°$, and $E_{\boldsymbol{t}} < 1$ m), the success rate of our method becomes 100%.

C. Advantages and Limitations

As a highly practical method, the advantages of the proposed method are as follows.

1) *High Robustness:* Our method is robust against over 99% outliers. The one-point RANSAC inherits the advantages of RANSAC, such as robustness and interpretability. Moreover, the proposed graduated optimization strategy large improves the robustness of IRLS.
2) *High Efficiency:* The time complexity of our method is $O(N^2)$. We decompose the full seven-parameter registration problem into three subproblems. The parameter space is largely reduced. Thus, one-point RANSAC can be very efficient even at extremely high outlier rates.
3) *Good Scalability:* Our method solves the full seven-parameter registration problem. It is suitable for both same-source data, such as LIDAR PCR, and cross-source data, such as MVS-LIDAR registration.

The limitations of the proposed method are twofold.

1) *Dependence on Correspondences:* Our method needs to establish an initial correspondence set M. It cannot take original point clouds as input. Thus, its registration accuracy relies on the location accuracy of 3-D keypoint detectors to some extent. Moreover, too few inliers (e.g., < 10) in M may cause the proposed method to fail.
2) *Maximum Consensus Assumption:* In our method, the scale and translation estimation rely on an assumption that the optimal solution corresponds to the maximum consensus. However, if point clouds contain multiple geometric structures (generally occur in nonrigid point clouds), this assumption may not hold anymore. For example, if point clouds contain moving

objects, the proposed method may align the moving objects instead of a static environment.

V. Conclusion

In this article, we propose a highly practical seven-parameter PCR method that is initial-guess free, accurate, fast, and robust. Compared with most of the current methods, the proposed approach is suitable for both same-source data and cross-source data. To reduce parameter space and improve efficiency, we decompose the full seven-parameter registration into three subproblems based on the concept of extended line vector. We propose a fast one-point RANSAC algorithm for scale and translation estimations. We also propose a scale-annealing biweight estimator for rotation estimation and introduce a graduated optimization strategy to largely alleviate the effect of local minima. These improvements ensure the high robustness and high efficiency of our method. Both the same-source and cross-source registration experiments demonstrate that our method is much superior to current methods, i.e., it is robust against over 99% outliers and much faster than its competitors.

References

[1] J. Heo *et al.*, "Productive high-complexity 3D city modeling with point clouds collected from terrestrial LiDAR," *Comput., Environ. Urban Syst.*, vol. 41, pp. 26–38, Sep. 2013.

[2] C. Hladik and M. Alber, "Accuracy assessment and correction of a LIDAR-derived salt marsh digital elevation model," *Remote Sens. Environ.*, vol. 121, pp. 224–235, Jun. 2012.

[3] J. Zhang and S. Singh, "Loam: Lidar odometry and mapping in real-time," in *Robotics, Science and Systems*, vol. 2, no. 9, 2014.

[4] J. Li, R. Zhong, Q. Hu, and M. Ai, "Feature-based laser scan matching and its application for indoor mapping," *Sensors*, vol. 16, no. 8, p. 1265, Aug. 2016.

[5] A. F. Chase, D. Z. Chase, C. T. Fisher, S. J. Leisz, and J. F. Weishampel, "Geospatial revolution and remote sensing LiDAR in mesoamerican archaeology," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 32, pp. 12916–12921, Aug. 2012.

[6] J. Li, P. Zhao, Q. Hu, and M. Ai, "Robust point cloud registration based on topological graph and cauchy weighted lq-norm," *ISPRS J. Photogramm. Remote Sens.*, vol. 160, pp. 244–259, Feb. 2020.

[7] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[8] X. Huang, J. Zhang, Q. Wu, L. Fan, and C. Yuan, "A coarse-to-fine algorithm for matching and registration in 3D cross-source point clouds," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 10, pp. 2965–2977, Oct. 2018.

[9] X. Huang, J. Zhang, L. Fan, Q. Wu, and C. Yuan, "A systematic approach for cross-source point cloud registration by preserving macro and micro structures," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3261–3276, Jul. 2017.

[10] R. Bogdan Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 3212–3217.

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

[12] Y. Guo, M. Bennamoun, F. Sohel, M. Lu, J. Wan, and N. M. Kwok, "A comprehensive performance evaluation of 3D local feature descriptors," *Int. J. Comput. Vis.*, vol. 116, no. 1, pp. 66–89, 2016.

[13] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[14] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.

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

[16] A. P. Bustos and T.-J. Chin, "Guaranteed outlier removal for point cloud registration with correspondences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 12, pp. 2868–2882, Dec. 2018.

[17] L. Cheng *et al.*, "Registration of laser scanning point clouds: A review," *Sensors*, vol. 18, no. 5, p. 1641, May 2018.

[18] Z. Dong *et al.*, "Registration of large-scale terrestrial laser scanner point clouds: A review and benchmark," *ISPRS J. Photogramm. Remote Sens.*, vol. 163, pp. 327–342, May 2020.

[19] H. Chen and B. Bhanu, "3D free-form object recognition in range images using local surface patches," *Pattern Recognit. Lett.*, vol. 28, no. 10, pp. 1252–1262, Jul. 2007.

[20] Y. Zhong, "Intrinsic shape signatures: A shape descriptor for 3D object recognition," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops, ICCV Workshops*, Sep. 2009, pp. 689–696.

[21] A. Zaharescu, E. Boyer, K. Varanasi, and R. Horaud, "Surface feature detection and description with applications to mesh matching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 373–380.

[22] S. Suwajanakorn, N. Snavely, J. J. Tompson, and M. Norouzi, "Discovery of latent 3d keypoints via end-to-end geometric reasoning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 2059–2070.

[23] J. Li and G. Hee Lee, "USIP: Unsupervised stable interest point detection from 3D point clouds," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Nov. 2019, pp. 361–370.

[24] G. Tinchev, A. Penate-Sanchez, and M. Fallon, "SKD: Unsupervised keypoint detection for point clouds using saliency estimation," 2019, *arXiv:1912.04943*. [Online]. Available: http://arxiv.org/abs/1912.04943

[25] F. Tombari, S. Salti, and L. Di Stefano, "Unique shape context for 3D data description," in *Proc. ACM Workshop 3D Object Retr.*, 2010, pp. 57–62.

[26] Z. Dong, B. Yang, Y. Liu, F. Liang, B. Li, and Y. Zang, "A novel binary shape context for 3D local surface description," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 431–452, Aug. 2017.

[27] Z. Dong, B. Yang, F. Liang, R. Huang, and S. Scherer, "Hierarchical registration of unordered TLS point clouds based on binary shape context descriptor," *ISPRS J. Photogramm. Remote Sens.*, vol. 144, pp. 61–79, Oct. 2018.

[28] Y. Guo, F. Sohel, M. Bennamoun, M. Lu, and J. Wan, "Rotational projection statistics for 3D local surface description and object recognition," *Int. J. Comput. Vis.*, vol. 105, no. 1, pp. 63–86, 2013.

[29] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser, "3dmatch: Learning local geometric descriptors from rgb-d reconstructions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1802–1811.

[30] H. Deng, T. Birdal, and S. Ilic, "PPFNet: Global context aware local features for robust 3D point matching," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 195–205.

[31] Z. Gojcic, C. Zhou, J. D. Wegner, and A. Wieser, "The perfect match: 3d point cloud matching with smoothed densities," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5545–5554.

[32] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, 2000.

[33] O. Chum, J. Matas, and J. Kittler, "Locally optimized ransac," in *Proc. Joint Pattern Recognit. Symp.* Berlin, Germany: Springer, 2003, pp. 236–243.

[34] K. Ni, H. Jin, and F. Dellaert, "GroupSAC: Efficient consensus in the presence of groupings," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 2193–2200.

[35] O. Chum and J. Matas, "Optimal randomized RANSAC," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 8, pp. 1472–1482, Aug. 2008.

[36] K. Lebeda, J. Matas, and O. Chum, "Fixing the locally optimized ransac–full experimental evaluation," in *Brit. Mach. Vis. Conf.*, 2012, pp. 1–11.

[37] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J.-M. Frahm, "USAC: A universal framework for random sample consensus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 2022–2038, Nov. 2012.

[38] J. Li, Q. Hu, and M. Ai, "Robust feature matching for geospatial images via an affine-invariant coordinate system," *Photogramm. Rec.*, vol. 32, no. 159, pp. 317–331, Sep. 2017.

[39] D. Barath, J. Noskova, M. Ivashechkin, and J. Matas, "MAGSAC++, a fast, reliable and accurate robust estimator," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1304–1312.

[40] E. Brachmann *et al.*, "Dsac-differentiable ransac for camera localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 6684–6692.

[41] E. Brachmann and C. Rother, "Neural-guided RANSAC: Learning where to sample model hypotheses," 2019, *arXiv:1905.04132*. [Online]. Available: http://arxiv.org/abs/1905.04132
[42] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Commun. Statist.-Theory Methods*, vol. 6, no. 9, pp. 813–827, 1977.
[43] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, vol. 589. Hoboken, NJ, USA: Wiley, 2005.
[44] P. J. Huber, *Robust Statistics*, vol. 523. Hoboken, NJ, USA: Wiley, 2004.
[45] P. Meer, "Robust techniques for computer vision," in *Emerging Topics in Computer Vision*, G. Medioni and S. B. Kang, Eds. Englewood Cliffs, NJ, USA: Prentice-Hall, 2004, pp. 107–190.
[46] S. Mittal, S. Anand, and P. Meer, "Generalized projection-based M-estimator," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 12, pp. 2351–2364, Dec. 2012.
[47] Q.-Y. Zhou, J. Park, and V. Koltun, "Fast global registration," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 766–782.
[48] J. Li, Q. Hu, and M. Ai, "Robust feature matching for remote sensing image registration based on L_q-estimator," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1989–1993, Dec. 2016.
[49] H. Yang, J. Shi, and L. Carlone, "TEASER: Fast and certifiable point cloud registration," 2020, *arXiv:2001.07715*. [Online]. Available: http://arxiv.org/abs/2001.07715
[50] J. Li, Q. Hu, and M. Ai, "Robust geometric model estimation based on scaled welsch Q-norm," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5908–5921, Aug. 2020.
[51] R. Bogdan Rusu and S. Cousins, "3D is here: Point cloud library (PCL)," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 1–4.
[52] R. Schnabel and R. Klein, "Octree-based point-cloud compression," in *Proc. Eurographics Symp. Point-Based Graph.*, vol. 6, Jul. 2006, pp. 111–120.
[53] A. Censi, "An ICP variant using a point-to-line metric," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2008, pp. 19–25.
[54] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image Vis. Comput.*, vol. 10, no. 3, pp. 145–155, Apr. 1992.
[55] A. V. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," *Robot. Sci., Syst.*, vol. 2, no. 4, p. 435, Jun. 2009.
[56] D. Chetverikov, D. Stepanov, and P. Krsek, "Robust Euclidean alignment of 3D point sets: The trimmed iterative closest point algorithm," *Image Vis. Comput.*, vol. 23, no. 3, pp. 299–309, 2005.
[57] L. Maier-Hein *et al.*, "Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1520–1532, Aug. 2012.
[58] S. Bouaziz, A. Tagliasacchi, and M. Pauly, "Sparse iterative closest point," *Comput. Graph. Forum*, vol. 32, no. 5, pp. 113–123, Aug. 2013.
[59] J. Yang, H. Li, D. Campbell, and Y. Jia, "Go-ICP: A globally optimal solution to 3D ICP point-set registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2241–2254, Nov. 2016.
[60] D. Aiger, N. J. Mitra, and D. Cohen-Or, "4-points congruent sets for robust pairwise surface registration," *ACM Trans. Graph.*, vol. 27, no. 3, p. 85, Aug. 2008.
[61] N. Mellado, D. Aiger, and N. J. Mitra, "Super 4PCS fast global pointcloud registration via smart indexing," *Comput. Graph. Forum*, vol. 33, no. 5, pp. 205–215, Aug. 2014.
[62] X. Ge, "Non-rigid registration of 3D point clouds under isometric deformation," *ISPRS J. Photogramm. Remote Sens.*, vol. 121, pp. 192–202, Nov. 2016.
[63] P. W. Theiler, J. D. Wegner, and K. Schindler, "Keypoint-based 4-Points congruent sets—Automated marker-less registration of laser scans," *ISPRS J. Photogramm. Remote Sens.*, vol. 96, pp. 149–163, Oct. 2014.
[64] J.-K. Lee, Y.-K. Baik, H. Cho, K. Kim, and D. H. Kim, "1-point RANSAC-based method for ground object pose estimation," 2020, *arXiv:2008.03718*. [Online]. Available: http://arxiv.org/abs/2008.03718
[65] D. Scaramuzza, "1-Point-RANSAC structure from motion for vehicle-mounted cameras by exploiting non-holonomic constraints," *Int. J. Comput. Vis.*, vol. 95, no. 1, pp. 74–85, Oct. 2011.
[66] J. Civera, O. G. Grasa, A. J. Davison, and J. M. M. Montiel, "1-point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry," *J. Field Robot.*, vol. 27, no. 5, pp. 609–631, Sep. 2010.
[67] D. Scaramuzza, "Performance evaluation of 1-point-RANSAC visual odometry," *J. Field Robot.*, vol. 28, no. 5, pp. 792–811, Sep. 2011.
[68] G. D. Evangelidis, D. Kounades-Bastian, R. Horaud, and E. Z. Psarakis, "A generative model for the joint registration of multiple point sets," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 109–122.
[69] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "G^2o: A general framework for graph optimization," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3607–3613.
[70] Z. Cai, T.-J. Chin, A. P. Bustos, and K. Schindler, "Practical optimal registration of terrestrial LiDAR scan pairs," *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 118–131, Jan. 2019.
[71] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010.