

RIDF: A ROBUST ROTATION-INVARIANT DESCRIPTOR FOR 3D POINT CLOUD REGISTRATION IN THE FREQUENCY DOMAIN

R. Huang¹, W. Yao^{2,*}, Z. Ye^{1,3}, Y. Xu¹, U. Stilla¹

¹ Photogrammetry and Remote Sensing, Technical University of Munich (TUM), Munich, Germany
- (rong.huang, yusheng.xu, stilla@tum.de

² Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
- wei.hn.yao@polyu.edu.hk

³ College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China
- 89.yezhen@tongji.edu.cn

KEY WORDS: 3D descriptor, Rotation-invariance, Fourier analysis, Point cloud registration

ABSTRACT:

Registration of point clouds is a fundamental problem in the community of photogrammetry and 3D computer vision. Generally, point cloud registration consists of two steps: the search of correspondences and the estimation of transformation parameters. However, to find correspondences from point clouds, generating robust and discriminative features is of necessity. In this paper, we address the problem of extracting robust rotation-invariant features for fast coarse registration of point clouds under the assumption that the pairwise point clouds are transformed with rigid transformation. With a Fourier-based descriptor, point clouds represented by volumetric images can be mapped from the image to feature space. It is achieved by considering a gradient histogram as a continuous angular signal which can be well represented by the spherical harmonics. The rotation-invariance is established based on the Fourier-based analysis, in which high-frequency signals can be filtered out. This makes the extracted features robust to noises and outliers. Then, with the extracted features, pairwise correspondence can be found by the fast search. Finally, the transformation parameters can be estimated by fitting the rigid transformation model using the corresponding points and RANSAC algorithm. Experiments are conducted to prove the effectiveness of our proposed method in the task of point cloud registration. Regarding the experimental results of the point cloud registration using two TLS benchmark point cloud datasets, featuring with limited overlaps and uneven point densities and covering different urban scenes, our proposed method can achieve a fast coarse registration with rotation errors of less than 1 degree and translation errors of less than 1m.

1. INTRODUCTION

Registration of point clouds is a fundamental problem in the community of photogrammetry and 3D computer vision. To some extent, the result of point cloud registration serve as prerequisite for many applications, such as, construction monitoring (Bosché et al., 2015, Tuttas et al., 2017), forestry monitoring (Polewski et al., 2019), urban planning (Vosselman, Maas, 2010), and 3D modeling (Lafarge, Mallet, 2012, Yang et al., 2013). Briefly, the objective of point cloud registration is to unite the data (i.e., 3D points) acquired in various views, platforms, or times into a common coordinate system.

Generally, the registration methods mainly consist of two essential categories: coarse registration and fine registration. For fine registration, iterative closest point (ICP) (Besl, McKay, 1992, Chen, Medioni, 1991, Zhang, 1994) and its variants (Habib et al., 2010, Yang et al., 2015) are the representative algorithms and have been proved to be effective and efficient. Additionally, normal distribution transform (NDT) (Biber, Straßer, 2003) is also a commonly used standard algorithm for fine registration. However, in the step of fine registration, proper initial transformation values are of great importance for fine registration to avoid local optimum. In this paper, we address the problem of how to provide good initial transformation parameters to find registration, namely the problem of coarse registration. To achieve coarse registration of point clouds, at least two steps are required: finding the correspondences and estimating the

transformation parameters. Among the two steps, finding the correspondences between point clouds is important and is the basis for estimating transformation parameters.

For most general registration tasks, one commonly used strategy of finding correspondences is to use a feature-based strategy, which can be divided into three steps:

- Detecting of key geometric points or primitives;
- Constructing features for the detected key points or primitives;
- Determining the correspondences using the extracted features.

However, the estimation of correspondences of unordered and unstructured point clouds is still challenging due to the following reasons:

- Uneven point densities caused the varying observation distances of the scanners;
- Noise and outliers resulting from temporary or moving objects;
- Low-overlapping rate between point clouds;
- High-similarity and symmetry, especially for point clouds acquired in urban areas;

* Corresponding author

- Large data amount of point clouds.

To tackle the aforementioned problems, we present a robust and fast registration workflow. The workflow consists of four steps: representing unordered 3D point clouds by 3D volumetric images, constructing rotation-invariant features for each voxel in volumetric images, searching for corresponding voxels in feature domain, and estimating the transformation parameters with a robust estimator. Differing from the commonly used feature-based strategy, as mentioned before, in the proposed workflow, we skip the step of detecting key geometric points or primitives to eliminate the errors or inaccuracies caused in the step of detecting edges, planes or other primitives. For most of the point-wise or voxel-wise rotation-invariant descriptors, the rotation invariance is achieved by aligning the local reference frame (LRF) by estimating the dominant gradient direction. However, the descriptor used in this paper for constructing rotation-invariance can map arbitrary voxels in volumetric images to a feature field, which is also the reason why detecting key points or primitives can be skipped in the proposed workflow. Besides, the rotation-invariant features are robust to noise without the step of estimating the dominant gradient direction. With these factors taken into consideration, the main concept is to represent point clouds with robust and rotation-invariant features in the frequency domain, so that a robust alignment between point clouds can be achieved.

2. RELATED WORK

In this paper, we mainly address the problem of coarse registration. Generally, many kinds of research have reported different solutions for coarse point cloud registration. The registration methods can be primarily categorized into two classes: Feature description-based registration and geometric constraint-based registration.

2.1 Feature description-based registration

For the feature description-based registration methods, as mentioned before, feature descriptor plays an important role in feature matching. Corresponding points are identified via the similarity measurement between features. In the literature, many different descriptors have been reported to be useful in the feature matching, such as scale-invariant feature transform (SIFT) (Flitton et al., 2010), fast point feature histogram (FPFH) (Rusu et al., 2009), rotational projection statistics (RoPS) (Guo et al., 2013) and signature of histogram of orientations (SHOT) (Tombari et al., 2010) and so on. Generally, two characteristics of a sufficient descriptor have to be addressed, which are high descriptiveness and rotation-invariance. However, descriptors, such as SIFT, highly depend on intersecting point detectors. Thus, it is impossible for dense feature extraction using this kind of descriptor. Additionally, the strategy used for achieving rotation in-variance is based on pose normalization. For example, SIFT produces rotation-invariant features by aligning LRF to the dominant gradient orientation. However, most of the LRF-based feature descriptors are not robust to surface noise and outliers. The other way is to obtain the statistics of the local geometry. It is of easy conduction and fast computation. But the only problem for this strategy is that the features may suffer from low descriptiveness.

Apart from the point-based feature matching methods, many other researches utilized geometric primitives as features instead of points, such as lines (Habib et al., 2005, Polewski, Yao,

2019), curves (Yang, Zang, 2014), planes (Xiao et al., 2013) and plane pairs (Chen et al., 2019). On the other hand, the combination of different kinds of primitives is also an effective direction for the task of registration and simultaneously improve the robustness of these geometric features (Xu et al., 2017). No matter the process of detecting keypoints or forming geometric primitives, the accuracy of the registration highly relies on the accuracy of the estimation of these geometric representations. Sometimes, outliers will be induced in the process of finding edges, lines, or planes.

2.2 Geometric constraint-based registration

In addition to the feature description-based method, some methods follow a different registration scheme. In this type of method, the registration is not achieved by generating features and find feature correspondences; instead, they will identify corresponding points by predefined geometric constraints. Here, 4-points congruent set (4PCS) (Aiger et al., 2008), as well as its variants, such as Super4PCS (Mellado et al., n.d.), keypoint-based 4PCS (K4PCS) (Theiler et al., 2014), and semantic keypoint-based 4PCS (SK4PCS) (Ge, 2017) are representative registration methods using this strategy, in which pairwise congruent point sets are determined by exploiting the rule of intersection ratios. For point clouds, intersection ratios of congruent points are invariant under the affine transformation, and they are used to determine whether two point sets are matched or not. Compared with feature description-based registration, these algorithms are more robust to noise, occlusions, and uneven densities. Additionally, in this method, the candidates for the transformation parameters can be largely reduced, which releases the pressure on computation. Meanwhile, some other strategies are based on the 4PCS strategy but conducted using different primitives. Representative extensions of this strategy are the ones using planes replacing points. Voxel-based 4-plane congruent sets (V4PCS) (Xu et al., 2019) and 4-planes congruent sets (4-PICS) (Bueno et al., 2017) are symbolic of the implementation of structured planes as congruent sets with angle ratio constraints.

The assembly or extracting of planes from points will powerfully overcome the uneven dot density and outlines that ensure consistent alignment of planes of specific angles. The distance measured of points used by conventional 4PCS approaches, as a comparison, is more sensitive to outliers and noises. Another interesting example is the volumetric 4PCS (V-4PCS) (Huang et al., 2017), which extends the four coplanar points to volumetric non-coplanar data and theoretically shows a significant reduction in computational complexity.

2.3 Our contributions

Although the literature mentioned above is not comprehensive, a few pieces of research have addressed the characteristics of 3D point clouds in the frequency domain. In 2D image matching, methods based on Fourier transforms are commonly utilized, such as phase correlation (Nagashima et al., 2006). In (Huang et al., 2019), the strategy of Fourier-based 2D image matching was successfully applied to 3D point cloud registration. In (Liu et al., 2014), a 3D rotation-invariant descriptor was proposed and tested in the task of shape retrieval. In our work, we embed the concept of achieving rotation-invariance in the frequency domain in the framework of point cloud registration. The major contributions of this paper are abstracted as follows:

- We proposed an efficient framework for fast coarse registration of point clouds, which is sufficient in the low-overlapping dataset. In the proposed workflow, point clouds are represented using volumetric images and mapped to feature space using a rotation-invariant descriptor, spherical harmonic HoG (SHHOG), in the frequency domain.
- The descriptor has two main benefits. The descriptor is not restricted by geometric shape extraction. On the other hand, the descriptor is robust to noise and outliers since high-frequency signals are filtered out when representing the corresponding signal with spherical harmonics.

3. METHODOLOGY

The proposed methodology for point cloud registration comprises four essential steps: voxelization of point clouds, local rotation-invariant feature extraction from two volumetric images, finding pairwise correspondences, and estimation of transformation parameters, as illustrated in Fig. 1. In the first step, point clouds are individually transformed into 3D volumetric images, in which each voxel in the 3D grids represents the number of points located within the corresponding spatial boundaries. Second, features for each voxel are constructed using Fourier-based HoG features in a continuous way. Subsequently, the corresponding voxels are matched with a fast search method using the features extracted in the previous step. Finally, the transformation parameters are calculated using the corresponding voxels and a RANSAC-based strategy for robust estimation. The detailed explanation of each step in the workflow is introduced in the following sub-sections.

3.1 Local rotation-invariant feature extraction in the frequency domain

The local rotation-invariant feature extraction consists of three steps: voxel-wise HoG feature extraction with spherical harmonic representations, computation of regional features, and the final feature extraction. In Fig. 2, the key parts in feature extraction are illustrated.

3.1.1 Voxel-wise feature extraction The first step is to create a distribution for each voxel. Let a 3D location of any voxel in a volumetric image \mathbf{C} be (x, y, z) in Cartesian coordinates, and the corresponding 3D gradient to be $\mathbf{d}(x, y, z)$ denoted as the gradient in each direction. Thus, the SH coefficients will be:

$$\mathbf{F}^l(x, y, z) = \frac{2l+1}{4\pi} \|\mathbf{d}(x, y, z)\| \mathbf{Y}^l(\mathbf{d}(x, y, z)) \quad (1)$$

where the Schmidt semi-normalized SHs are defined as:

$$\mathbf{Y}_m^l = \sqrt{\frac{(l-m)!}{(l+m)!}} \mathbf{P}_m^l(\cos(\theta)) e^{im\phi}, \quad (2)$$

where \mathbf{P}_m^l are Legendre polynomials, which can be calculated by given functions in Matlab. With the issue of noise and outliers considered, only low-frequency degrees in the SH representation are computed, which has the same effect as low-pass filtering.

Then, with spatial aggregation and normalization, the SHHOG field can be expressed as:

$$\tilde{\mathbf{F}}_m^l = \mathbf{F}_m^l / \sqrt{\|\mathbf{d}^2\| * K} \quad (3)$$

Assuming that the 3D signal is rotated with l -th degree, the feature behavior under the rotation can be derived as:

$$g\tilde{\mathbf{F}}^l = [\|\mathbf{d}\| \mathcal{D}_g^l \mathbf{Y}^l(\mathbf{d})] \circ \mathbf{T}^g = \mathcal{D}_g^l [\tilde{\mathbf{F}}^l \circ \mathbf{T}_g] \quad (4)$$

Thus, the spherical tensor operations are needed to create rotation-invariant regional descriptors from the SH HOG field.

3.1.2 Regional feature extraction To obtain rotation-invariant features and improve the descriptiveness of the features, multiple concentric voxels around a selected voxel are used to describe the surroundings of the corresponding voxel.

Before the identification of rotation-invariance of features, it should be noted that the inner-product of two spherical tensors are rotation-invariant, which can be proved by

$$\begin{aligned} g(\mathbf{F}_1^\dagger \mathbf{F}_2) &= g(\mathbf{F}_1)^\dagger g(\mathbf{F}_2) = (\mathcal{D}_g^l \mathbf{F}_1)^\dagger (\mathcal{D}_g^l \mathbf{F}_2) \\ &= \mathbf{F}_1 \mathcal{D}_g^{l\dagger} \mathcal{D}_g^l \mathbf{F}_2 = \mathbf{F}_1^\dagger \mathbf{F}_2 \end{aligned} \quad (5)$$

Thus, we can produce rotation-invariant features by coupling the filtering output of the same rank as the inner product. Finally, the descriptor can be calculated by

$$\mathbf{f}_i = \int \text{tensor_product}(\tilde{\mathbf{F}}^l(\mathbf{x} - \mathbf{x}'), \mathbf{G}_n(\mathbf{x}'), k) d\mathbf{x}' \quad (6)$$

where \mathbf{G}_n denotes the spherical Gaussian derivatives.

3.1.3 Procedures of local rotation-invariant feature extraction In summary, the procedures for extracting rotation-invariant features are given below:

- By applying Fourier transformation in spherical coordinates on each voxel in the volumetric image, the magnitude features can be obtained given the corresponding Fourier order numbers, which can be denoted as $\mathbf{f}_m^1 = \|\mathbf{D}_m\|, m = 0, 1, 2, \dots, m$. This part is called the magnitude part.
- Using Eq. 5, rotation behavior can be compensated. Thus, the second part of features is absolute rotation-invariant features, which can be presented as $\mathbf{F}_m \mathbf{F}_m$.
- In the former two parts, the phase information is completely removed. Aiming at reducing the loss of phase information, we derive the relative rotation-invariant features by the convolution of two spherical tensor fields using the tensor product using Eq. 6.

Combining the aforementioned parts, we can finally derive the rotation-invariant features. In general, there are three main advantages of this Fourier-based rotation-invariant features. First, continuous dense feature computation can be achieved using this descriptor. Second, a low-frequency filtering is conducted with the spherical harmonic representation, which makes the descriptor robust to noise and outliers. Third, the final features contain not only the information of the corresponding voxels but also the surroundings, which makes this descriptor of high descriptiveness.

3.2 Generating pairwise correspondences

The identification of corresponding voxels can be conducted in a fast way using the estimated features. The correspondences

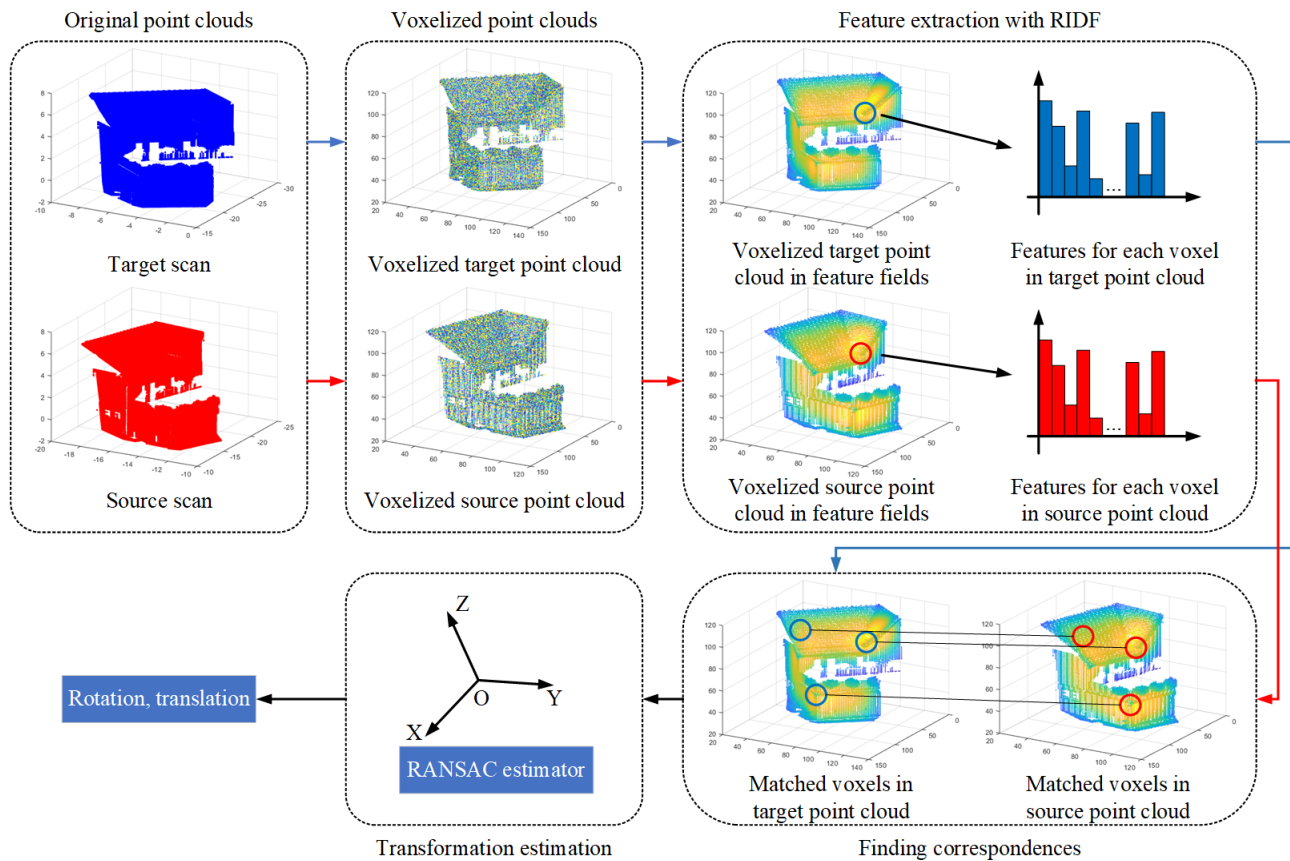


Figure 1. Workflow of the proposed coarse registration strategy.

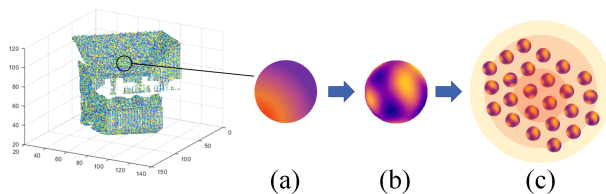


Figure 2. RIDF descriptor. (a) An individual gradient vector is represented by a distribution in sphere. (b) Signals on sphere are reconstructed from spherical harmonics. (c) Regional convolution obtained by spherical tensor operations considering the corresponding surroundings.

between features are estimated by the criteria set with Euclidean distances in the feature domain. The selection of the threshold for feature matching varies for different datasets and is set empirically. It is set by checking the distribution of the distances between each voxel and its corresponding nearest voxel and also the number of voxel-pairs selected to the matched voxel list. The specific workflow for finding correspondences are provided in Algorithm 1.

3.3 Estimation of transformation parameters

After finding the corresponding voxel pairs, the transformation parameters of the rigid transformation, including 3D rotations and translation, between point clouds can be estimated. For achieving a robust estimation for the parameters, a RANSAC-based strategy is applied. In each round of RANSAC, a set of transformation parameters can be obtained. Then the estimated transformation parameters are utilized to transform all

Algorithm 1: Finding feature correspondences: Given feature representations of volumetric images of two scans, find the matched voxels from the two scans using the features.

Require: Volumetric images I_1 and I_2 and corresponding feature fields F_1 and F_2

Ensure: Matched voxel pairs denoted by their corresponding positions

- 1: For each voxel in I_1 , search for the nearest points in feature domain using the Tree(F_2) created with the corresponding features
- 2: Judge whether the distance between F_1 and the matched F_2 are smaller than the threshold σ
- 3: If not, jump to the next voxel
- 4: If yes, add the voxel pairs to the matched voxel list

the matched voxels in the target scan. Here, the thresholds for the RANSAC trials are set to be two-fold. One should be the maximum number of trials, and the other one is the minimum threshold of the registration score. For any trials, if the registration score is lower than the threshold, the transformation parameters will be the final estimated transformation parameters, and the RANSAC process will be stopped. The other case is that the RANSAC process will be stopped when the number of trial reaches the maximum amount. In this case, the transformation parameters in the last trial will be saved.

4. EXPERIMENTS AND ANALYSIS

4.1 Experimental data

The proposed method for performing the registration is tested using two TLS point cloud datasets, which are acquired with

different scanners and from different urban scenes, as illustrated in Fig. 3. The first one is a point cloud pair from the ThermalMapper project acquired by the Jacobs University Bremen (Borrmann et al., 2013). The second dataset is one pair of point clouds chosen from the Real-world Scans with Small Overlap (RESSO) dataset (Chen et al., 2019). Compared with the first dataset, the overlapping rate is relatively lower. For the first dataset, the source and target point clouds are manually aligned as ground truth. For the second one, it is provided as a benchmark dataset, in which the accurate transformation parameters between the source and target point clouds are provided. All the experiments were conducted on a 64 Bit Windows 10 PC with 8 GB RAM and an Intel (R) Core (TM) i7-6700 @ 3.4 GHz CPU.

Table 1. Information of experimental datasets.

Dataset	Number of points and overlaps		
	Source	Target	Overlap ratio
Bremen	15033238	15246878	0.85
Resso Pair2-3	825915	618672	0.60

4.2 Evaluation metric

For evaluation, we use the following evaluation metrics: the rotation errors and the translation errors, which are computed using the references provided by manually registration or ground truth. Providing the reference transformation matrix \mathbf{T}_{gt} and the estimated transformation matrix \mathbf{T}_e , the rotation error e^r and translation error e^t can be computed by:

$$\Delta \mathbf{T} = \mathbf{T}_{gt}(\mathbf{T}_e)^{-1} = \begin{bmatrix} \Delta \mathbf{R} & \Delta \mathbf{t} \\ 0 & 1 \end{bmatrix} \quad (7)$$

$$e^r = \arccos\left(\frac{\text{tr}(\Delta \mathbf{R}) - 1}{2}\right) \quad (8)$$

$$e^t = \|\Delta \mathbf{t}\| \quad (9)$$

wherein $\text{tr}(\cdot)$ denotes the trace.

4.3 Experimental results

The coarse registration results using two datasets are illustrated in Fig. 4. In the experiments, the voxel size used for the Bremen dataset is set to 2m, and the voxel size for the Resso dataset is set as 1m. The thresholds for the identification of feature correspondence are both set to be 0.5, and the maximum iteration number of RANSAC is set as 1 million times. In Table 2, the number of matched voxels are provided. From the visualization aspect, it is shown that the point clouds are well aligned. Meanwhile, the quantitative results for the registration are given in Table 3 and 4. For the two scenes of Bremen and Resso, the rotation errors are less than 1 degree. As for translation errors, they are both less than 1m. The results reveal that the proposed method can achieve successful registration among different point cloud datasets, even for the datasets which have a comparatively low overlapping rate. On the other hand, based on the coarse registration results, the registration can be further improved by applying fine registration.

To further validate the effectiveness, the experimental results are compared with three benchmark methods using the Bremen dataset. The baseline methods we used for comparison are FPFHSAC (Holz et al., 2015), K4PCS (Theiler et al., 2014), and

Table 2. Matched voxels in the experimental datasets.

Dataset	Number of matched voxels		
	Source	Target	Matched
Bremen	12712	13197	142
Resso Pair2-3	13428	19424	382

V4PCS (Xu et al., 2019). The FPFHSAC method registered point clouds using FPFH features and a RANSAC process for the robust estimation of transformation parameters. The K4PCS method is under the framework of 4PCS. The difference is that this method using keypoints instead of random points, which largely reduces the number of candidates and improves the efficiency and robustness of the algorithm. V4PCS provided a different solution for utilizing the 4PCS framework. The 4PCS are replaced by voxel-based 4-planes congruent sets. In this method, more geometric constraints are added, which improved the robustness and has been proved to be a fast and effective way for point cloud registration. The registration results for these methods are listed in Table 3. It is shown that all com-

Table 3. Results of four registration methods using the Bremen dataset.

Methods	Rotation error (deg)	Translation error (m)	Time (s)
FPFHSAC	0.3601	0.0692	318
K4PCS	0.3682	0.4826	256
V4PCS	0.1916	0.6312	78
Ours	0.9218	0.6054	104

pared methods provide satisfying performance. Comparatively, our proposed method offers a weaker performance considering the rotation error and translation error. The main reason lies in the two factors which highly influence the registration accuracy. The first one is the geometric resolution of each voxel. In our experiments, considering the limitation of computation memory and the covered areas of the point clouds, the voxel size can not be set to a small value, which largely limits the representation of details of the scene. In future work, the representation of voxels and the processing way for extracting the features can be changed to reduce the required memory. On the other hand, the criteria for estimating correspondences is also an important factor. The smaller the threshold is set, the fewer correspondences are extracted. The time spent on the RANSAC process will be largely reduced. However, a strict setting for the threshold will also decrease the flexibility of the registration framework in considering noise, outliers, or even artifacts caused by our former voxelization processing. As for the running time, although the time performance of our proposed method is slightly weaker than V4PCS, it is obviously better than the other two feature-based methods. However, it is not fair to draw a conclusion for the time performance since our proposed method was conducted on Matlab while others were conducted on C++. Additionally, as stated before, the voxel size and the criteria for estimating correspondence are two important factors influencing time performance. The time spent on the RANSAC process can be largely reduced with larger voxel size and smaller threshold.

Table 4. Results of two registration methods using the Resso dataset.

Methods	Rotation error (deg)	Translation error (m)
PLADE	0.0803	0.0854
Ours	0.4130	0.4619

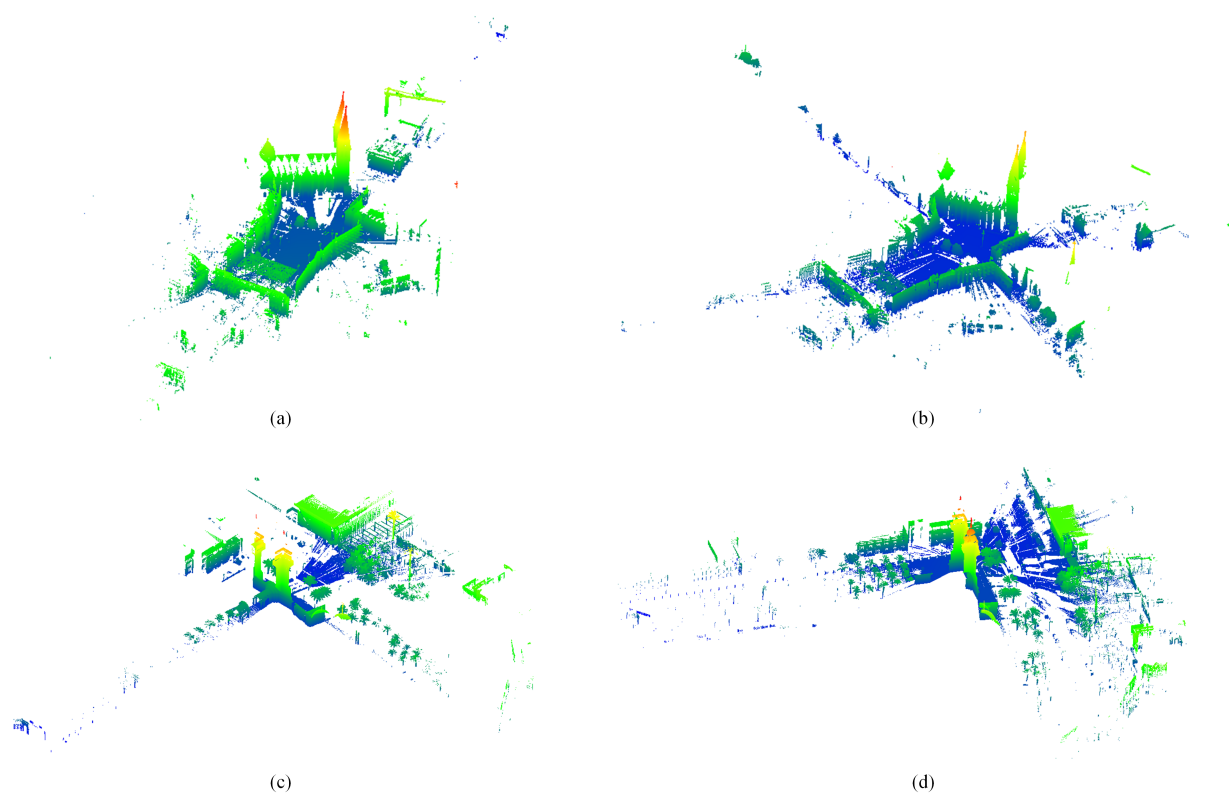


Figure 3. The experimental datasets. (a) and (b) are the source and target point clouds in the Bremen dataset. (c) and (d) are the source and target point clouds in the Resso dataset. (color coded with intensities).

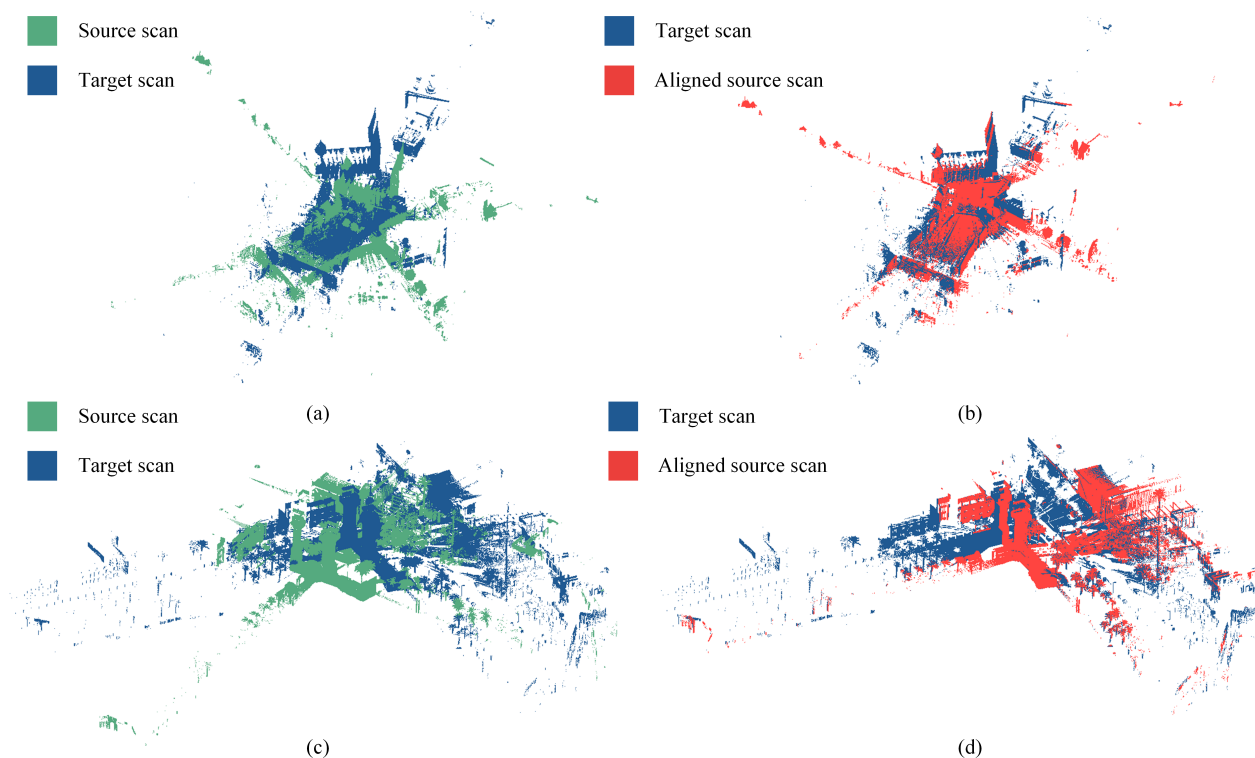


Figure 4. Coarse registration results. (a) and (b) are the source and target point clouds before registration. (c) and (d) are the coarse registration results.

Additionally, we also compare our registration results with the baseline coarse registration method, PLADE, provided by the data provider of Resso. The PLADE method offers a solution for the small-overlapping dataset, which utilizes strong constraints between different planes as descriptors for point clouds. As shown in Table 4 and Fig. 4, our proposed method can also achieve acceptable results for the dataset with a small overlap. Although our result is not as satisfying as the result provided by PLADE, our result is at an acceptable level of coarse registration. It proves that our method is also adaptive to low-overlapping cases.

5. SUMMARY

In this work, we present a marker-free registration method which combines a robust rotation-invariant descriptor with a RANSAC estimator. Specifically, the Fourier-based descriptor enables us to describe each point cloud with rotation-invariant features in a continuous dense way. The main benefit of this descriptor can be concluded as high-descriptiveness, rotation-invariance, robustness, and independent on shape feature extraction. The RANSAC-based strategy provides the possibilities for excluding some outliers in the process of finding correspondences. The experimental results using TLS benchmark datasets have proved the effectiveness of our proposed method.

Meanwhile, the proposed method also provides satisfying results for datasets with small overlaps. However, there are also some limitations of the proposed methods. The extraction of features using the rotation-invariant descriptor relies on the representation of point clouds in a volumetric way, which has a high requirement for the computation memory. In the future, we will try to embed a more efficient and less computation-expensive data structure to this framework. On the other hand, since this descriptor is less dependent on geometric feature extraction, such as edges, lines, or planes, it may be sufficient to utilize this descriptor for the registration in irregular-shaped areas (i.e., forested areas, non-urban areas).

REFERENCES

- Aiger, D., Mitra, N. J., Cohen-Or, D., 2008. 4-points congruent sets for robust pairwise surface registration. *ACM Transactions on Graphics (TOG)*, 27(3), 85.
- Besl, P. J., McKay, N. D., 1992. Method for registration of 3-D shapes. P. S. Schenker (ed.), *Sensor Fusion IV: Control Paradigms and Data Structures*, 1611, International Society for Optics and Photonics, SPIE, 586 – 606.
- Biber, P., Straßer, W., 2003. The normal distributions transform: A new approach to laser scan matching. *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3, IEEE, 2743–2748.
- Borrmann, D., Elseberg, J., Nüchter, A., 2013. Thermal 3d mapping of building façades. *Intelligent Autonomous Systems 12*, Springer, 173–182.
- Bosché, F., Ahmed, M., Turkan, Y., Haas, C., Haas, R., 2015. The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction*, 49, 201–213.
- Bueno, M., González-Jorge, H., Martínez-Sánchez, J., Lorenzo, H., 2017. Automatic point cloud coarse registration using geometric keypoint descriptors for indoor scenes. *Automation in Construction*, 81, 134–148.
- Chen, S., Nan, L., Xia, R., Zhao, J., Wonka, P., 2019. PLADE: A Plane-Based Descriptor for Point Cloud Registration With Small Overlap. *IEEE Transactions on Geoscience and Remote Sensing*.
- Chen, Y., Medioni, G., 1991. Object modeling by registration of multiple range images. *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, 3, 2724–2729.
- Flitton, G., Breckon, T., Megherbi Bouallagu, N., 2010. Object recognition using 3d sift in complex ct volumes. *Proceedings of the British Machine Vision Conference*, BMVA Press, 11.1–11.12.
- Ge, X., 2017. Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 344–357.
- Guo, Y., Sohel, F., Bennamoun, M., Lu, M., Wan, J., 2013. Rotational projection statistics for 3D local surface description and object recognition. *International Journal of Computer Vision*, 105(1), 63–86.
- Habib, A., Datchev, I., Bang, K., 2010. A comparative analysis of two approaches for multiple-surface registration of irregular point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(1).
- Habib, A., Ghanma, M., Morgan, M., Al-Ruzouq, R., 2005. Photogrammetric and LiDAR data registration using linear features. *Photogrammetric Engineering & Remote Sensing*, 71(6), 699–707.
- Holz, D., Ichim, A. E., Tombari, F., Rusu, R. B., Behnke, S., 2015. Registration with the point cloud library: A modular framework for aligning in 3-D. *IEEE Robotics & Automation Magazine*, 22(4), 110–124.
- Huang, J., Kwok, T.-H., Zhou, C., 2017. V4PCS: Volumetric 4PCS algorithm for global registration. *Journal of Mechanical Design*, 139(11).
- Huang, R., Ye, Z., Boerner, R., Yao, W., Xu, Y., Stilla, U., 2019. Fast pairwise coarse registration between point clouds of construction sites using 2D projection based phase correlation. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W13, 1015–1020.
- Lafarge, F., Mallet, C., 2012. Creating large-scale city models from 3D-point clouds: a robust approach with hybrid representation. *International Journal of Computer Vision*, 99(1), 69–85.
- Liu, K., Skibbe, H., Schmidt, T., Blein, T., Palme, K., Brox, T., Ronneberger, O., 2014. Rotation-invariant HOG descriptors using Fourier analysis in polar and spherical coordinates. *International Journal of Computer Vision*, 106(3), 342–364.
- Mellado, N., Aiger, D., Mitra, N. J., n.d. Super 4PCS Fast Global Pointcloud Registration via Smart Indexing. *Computer Graphics Forum*, 33(5), 205–215.

Nagashima, S., Aoki, T., Higuchi, T., Kobayashi, K., 2006. A subpixel image matching technique using phase-only correlation. *Intelligent Signal Processing and Communications, 2006. ISPACS'06. International Symposium on*, IEEE, 701–704.

Polewski, P., Yao, W., 2019. Scale invariant line-based co-registration of multimodal aerial data using L1 minimization of spatial and angular deviations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 79–93.

Polewski, P., Yao, W., Cao, L., Gao, S., 2019. Marker-free coregistration of UAV and backpack LiDAR point clouds in forested areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 307 - 318.

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

Theiler, P. W., Wegner, J. D., Schindler, K., 2014. Keypoint-based 4-Points Congruent Sets—Automated marker-less registration of laser scans. *ISPRS Journal of Photogrammetry and Remote Sensing*, 96, 149–163.

Tombari, F., Salti, S., Di Stefano, L., 2010. Unique signatures of histograms for local surface description. *European Conference on Computer Vision*, Springer, 356–369.

Tuttas, S., Braun, A., Borrmann, A., Stilla, U., 2017. Acquisition and consecutive registration of photogrammetric point clouds for construction progress monitoring using a 4D BIM. *PFG—Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 85(1), 3–15.

Vosselman, G., Maas, H.-G., 2010. *Airborne and terrestrial laser scanning*. Whittles Publishing.

Xiao, J., Adler, B., Zhang, J., Zhang, H., 2013. Planar segment based three-dimensional point cloud registration in outdoor environments. *Journal of Field Robotics*, 30(4), 552–582.

Xu, Y., Boerner, R., Yao, W., Hoegner, L., Stilla, U., 2017. Automated coarse registration of point clouds in 3D urban scenes using voxel based plane constraint. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 4.

Xu, Y., Boerner, R., Yao, W., Hoegner, L., Stilla, U., 2019. Pairwise coarse registration of point clouds in urban scenes using voxel-based 4-planes congruent sets. *ISPRS Journal of Photogrammetry and Remote Sensing*, 151, 106–123.

Yang, B., Xu, W., Dong, Z., 2013. Automated extraction of building outlines from airborne laser scanning point clouds. *IEEE Geoscience and Remote Sensing Letters*, 10(6), 1399–1403.

Yang, B., Zang, Y., 2014. Automated registration of dense terrestrial laser-scanning point clouds using curves. *ISPRS Journal of Photogrammetry and Remote Sensing*, 95, 109–121.

Yang, J., Li, H., Campbell, D., Jia, Y., 2015. Go-ICP: A globally optimal solution to 3D ICP point-set registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(11), 2241–2254.

Zhang, Z., 1994. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2), 119–152.