

Robust symmetric iterative closest point

Jiayuan Li, Qingwu Hu^{*}, Yongjun Zhang^{*}, Mingyao Ai

School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China



ARTICLE INFO

Keywords:

Point cloud registration
Simultaneous Localization and Mapping (SLAM)
Iterative Closest Point (ICP)
Symmetric distance metric
Robust loss
3D alignment

ABSTRACT

Point cloud registration (PCR) is an important technique of 3D vision, which has been widely applied in many areas such as robotics and photogrammetry. The iterative closest point (ICP) is a de facto standard for PCR. However, it mainly suffers from two drawbacks: small convergence basin and the sensitivity to outliers and partial overlaps. In this paper, we propose a robust symmetric ICP (RSICP) to tackle these drawbacks. First, we present a new symmetric point-to-plane distance metric whose functional zero-set is a set of locally-second-order surfaces. It has a wider convergence basin and higher convergence speed than the point-to-point metric, point-to-plane metric, and even original symmetric metric. Second, we introduce an adaptive robust loss to construct our robust symmetric metric. This robust loss bridges the gap between the non-robust ℓ_2 cost and robust M-estimates. In the optimization, we gradually improve the degrees of robustness via the decay of a robustness control parameter. This loss has a high “breakdown” point or low computational overhead compared with recent work (e.g., Sparse ICP and Robust ICP). We also present a simple but effective linearization for the alignment function based on Rodrigues rotation parameterization with the small incremental rotation assumption. Extensive experiments on challenging datasets with noise, outliers or partial overlaps show that the proposed algorithm significantly outperforms Sparse ICP and Robust ICP in terms of both accuracy and efficiency. Our source code will be publicly available in <https://ljjy-rs.github.io/web>.

1. Introduction

Point cloud registration (PCR) is an important and fundamental technique of 3D vision, with wide applications in the fields of robotics, photogrammetry, and computer graphics, such as simultaneous localization and mapping (SLAM), scene perception, and 3D modelling. Current point cloud acquisition device can only capture a part of the object at a single frame. Thus, to cover the whole object or scene, a sequence of frames with overlaps should be taken from different positions. PCR is a technique to merge this sequence into a panorama, in a way that seeks the best rigid model to align each pair of adjacent point clouds (Li, 2021).

Iterative closest point (ICP) (Besl and McKay, 1992) is the standard method for PCR problem, which consists of two main steps, i.e., correspondence step and alignment step. The first step searches a closest point from the target set for each source point to establish correspondences; then, the alignment step estimates an optimal transformation to register the correspondences. These two steps are alternately performed until a locally optimal alignment is achieved. However, ICP suffers from two major drawbacks: small convergence basin and the sensitivity to outliers

and partial overlaps (Zhang et al., 2021). Classical ICP has a slow convergence speed because of its linear convergence rate (Pottmann et al., 2006). Although the improved point-to-plane ICP (Chen and Medioni, 1992) can converge faster, its convergence basin is small. The functional zero-set of point-to-plane metric is only planar patches. Another problem with ICP lies in the alignment step, which estimates a rigid model based on the least-squares optimization. As known, the ℓ_2 loss is not a robust cost. If many source points have no true correspondences in the target set, ICP will converge to a bad local optimum. Therefore, ICP is sensitive to outliers and partial overlaps, which are common for real datasets. To tackle this issue, several methods use distance or normal constraints to filter potential outliers (Rusinkiewicz and Levoy, 2001; Rusinkiewicz, 2019). These methods are heuristics and the constraint parameters are difficult to tune. Recently, Sparse ICP (Bouaziz et al., 2013) and Robust ICP (Zhang et al., 2021) show superiority to other ICP variants, which introduce robust cost functions instead of the ℓ_2 loss to enhance the alignment step.

In this paper, we aim to design a new ICP-based approach with following properties compared with classical ICP: better registration accuracy, wider convergence basin, higher convergence speed, and

* Corresponding authors.

E-mail addresses: huqw@whu.edu.cn (Q. Hu), zhangyj@whu.edu.cn (Y. Zhang).

better robustness to outliers and partial overlaps. Our key observations are twofold: (1) Symmetric point-to-plane distance metric minimizes point-to-surface error based on normals of both the source and target points, which has a wider convergence basin and higher convergence speed, since its functional zero-set is locally-second-order surfaces while the one of point-to-plane metric is only local plane patches. (2) Using robust loss instead of the least-squares cost of classical ICP can largely improve the robustness to outliers and partial overlaps. The least-squares is the best fitting for observations with Gaussian noise. However, it is not a robust loss and very sensitive to outliers. This conclusion had been evidenced in several work such as the Sparse ICP and Robust ICP. Based on these observations, we propose a robust symmetric ICP (RSICP) method to achieve our goals. We first present a rigorous mathematical model for the symmetric point-to-plane distance metric. Compared with the original approximate model, our mathematical model also considers the rotation of normals without increasing the computational complexity, which does not suffer from the loss of registration accuracy. We then introduce an adaptive robust loss (Li et al., 2021) to deal with outliers and partial overlaps, which bridges the gap between the non-robust ℓ_2 cost and robust M-estimates. This loss is optimized based on a dynamic strategy so that it acts as functions with different degrees of robustness along with iterations. Compared with the loss of Sparse ICP, our loss has a much lower computational overhead and higher efficiency. Our loss has a higher “breakdown” point than the traditional Welsch (Holland and Welsch, 1977) loss used in Robust ICP. Extensive experiments on both synthetic and real-world datasets demonstrate that our RSICP is accurate, fast, and robust. It gets better performance in terms of both accuracy and efficiency than several baselines and state-of-the-arts, including point-to-point ICP (Besl and McKay, 1992), point-to-plane ICP (Chen and Medioni, 1992), Anderson acceleration ICP (Pavlov et al., 2018), Fast ICP (Zhang et al., 2021), Sparse ICP (Bouaziz et al., 2013), symmetric ICP (Rusinkiewicz, 2019), and robust ICP (Zhang et al., 2021).

Our main contributions are summarized as follows:

- We propose a new ICP-type method called RSICP, which has better registration accuracy, wider convergence basin, higher convergence speed, or better robustness to outliers than current methods.
- We present a new rigorous mathematical model for the symmetric metric, which also considers the rotation of normals without increasing the computational complexity. It is more accurate than the original one, with a wider convergence basin and higher convergence speed.
- We propose a robust registration objective based on an adaptive loss, which bridges the gap between the non-robust ℓ_2 cost and robust M-estimates. It has a higher “breakdown” point compared with the Welsch loss.
- We present a simple but effective linearization for the alignment function based on Rodrigues rotation parameterization with the small incremental rotation assumption.

2. Related work

PCR methods can be roughly classified into two categories, i.e., feature-based approaches and point-based ones. In this section, we briefly review both the two types of methods. The readers are referred to (Tam et al., 2013; Huang et al., 2021) for a comprehensive review of PCR techniques.

2.1. Feature-based registration

Feature-based registration is the 3D extension of image matching, which also contains two major steps: feature matching and geometric estimation (Li et al., 2020c). In the feature matching, 3D distinctive keypoints are first detected from point clouds via feature detectors (e.g., intrinsic shape signatures (ISS) (Zhong, 2009), MeshDoG (Zaharescu

et al., 2009), KeypointNet (Suwajanakorn et al., 2018), and USIP (Li and Lee, 2019)). Then, each keypoint is encoded to a compact feature vector by analysing its local surface based on descriptors (e.g., fast point feature histogram (FPFH) (Rusu et al., 2009), 3DMatch (Zeng et al., 2017), 3DSmoothNet (Gojcic et al., 2019), fully convolutional geometric features (FCGF) (Choy et al., 2019), and SpinNet (Ao et al., 2021)), so that keypoints have a high degree of discrimination. Finally, one-to-one corresponding relationship is established via computing pairwise similarity scores (e.g., nearest neighbor distance ratio (NNDR) (Lowe, 2004) and chi-square test (Zhong, 2009)).

In the geometric estimation, six degrees of freedom (DoF) rigid transformation is estimated based on robust fitting techniques, among which random sample consensus (RANSAC) (Fischler and Bolles, 1981) and its variants (Tordoff and Murray, 2005; Chum and Matas, 2008; Raguram et al., 2012; Barath and Matas, 2018; Brachmann and Rother, 2019; Barath et al., 2020) are the most widely used methods. RANSAC alternates between a random sampling step and a model fitting step. However, the computational complexity of RANSAC-type methods grows exponentially with outlier rate (Li et al., 2020a). Compared with 2D image matching (e.g., scale-invariant feature transform (SIFT) (Lowe, 2004) and radiation-variation insensitive feature transform (RIFT) (Li et al., 2020b)), 3D feature matching is much more difficult due to uneven densities, lack of texture, and noise. Hence, initial correspondence set contains a large portion of outliers (often > 95%). In this case, RANSAC-type methods may cost dozens of minutes or hours to find an approximate solution, which are too slow to be practical.

Recently, several approaches are proposed for extremely high outlier rate problem. For example, guaranteed outlier removal (GORE) (Bustos and Chin, 2017) identifies true outliers based on the conflict between upper bounds and lower bounds, which is robust to 99% of outliers. Based on the idea of GORE, polynomial-time GORE (Li, 2021) uses two novel concepts (correspondence matrix and augmented correspondence matrix) to decide the tight bounds, which improves the efficiency from exponential complexity $O(2^N)$ to polynomial complexity $O(N^2)$. Several methods (e.g., TEASER (Yang and Carlone, 2019; Yang et al., 2020b), 1pt-RANSAC (Li et al., 2021), and weighted q-norm estimation (Li et al., 2020c)) decompose the 7-DoF/6-DoF registration problem into subproblems (scale, rotation, and translation estimation) to decrease the parameter space, and solve these subproblems based on truncated least-squares (TLS), 1-point RANSAC, or improved M-like robust estimators. Although these methods achieve promising performance, their registration accuracies are lower than point-based registration. Therefore, feature-based registration is generally referred as coarse registration, which provides good initializations for point-based registration for refinement.

2.2. Point-based registration

The most widely used point-based registration method is the ICP (Besl and McKay, 1992), which is a simultaneous correspondence and alignment technique. Compared with feature-based methods, ICP uses raw points as input instead of features, which does not require corresponding relationship between points. Thus, ICP has one more step called correspondence update step, in which closest points are established as correspondences. Different from feature correspondences, the correspondences in ICP are dynamic. They are changed along with iterations.

ICP has many variants (Chen and Medioni, 1992; Rusinkiewicz and Levoy, 2001; Chetverikov et al., 2005; Bouaziz et al., 2013; Pavlov et al., 2018; Rusinkiewicz, 2019; Wang and Solomon, 2019; Zhang et al., 2021), which improve the sampling and matching, distance metric, or outlier rejection strategies of the original ICP. (1) Sampling and matching. Generally, the sizes of point clouds captured by 3D laser scanners are very large ($10^4 \sim 10^7$). Direct registration of the whole point sets is time-consuming and not acceptable in practice. Proper sampling can largely improve the efficiency without losing registration

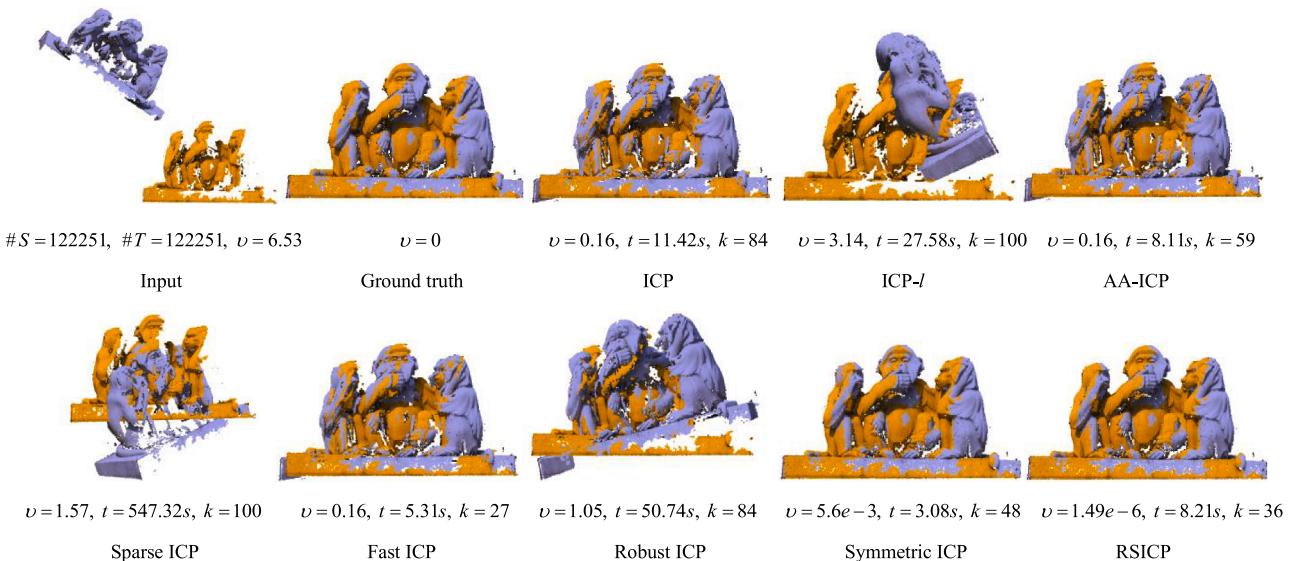


Fig. 1. Registration results of different methods (see Section 5.1.2) on a point cloud pair with partial overlap, obtained from the monkey model of the EPFL statue dataset. $\#S$ and $\#T$ are the point cloud sizes of the source and the target sets, respectively. The RMSE v , running time t , and number of iterations k are reported below each result. Our RSICP and Symmetric ICP are the only two methods that achieve good registrations, while our RSICP being much more accurate than Symmetric ICP. Our RSICP can also be faster than Symmetric ICP if we limit the maximum number of correspondences as in the Symmetric ICP, since our iteration number k is smaller than Symmetric ICP.

accuracy. Typical sampling strategies are random sampling, Voxel Grid filtering (Rusu and Cousins, 2011), and octree compression (Schnabel and Klein, 2006), etc. ICP uses nearest search for correspondence establishment. Some variants improve this step based on additional properties such as invariant features (Feldmar and Ayache, 1996) and normals (Serafin and Grisetti, 2015). (2) Error metric. Classical ICP uses the Euclidean distance between the transformed point and its closest point as the distance metric, called point-to-point metric. Chen and Medioni (Chen and Medioni, 1992) proposed a point-to-plane metric that measures the distance between the transformed point and the tangent plane of its correspondence. As evidenced, point-to-plane metric has a higher convergence speed than the point-to-point metric. Segal et al. (Segal et al., 2009) associated a probabilistic model in the ICP to construct a soft plane-to-plane metric. Recently, symmetric point-to-plane (Rusinkiewicz, 2019) shows great superiority in both convergence speed and convergence basin compared with the point-to-point and point-to-plane metrics. The functional zero-set of symmetric metric is a set of locally-second-order surfaces instead of just planar patches of the point-to-plane metric. (3) Outlier rejection. Some methods are heuristics, which use additional geometric constraints between distance and normals to reject possible outliers (Zhang, 1994; Rusinkiewicz and Levoy, 2001; Rusinkiewicz, 2019). However, as mentioned earlier, these methods are difficult to tune for different datasets. Robust estimation techniques are preferred for this problem. For example, least trimmed squares cost is used in Trimmed ICP (Chetverikov et al., 2005) and Anisotropic ICP (Maier-Hein et al., 2011) for alignment step. Robust ICP (Zhang et al., 2021) adapts the Welsch function as robust kernel. The common drawback of these M-like robust cost is the sensitive to high outlier rates. Sparse ICP (Bouaziz et al., 2013) introduces a sparse norm, i.e., ℓ_p -norm ($0 < p < 1$), to achieve robustness to outliers. However, optimization of the sparse norm suffers from high memory footprint and slow convergence. Besides the above improvements, several variants try to achieve globally optimal, such as the globally optimal ICP (Go-ICP) (Yang et al., 2015), which uses a nested Branch-and-Bound scheme for rigid transformation estimation. A comprehensive survey of ICP-type methods can be found in (Huang et al., 2021).

Apart from ICP-type registration, probabilistic-based registration is also a branch of point-based methods, such as the normal distributions

transform (NDT) (Biber and Straßer, 2003), Gaussian mixture models (GMMs) (Jian and Vemuri, 2010), and coherent point drift (CPD) (Myronenko and Song, 2010). Probabilistic-based approaches have lower dependencies on initializations than ICP-family. However, the results of ICP are more predictable than the ones of probabilistic-based registration (Magnusson et al., 2009).

3. Classical ICP revisited

Given a pair of point clouds $\mathcal{X} = \{x_1, \dots, x_M\}$ and $\mathcal{Y} = \{y_1, \dots, y_N\}$ in \mathbb{R}^3 , the goal of registration is to estimate a 6-DoF rigid transformation (represented by a 3-DoF rotation matrix \mathbf{R} and a 3-DoF translation vector \mathbf{t}) to align the source set \mathcal{X} with the target set \mathcal{Y} :

$$\min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^M (r_i(\mathbf{R}, \mathbf{t}))^2 + I_{SO(3)}(\mathbf{R}), \quad (1)$$

where $r_i(\mathbf{R}, \mathbf{t}) = \min_{y_j \in \mathcal{Y}} \|\mathbf{R}x_i + \mathbf{t} - y_j\|$ is the residual distance between the transformed point $\mathbf{R}x_i + \mathbf{t}$ and its closest point y_j in \mathcal{Y} , and $I_{SO(3)}(\cdot)$ provides a constraint on the rotation matrix \mathbf{R} , which should belong to the orthogonal group $SO(3)$. Exactly, $I_{SO(3)}(\cdot)$ acts as an indicator function:

这个约束好像加和没加类似

$$I_{SO(3)}(\mathbf{R}) = \begin{cases} 0, & \text{if } \mathbf{R}^T \mathbf{R} = \mathbf{I} \text{ and } \det(\mathbf{R}) = 1, \\ +\infty, & \text{otherwise.} \end{cases} \quad (2)$$

where $\mathbf{I} \in \mathbb{R}^{3 \times 3}$ is an identity matrix and $\det(\cdot)$ is the determinant of a matrix. Problem (1) is generally solved by the well-known ICP algorithm (Besl and McKay, 1992), which alternates between two major stages until convergence:

- *Correspondence update step:* seek the corresponding point $\hat{y}_i^{(k)} \in \mathcal{Y}$ for each point $x_i \in \mathcal{X}$ with fixed transformation parameters $(\mathbf{R}^{(k)}, \mathbf{t}^{(k)})$:

$$\hat{y}_i^{(k)} = \operatorname{argmin}_{y_j \in \mathcal{Y}} \|\mathbf{R}^{(k)}x_i + \mathbf{t}^{(k)} - y_j\|. \quad (3)$$

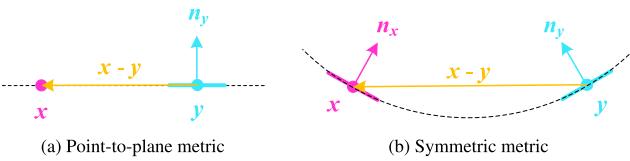


Fig. 2. Comparison between point-to-plane and symmetric point-to-plane metrics. **Left:** The point-to-plane error (Eq. 6) is zero only if point x locates on the tangent plane of y , regardless of n_x . **Right:** The symmetric point-to-plane error (Eq. 7) is zero whenever the local surface containing (x, y) is an arc of some circle, since the sum of normals n_x and n_y must be orthogonal to the vector $x - y$.

- *Transformation update step:* optimize the transformation parameters $(\mathbf{R}^{k+1}, \mathbf{t}^{k+1})$ based on Eq. (1) with known correspondence points $\{\hat{\mathbf{y}}_i^{(k)}, \mathbf{x}_i\}_1^M$:

$$\left(\mathbf{R}^{(k+1)}, \mathbf{t}^{(k+1)}\right) = \underset{\mathbf{R}, \mathbf{t}}{\operatorname{argmin}} \sum_{i=1}^M \|\mathbf{R}\mathbf{x}_i + \mathbf{t} - \hat{\mathbf{y}}_i^{(k)}\|^2 + I_{SO(3)}(\mathbf{R}), \quad (4)$$

where the superscript k is an iteration indicator. The alignment (4) is an absolute orientation problem, which can be solved by the Horn's method (Horn, 1987) in a closed-form. Apart from the above point-to-point distance metric, point-to-plane distance $r_i^l(\cdot)$ is another popular metric for ICP,

$$r_i^l(\mathbf{R}, \mathbf{t}) = (\mathbf{R}\mathbf{x}_i + \mathbf{t} - \hat{\mathbf{y}}_i) \cdot \mathbf{n}_{\hat{\mathbf{y}}_i}, \quad (5)$$

where symbol \mathbf{n} represents a normal vector, e.g., $\mathbf{n}_{\hat{\mathbf{y}}_i}$ is the normal of $\hat{\mathbf{y}}_i$. This metric measures the distance between the transformed point $\mathbf{R}\mathbf{x}_i + \mathbf{t}$ and the tangent plane of $\hat{\mathbf{y}}_i$. Compared with the point-to-point ICP, point-to-plane ICP has faster convergence speed.

4. Robust symmetric ICP

4.1. Symmetric point-to-plane metric

The symmetric point-to-plane metric proposed by Rusinkiewicz (Rusinkiewicz, 2019) had shown to have a wider convergence basin and faster convergence speed than other distance metrics. Its basic idea is to minimize point-to-surface error based on normals of both the source and the target points.

To show the superiority of symmetric metric, let's first compare the zero-sets of the objective functions of point-to-plane and symmetric point-to-plane distances. In the overlapping regions, the surfaces of point clouds \mathcal{X} and \mathcal{Y} are the same up to noise. Assume that the surfaces are perfectly aligned, a correspondence (x, y) becomes nearby points on the surface due to the discrete point cloud sampling step. Then, the point-to-plane error of (x, y) is:

$$r = (x - y) \cdot \mathbf{n}_y. \quad \text{这个考虑的是 } x \text{ 在相当靠近的情况下, 二者只有微小的Noise扰动, (6) 这个时候应该可以是做二者的residual为0, 已经收敛,}$$

Therefore, the residual is zero only if point x locates on the tangent plane of y . It means that the local surface containing (x, y) is perfectly flat, which is intuitively illustrated in Fig. 2(a). Considering the symmetric objective, the error is:

$$r = (x - y) \cdot (\mathbf{n}_x + \mathbf{n}_y). \quad (7)$$

Fig. 2(b) gives an illustration of the behavior of the objective (7) in 2D case. As shown, the residual reaches zero whenever the local surface

containing (x, y) is an arc of some circle, since the sum of normals n_x and n_y must be orthogonal to the vector $x - y$. This property can be easily extended to 3D case: function (7) gets zero error as long as x and y are sampled from some local cylinder. Further, Rusinkiewicz had proved that the symmetric metric can be minimized whenever the local surface is quadratic or planar (Rusinkiewicz, 2019). In summary, the zero-set of point-to-plane function is a class of local planar patches while the one of symmetric function is a set of locally-second-order surfaces. Therefore, the symmetric function has a wider convergence basin in theory.

The original symmetric metric in (Rusinkiewicz, 2019) is:

$$r_i^s(\mathbf{R}, \mathbf{t}) = (\mathbf{R}\mathbf{x}_i + \mathbf{t} - \mathbf{R}^{-1}\hat{\mathbf{y}}_i) \cdot (\mathbf{n}_{x_i} + \mathbf{n}_{\hat{\mathbf{y}}_i}). \quad (8)$$

In this function, both the point clouds \mathcal{X} and \mathcal{Y} are moved to a neutral coordinate system based on the symmetric split of the transformation. Namely, the optimized parameters (\mathbf{R}, \mathbf{t}) in (8) is a neutral transformation. However, the common process of the ICP-family is to transform one point cloud to another. Importantly, this function ignores the rotation of normals. According to (7), the normals should be in the registration coordinate system instead of its original coordinate system. Thus, it may cause loss of registration accuracy. Although Rusinkiewicz also presented an exact formulation for symmetric metric, it is complex to linearize and implement. 难道法向量是只计算一次就不更新了？

In this paper, we present a new symmetric point-to-plane metric:

$$r_i^s(\mathbf{R}, \mathbf{t}) = (\mathbf{R}\mathbf{x}_i + \mathbf{t} - \hat{\mathbf{y}}_i) \cdot (\mathbf{R}\mathbf{n}_{x_i} + \mathbf{n}_{\hat{\mathbf{y}}_i}). \quad (9)$$

Similar to the most ICP variants, this function performs a rigid transformation to only the source set. Since our symmetric metric almost does not increase the amount of calculation, it achieves the efficiency and simplicity properties of the point-to-plane ICP. Moreover, our symmetric metric has a wider convergence basin, faster convergence speed, and better registration accuracy than the original one (See Section 5.1.1).

4.2. Robust loss for symmetric ICP

Traditional ICP uses the ℓ_2 norm as the cost in the alignment step (transformation update step), which relies an assumption of Gaussian noise of the observations. The behavior of the ℓ_2 cost is to penalize correspondences with large residuals established in the first step. However, it may cause erroneous alignment due to outliers and partial overlaps, in which a portion of source points in \mathcal{X} have no correct correspondences in the target set \mathcal{Y} . In these cases, the Gaussian noise assumption is violated and the ℓ_2 cost produces a large error under the ground-truth alignment. This problem is generally solved by replacing the ℓ_2 function with a robust loss, which gives high confidences to the true correspondences, while excluding the outlier correspondences in the optimization by assigning very small weights. Therefore, the behavior of a robust loss is to induce sparsity of the distance metric for the ICP. For example, the ℓ_p -norm ($p \in (0, 1)$) cost is introduced in Sparse ICP, which minimizes the function $\sum_{i=1}^M (r_i(\mathbf{R}, \mathbf{t}))^p$. The only difference with problem (1) is to use the ℓ_p -norm instead of the ℓ_2 -norm. However, the ℓ_p based alignment step is a non-convex and non-smooth function, which is optimized by the alternating direction method of multipliers (ADMM) since no closed-form solution is available. The drawbacks of sparse ICP are the high computational cost and memory footprint. Another recently proposed example is the Robust ICP that adapts the Welsch robust loss. Robust ICP shows great improvement compared with classical ICP in cases of outliers and partial overlaps. The limitation of Robust ICP is the sensitivity to high outlier rates.

In this paper, we introduce a different robust loss for symmetric metric, which has a high “breakdown” point and low computational overhead. The objective function of our RSICP is:

$$\min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^M \rho_{(\alpha, \beta)}(r_i^s(\mathbf{R}, \mathbf{t})) + I_{SO(3)}(\mathbf{R}), \quad (10)$$

where $\rho_{(\alpha, \beta)}$ is our robust model (Li et al., 2021):

$$\rho_{(\alpha, \beta)}\left(\begin{array}{c} r \\ \mathbf{R} \end{array}\right) = \begin{cases} \frac{\beta^2}{2} \ln\left(1 + \left(\frac{r}{\beta}\right)^2\right) & \alpha = 0 \\ \frac{\beta^2}{\alpha} \left(\left(1 + \left(\frac{r}{\beta}\right)^2\right)^{\frac{\alpha}{2}} - 1\right) & \alpha \neq 0 \end{cases}, \quad (11)$$

α is a robustness-control parameter, and $\beta > 0$ is a scale factor. This robust model acts as different functions with different parameter α . For instance, it is the ℓ_2 loss when $\alpha = 2$, the ℓ_1 - ℓ_2 loss when $\alpha = 1$, the Cauchy loss when $\alpha = 0$, and the Geman-McClure loss when $\alpha = -2$. Our model is able to connect the non-robust ℓ_2 cost and robust M-estimation cost. Hence, we can gradually improve the degrees of robustness via the decay of α in the optimization. As α decreases, many outliers are assigned ≈ 0 weights. In other words, the point-wise distance between the source set and the target set becomes sparse.

Similar to Robust ICP, our formulation is also non-convex. Fortunately, based on the Black-Rangarajan duality (Black and Rangarajan, 1996) (See Appendix A), we can reformulate it as:

$$\min_{\mathbf{R}, \mathbf{t}, w_i \in [0, 1]} \sum_{i=1}^M w_i \cdot (r_i^s(\mathbf{R}, \mathbf{t}))^2 + \chi_\rho(w_i) + I_{SO(3)}(\mathbf{R}), \quad (12)$$

where $\chi_\rho(w_i)$ is a penalty on the weight $w_i = w_{(\alpha, \beta)}(r_i^s(\mathbf{R}, \mathbf{t}))$, whose formulation is as follows:

$$\chi_\rho\left(\begin{array}{c} w_i \\ \mathbf{R} \end{array}\right) = \begin{cases} \beta^2(w_i - \ln w_i) & \alpha = 0 \\ \beta^2 w_i \left(1 - \frac{\alpha - 2}{\alpha} w_i^{\frac{2}{\alpha-2}}\right) & \alpha \neq 0 \end{cases}, \quad (13)$$

and the weight function $w_{(\alpha, \beta)}(r)$ is:

$$w_{(\alpha, \beta)}\left(\begin{array}{c} r \\ \mathbf{R} \end{array}\right) = \frac{\partial \rho_{(\alpha, \beta)}}{\partial r} / r = \begin{cases} \frac{\beta^2}{\beta^2 + r^2} & \alpha = 0 \\ \left(1 + \left(\frac{r}{\beta}\right)^2\right)^{\frac{\alpha}{2}-1} & \alpha \neq 0 \end{cases}. \quad (14)$$

In our robust model, parameter α is important and controls the robustness. It begins with the least-squares cost ($\alpha = 2$) and gradually becomes a robust M-estimate along with iterations. A larger α allows more observations to take part in the alignment (e.g., giving equal weights to all correspondences at the beginning.), which helps to avoid undesirable local minima. Meanwhile, a smaller α largely discounts the influence from outlier observations (e.g., assigning large weights to point pairs with small transformed distances, while giving small weights to the ones with large residuals.). Therefore, our model performs more global optimization with more observations and gradually reduces the effect of observations with large residuals to obtain high-precise alignment.

Several ICP variants achieve robustness through discarding the correspondences with large position or normal errors. However, as point out in (Zhang et al., 2021), it is difficult to tune these methods to achieve good results. Further, these strategies may increase the possibility of getting stuck in local minima. Differently, our method acts as a dynamic thresholding approach which gradually penalizes outliers. Compared with Robust ICP, our registration model has better performance in ac-

curacy and robustness, which benefits from our symmetric point-to-plane metric and the adaptive robust model.

4.3. Main algorithm and linearization

Similar to the classical ICP, our RSICP also has a correspondence step and an alignment step. Differently, our alignment is an iteratively reweighted least squares (IRLS) problem instead of the least-squares, which contains a weighted least-squares estimation and a weight update step. The main procedure is as follows:

- *Correspondence update*: seek the correspondence $\hat{y}_i^{(k)} \in \mathcal{Y}$ for each point $x_i \in \mathcal{X}$ with fixed transformation parameters $(\mathbf{R}^{(k)}, \mathbf{t}^{(k)})$:

$$\hat{y}_i^{(k)} = \operatorname{argmin}_{y_j \in \mathcal{Y}} \|(\mathbf{R}^{(k)}x_i + \mathbf{t}^{(k)} - y_j) \cdot (\mathbf{R}^{(k)}n_{x_i} + n_{y_j})\|. \quad (15)$$

This step can be approximated by the step (3) or we can search the closest $m \ll N$ points in \mathcal{Y} for x_i and optimize (15) on these selected points.
matching — neighbor search

- *Transformation update*: optimize the transformation parameters $(\mathbf{R}^{(k+1)}, \mathbf{t}^{(k+1)})$ based on Eq. (12) with known correspondences $\{(\hat{y}_i^{(k)}, x_i)\}_1^M$ and weights $\{w_i^{(k)}\}_1^M$:

$$\left(\mathbf{R}^{(k+1)}, \mathbf{t}^{(k+1)}\right) = \operatorname{argmin}_{\mathbf{R} \in SO(3), \mathbf{t}} \sum_{i=1}^M w_i^{(k)} \|(\mathbf{R}x_i + \mathbf{t} - \hat{y}_i^{(k)}) \cdot (\mathbf{R}n_{x_i} + n_{\hat{y}_i^{(k)}})\|^2, \quad (16)$$

where the item $\chi_\rho(w_i)$ is dropped since it does not depend on the transformation parameters. It is difficult to directly optimize the problem (16), since our symmetric point-to-plane metric is a quadratic function of \mathbf{R} . Fortunately, the rotation is gradually recovered along with iterations, which means that the increment of rotation $\Delta\mathbf{R}$ is small in each iteration. Thus we can make an approximation on the rotated normal with the known rotation \mathbf{R} of the last iteration. Then, problem (16) becomes:

$$\left(\mathbf{R}^{(k+1)}, \mathbf{t}^{(k+1)}\right) = \operatorname{argmin}_{\mathbf{R} \in SO(3), \mathbf{t}} \sum_{i=1}^M w_i^{(k)} \|(\mathbf{R}x_i + \mathbf{t} - \hat{y}_i^{(k)}) \cdot (\mathbf{R}^{(k)}n_{x_i} + n_{\hat{y}_i^{(k)}})\|^2. \quad (17)$$

The term $\mathbf{R}^{(k)}$ is a known value. Hence, it becomes a linear function of \mathbf{R} instead of the quadratic one in the problem (16). Then, we can easily optimize this problem in the same way as the point-to-plane metric (If we set the known term $(\mathbf{R}^{(k)}n_{x_i} + n_{\hat{y}_i^{(k)}})$ to be $n_{\hat{y}_i^{(k)}}$, problem (17) exactly becomes the point-to-plane metric).

- *Weight update*: optimize weights $\mathbf{w}^{(k+1)} = \{w_i^{(k+1)}\}_1^M$ based on Eq. (12) with fixed transformation parameters $(\mathbf{R}^{(k+1)}, \mathbf{t}^{(k+1)})$ and known correspondences $\{(\hat{y}_i^{(k)}, x_i)\}_1^M$:
weight function!

$$\mathbf{w}^{(k+1)} = \operatorname{argmin}_{w_i \in [0, 1]} \sum_{i=1}^M w_i \|r_i^s(\mathbf{R}^{(k+1)}, \mathbf{t}^{(k+1)})\|^2 + \chi_\rho(w_i), \quad (18)$$

w既是一个值，也是一个函数…

where r_i^s is a constant with a given $(\mathbf{R}^{(k+1)}, \mathbf{t}^{(k+1)})$. The solution of (18) is exactly equivalent to (14).

- *α update*: decrease α by a step-size τ , i.e., $\alpha^{(k+1)} = \alpha^{(k)} - \tau$, to change the weight/cost function.

These steps are iteratively performed until convergence. The details are summarized in Algorithm 1.

Algorithm 1. Robust symmetric ICP**Algorithm 1:** Robust symmetric ICP

```

Input: The source set  $\mathcal{X}$  and the target set  $\mathcal{Y}$ .
// Transformation  $\mathbf{T} = (\mathbf{R}, \mathbf{t})$ 
Output: Transformation  $\mathbf{T}$  to align  $\mathcal{X}$  with  $\mathcal{Y}$ .
1 Initialization:  $\mathbf{T} = \mathbf{T}^{(0)}$ ,  $\alpha^{(0)} = 2$ ,  $\beta = \mu$ ,  $n = 100$ ,  $\epsilon = 10^{-5}$ ,
 $k = 1$ ,  $\tau = 0.5$ ;
//  $\mu$ : point cloud resolution;
//  $n$ : maximum number of iterations;
//  $\epsilon$ : convergence threshold;
//  $\tau$ : a decay step for  $\alpha$ 

2 while TRUE do
3    $k_{start} = k - 1$ ;
4   while  $k - k_{start} \leq n$  do
5     // Correspondence update (Eq. (15))
6      $\widehat{\mathcal{Y}}^{(k)} = correspondence(\mathbf{T}^{(k-1)})$ ;
7     // Weight update via Eq. (18)
8      $\mathbf{w}^{(k)} = weight(\widehat{\mathcal{Y}}^{(k)}, \mathbf{T}^{(k-1)})$ ;
9     // Transformation update (Eq. (17))
10     $\mathbf{T}^{(k)} = alignment(\widehat{\mathcal{Y}}^{(k)}, \mathbf{T}^{(k-1)}, \mathbf{w}^{(k)})$ ;
11    // Check convergence
12    if  $\|\mathbf{T}^k - \mathbf{T}^{k-1}\|_F < \epsilon$  then break ;
13     $k = k + 1$ ;
14 end
15 // Check stop criteria
16 if  $\alpha < -2$  then return  $\mathbf{T} = \mathbf{T}^{(k)}$ ;
17 //  $\alpha$  update step
18  $\alpha^{(k)} = \alpha^{(k-1)} - \tau$ ;  $k = k + 1$ ;
19 end

```

The transformation update (alignment) step is a non-linear problem. To make the optimization easier, we first transfer it to a linear least-squares problem based on the small incremental rotation assumption. Specifically, we first use the Rodrigues rotation formula to parameterize the rotation:

$$\mathbf{Rx}_i = \mathbf{x}_i \cos\theta + (\mathbf{v} \times \mathbf{x}_i) \sin\theta + \mathbf{v}(\mathbf{v} \cdot \mathbf{x}_i)(1 - \cos\theta). \quad (19)$$

Under the small incremental rotation assumption, we have $\cos\theta \approx 1$ and $\sin\theta \approx \theta$. Then, the above equation can be approximated by:

$$\begin{aligned} \mathbf{Rx}_i &\approx \mathbf{x}_i + (\mathbf{v} \times \mathbf{x}_i)\theta \\ &= \mathbf{x}_i + (\mathbf{v} \times \mathbf{x}_i), \end{aligned} \quad (20)$$

where \mathbf{v} represents the axis of \mathbf{R} , θ is the angle of \mathbf{R} , and $\mathbf{v}' = \mathbf{v}\theta$. Substituting it into Eq. (17):

$$\underset{\mathbf{v}', \mathbf{t}}{\operatorname{argmin}} \sum_{i=1}^M w_i^{(k)} \left((\mathbf{x}_i - \widehat{\mathbf{y}}_i^{(k)}) \cdot \mathbf{n}_i + (\mathbf{x}_i \times \mathbf{n}_i) \cdot \mathbf{v}' + \mathbf{n}_i \cdot \mathbf{t} \right)^2, \quad (21)$$

where $\mathbf{n}_i = \mathbf{R}^{(k)} \mathbf{n}_{x_i} + \mathbf{n}_{y_i}$. This weighted linear least-squares can be written in the following form:

$$\underset{\xi}{\operatorname{argmin}} \mathbf{e}^T \mathbf{W} \mathbf{e} = \underset{\xi}{\operatorname{argmin}} (\mathbf{A} \xi - \mathbf{b})^T \mathbf{W} (\mathbf{A} \xi - \mathbf{b}), \quad (22)$$

where $\xi = [\mathbf{v}'; \mathbf{t}]$ is the variable to be estimated, $\mathbf{e} = (\mathbf{A} \xi - \mathbf{b})$ is the error equation, diagonal matrix $\mathbf{W} = diag(w_1 \cdots w_M)$ is a $M \times M$ weight matrix, $\mathbf{A} = [A_1^T \cdots A_i^T \cdots A_M^T]^T$ is a $M \times 6$ coefficient matrix and $\mathbf{b} = [b_1 \cdots b_i \cdots b_M]^T$ is a $M \times 1$ observation vector, where:

Table 1

Detailed settings of each compared algorithms (MNII and MNOI represent maximum number of inner ICP and outer robust iterations, respectively; v_{\max} and v_{\min} are the maximum and minimal values for the Welsch cost, respectively. All methods run in single thread.).

Method	Parameters	Implementations
ICP	Error metric: point-to-point; convergence threshold: $1e-5$; MNII: 100.	C++ code https://github.com/ OpenGP/sparseicp
ICP-l	Error metric: point-to-plane; convergence threshold: $1e-5$; MNII: 100.	C++ code https://github.com/ OpenGP/sparseicp
AA-ICP	Error metric: point-to-point; alpha limit: 10; history length for Anderson: 5; convergence threshold: $1e-5$; MNII: 100.	C++ code https://github.com/yaoxy689/Fast-Robust-ICP
Sparse ICP	Error metric: point-to-plane; p -norm: $p = 0.5$; convergence threshold: $1e-5$; MNII: 100.	C++ code https://github.com/ OpenGP/sparseicp
Fast ICP	Error metric: point-to-point; convergence threshold: $1e-5$; history length for Anderson: 5; MNII: 100.	C++ code https://github.com/yaoxy689/Fast-Robust-ICP
Robust ICP	Error metric: point-to-plane; history length for Anderson: 5; $v_{\max} = 3\bar{D}^{(0)}$; $v_{\min} = \bar{E}_Q/3\sqrt{3}$; convergence threshold: $1e-5$; MNII: 100; MNOI: 100.	C++ code https://github.com/yaoxy689/Fast-Robust-ICP
Symmetric ICP	Thresholds {distance: $0.5 \times$ centroid distance; angle: 60° ; convergence: $1e-5$ }; error metric: symmetric point-to-plane; number of correspondences: 5×10^4 ; MNII: 100.	C++ code https://gfx.cs.princeton.edu/proj/trimesh2/
RSICP (ours)	Error metric: improved symmetric point-to-plane; step-size: $\alpha = 0.5$; convergence threshold: $1e-5$; MNII: 100; MNOI: 100.	MATLAB code https://ljy-rs.github.io/web/

$$\begin{cases} \mathbf{A}_i = [(\mathbf{x}_i \times \mathbf{n}_i)^T \quad \mathbf{n}_i^T]^T \\ b_i = (\widehat{\mathbf{y}}_i^{(k)} - \mathbf{x}_i) \cdot \mathbf{n}_i \end{cases}. \quad (23)$$

Then, the solution of (22) is:

$$\begin{aligned} \widehat{\xi} &= (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{W} \mathbf{b}) \\ &= \left(\sum_{i=1}^M \mathbf{A}_i^T w_i \mathbf{A}_i \right)^{-1} \left(\sum_{i=1}^M \mathbf{A}_i^T w_i b_i \right) \end{aligned} \quad (24)$$

5. Results

In this section, both synthetic and real experiments are designed to comprehensively evaluate the proposed algorithm (denoted as RSICP). We compare our performance with seven ICP-based baselines or state-of-the-arts, including point-to-point ICP (denoted as ICP) (Besl and McKay, 1992), point-to-plane ICP (denoted as ICP-l) (Chen and Medioni, 1992), Anderson acceleration ICP (denoted as AA-ICP) (Pavlov et al., 2018), Sparse ICP (Bouaziz et al., 2013), Fast ICP (Zhang et al., 2021), Robust ICP (Zhang et al., 2021), and Symmetric ICP (Rusinkiewicz, 2019). For Sparse ICP and Robust ICP, we use the point-to-plane distance metric, since its convergence speed is faster than the point-to-point metric. Then, it is fairer to compare them with Symmetric ICP and our RSICP that are based on improved point-to-plane metrics. The parameter settings and implementation details of each compared algorithm are summarized in Table 1. In the following experiments, input point clouds are first normalized by eliminating their centroids and scaling them such

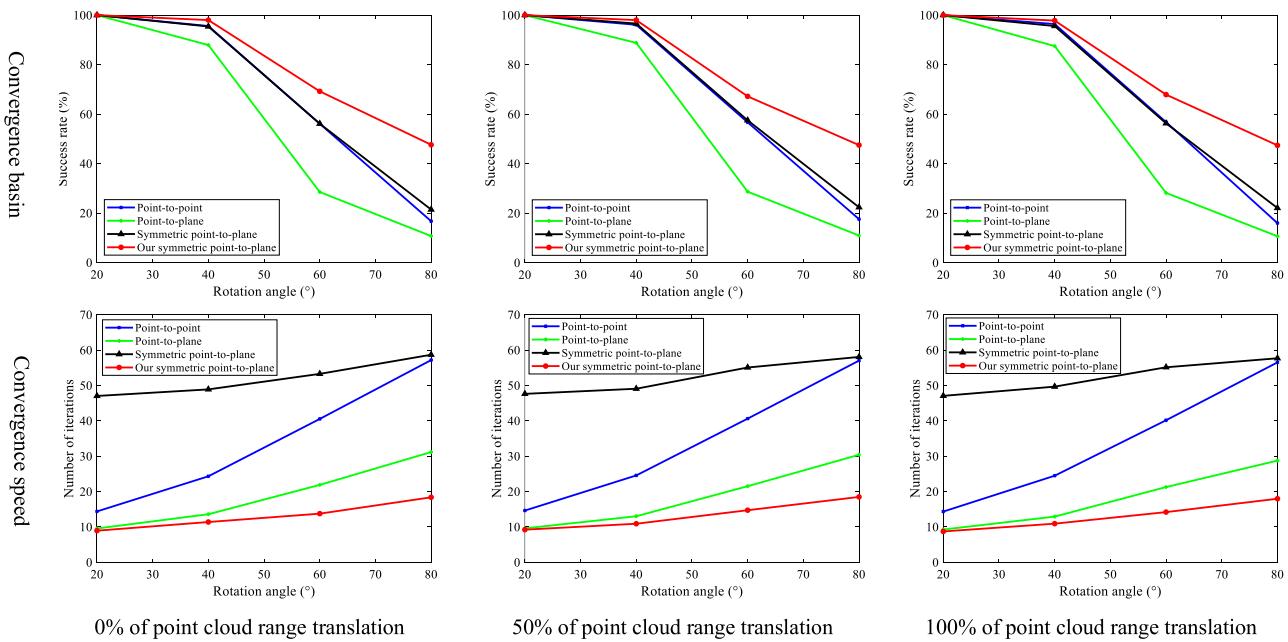


Fig. 3. Comparison of the convergence basin and convergence speed between different distance metrics on the FGR synthetic dataset, where $\{[0^\circ, 20^\circ], [20^\circ, 40^\circ], [40^\circ, 60^\circ], [60^\circ, 80^\circ]\}$ are represented as points $\{20^\circ, 40^\circ, 60^\circ, 80^\circ\}$ in the figure.

that the diagonal length of their bounding box is 1. Since the point-to-plane and symmetric point-to-plane metrics require normals at the points, we use the Point Cloud Toolbox¹ to estimate normals if they are unknown. The same as in (Zhang et al., 2021), we adapt the registration root mean square error (RMSE) for quantitative evaluation,

$$v = \sqrt{\frac{1}{M} \sum_{i=1}^M \| \mathbf{R}^* \mathbf{x}_i + \mathbf{t}^* - \mathbf{R}\mathbf{x}_i - \mathbf{t} \|_2^2} \quad (25)$$

where (\mathbf{R}, \mathbf{t}) is an estimated rigid transformation and $(\mathbf{R}^*, \mathbf{t}^*)$ is the ground truth alignment. We also report the iteration count and running time of each compared method. All the methods are performed on a laptop with single CPU Core i7-8550U at 1.8 GHz and 8 GB of RAM.

5.1. Synthetic data

Here, we first compare the convergence basin and convergence speed of our proposed symmetric point-to-plane with classical point-to-point, point-to-plane, and the original symmetric point-to-plane distance metrics. Then, We comprehensively evaluate our robustness to outliers and partial overlaps.

5.1.1. Convergence basin and speed

We use the dataset constructed in the fast global registration (FGR) (Zhou et al., 2016) for simulations, which contains a total of 50 point clouds. For each point cloud, we apply a randomly generated rigid transformation (φ, t) to obtain its target point cloud, where φ is a 3×1 rotation angle vector. This experiment only studies the convergence properties of different distance metrics. Therefore, we do not simulate any outliers, partial overlaps, and initializations. We only add Gaussian noise to the target points along normal directions with the standard deviation being the point cloud resolution μ . In our experiment, φ is generated in four categories, i.e., $\{[0^\circ, 20^\circ], [20^\circ, 40^\circ], [40^\circ, 60^\circ], [60^\circ, 80^\circ]\}$, and t is set to be one of the three configurations, i.e., $\{0, 50\% \times d, 100\% \times d\}$, where d is the diagonal length of the bounding box of source set. For each (φ, t) category configuration, we perform 100 independent

tests. We use the success rate and number of iterations to reflect the convergence basin and convergence speed, respectively. If the RMSE v of a method in a test is smaller than triple of the noise standard deviation, this registration is accepted as a good registration. Thus, the success rate of a method is the number of good registrations that the method achieves in total 100 registrations. Note that the number of iterations for each method refers to its average value of all good registrations.

Fig. 3 plots the quantitative results. When the rotation angles are small ($< 20^\circ$), all the distance metrics achieve a 100% success rate. However, the success rate decreases as the rotation angle increases. Compared with rotation, translation has little influence on the performance. Our symmetric point-to-plane metric has the widest convergence basin while point-to-plane has the narrowest one. When the rotation is within $[60^\circ, 80^\circ]$, our metric (47.6%) gains a growth rate of 26% compared with original symmetric point-to-plane (21.4%), which ranks the second. In terms of convergence speed, original symmetric point-to-plane is the slowest. The reason may be that its implementations provided by the author limit the number of correspondences and use an outlier rejection module based on correspondence distance and normal angle. Reasonably, point-to-plane converges much faster than point-to-point metric. Again, our symmetric metric gets the fastest convergence speed. It is 200% faster than point-to-point and symmetric point-to-plane, and 70% faster than point-to-plane if the rotation is within $[60^\circ, 80^\circ]$.

5.1.2. Outliers and partial overlaps

Fig. 1 compares registration results on a point cloud pair with partial overlap, which is obtained from the monkey model of the EPFL statue dataset². Specifically, the first 60% of the points from the full model are taken to produce the source point cloud, and the last 60% are regarded as the target one. The overlap of this pair is only about 33%. Then, a randomly generated rigid transformation $(\varphi \in [0^\circ, 45^\circ], t \in [0, d])$ is applied to the target set. As shown, our RSICP achieves the highest RMSE accuracy among eight compared methods, while being comparable with AA-ICP and Fast ICP in terms of running time (Note that our RSICP is implemented in MATLAB). Sparse ICP and Robust ICP, the two typical

¹ <https://www.mathworks.com/help/vision/ref/pcnormals.html>.

² https://lgg.epfl.ch/statues_dataset.php.

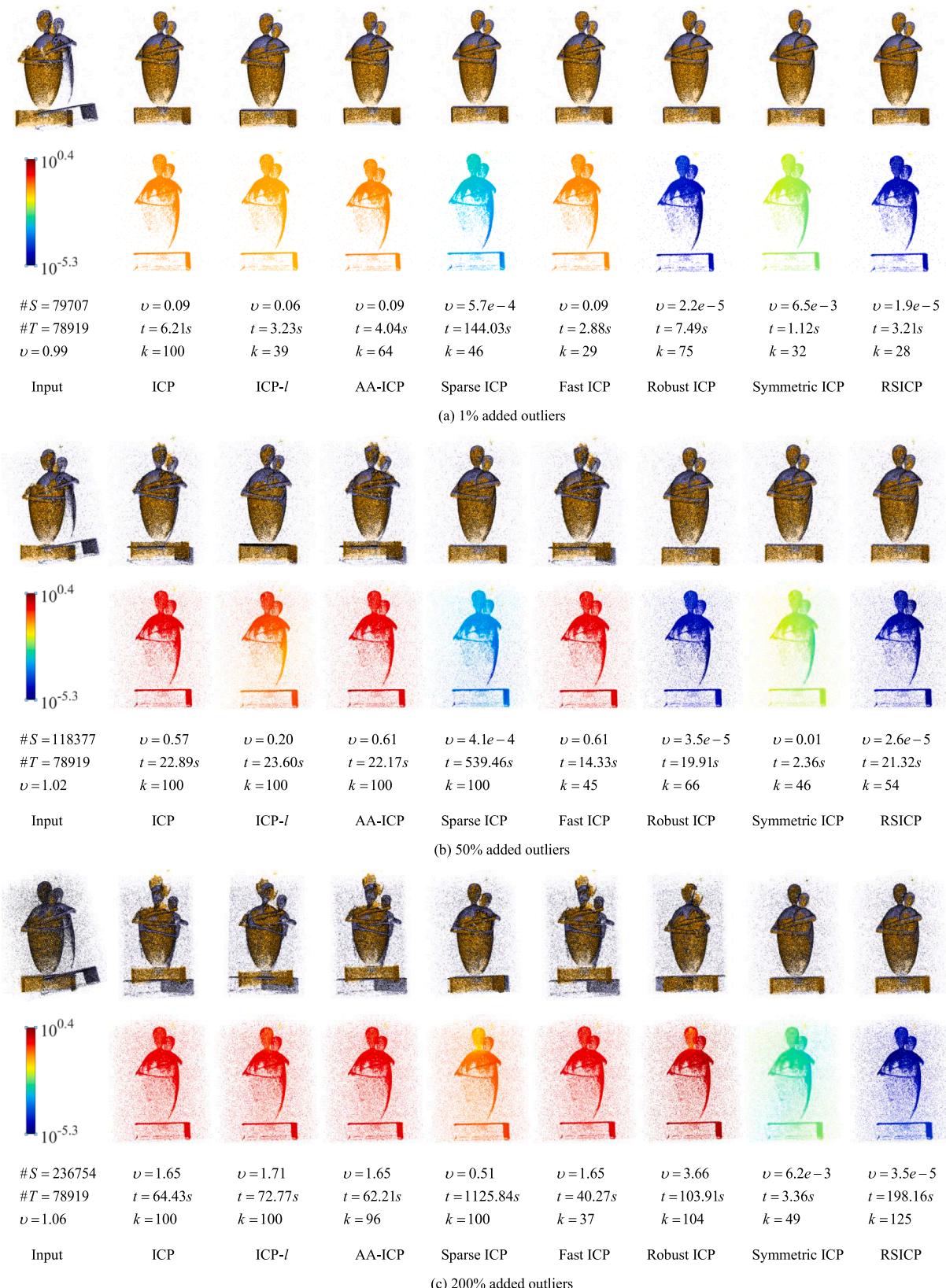


Fig. 4. Comparison between different methods on point clouds with outliers and partial overlaps, generated from the femme enfant model of the EPFL dataset. In each subfigure, the first row provides registration results and the second raw shows error maps between the estimated alignment and the ground truth alignment of the source set using log-scale color-coding.

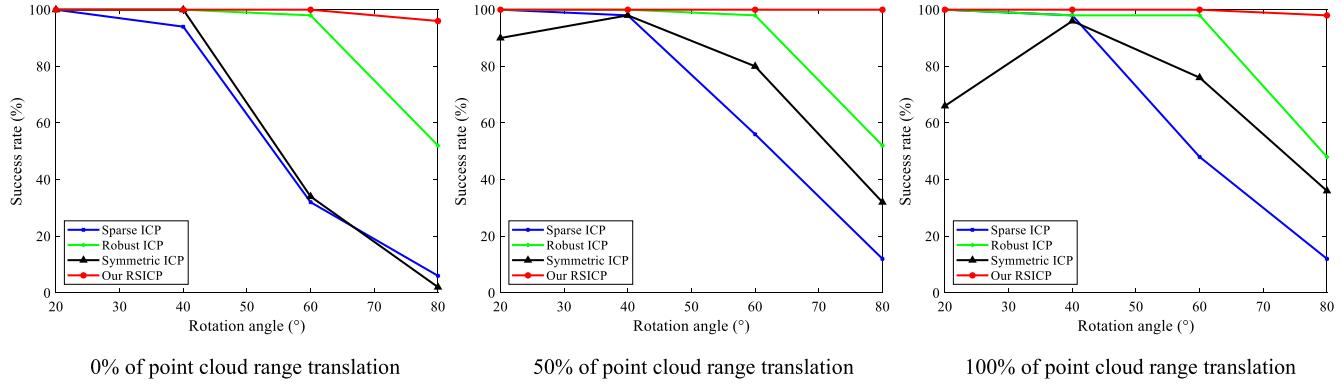


Fig. 5. Comparison of the sensitivity to initializations between different methods, where $\{[0^\circ, 20^\circ], [20^\circ, 40^\circ], [40^\circ, 60^\circ], [60^\circ, 80^\circ]\}$ are represented as points $\{20^\circ, 40^\circ, 60^\circ, 80^\circ\}$ in the figure.

ICP methods with high degrees of robustness, fail to achieve good registrations, since their convergence basins are narrow. In this registration, no initialization information is provided. Symmetric ICP performs much better than ICP, ICP-*l*, and their accelerated variants. It uses distance and normal constraints to reject outliers. Hence, it has good robustness to outliers and partial overlaps. Symmetric ICP is the fastest among all methods, which benefits from the limit of correspondences. It only samples no more than 5×10^4 correspondences for optimization. Actually, it will be slower than our RSICP if all correspondences are taken in optimization, since its iteration number is larger than ours.

Fig. 4 compares the methods on point clouds with outliers and partial overlaps, which are generated from the femme enfant model of the EPFL dataset. Similarly, the first 55% of the points from the full model are taken to produce the source point cloud, and the last 55% are regarded as the target one. Then, a random transformation ($\varphi \in [0^\circ, 10^\circ], t \in [0, d/10]$) is applied on the target set. In this test, we want to eliminate the impact of initializations on different methods. Therefore, we only apply small rotation angles and offsets. To simulate outliers, $\eta \cdot M$ random points are uniformly generated from the bounding box of source set, where η is set to be 1%, 50% and 200%, respectively. From the results, ICP, ICP-*l*, and there accelerated variants are very sensitive to outliers. This is reasonable, since there are no outlier process strategies in these methods. For case with up to $\eta = 50\%$ added outliers, our RSICP and Robust ICP achieve better registration accuracies than other robust-type ICPs (Sparse ICP and Symmetric ICP). For 200% added outliers, Sparse ICP and Robust ICP can no longer obtain good registrations. Symmetric ICP uses distance and normal constraints to reject outliers. It has very good robustness when good initializations are available. However, this outlier rejection strategy may cause registration accuracy loss. Our RSICP achieves the best RMSE among all methods. This example shows that our RSICP has better robustness than Sparse ICP and Robust ICP, and has better accuracy than Symmetric ICP.

Fig. 5 shows the sensitivity to initializations of different methods. In Section 5.1.1, we report the convergence basin of different distance metrics. Hence, we only compare with the three methods with robustness to outliers and partial overlaps in this example, i.e., Sparse ICP, Robust ICP, and Symmetric ICP. The first 60% of the points from the femme enfant model are taken as the source set, and the last 60% are the target set. The configurations of noise, random rotation angles φ , and random translations t are the same as in Section 5.1.1. For problems with small rotation angles, all methods achieve good performance. However, as the rotation increases, Sparse ICP and Symmetric ICP become very unreliable, i.e., their success rates dramatically drop. Robust ICP has much better robustness to large rotations. But it is still much worse than our RSICP. Our success rate is always close to 100%, which gains a growth rate of $\approx 45\%$ compared with Robust ICP, the method that ranks second, when the rotation is within $[60^\circ, 80^\circ]$. For problems with translations, the behavior of Symmetric ICP looks a bit strange. Its

performance under $[0^\circ, 20^\circ]$ rotations is lower than the one under $[20^\circ, 40^\circ]$ rotations. The reason may be that the correspondence picking and outlier rejection strategies establish wrong correspondence relationship in this example with large translation offsets. This example shows that our RSICP has less sensitivity to initializations than current state-of-the-arts.

Our method is further evaluated on the FGR dataset, which consists of five point cloud models with partial overlaps, i.e., Bimba, Children, Chinese Dragon, Angle, and Bunny. Each model contains five point cloud pairs, where each pair only suffers from rotation changes. As shown in the above, when rotation angle is smaller than 20° , all robust methods can achieve good results. Hence, we add a random rotation within $[-20^\circ, 20^\circ]$ to the ground truth to obtain the initializations. One problem instance from each model is selected for visual comparisons in Fig. 6, which shows the RMSE, running time, and iteration number of each compared method, where the RMSE is reflected by the error map between the estimated registration and the ground truth registration of the source set. As shown, only our RSICP and Robust ICP get good registrations on all the five pairs. Sparse ICP and Symmetric ICP fail on the first pair and the fourth pair, respectively. Quantitative results are reported in Table 2. Again, our RSICP and Robust ICP are the only two methods that achieve good results on all pairs. Our RSICP gets comparable RMSE accuracy to Robust ICP, while being much faster. Our running time ranks second, which is 50% of Robust ICP, 33% of Symmetric ICP, and only 2% of Sparse ICP.

5.2. Real-world data

To test our RSICP on real-world problem instances, we compare the methods on the large-scale KITTI dataset³. We select five sequence (03 ~ 07) with a total of 6035 laser scans from the KITTI training dataset for evaluations, where sequences 03 and 04 are taken in country scenery, and 05 ~ 07 are captured in urban scenery. A ground truth pose file for each sequence is provided by a high precision GPS-INS equipment. For each sequence, we use a registration method to align each two adjacent laser scans and recover their 6-DoF relative pose information. This is known as the laser odometry problem. Here, we only perform scan-to-scan matching; namely, scan-to-model registration is not adapted. There are a total of 6030 registration instances in this experiment. For each registration instance, we use the relative transformation between the previous scan pair as the initialization. The same as in (Zhang and Singh, 2014), the average relative translation error (RTE) is used as the evaluation metric.

Fig. 7 plots the estimated trajectories of four methods with good robustness. Robust ICP and Sparse ICP get very bad results on the

³ http://www.cvlibs.net/datasets/kitti/eval_odometry.php.



Fig. 6. Registration results on several example instances of the FGR dataset with partial overlaps. The registration RMSE, running time, and iteration number are reported below each result, where the RMSE is reflected by the error map using log-scale color-coding.

sequence 04, where their estimated trajectory lengths are shorter than a half of the ground truth. The reason is that they fail to recover the true translations on many instances, since the driving speed of 04 sequence is the fastest among these five sequences. Our RSICP achieves the best results, whose trajectories are very close to the ground truth.

Table 3 reports the average RTE and total running time of each sequence. For this laser odometry problem, the performance of Robust

ICP is much worse than that on the synthetic data. Its average RTE of all sequences is even larger than non-robust ICP variants. The reasons are twofold: First, the overlapping ratios between two adjacent laser scans are much higher than the ones in the synthetic experiments. Hence, non-robust ICP variants can achieve much better results. Second, Robust ICP is less stable than non-robust ICP variants. Although Robust ICP obtains better registration accuracy than non-robust ICP methods on most of the

Table 2

Average RMSE ($\times 10^{-3}$), running time (in seconds), and number of iterations for each method on the partially overlapping FGR dataset constructed by five models. The best results are highlighted in bold fonts, and the second best results are underlined.

Dataset	Bimba			Children			Dragon			Angle			Bunny		
	RMSE	Time	k												
ICP	83	0.81	88	57	0.41	49	64	0.88	74	29	0.66	59	64	0.63	62
ICP-l	34	0.16	<u>12</u>	15	0.14	<u>11</u>	92	0.56	<u>30</u>	17	0.20	<u>12</u>	23	0.38	<u>31</u>
AA-ICP	79	0.74	75	57	0.31	33	64	0.72	59	29	<u>0.41</u>	<u>25</u>	64	0.57	51
Sparse ICP	100	17.54	43	2.1	19.98	55	2.0	24.84	53	1.9	18.5	33	1.9	24.70	65
Fast ICP	79	0.55	38	57	0.28	27	61	<u>0.54</u>	33	29	0.46	34	64	<u>0.42</u>	28
Robust ICP	<u>1.5</u>	0.78	64	2.0	0.68	51	2.0	1.00	57	<u>1.9</u>	0.87	49	1.8	0.78	60
Symmetric ICP	1.5	0.95	44	2.1	1.10	50	2.1	1.03	47	18	1.35	54	1.9	1.01	48
Our RSICP	1.5	<u>0.37</u>	<u>28</u>	<u>2.0</u>	<u>0.28</u>	<u>25</u>	<u>2.0</u>	<u>0.32</u>	<u>26</u>	<u>1.9</u>	0.44	29	<u>1.9</u>	0.49	34

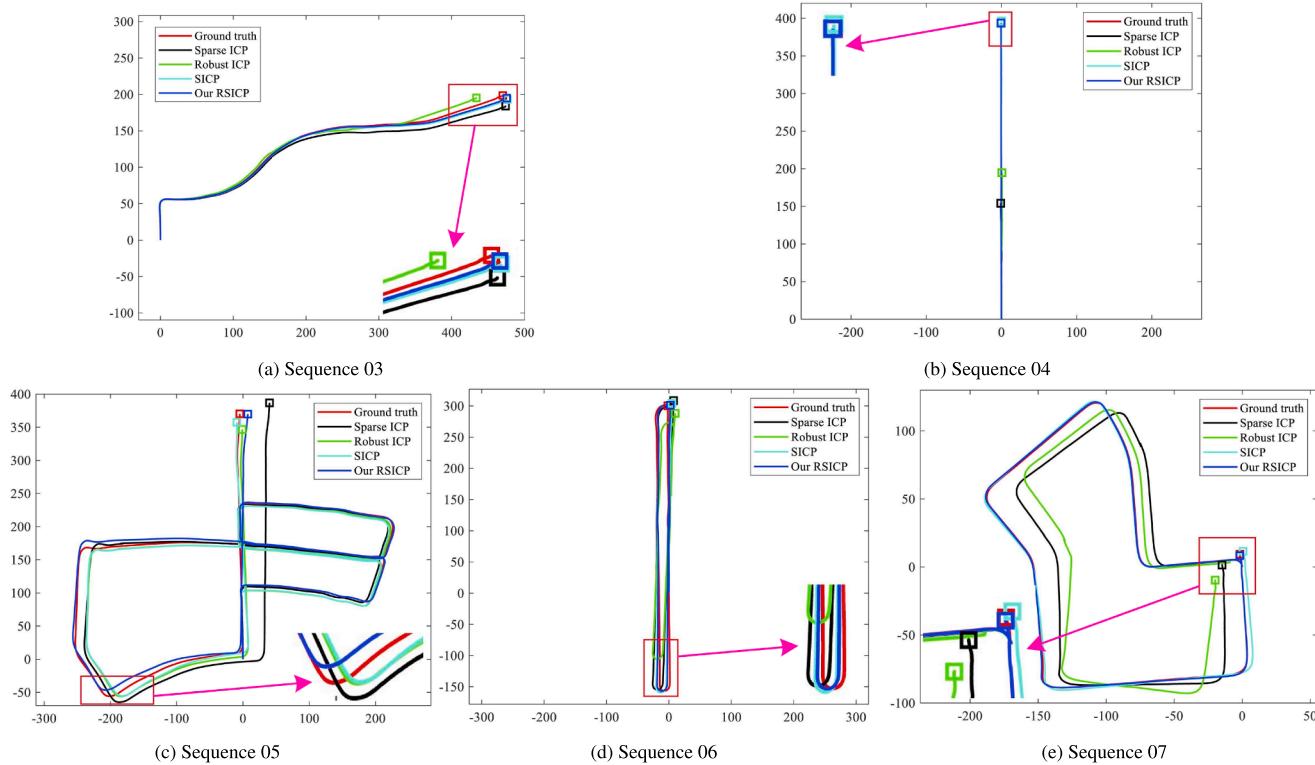


Fig. 7. The estimated trajectories ended with squares of different methods on the KITTI dataset.

Table 3

Quantitative evaluation (metric: RTE (%) \downarrow and total running time (in seconds) \downarrow) on the KITTI Dataset. The best results are highlighted in bold fonts, and the second best results are underlined.

Sequence	03 (801 scans)		04 (271 scans)		05 (2761 scans)		06 (1101 scans)		07 (1101 scans)		Mean	
	RTE	Time	RTE	Time	RTE	Time	RTE	Time	RTE	Time	RTE	Time
ICP	14.93	100.10	14.26	51.98	4.46	208.95	4.08	146.74	5.16	77.44	8.58	117.04
ICP-l	5.75	74.45	5.10	<u>33.23</u>	5.00	171.40	18.90	107.28	4.13	<u>77.34</u>	7.78	92.74
AA-ICP	15.81	116.48	16.15	56.51	4.17	246.91	4.02	178.04	4.92	96.15	9.01	138.82
Sparse ICP	2.64	2.77e4	72.92	1.03e4	<u>3.54</u>	7.24e4	2.03	3.69e4	7.38	2.49e4	17.70	3.44e4
Fast ICP	18.09	<u>84.17</u>	20.62	<u>38.08</u>	4.71	<u>193.46</u>	4.27	<u>120.09</u>	5.76	71.98	10.69	<u>101.56</u>
Robust ICP	9.17	359.41	60.47	149.53	3.79	882.24	9.80	559.41	11.86	326.95	19.02	455.51
Symmetric ICP	<u>1.08</u>	612.65	<u>0.74</u>	207.79	3.63	950.60	<u>0.69</u>	821.37	<u>1.05</u>	737.67	<u>1.44</u>	666.02
Our RSICP	<u>0.96</u>	138.14	<u>0.32</u>	64.77	0.72	398.25	<u>0.59</u>	202.62	<u>0.73</u>	149.94	<u>0.66</u>	190.74

instances, it still gets a larger RTE if it totally fails on some instances. Thus, laser odometry problem can not only reflect the registration accuracy of a method, but also reflect its stability and robustness. Sparse ICP suffers from similar problem with Robust ICP. It gets the worst result on the sequence 04. Our RSICP achieves the best performance. Its RTE error is smaller than 50% of Symmetric ICP. In terms of running time, our RSICP is 2 + times faster than Symmetric ICP, 1 + times faster than Robust ICP, and about 180 times faster than Sparse ICP, although RSICP is implemented in MATLAB code.

5.3. Limitations

Similar to other ICP-type methods, our RSICP also converges to local minima. Although it has a wider convergence basin, it still requires initializations. If the initialization is far from the ground truth, it may perform poorly (See Fig. 3). The second limitation is that our convergence speed becomes much slower if the outlier rate is very high. In Fig. 4, the iteration number of RSICP is much larger than that of Symmetric ICP in the case with 200% added outliers.

6. Conclusion and future work

In this paper, we proposed an accurate, fast, and robust method for registration problem. We first propose an improved version of the symmetric distance metric by considering the rotation of normals, together with a simple but effective linearization based on Rodrigues rotation parameterization. The functional zero-set of our metric is a set of locally-second-order surfaces instead of just planar patches of the point-to-plane metric. Compared with original symmetric metric, point-

to-point, and point-to-plane metrics, our metric has a wider convergence basin and higher convergence speed. We also develop a robustified symmetric point-to-plane ICP (RSICP) formulation based on an adaptive robust loss with a high breakdown point and the Black-Rangarajan duality. This problem is solved using the IRLS method with a robustness-parameter decay strategy, so that our loss acts as different functions with different degrees of robustness. Our RSICP achieves comparable or better registration accuracy than several state-of-the-arts (including Sparse ICP, Robust ICP, and Symmetric ICP), while being much less sensitive to initializations and significantly faster. Thus, it has better stability and is more suitable for large-scale problems, e.g., laser odometry.

To tackle the limitations of our RSICP, further improvements can be achieved in two directions. First, a correspondence-based registration method can be adopted for determining good initializations. Second, soft constraints on correspondence distance and normals can be applied to decrease outlier rate and correspondence selection strategy can be used to improve the efficiency.

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.

Acknowledgments

This work was supported by National Natural Science Foundation of China (No. 42030102 and 41901398).

Appendix A. B&R duality

Black-Rangarajan Duality (Black and Rangarajan, 1996; Yang et al., 2020a): Given a robust cost function $\rho(\cdot)$, define $\phi(z) = \rho(\sqrt{z})$. If $\phi(z)$ satisfies $\lim_{z \rightarrow 0} \phi'(z) = 1, \lim_{z \rightarrow \infty} \phi'(z) = 0$, and $\phi''(z) < 0$, then the robust estimation problem $\min_{\delta} \sum_{i=1}^M \rho(r_i(\delta))$ is equivalent to the following optimization with outlier process:

$$\min_{\delta, w_i \in [0, 1]} \sum_{i=1}^M w_i \cdot (r_i(\delta))^2 + \chi_\rho(w_i), \quad (\text{A.1})$$

where $w_i \in [0, 1]$ ($i = 1, \dots, M$) are slack variables (or weights) associated to each measurement residual r_i, δ are parameters to be optimized, and the function $\chi_\rho(w_i)$ (the so called outlier process) defines a penalty on the weight w_i . The expression of $\chi_\rho(w_i)$ depends on the choice of robust cost function $\rho(\cdot)$.

References

- Ao, S., Hu, Q., Yang, B., Markham, A., Guo, Y., 2021. Spinnet: Learning a general surface descriptor for 3d point cloud registration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11753–11762.
- Barath, D., Matas, J., 2018. Graph-cut ransac. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6733–6741.
- Barath, D., Noskova, J., Ivashechkin, M., Matas, J., 2020. Magsac++, a fast, reliable and accurate robust estimator. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1304–1312.
- 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), vol. 3, IEEE, pp. 2743–2748.
- Black, M.J., Rangarajan, A., 1996. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *Int. J. Comput. Vision* 19, 57–91.
- Bouaziz, S., Tagliasacchi, A., Pauly, M., 2013. Sparse iterative closest point. In: Computer Graphics Forum, vol. 32, Wiley Online Library, pp. 113–123.
- Brachmann, E., Rother, C., 2019. Neural-guided ransac: Learning where to sample model hypotheses. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 4322–4331.
- 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.
- Chen, Y., Medioni, G., 1992. Object modelling by registration of multiple range images. *Image Vis. Comput.* 10, 145–155. [https://doi.org/10.1016/0262-8856\(92\)90066-c](https://doi.org/10.1016/0262-8856(92)90066-c).
- Chetverikov, D., Stepanov, D., Krsek, P., 2005. Robust euclidean alignment of 3d point sets: the trimmed iterative closest point algorithm. *Image Vis. Comput.* 23, 299–309.
- Choy, C., Park, J., Koltun, V., 2019. Fully convolutional geometric features. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 8958–8966.
- Chum, O., Matas, J., 2008. Optimal randomized ransac. *IEEE Trans. Pattern Anal. Mach. Intell.* 30, 1472–1482.
- 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.
- 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.
- Holland, P.W., Welsch, R.E., 1977. Robust regression using iteratively reweighted least-squares. *Commun. Stat.-theory Methods* 6, 813–827.
- Horn, B.K., 1987. Closed-form solution of absolute orientation using unit quaternions. *JOSA A* 4, 629–642.
- Huang, X., Mei, G., Zhang, J., Abbas, R., 2021. A comprehensive survey on point cloud registration. *arXiv preprint arXiv:2103.02690*.

- Jian, B., Vemuri, B.C., 2010. Robust point set registration using gaussian mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 1633–1645.
- Li, J., 2021. A practical o(n2) outlier removal method for point cloud registration. *IEEE Trans. Pattern Anal. Mach. Intell.* <https://doi.org/10.1109/TPAMI.2021.3065021>.
- Li, J., Hu, Q., Ai, M., 2020a. Gesac: Robust graph enhanced sample consensus for point cloud registration. *ISPRS J. Photogramm. Remote Sens.* 167, 363–374.
- Li, J., Hu, Q., Ai, M., 2020b. Rift: Multi-modal image matching based on radiation-variation insensitive feature transform. *IEEE Trans. Image Process.* 29, 3296–3310.
- Li, J., Hu, Q., Ai, M., 2021. Point cloud registration based on one-point ransac and scale-annealing biweight estimation. *IEEE Trans. Geosci. Remote Sens.*
- Li, J., Lee, G.H., 2019. Usip: Unsupervised stable interest point detection from 3d point clouds. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 361–370.
- Li, J., Zhang, Y., Hu, Q., 2021. Robust estimation in robot vision and photogrammetry: A new model and its applications. *ISPRS Ann. Photogramm., Remote Sens. Spatial Informat. Sci.* 1, 137–144.
- Li, J., Zhao, P., Hu, Q., Ai, M., 2020c. Robust point cloud registration based on topological graph and cauchy weighted lq-norm. *ISPRS J. Photogramm. Remote Sens.* 160, 244–259.
- 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.
- Myronenko, A., Song, X., 2010. Point set registration: Coherent point drift. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 2262–2275.
- Pavlov, A.L., Ovchinnikov, G.W., Derbyshev, D.Y., Tsetserukou, D., Oseledets, I.V., 2018. Aa-icp: Iterative closest point with anderson acceleration. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 3407–3412.
- Pottmann, H., Huang, Q.-X., Yang, Y.-L., Hu, S.-M., 2006. Geometry and convergence analysis of algorithms for registration of 3d shapes. *Int. J. Comput. Vision* 67, 277–296.
- 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.
- Rusinkiewicz, S., 2019. A symmetric objective function for icp. *ACM Trans. Graphics (TOG)* 38, 1–7.
- Rusinkiewicz, S., Levoy, M., 2001. Efficient variants of the icp algorithm. In: Proceedings Third International Conference on 3-D Digital Imaging and Modeling. IEEE, pp. 145–152.
- Rusu, R.B., Blodow, N., Beetz, M., 2009. Fast point feature histograms (fpfh) for 3d registration. In: IEEE International Conference on Robotics and Automation, pp. 3212–3217.
- Rusu, R.B., Cousins, S., 2011. 3d is here: Point cloud library (pcl). In: 2011 IEEE International Conference on Robotics and Automation. IEEE, pp. 1–4.
- Schnabel, R., Klein, R., 2006. Octree-based point-cloud compression. *Spgb* 6, 111–120.
- Segal, A., Haehnel, D., Thrun, S., 2009. Generalized-icp. In: Robotics: Science and Systems, vol. 2, Seattle, WA, p. 435.
- 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.
- 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.
- Tam, G.K., Cheng, Z.-Q., Lai, Y.-K., Langbein, F.C., Liu, Y., Marshall, D., Martin, R.R., Sun, X.-F., Rosin, P.L., 2013. Registration of 3d point clouds and meshes: A survey from rigid to nonrigid. *IEEE Trans. Visual Comput. Graphics* 19, 1199–1217. <https://doi.org/10.1109/TVCG.2012.310>.
- Tordoff, B.J., Murray, D.W., 2005. Guided-mlesac: Faster image transform estimation by using matching priors. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 1523–1535.
- Wang, Y., Solomon, J.M., 2019. Deep closest point: Learning representations for point cloud registration. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3523–3532.
- Yang, H., Antonante, P., Tzoumas, V., Carbone, L., 2020a. Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection. *IEEE Robot. Automat. Lett.* 5, 1127–1134.
- Yang, H., Carbone, L., 2019. A polynomial-time solution for robust registration with extreme outlier rates. In: Robotics: Science and Systems.
- Yang, H., Shi, J., Carbone, L., 2020b. Teaser: Fast and certifiable point cloud registration. *IEEE Trans. Rob.* 37, 314–333.
- 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: IEEE Conference on Computer Vision and Pattern Recognition, 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: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1802–1811.
- Zhang, J., Singh, S., 2014. Loam: Lidar odometry and mapping in real-time. In: Robotics: Science and Systems, vol. 2.
- Zhang, J., Yao, Y., Deng, B., 2021. Fast and robust iterative closest point. *IEEE Trans. Pattern Anal. Mach. Intell.* <https://doi.org/10.1109/TPAMI.2021.3054619>.
- Zhang, Z., 1994. Iterative point matching for registration of free-form curves and surfaces. *Int. J. Comput. Vision* 13, 119–152.
- Zhong, Y., 2009. Intrinsic shape signatures: A shape descriptor for 3d object recognition. In: IEEE International Conference on Computer Vision Workshops, pp. 689–696.
- Zhou, Q.-Y., Park, J., Koltun, V., 2016. Fast global registration. In: European Conference on Computer Vision. Springer, pp. 766–782.