# Point cloud registration based on global compatibility feature

*S.X. Liu [1,2], G.J. Ji [1],C.C. Shi [1]*
*[1] School of Electrical Engineering, Shanghai Dianji University, 201306, Shanghai, China, Shuihua Road 300;*
*[2] The Key Laboratory of Cognitive Computing and Intelligent Information Processing of Fujian Education Institutions, Wuyi University, 354300, Fujian, China, Wuyi Avenue 16*

***Abstract***

In this paper, we present a point cloud registration method that utilizes a global point cloud compatibility feature. We introduce an evaluation technique called global compatibility, which helps distinguish between correct and incorrect feature point pairs by calculating the corresponding compatibility weights. To begin, we employ a spectral matching technique to select reliable seed points, allowing us to construct a consistent point set in the vicinity of these seed points. We then design a consistent filter to eliminate outliers from the obtained set. Our approach includes proposing optimal weight matching based on the characteristics of each compatible point set, alongside spectral matching for decomposing the constructed multiple compatible point sets. We assign smaller weights for points affected by larger noise, which aids in generating the corresponding rigid transformation. Ultimately, we select the best transformation as the final result. Notably, our method does not require retrieving all features from the entire point set, and it effectively removes discrete points, thereby constructing a more efficient and robust consistent point set. Experimental results demonstrate that our method performs very well on both indoor and outdoor datasets, as well as on datasets with low overlap.

*Keywords*: steric compatibility, point cloud registration, consistent point set, rigid transformation, optimal weight matching.

*Citation*: Liu SX, Ji GJ, Shi CC. Point cloud registration based on global compatibility feature. Computer Optics 2025; 49(6): 1012-1022. DOI: /2412-6179-CO-1616.

## *Introduction*

Point cloud registration involves aligning two point clouds that share a common overlap area through a combination of rotation and translation. This technique has significant applications in fields such as synchronous positioning mapping[1, 2], autonomous driving, augmented reality[3, 4], and robotic automatio[5]. To begin the registration process, corresponding features between the two point clouds are established. Multiple iterations are then conducted to estimate the optimal 3D rotation and translation that achieves the best match. However, issues arising from the scanning equipment and environmental factors can lead to improper overlaps or sparse and fuzzy features. These challenges can introduce corresponding outliers during model estimation, resulting in incorrect matching pairs and ultimately causing mismatches in point cloud registration. RANSAC (Random Sample Consensus) [6]was the first method to introduce iterative sampling for model estimation. However, it tends to include a significant number of invalid point cloud matches, which can prolong convergence time. Sometimes, the presence of outliers can lead to erroneous registration results. In recent years, a similarity measure known as spatial compatibility has been proposed to enhance the stability and accuracy of rigid transformation estimations[7, 8, 9]. This approach operates on the assumption that if the differences between corresponding spatial distances are small, the evaluation score is likely to be high, as shown in the notation $|D1-D1'|$, $|D2-D2'|$, which are expected to be very similar. By increasing the probability of sampling valid points, this method aims to reduce errors. However, due to its locality and reliance on one-dimensional calculations, the resulting data may not accurately represent direction, leading to first-order measurement errors. In the figure1, points a1 to a6 represent correct matching values, while outliers such as a7 and a8 can occur occasionally. These incorrect matches can adversely affect the iterative calculations, ultimately impacting the final matching results.

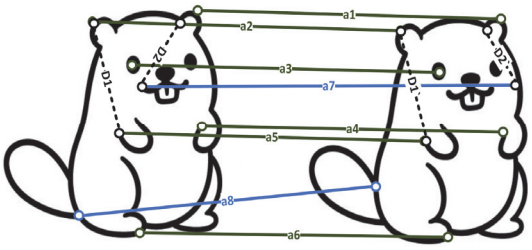


*Fig. 1. Illustration of point cloud matching:The illustration represents a 3D global point cloud structure distribution feature, and the selected points on the illustration correspond to the feature descriptor of the 3D point cloud. The labeled a7, a8 on the illustration are the error compatible matches corresponding to the point cloud*

In this paper, we tackle the challenge of eliminating mismatches in 3D point cloud registration. We propose a new method based on multidimensional spatial compatibility features. Specifically, our algorithm introduces a multi-dimensional compatible common feature to efficiently identify matching pairs between the target registration point cloud and the source point cloud. Instead of focusing solely on local consistency, we consider the overall compatibility of matching pairs

within the entire point cloud block. This approach helps reduce the number of false matches and yields more accurate matching results. As illustrated in the figure2, we first extract a set of corresponding features, referred to as group A. Next, we calculate and binarize their compatible corresponding scores within this group. We compute the number of common compatible scores in their respective groups to establish a new compatibility score. We then extract corresponding features with matching scores again and designate this new set as group B, which may include features originally from group A. This process continues, gene pairs. Although incorrect matches may occasionally have correct compatibility relations resulting in a value of 1, most other matching feature point relations will yield a score of 0. This is further illustrated in Fig. 2.

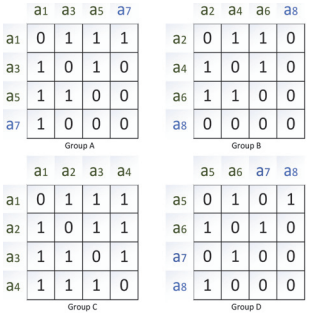


*Fig. 2. Diagram of the compatibility relationship*

As shown in Fig. 3, the compatibility relation value for correctly matched pairs is no less than 4, while the compatibility for incorrectly matched pairs does not exceed 2. Therefore, the binary matrix representing the compatibility relationship, calculated from multiple groups of extracted features, can effectively distinguish between correct and incorrect point pairs. Ultimately, the multi-dimensional spatial consistency feature is expressed as a grouping calculation of the second-order spatial consistency based on the traditional first-order metric.

The proposed multi-dimensional spatial consistency feature has several advantages:

1. It can efficiently separate correct matching pairs from incorrect ones. For instance, if there are m points in a dataset with n correct point pairs, the compatibility value for a single group will be at least n-2. This value will be significantly higher than that of incorrect pairs after aggregating the total statistics. Since an incorrect matching pair closely resembles at most one correct pair, it is unlikely for multiple corresponding pairs to yield a high compatibility score.
2. Unlike traditional methods such as RANSAC and its variants, which require extensive random calculations to robustly estimate a model from a group of points, our proposed method does not necessitate traversing all point pairs in the entire point cloud for compatibility calculations. Instead, we only need to extract the relevant point pairs for repeated grouping. This approach substantially reduces computational effort and efficiently provides the correct number of points, leading to more accurate and efficient model estimation.
3. Our method primarily focuses on processing point cloud data features to differentiate between correct and incorrect point pairs. It can serve as a preprocessing step for popular deep learning frameworks, significantly decreasing the computational load and extracting richer information compared to first-order measures.

The remaining sections of the paper are organized as follows: First, we will discuss related work in point cloud registration, focusing on 3D feature matching, traditional model fitting, and learning-based model fitting. Section 2 provides an overview of solutions to the registration problem, covering multidimensional space compatibility, feature point selection, handling of discrete outliers, and optimal weight matching. Section 3 presents experimental verification using the 3DMatch, 3DLoMatch, and KITTI datasets. The final section concludes the paper and outlines possible future research directions.

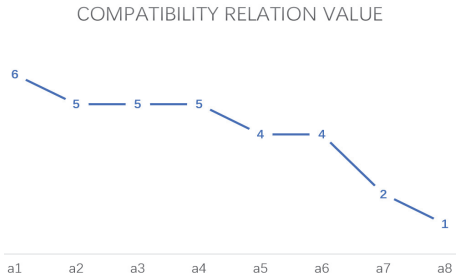


*Fig. 3. Diagram of the compatibility relationship*

## *1. Related work*

Iterative closest point is widely used in 3D Feature Matching [11] and its many variants[12, 13], which iteratively establishes correspondence by searching for the closest point in the coordinate space. The purpose of local feature descriptor is to establish the corresponding matching relationship in the feature space. Handcrafted descriptors encoding spatial distribution histograms [14, 15], geometric histograms [16, 17], and more recently deep learning techniques [18, 19] can also greatly optimize the extraction of feature descriptors. Although this kind of method can achieve significant performance improvement, it is still inevitable to produce false matches.

Traditional model fitting methods estimate geometric models from noisy correspondence sets. RANSAC mentioned above uses a generation and validation pipeline to effectively remove outliers, and its multi-class variant introduces new sampling strategies to speed up estimation and improve stability. For specific 3D model fitting, FGR [20] uses a special cost function to fit and estimate the model, and TEASER [21]reformulates the labeling problem by using the method of minimum truncation, and solves it by using the graph theory framework.

In recent studies, many deep learning techniques have been used to establish the fitting task, and the most widely used is the two-dimensional selective network CN-Net [22], which constructs the model fitting as a combination of the corresponding classification module and the model estimation module. Deep learning modules are also introduced to perform corresponding 3D pruning. 3DRegNet [23] reformulates CN-Net into a 3D form and designs a regression module to solve the rigid transformation problem. PointDSC develops a non-local module based on spatial consistency and a neural spectral matching to accelerate model generation and selection.

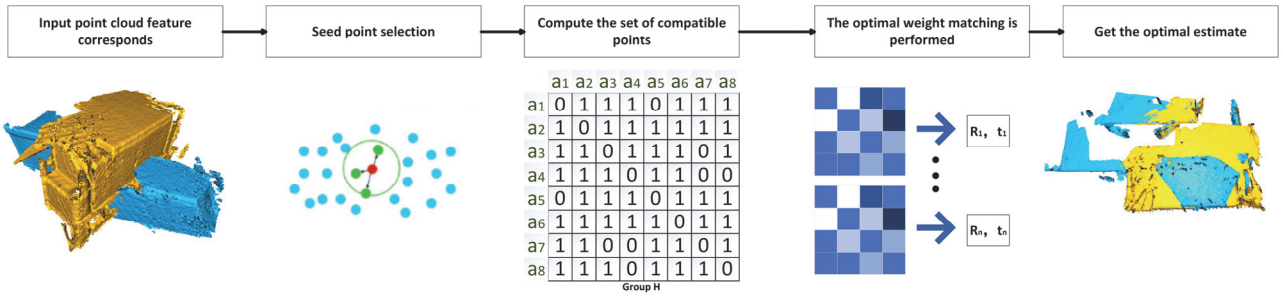


*Fig. 4. Pipeline of our method:1. Input point cloud and extract features. 2. Retrieve reliable seed points in neighborhood. 3. Calculate the corresponding compatible point set. 4. The optimal weight matching is performed for each seed to generate the corresponding R and t. 5. The best estimate is selected as the final resultglobal point cloud compatibility*

## 2. Method

### 2.1. Problem concept description

Given two point clouds that need to be matched: source point cloud $X=\{x_i \in R^3,\ i=1,2,3,n\}$, target point cloud $Y=\{y_i \in R^3,\ i=1,2,3,n\}$, where R3 is the set of three-dimensional points. Firstly, the data is preprocessed to remove most of the noise, and then the local features of the two are extracted. For each feature point in the source point cloud, the corresponding feature point pair is found in the target point cloud to form the assumed correspondence. The global compatibility of the generated correspondence is calculated, and the best weight matching is selected to generate the rigid transformation matrix of the two point clouds.The rigid transformation matrix can be computed into a rotation matrix(R) and a translation vector(t) — and then the point cloud registration is completed by the transformation calculation. A rough flowchart is shown below Fig. 4.

### 2.2. Global compatibility feature

To prove and analyze the consistency in multidimensional space, we need to set the consistency events in point cloud feature extraction to be represented by a set of probability functions :

$$\mathrm{P}(\mathrm{M}) = \mathrm{P}(\mathrm{M}_{in,out} > \mathrm{M}_{in,in}) = \mathrm{P}(\mathrm{N}_{out} > \mathrm{N}_{in}), \qquad (1)$$

where M represent the corresponding compatibility measure, while P(A) denotes the probability of practice A. In this context, all formulas with subscripts"in" and"out" refer to correct data inliers described by the model and abnormal data outliers, respectively – data that deviate from normal values. Specifically, $N_{out}$ represents the compatibility between point pairs formed by the inlier points of the function and the outlier points. Conversely, $N_{in}$ signifies the compatibility score between two inlier points. It is noteworthy that a substantial number of compatibility scores between inliers and outliers exhibit high values (i.e., $N_{out} > N_{in}$). This phenomenon indicates that the point cloud possesses insufficient local features, leading to a limited number of correct matches. Consequently, the generation of accurate rotation and translation matrices becomes challenging, resulting in a high likelihood of failure in meter-based sampling. In contrast, when $N_{out} > N_{in}$, the probability of event P diminishes, thereby yielding a more stable sampling metric. According to the widely used first-order spatial consistency method, the spatial consistency between corresponding i and j is defined as follows:

$$SC_{ij} = \Phi(d_{ij}), d_{ij} = \left| d(x_i, x_j) - d(y_i, y_j) \right|, \qquad (2)$$

where ($x_i$, $y_i$) and ($x_j$, $y_j$) denote the corresponding matching points of indices i and j, respectively. Here, **Φ** represents a monotonically decreasing kernel function, and d(,)signifies the Euclidean distance. Given the nature of rigid transformations, the distance difference between two inliers, denoted as $d_{in,in}$, should ideally equal zero. However, due to the inherent presence of error noise in the collected experimental data, $d_{in,in}$ is not exactly zero; rather, it is typically less than a specified threshold $X_{thr}$. To address this issue, it becomes essential to define a probability density function $T_{in}$. For the sake of simplicity, we will assume that $d_{in,in}$ is uniformly distributed over the range defined by $X_{thr}$. Consequently, when the distance difference between the inlier points is less than $X_{thr}$, it is assumed that they represent compatible matches:

$$T_{in}(l) = 1 / X_{thr}, 0 < l < X_{thr}. \qquad (3)$$

Additionally, it is crucial to consider the distinct probability density functions that exist between inlier points and outlier points, as well as among outlier points themselves. These points do not exhibit similar length consistency; hence, the distance differences between two

unrelated points are identically distributed. This observation leads to the formulation of the assumed probability density function $T_{out}$:

$$T_{out}(l) = F(l), 0 < l < X_r, \quad (4)$$

where $X_r$ denote the distance range between inlier points and outlier points, as well as between outlier points themselves. Furthermore, due to the compatibility considerations of inlier and outlier points, it follows that $X_r$ must be significantly larger than $X_{thr}$. Consequently, F(l) can be approximated as a constant $l_0$ within the context of $T_{in}$ (l):

$$F(l) = l_0, 0 < l < X_{thr}. \quad (5)$$

Subsequently, the first-order spatial compatibility is calculated as follows:

$$\begin{aligned} &\mathrm{P(M)} = \mathrm{P}(\mathrm{M}_{in,out} > \mathrm{M}_{in,in}) = \mathrm{P}(\mathrm{N}_{out} > \mathrm{N}_{in}) = \\ &= \mathrm{P}(\mathrm{d}_{\mathrm{in,out}} < \mathrm{d}_{\mathrm{in,in}}) = \int_0^{X_{thr}} \int_0^{l} T_{in}(l) T_{out}(x)\, dxdl = \\ &= \int_0^{X_{thr}} \int_0^{l} \frac{1}{X_{thr}} l_0 dxdl = \frac{X_{thr} l_0}{2}. \end{aligned} \quad (6)$$

The experimental results obtained from the 3DMatch dataset indicate thathen the critical threshold is set to 0.1 m, the first-order spatial compatibility ambiguity probability is 0.1. Notably, the presence of outliers can comprise a significant portion of the two-point clouds, which may still lead to registration errors. To address this issue, we aim to improve first-order spatial consistency by introducing a multi-dimensional spatial consistency approach. This method is designed to extract a more comprehensive set of feature point data from the point cloud. To initiate this process, we must first construct a binarized consistency matrix E:

$$E_{ij} = \begin{cases} 1; d_{ij} < X_{thr} \\ 0; d_{ij} > X_{thr} \end{cases}. \quad (7)$$

Matrix E indicates that two corresponding point pairs that satisfy the length consistency criterion are considered compatible, represented by a value of 1. Simultaneously, the number of compatible point pairs will be recorded during the retrieval process:

$$E'_{ij} = E_{ij} \sum_{k=1}^{N/2} E_{ik} E_{kj}. \quad (8)$$

To calculate the number of point pairs in each matrix that satisfy the compatibility criteria, it is essential to ensure that both the corresponding row and column of the matrix are non-zero simultaneously; only then can an effective count be conducted. The numerical values that meet these conditions are recorded in the matrix as instances of correct compatibility. In this process, we divide the entire set of feature points into four parts, specifically separating them into two groups based on their natural number order and parity (odd and even). This division aims to expedite the matching process between the two groups of point pairs while also enhancing the global compatibility calculations. Such an approach increases the discrete probability of correctly matching point pairs, thereby significantly improving the final accuracy of point cloud matching. The objective is to repeat the global compatibility calculations for both odd and even point pairs, subsequently aggregating these results to achieve global compatibility stability, as illustrated in Fig. 3. Importantly, the presence of zeros does not affect the overall result, regardless of the number of times they are superimposed. Ultimately, We have compiled four sets of corresponding points, which are logged and summed for analysis ;

$$\begin{aligned} E_{all} &= E_{ij} \sum_{k=1}^{N1} E_{ik} E_{kj} + E_{ij} \sum_{k=1}^{N2} E_{ik} E_{kj} + \\ &+ E_{ij} \sum_{k=1}^{N3} E_{ik} E_{kj} + E_{ij} \sum_{k=1}^{N4} E_{ik} E_{kj}. \end{aligned} \quad (9)$$

$$\begin{aligned} &\mathrm{P(M)} = \mathrm{P}(\mathrm{M}_{in,out} > \mathrm{M}_{in,in}) = \mathrm{P}(\mathrm{N}_{out} > \mathrm{N}_{in}) = \\ &= \mathrm{p} \cdot \mathrm{P}(\mathrm{X} > (2\mathrm{N} \cdot \mathrm{a})), \\ &\mathrm{X} \sim \mathrm{S}\left((2\mathrm{Na} - 4)\mathrm{p} + (2\mathrm{N}(2-\mathrm{a}) - 1)\mathrm{P}^2, \mathrm{N}(1-\mathrm{a})\mathrm{P}^2\right), \\ &p = 2X_{thr} l_0. \end{aligned} \quad (10)$$

In the formula, S (,) follows the Skellam distribution. At the same time, according to its property, as the value of p gets closer to 0, the proportion of interior points will also increase. Therefore, compared with the consistency of first-order space, the multidimensional space can still better remove discrete points in the case of fewer correct points, and complete efficient and more robust registration sampling.

### *2.3. Selection of feature points*

Utilizing the global compatibility feature of inlier and outlier feature points significantly improves the accuracy of identifying correct and incorrect point pairs. Once an inlier correspondence is established, we can construct a consistent point set by identifying its k nearest neighbors within the global compatibility space. By iterating over all feature points, we are likely to find at least one inlier to form a consistent spatial point set in the subsequent steps. However, this exhaustive approach can consume considerable computational resources.To optimize the retrieval process, we propose selecting reliable seed points through commonly employed spectral matching techniques. Specifically, we first establish a compatibility matrix using all obtained feature point pairs. For ease of subsequent processing, we binarize this matrix by calculating the correlation between the corresponding point pairs and the leading feature vector, which serves as a measure of confidence for each point pair. Additionally, to enhance the efficiency of seed point retrieval, we incorporate a priori conditions and adjust the retrieval

radii based on the number of domain points. This strategy aims to improve both the effectiveness and robustness of the retrieval method.

### 2.4. H ndling of discrete outliers

Once a seed point is established to create a compatible point set, it is crucial to screen for discrete outliers within that set. When the seed corresponds to correct matches, the resulting compatible point set will also comprise correct points. Initially, we collect all compatible point sets in the surrounding neighborhood to capture global information rather than solely focusing on local consistency. To facilitate the estimation of the final rigid transformation, it is essential to gather a consistent point set with a uniform distribution within the retrieval radius of the k nearest neighbors. Following this, we must filter out outliers from the obtained sets. This involves constructing a spatial consistency matrix for each set separately, allowing us to distinguish discrete and abnormal values. Given that the fuzzy probability of spatial consistency

$$\mathrm{P}(\mathrm{M}) = \mathrm{P}(\mathrm{M}_{in,out} > \mathrm{M}_{in,in}) \tag{11}$$

is low, this approach effectively removes outliers. Additionally, it is important to note that a discrete point itself may serve as a seed point, which could lead to the formation of a smaller local compatible point set, particularly in cases where numerous incorrect point sets exist within the point cloud matching.Even if these smaller sets generate corresponding compatible point sets, they will ultimately be filtered out in the final best-estimate transformation, eliminating the need for additional computations to remove them.

### 2.5. Optimal weight matching

In this step, we can efficiently generate the corresponding rigid transformation estimate using Singular Value Decomposition (SVD) to complete our registration task. However, while the previously proposed strategy for constructing the compatible point set allows for the identification of a corresponding point set devoid of outliers, incorporating varying weights for optimal matching can yield even better results. To achieve this, we employ spectral matching to analyze and compute the global compatibility feature. Notably, as the level of noise increases, a smaller weight is assigned to correspondences, reflecting their reduced reliability. Each correspondence is thus weighted according to the following formula:

$$\tilde{\mathrm{E}}_{\mathrm{ij}} = \mathrm{Sigmoid}\left(1 - \mathrm{d}_{\mathrm{ij}} / \mathrm{X}_{\mathrm{thr}}\right), \left(1 \le i, j \le n\right), \tag{12}$$

where i and j are indices greater than 1 and less than the total number of point sets n, the weight q is determined through local spectral decomposition of the $E_{ij}$ matrix. Subsequently, the rotation matrix R and translation vector t of the corresponding rigid transformation are derived by applying weighted SVD to the compatible point set formed by the seed point. From all the rigid transformations generated by the compatible point set, the best estimate is selected. The selection criterion is based on whether the error produced by the KTH seed calculation—using the corresponding R and t transformations—falls below a predefined error threshold e.

$$\mathrm{number}_{\mathrm{k}} = \sum_{\mathrm{i}=1} [] \left| \left( R_k x_i + t_k \right) - x \right| < e] , \tag{13}$$

where [] is the Iverson bracket, and the rotation matrix (R) and translation vector (t) of the rigid transformation with the highest number value is finally selected as the optimal result. The overall algorithm flowchart is shown in the Algorithm 1:

**Algorithm 1.** Point cloud registration based on global compatibility feature

**Require:** $x_i$, $y_i$: Input point clouds

**Ensure:** $T$: Predicted transformation matrix (Rotation $R$ and Translation $t$)

1. **Input:** source point cloud $x_i$, target point cloud $y_i$
2. Extract local features for both of them
3. Extract initial correspondences from $x_i$ and $y_i$
4. Compute pairwise distances: $D_1 \leftarrow \mathrm{dist}(x_i)$, $D_2 \leftarrow \mathrm{dist}(y_i)$
5. Compute cross-distance matrix: $D_{cross} \leftarrow |D_1 - D_2|$
6. Set first-order threshold $X_{thr}$
7. **for** each element $d \in D_{cross}$ **do**
8. Compute $hard\_SC \leftarrow (d < X_{thr})$ as float
9. **end for**
10. Use spectral analysis to select seeds $S$ with high confidence
11. Expand each seed to consensus set using global compatibility:
12. **for** each correspondence $N \in S$ **do**
13. Compute compatibility measure:

$$\mathrm{E}(ij) = \sum_{k \in \mathrm{neighbors}(N)} E[i\ k] \cdot E[k, j]$$

14. **end for**

Compute weights via power iteration on global compatibility matrix
**for** each seed in $S$ **do**
    Estimate transformation $T_i$ using weighted SVD on the consensus set
**while** $T$ is not converged **do**
$R \leftarrow$ empty matrix
$t \leftarrow$ empty vector
**for** $i = 1 \in \{0, ..., N\}$ **do**
$R \leftarrow$ append($R$, computeJacobian($S[i]$))
    $t \leftarrow$ append($t$, computeErrorVector($S[i]$, closestPoint))
**end for**
$\Delta T \leftarrow$ solve($R^{\top}R$, $R^{\top}t$)
$T \leftarrow T + \Delta T$
**if** $\|\Delta T\| < 1 \times 10^{-6}$ **then**
**break** ▷ Exit loop if convergence criterion met
**end if**
**end while**
**end for**
Select the best transformation $T \leftarrow \text{argmin}_{T_i} \text{Error}(T_i)$
Refine $T$ by re-estimating over the whole set of correspondences
**return** $T$

## *3. Experiment*

### *3.1. Dataset selection and experiment deployment*

This paper selected the 3DMatch dataset, 3DLoMatch dataset, and KITTI dataset for indoor, outdoor, and low overlap point cloud feature extraction. It also provides a detailed analysis of the characteristics and data collection methods of these datasets:

1. 3DMatch is a collection of 62 scenes, the official benchmark splits the data into 54 scenes for training and 8 for testing. Individual scenes are not only captured in different indoor spaces (e.g., bedrooms, offices, living rooms, restrooms) but also with different depth sensors (e.g., Microsoft Kinect, Structure Sensor, AsusXtion ProLive, and Intel RealSense). 3DMatch provides great diversity and allows our model to generalize across different indoor spaces. Individual scenes of 3DMatch are split into point cloud fragments, which are generated by fusing 50 consecutive depth frames using TSDF volumetric fusion. As a preprocessing step, we apply voxel-grid downsampling to all point clouds, and if multiple points fall into the same voxel, we randomly pick one.
2. 3DLoMatch dataset containing the previously ignored scan pairs of 3DMatch that have low overlap. It is typically employed to validate the stitching effect of low-overlap point clouds.
3. KITTI dataset was jointly founded by the Karlsruhe Institute of Technology and the Toyota Institute of Technology, and is currently the largest computer vision algorithm evaluation dataset in the world. This dataset is used to evaluate the performance of computer vision techniques such as stereo, optical flow, visual odometry, object detection and tracking in vehicular environments. KITTI contains real image data collected from urban, rural and highway scenes, with up to 15 vehicles and 30 pedestrians in each image, and various degrees of occlusion and truncation.

We conducted experiments on indoor point cloud scene reconstruction using the publicly available 3DMatch dataset [24]. During preprocessing, we applied voxel downsampling to reduce the point cloud density, thereby accelerating the matching process. Subsequently, we extracted local feature descriptors for registration to establish correspondences, employing both the FPFH and FCGF descriptors in our experiments. To evaluate experimental performance under conditions of low overlap, we selected the 3DLoMatch dataset[25] for further experimentation, utilizing FCGF and Predator as feature descriptors. Additionally, recognizing the diversity of indoor and outdoor scenes, we included the KITTI dataset[26] to assess the matching efficacy in outdoor environments. The KITTI dataset contains real image data collected from urban, rural, and highway scenes, and is commonly used for evaluating 3D object detection, point cloud registration, and stereo image processing. Similar to the previous datasets, voxel downsampling was employed for sparsification, and both FPFH and FCGF were selected as feature descriptors. For evaluation, we adopted widely used error statistics, outputting the registration and recovery (RR) metrics under specified error thresholds. These thresholds were set at (15 degrees, 20 cm) for the 3DMatch dataset and (10 degrees, 40 cm) for the KITTI dataset. We calculated and analyzed the translation and rotation errors for the registered point clouds, focusing on rotation error (RE) and translation error (TE) [27]. To assess the results concerning outliers, we also calculated inlier precision (IP), inlier recall (IR), and the F1-measure (F1). In calculating the spatial compatibility matrix, we set the critical threshold $K_{thr}$ to twice the voxel downsampling

value, resulting in 20 cm for indoor scenes and 40 cm for outdoor scenes. The number of selected seeds was established as half the logarithm of the total number of points within the corresponding point set. During sampling and screening, the radius of the field was set to one-fifth of the number of corresponding point sets, ultimately selecting the first 80 percent of the data to construct a consistent point set.

### 3.2. 3DMatch dataset

We will select 12 commonly used benchmark experimental methods for comparison DCP, 3DRegNet, DGR[28], DHVR, PointDS, SM, ICP, FGR, TEASER, GC-RANSAC[29], RANSA[30], SC2-PCR[31]. The first five methods utilize deep learning technology, while the last seven employ traditional techniques. Notably, DCP and ICP do not have corresponding outlier screening methods, which is why their results exclude this aspect. For the FPFH descriptor, we computed the mean rotation error (RE) and translation error (TE) associated with successful point cloud matching for each method. The registration recall (RR) of the registration results serves as the most accurate measurement index in 3D registration, and our method significantly outperforms the mainstream techniques. However, it is important to acknowledge that we cannot fully demonstrate the superiority of our method in eliminating false matches without considering all false registration results. Nevertheless, our approach still achieves a relatively favorable index in terms of outlier exclusion.as shown in Tab. 8.

*Tab. 1. Table of experimental results of FPFH feature descriptor in 3DMatch dataset*

| FPFH | RR(%) | RE(。) | TE(cm) | IP(%) | IR(%) | F1(%) |
|---|---|---|---|---|---|---|
| DCP | 3.22 | 8.42 | 21.40 | - | - | - |
| DRegNet | 26.31 | 3.75 | 9.61 | 28.21 | 8.91 | 11.63 |
| DGR | 32.84 | 2.45 | 7.53 | 29.51 | 16.78 | 21.35 |
| DHVR | 67.10 | 2.79 | 7.84 | 60.19 | 64.91 | 62.11 |
| PointDSC | 77.54 | 2.03 | 6.38 | 68.45 | 71.56 | 69.75 |
| SM | 55.98 | 2.94 | 8.51 | 47.96 | 70.96 | 50.71 |
| ICP | 5.79 | 7.96 | 17.56 | - | - | - |
| FGR | 40.91 | 4.96 | 10.25 | 6.84 | 38.90 | 11.53 |
| TEASER | 75.48 | 2.48 | 7.31 | 73.01 | 62.63 | 66.94 |
| GC-RANSAC | 67.65 | 2.33 | 6.84 | 48.56 | 69.39 | 56.79 |
| RANSAC | 64.21 | 4.03 | 11.36 | 63.96 | 57.90 | 61.02 |
| SC2PCR | 78.03 | 2.41 | 6.98 | 72.46 | 78.33 | 75.11 |
| Ours | 82.13 | 2.16 | 6.53 | 71.95 | 77.46 | 76.39 |

To further validate the performance of our approach, we selected FCGF as the feature descriptor for our experiments. The FCGF is employed to generate correspondence estimates and to report the registration results. Notably, the initialization rate with FCGF is significantly higher than that achieved with the FPFH descriptor, leading to substantial improvements in the performance of all methods. Furthermore, our method consistently outperforms the sampling method in terms of the registration recall (RR) measurement index. Additionally, our approach requires the construction of compatible point sets only for a limited number of seed points. This significantly reduces both the computational load and registration time, thereby rendering our method a practical solution for 3D registration.

*Tab. 2. Table of experimental results of FCGF feature descriptor in 3DMatch dataset*

| FCGF | RR(%) | RE(。) | TE(cm) | IP(%) | IR(%) | F1(%) | Times |
|---|---|---|---|---|---|---|---|
| DCP | 3.22 | 8.42 | 21.40 | - | - | - | 0.07 |
| DRegNet | 77.46 | 2.76 | 8.16 | 67.64 | 59.39 | 58.63 | 0.12 |
| DGR | 88.69 | 2.47 | 7.06 | 69.15 | 79.69 | 73.26 | 1.53 |
| DHVR | 91.89 | 2.26 | 7.09 | 79.56 | 78.56 | 78.69 | 3.93 |
| PointDSC | 92.67 | 2.09 | 6.52 | 78.91 | 86.24 | 82.16 | 0.46 |
| SM | 86.59 | 2.25 | 7.06 | 81.46 | 45.52 | 48.21 | 0.23 |
| ICP | 5.49 | 7.96 | 17.58 | - | - | - | 0.04 |
| FGR | 78.49 | 2.93 | 8.39 | 25.64 | 53.93 | 33.69 | 0.25 |
| TEASER | 85.94 | 2.71 | 8.46 | 82.45 | 69.18 | 74.86 | 0.98 |
| GCRANSAC | 91.56 | 2.69 | 7.17 | 69.46 | 89.26 | 76.46 | 0.55 |
| RANSAC | 88.43 | 3.06 | 9.46 | 77.96 | 79.87 | 78.56 | 0.97 |
| SC2PCR | 89.63 | 2.42 | 7.69 | 75.32 | 85.62 | 79.49 | 0.21 |
| Ours | 93.56 | 2.01 | 6.59 | 78.64 | 86.43 | 82.31 | 0.13 |

### 3.3. 3DLoMatch dataset

To assess the robustness of our method with low-overlap datasets, we conducted experiments using the 3DLoMatch dataset. For these experiments as shown in Table 3:, we selected two feature descriptors: FCGF and Predator. The measurement indices used for evaluation were registration recall (RR), mean rotation error (RE), and translation error (TE). The experimental results demonstrate that our method achieves excellent

performance regardless of whether the FCGF or Predator descriptor is employed.

*Tab. 3. Table of experimental results of FCGF and Predator feature descriptor in 3DLoMatch dataset*

| FCGF | RR(%) | RE(°) | TE(cm) | Predator | RR(%) | RE(°) | TE(cm) |
|---|---|---|---|---|---|---|---|
| DHVR | 54.43 | 4.14 | 12.56 | DHVR | 65.41 | 4.97 | 12.33 |
| DGR | 43.80 | 4.17 | 10.82 | DGR | 59.46 | 3.19 | 10.11 |
| PointDSC | 56.09 | 3.87 | 10.39 | PointDSC | 68.89 | 3.43 | 9.63 |
| FGR | 19.99 | 5.28 | 12.98 | FGR | 35.99 | 4.77 | 11.65 |
| RANSAC | 46.38 | 5.00 | 13.11 | RANSAC | 64.89 | 4.28 | 11.04 |
| SC2-PCR | 57.83 | 3.77 | 10.46 | SC2-SAC | 69.46 | 3.46 | 9.58 |
| Ours | 59.33 | 3.89 | 11.23 | Ours | 72.98 | 3.28 | 10.65 |

### *3.4. KITTI dataset*

Additionally, we conducted experiments on outdoor datasets, specifically utilizing the KITTI dataset. We selected several traditional registration methods for comparison, including DHVR, DGR, PointDSC, FGR, RANSAC, and SC2-PCR. The experimental results indicate that our method outperforms these traditional registration techniques, with mean rotation error (RE) and translation error (TE) significantly lower than those of RANSAC, placing our performance on par with deep learning methods. For these experiments, both FPFH and FCGF were chosen as feature descriptors: experiments as shown in Table 4:

*Tab. 4. Table of experimental results of FPFH and FCGF feature descriptor in 3DLoMatch dataset*

| FPFH | RR(%) | RE(°) | TE(cm) | FCGF | RR(%) | RE(°) | TE(cm) |
|---|---|---|---|---|---|---|---|
| DHVR | - | - | - | DHVR | 98.67 | 0.29 | 19.83 |
| DGR | 77.12 | 1.64 | 33.12 | DGR | 96.32 | 0.46 | 21.76 |
| PointDSC | 97.64 | 0.37 | 8.46 | PointDSC | 98.03 | 0.34 | 21.04 |
| FGR | 5.24 | 0.78 | 43.89 | FGR | 89.56 | 0.49 | 25.16 |
| RANSAC | 75.46 | 1.55 | 30.20 | RANSAC | 80.37 | 0.76 | 27.83 |
| SC2-PCR | 76.46 | 0.74 | 14.12 | SC2-SAC | 96.43 | 0.57 | 22.97 |
| Ours | 99.46 | 0.43 | 7.46 | Ours | 98.42 | 0.34 | 20.95 |

### *3.5. Ablation experiment*

To demonstrate the effectiveness of the proposed method, we conducted experiments using the 3DMatch dataset. For feature extraction, we selected FPFH and FCGF descriptors, and RANSAC was employed as the baseline method. We systematically added new modules to the RANSAC framework and recorded the results.

First, we transformed the sampling approach of RANSAC into a multi-dimensional spatial compatible point set (denoted as A). By identifying neighboring points in the metric space, we constructed a spatial compatible point set for registration from each point set. The results indicated a significant improvement in registration recall (RR), while also effectively mitigating the influence of erroneous point sets and eliminating numerous false matches. This led to enhancements in both accuracy and computational efficiency.

Next, we introduced a more efficient seed selection method (denoted as B), utilizing commonly employed spectral matching techniques to identify reliable seed points, thereby accelerating the retrieval process. With a sufficient number of initial correct seeds available for establishing correspondences, many false matches were naturally filtered out during the formation of the correspondence tolerance matrix, resulting in an improved recall rate.

Finally, we implemented optimal weight matching (denoted as C). By applying different weights to the compatible point sets, we achieved superior results in matching accuracy. In estimating the rigid body transformation, spectral matching was employed to analyze and compute the constructed multi-dimensional space tolerance matrix, with greater noise leading to smaller weights being assigned. This approach significantly reduced both mean rotation error (RE) and translation error (TE).Experiments as shown in Table 5:

*Tab. 5. Table of ablation experimental results*

| | A | B | C | RR(%) | RE(°) | TE(cm) | Times |
|---|---|---|---|---|---|---|---|
| FPFH | | | | 66.13 | 3.98 | 11.06 | 2.89 |
| | ✓ | | | 71.59 | 2.06 | 6.89 | 0.29 |
| | ✓ | ✓ | | 80.89 | 2.65 | 6.91 | 0.34 |
| | ✓ | ✓ | ✓ | 82.13 | 2.16 | 6.53 | 0.11 |
| FCGF | | | | 87.53 | 2.43 | 7.68 | 0.27 |
| | ✓ | | | 93.12 | 2.16 | 6.79 | 0.32 |
| | ✓ | ✓ | | 93.27 | 2.08 | 6.56 | 0.37 |
| | ✓ | ✓ | ✓ | 93.56 | 2.01 | 6.59 | 0.11 |

## *4. Conclusion*

In this paper, we propose a point cloud registration method based on a global compatibility feature. This approach accurately and efficiently classifies feature point pairs by distinguishing between correct and incorrect matches, while clustering the correct point clouds to eliminate errors. Consequently, this process facilitates the construction of a compatible point set, ultimately enabling successful point cloud registration. Initially, the source and target point clouds, having undergone complete feature extraction, are input into the

system. We utilize spectral matching techniques to identify high-confidence points as seed points within these features. Following this, neighborhood analysis is conducted to construct the compatible point set. For each compatible point set, we perform optimal weighted matching estimation, which generates corresponding rigid transformations. The optimal rotation matrix and translation vector, yielding the highest score, are selected in the final step. The transformation matrix is then applied to the source point cloud to achieve registration with the target point cloud. Through extensive experiments, we demonstrate that our method is both feasible and exhibits excellent performance and accuracy. Looking ahead, we aim to explore future research directions that further integrate deep learning technologies. By leveraging the efficient iterative capabilities of deep learning, we aspire to optimize our method and enhance its overall performance.

### *Acknowledgements*

This research supported by the Open Project Program of The Key Laboratory of Cognitive Computing and Intelligent Information Processing of Fujian Education Institutions, Wuyi University (KLCCIIP202203).

*Tab. 6. Terminology Definition Table*

| | |
|---|---|
| the corresponding feature point pair | The corresponding feature point pair refers to a set of two points, one from each of the two point clouds, that are identified as matching or corresponding to each other based on a specific criterion. |
| the global compatibility | Global compatibility refers to the degree to which a set of points, features, or transformations across multiple data sources align in a consistent and coherent manner when considered as a whole.Global compatibility ensures that the alignment or transformation applied to individual pairs of points or features maintains consistency across the entire point cloud, leading to an overall optimal match between the point clouds. |
| the first-order spatial consistency method | The first-order spatial consistency method typically refers to a technique used to ensure the alignment or registration of data (such as 3D point clouds or images) by maintaining consistency in local spatial relationships. |
| the first-order spatial compatibility | First-order spatial compatibility refers to the comparison of local geometric relationships between points or features in two datasets based on their immediate or direct spatial properties. It measures how similar the local arrangement or configuration of neighboring points is between two datasets |
| leading feature vector | A leading feature vector in 3D point cloud registration refers to the most important vector that captures key geometric features of the point cloud, which are used to align or match different point clouds during the registration process. |
| spatial consistency matrix | A spatial consistency matrix is a matrix that captures and represents the spatial relationships between points in a dataset. it helps ensure that points stay consistent in their relative positions when aligning or transforming the data. |

*Tab. 7. The experimental deepLearning methods introduction*

| | |
|---|---|
| DCP | DCP (Deep Closest Point) is a deep learning-based network for point cloud registration that automatically aligns different point cloud data by learning the geometric transformation between them, aiming to improve accuracy and efficiency, especially in scenarios with complex shapes or missing data. |
| DRegNet | 3DRegNet is a deep learning model for aligning 3D point clouds. It automatically finds the correct transformation to match different point clouds, even when there are large differences or noise, making the registration process more accurate and reliable. |
| DGR | DHVR (Deep Hierarchical Volumetric Registration) uses a hierarchical approach to progressively refine the alignment. It combines volumetric representations and deep learning to handle large deformations, partial overlaps, and noisy data, providing accurate and robust registration results. |
| DHVR | First-order spatial compatibility refers to the comparison of local geometric relationships between points or features in two datasets based on their immediate or direct spatial properties. It measures how similar the local arrangement or configuration of neighboring points is between two datasets |
| PointDSC | PointDSC (Point Cloud Deep Shape Consistency) is a deep learning method for 3D point cloud registration that focuses on achieving shape consistency between point clouds. It uses a deep network to align point clouds by minimizing shape differences, effectively handling partial overlaps and noise, and ensuring robust and accurate matching. |

### *References*

[1] Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): part II. IEEE Robot Autom Mag 2006; 13(3): 108-117. DOI: 10.1109/MRA.2006.1678144.

[2] Durrant-Whyte H, Bailey T. Simultaneous locaization and mapping: part I. IEEE Robot Autom Mag 2006; 13(2): 99-110. DOI: 10.1109/MRA.2006.1638022.

[3] Azuma RT. A survey of augmented reality. Presence-Teleop Virt 1997; 6(4): 355-385. DOI: 10.1162/pres.1997.6.4.355.

[4] Billinghurst M, Clark A, Lee G. A survey of augmented reality. Found Trends Hum Comput Interact 2015; 8(2-3): 73-272. DOI: 10.1561/1100000049.

[5] Kostavelis I, Gasteratos A. Semantic mapping for mobile robotics tasks: A survey. Robot Auton Syst 2015. 66: 86-103, DOI: 10.1016/j.robot.2014.12.006.

*Table. 8. The experimental traditional methods introduction*

| | |
|---|---|
| SM | SM (Shape Matching) is a 3D point cloud registration method that focuses on matching the shapes of two point clouds by aligning their geometric structures. It typically involves optimizing a transformation that minimizes the difference between the shapes, even in the presence of noise or partial overlap, aiming for accurate and consistent registration. |
| ICP | ICP (Iterative Closest Point) is an algorithm that aligns two 3D point clouds by repeatedly adjusting their positions to minimize the distance between matching points, refining the alignment step by step. It works best when the initial guess is already close. |
| FGR | FGR (Fast Global Registration) is a 3D point cloud registration method that aims to quickly and accurately align point clouds by minimizing the global transformation between them. It uses a coarse-to-fine strategy, leveraging feature matching and geometric constraints to handle large-scale deformations and noise efficiently. |
| TEASER | TEASER is a 3D point cloud registration method that uses a transformer-based network to accurately align point clouds, even with noise or partial overlap. It efficiently estimates the transformation (rotation and translation) between point clouds for precise matching. |
| GC-RANSAC | GC-RANSAC is an improved version of RANSAC for 3D point cloud registration. It helps align point clouds more accurately by handling outliers and mismatches better, making it effective for noisy or partially overlapping data. |
| RANSAC | RANSAC (Random Sample Consensus) is an algorithm used to find the best fit for a model (like a transformation in 3D point cloud registration) while ignoring outliers. It works by randomly selecting a subset of points, fitting a model to them, and then checking how many other points fit the model. It repeats this process to find the best model with the most inliers. |
| SC2PCR | SC2PCR is a 3D point cloud registration method that uses a convolutional neural network to find the best alignment. It learns from both local and global features of the point clouds, making it robust to noise and partial overlaps. |


[6] Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM 1981; 24(6): 381-395. DOI: 10.1145/358669.358692.

[7] Bai X, Luo Z, Zhou L, Chen H, Li L, Hu Z, Fu H, Tai C-L. PointDSC: Robust point cloud registration using deep spatial consistency. 2021 IEEE/CVF Conf on Computer Vision and Pattern Recognition (CVPR) 2021: 15859-15869. DOI: 10.1109/CVPR46437.2021.01560.

[8] Lee J, Kim S, Cho M, Park J. Deep hough voting for robust global registration. 2021 IEEE/CVF Int Conf on Computer Vision (ICCV) 2021: 15994-16003. DOI: 10.1109/ICCV48922.2021.01569.

[9] Chum O, Matas J. Matching with PROSAC – progressive sample consensus. 2005 IEEE Computer Society Conf on Computer Vision and Pattern Recognition (CVPR'05) 2005; 1: 220-226. DOI: 10.1109/CVPR.2005.221.

[10] Myatt DR, Torr PHS, Nasuto SJ, Bishop JM, Craddoc R. NAPSAC: High noise, high dimensional robust estimation – it's in the bag. British Machine Vision Conf (BMVC) 2002: 458-467. D01: 10.5244/C.16.44.

[11] Besl PJ, McKay ND. Method for registration of 3-D shapes. IEEE Trans Pattern Anal Machine Intell 1992; 14(2): 239-256. DOI: 10.1109/34.121791.

[12] Bouaziz S, Tagliasacchi A, Pauly M. Sparse iterative closest point. SGP '13: Proceedings of the Eleventh Eurographics/ACMSIGGRAPH Symposium on Geometry Processing 2013: 113-123. DOI: 10.1111/cgf.12178.

[13] Rusinkiewicz S, Levoy M. Efficient variants of the ICP algorithm. Proc Third Int Conf on 3-D Digital Imaging and Modeling 2001: 145-152. DOI: 10.1109/IM.2001.924423.

[14] Johnson AE, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Trans Pattern Anal Machine Intell 1999; 21(5): 433-449. DOI: 10.1109/34.765655.

[15] Tombari F, Salti S, Di Stefano L. Unique shape context for 3d data description. 3DOR '10: Proc ACM Workshop on 3D Object Retrieval 2010: 57-62. DOI: 10.1145/1877808.1877821.

[16] Chen H, Bhanu B. 3D free-form object recognition in range images using local surface patches. Pattern Recogn Lett 2007; 28(10): 1252-1262. DOI: 10.1016/j.patrec.2007.02.009.

[17] Rusu RB, Blodow N, Beetz M. Fastpoint feature histograms (FPFH) for 3D registration. 2009 IEEE Int Conf on Robotics and Automation 2009: 3212-3217. DOI: 10.1109/ROBOT.2009.5152473.

[18] Deng H, Birdal T, Ilic S. PPF-FoldNet: Unsupervised learning of rotation invariant 3D local descriptors. In Book: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, eds. Computer Vision – ECCV 2018. 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part V. Cham, Switzerland: Springer Nature Switzerland AG; 2018: 620-638. DOI: 10.1007/978-3-030-01228-1_37.

[19] Choy C, Park J, Koltun V. Fully convolutional geometric features. 2019 IEEE/CVF International Conference on Computer Vision (ICCV) 2019: 8958-8966. DOI: 10.1109/ICCV.2019.00905.

[20] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast global registration. In Book: Leibe B, Matas J, Sebe N, Welling M, eds. Computer Vision – ECCV 2016. 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II. Cham, Switzerland: Springer International Publishing AG; 2016: 766-782. DOI: 10.1007/978-3-319-46475-6_47.

[21] Yang H, Shi J, Carlone L. TEASER: Fast and certifiable point cloud registration. IEEE Trans Robot 2020; 37(2): 314-333. DOI: 10.1109/TRO.2020.3033695.

[22] Yi KM, Trulls E, Ono Y, Lepetit V, Salzmann M, Fua P. Learning to find good correspondences. 2018 IEEE/CVF Conf on Computer Vision and Pattern Recognition 2018: 2666-2674. DOI: 10.1109/CVPR.2018.00282.

[23] Pais GD, Ramalingam S, Govindu VM, Nascimento JC, Chellappa R, Miraldo P. 3DRegNet: A deep neural network for 3D point registration. 2020 IEEE/CVF Conf on Computer Vision and Pattern Recognition (CVPR) 2020: 7193-7203. DOI: 10.1109/CVPR42600.2020.00722.

[24] Zeng A, Song S, Nießner M, Fisher M, Xiao J, Funkhouser T. 3DMatch: Learning local geometric descriptors from RGB-D reconstructions. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 1802-1811. DOI: 10.1109/CVPR.2017.29.

[25] Huang S, Gojcic Z, Usvyatsov M, Wieser A, Schindler K. PREDATOR: Registration of 3D point clouds with low

overlap. 2021 IEEE/CVF Conf on Computer Vision and Pattern Recognition (CVPR) 2021: 4267-4276. DOI: 10.1109/CVPR46437.2021.00425.
[26] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite. 2012 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2012: 3354-3361. DOI: 10.1109/CVPR.2012.6248074.
[27] Ma Y, Soatto S, Košecká J, Sastry SS. An invitation to 3-D vision: From images to geometric models. New York: Springer Science+Business Media; 2004. ISBN: 978-1-4419-1846-8.
[28] Choy C, Dong W, Koltun V. Deep global registration. 2020 IEEE/CVF Conf on Computer Vision and Pattern Recognition (CVPR) 2020: 2514-2523. DOI: 10.1109/CVPR42600.2020.00259.
[29] Wang Y, Solomon JM. Deep closest point: Learning representations for point cloud registration. 2019 IEEE/CVF Int Conf on Computer Vision (ICCV) 2019: 3523-3532. DOI: 10.1109/ICCV.2019.00362.
[30] Barath D, Matas J. Graph-cut RANSAC. 2018 IEEE/CVF Conf on Computer Vision and Pattern Recognition 2018: 6733-6741. DOI: 10.1109/CVPR.2018.00704.
[31] Chen Z, Sun K, Yang F, Tao W. SC$^2$-PCR: A second order spatial compatibility for efficient and robust point cloud registration. 2022 IEEE/CVF Conf on Computer Vision and Pattern Recognition (CVPR) 2022: 13211-13221. DOI: 10.1109/CVPR52688.2022.01287.

### *Authors' information*

**Shuxin Liu** (b. 1977) is currently an Associate Professor at Shanghai University of Electric Power. He received his PhD from East China Normal University in September 2012, after pursuing his doctoral degree there from 2005 to 2012. Prior to that, he studied for his master's degree at Donghua University of Technology from 2003 to 2005. From 2006 to 2020, Dr. Liu worked at Minnan Normal University, starting as a teaching assistant and later serving as a lecturer and associate professor. Since December 2020, he has taken up the position of Associate Professor at Shanghai University of Electric Power, where he also serves as a graduate student supervisor. For over ten years, Dr. Liu's research has focused on key technologies in the fields of computer vision and artificial intelligence. He has undertaken and participated in more than ten science and technology projects sponsored by national, provincial and municipal funds. E-mail: *liusx@sdju.edu.cn*

**Guanjun Ji** (b. 2000) received his bachelor's degree from Hebei Normal University of Science and Technology and his master's degree from Shanghai Institute of Electrical Technology, where his research interests include 3D reconstruction and object detection. E-mail: *226001010507@st.sdju.edu.cn*

**Chengcheng Shi** (b. 1995) received his Bachelor's degree from Chongqing University of Science and Technology in 2020 and is currently pursuing his Master's degree in Electrical Engineering at Shanghai Dianji University. His research interests include deep learning and its applications to computervision,and human action recognition.
E-mail: *shicc1995@163.com*

*Received September 21, 2024. The final version – December 03, 2024.*