ITC 3/55
Information Technology and Control
Vol. 55 / No. 3/ 2026
pp. 1052-1073
DOI 10.5755/j01.itc.55.3.44470

Laser Radar 3D Target Recognition by Integrating Point Distribution Feature Descriptors and ISS Algorithm

Received 2026/03/05 | Accepted after revision 2026/06/03

HOW TO CITE: Li, X., Yang, J. (2026) Laser Radar 3D Target Recognition by Integrating Point Distribution Feature Descriptors and ISS Algorithm. *Information Technology and Control*, 55(3), 1049-1070. https://doi.org/10.5755/j01.itc.55.3.44470

Laser Radar 3D Target Recognition by Integrating Point Distribution Feature Descriptors and ISS Algorithm

Xihan Li

College of Land Science and Technology, China Agricultural University, Beijing, 100193, China; e-mail: xihan0316@sina.com

Jianyu Yang*

College of Land Science and Technology, China Agricultural University, Beijing, 100193, China;

Key Laboratory of Remote Sensing of Agricultural Disasters, Ministry of Agriculture and Rural Affairs, Beijing 100193, China; e-mail: JianyuYangjy@outlook.com

Corresponding author: JianyuYangjy@outlook.com

This study proposed a laser radar 3D target recognition method that integrates point distribution feature descriptors and intrinsic shape feature algorithms, aiming to solve the problems of insufficient robustness and low computational efficiency in point cloud recognition in complex scenes. This method constructed a multi-level feature representation with strong invariance to rotation, translation, and noise by combining globally aligned spatial distribution descriptors and point distribution feature descriptors. It utilized the intrinsic shape feature algorithm to stably extract feature points, and introduces improved region growth segmentation and Iterative Closest Point (ICP) registration algorithms to enhance recognition performance. Experiments on two public datasets showed that the research method achieved feature extraction accuracies of 93.68% and 91.46%, both of which were superior to existing mainstream descriptors. The proposed descriptor exhibited superior robustness with an SNR of 19.46 dB and an occlusion rate of 82.75% in noisy environments. In addition, the average feature extraction time was 52.37 ms, which improved efficiency by over 47% compared to methods such as rotational projection statistical feature extraction. In two real-world datasets, the proposed method achieved recognition rates of 89.86% and 80.72% when the noise deviation was 0.8 m, and 88.76% and 80.78% under low resolution conditions. Its recognition performance outper-

formed the compared methods. The results indicate that the research method has excellent performance in accuracy, robustness, and efficiency, providing an effective technical means for 3D object recognition in complex environments. It also has strong practical application value.

KEYWORDS: Intrinsic shape feature algorithm; Lidar; Feature descriptor; 3D object recognition; Point cloud.

1. Overview

In today's era of rapid technological development, LiDAR technology provides rich spatial information for object detection and recognition due to its high precision, high resolution, and ability to obtain rich 3D information in complex environments. LiDAR has been extensively utilized in many fields like autonomous driving, robot navigation, industrial automation, geographic information mapping, etc. [21]. Among them, 3D Object Recognition (3DOR), as a key link in LiDAR data processing, is of great significance for achieving intelligent decision-making and automated operations [7]. However, due to the complexity of LiDAR data, Point Cloud (PC) data suffers from factors such as viewpoint changes, occlusion, noise, and uneven density, resulting in poor accuracy and efficiency of 3DOR [23]. Therefore, achieving efficient and accurate 3DOR has become a hotspot in research. In recent years, the development of feature-based recognition methods has provided new ideas for solving the challenges in 3DOR [20]. In 3DOR, Feature Point Detection (FPD) and feature description are the two key factors in building robust recognition systems [6].

FPD can extract representative, informative, and noise and occlusion insensitive feature points from the original PC, reducing processing complexity [12]. Numerous scholars have researched it. Cui et al. designed a novel feature point representation method to address the problems of poor descriptive ability and low efficiency in 3D LiDAR PC feature matching and extraction. This method utilized the sparsity and complexity of PCs to provide strong constraints for describing feature points. This method only took about 20 ms to match two frames of PCs and had excellent real-time performance [4]. Romero Gonzalez et al. put forth a local descriptor and FPD built on 3D binary mode to address the high computational cost in 3D PC processing. This method replaced real valued calculations with binary encoding, reducing computational complexity. This method could maintain comparable performance to advanced methods while significantly improving processing efficiency [19]. Nagar proposed a reflection symmetry plane method based on statistical estimation, which transforms the problem into a fast convergence optimization problem by combining FPD with kernel-based feature matching, effectively decoupling the symmetry plane solution from the correspondence between feature points. This method achieved an average error comparable to advanced methods on the benchmark dataset and improved F1 scores by 4.5% [16]. Zhou et al. proposed a registration algorithm that combines Intrinsic Shape Signature (ISS) FPD with 3D shape context descriptors to address the issues of large data volume, time-consuming, and low accuracy in PC registration. This method extracted key points through ISS and integrated improved random sample consistency with iterative closest point algorithm to achieve coarse to fine matching. This algorithm significantly improved registration accuracy and efficiency [35]. Xiong et al. proposed a LiDAR exclusion-based 3D object detection method that combines 4D imaging radar with a camera. In the LiDAR exclusion tilt model, predicted image depth distribution maps and radar 3D occupancy grids are generated from image perspective view features and radar bird's-eye view features, respectively. The sampling strategy is optimized by introducing image depth and radar information. The results indicated that the proposed method was significantly superior to the current state-of-the-art 3D object detection methods [26]. Xu et al. proposed a novel 3D object detection model based on LiDAR. This model learns prior knowledge of object shape and estimates the complete shape of partially occluded objects in the PC. Integrate occupancy estimation into the proposal refinement module to generate accurate bounding boxes. It was found that the performance of the model in 3D detection of cars and cyclists far exceeded existing advanced methods [27].

In addition, recognition methods based on feature descriptors have received widespread attention due to their ability to effectively extract salient features of targets, thereby achieving fast and accurate recognition of targets in complex environments [18]. Liu et al. proposed an anisotropic wavelet energy decomposition descriptor that utilizes multi-scale anisotropic spectral manifold wavelet decomposition energy to collect local energy at each vertex to simultaneously encode multi-scale internal and external features. This descriptor was more discriminative and robust than the current advanced descriptors, and significantly improved feature recognition and intrinsic symmetry recognition capabilities [13]. Poiesi and Boscaini proposed a unique and universal deep learning descriptor to address the weak generalization ability and transformation sensitivity of 3D local descriptors. This method utilized deep neural networks to encode normalized PC patches, generating compact features that remain invariant to scale, rotation, and point arrangement. This descriptor far exceeded existing methods in terms of cross-domain generalization ability [17]. Yang et al. proposed a feature descriptor built on spatial voxel homogenization to address the issue of insufficient robustness in 3D feature description under complex interference. This method enhanced stability by improving local reference frame construction and encoded spatial features using voxel homogenization segmentation strategy. This descriptor had good descriptive performance on multiple public datasets [32]. Wang et al. developed a new framework that fully utilizes directional descriptors and local rotation estimation. This method estimated local rotation by designing a novel orientation descriptor, and developed a rotation guided detector, matching device, and random sample consistency optimization process. The framework achieved the most advanced performance on multiple datasets [24]. Meng et al. proposed a hybrid object detection and tracking method based on 3D LiDAR, which combines object detection features from state-of-the-art object detection algorithms with historical information from object tracking algorithms to infer objects. The results showed that the proposed method could effectively utilize target tracking information and achieve excellent object detection performance, with an AP 3.7% higher than current leading technologies [14].

In short, many scholars have deeply studied the 3D PC feature extraction and matching methods and achieved significant results. Nevertheless, existing 3DOR approaches still have problems including insufficient robustness, low computational efficiency, and insufficient utilization of local structural information when dealing with complex scenes. They ignore local feature description and extraction, causing a decrease in recognition precision. Given this, this study proposes a LiDAR 3DOR method that integrates Point Distribution Signature Histogram (PDSH) descriptors with ISS. The PDSH descriptor achieves rotation and translation invariance by constructing a local reference framework, and calculates the distribution information of point projections on three coordinate planes to form highly discriminative local feature vectors. It combines the Global Aligned Spatial Distribution (GASD) descriptor to extract global and local features of 3D objects. The ISS is utilized to detect feature points in PCs and combine it with an improved Region Growth Segmentation Algorithm (RGSA) to segment the scene PC. The ICP algorithm is used for fine registration to improve the efficiency and accuracy of 3DOR. The unique contributions of this study lie in three aspects. First, a joint optimization mechanism has been established between global and local descriptors, enabling iterative cross-refinement rather than simple feature concatenation. Second, an adaptive RGSA guided by ISS keypoints has been designed, where segmentation thresholds are dynamically adjusted based on local geometric constraints. Third, a feedback loop from feature matching residuals to ICP registration has been constructed, allowing closed-loop refinement between descriptor similarity and spatial alignment.

2. Methodology

The paper proposes a 3D Object Feature Extraction (OFE) method based on GASD and PDSH descriptors, targeting to simultaneously capture the global structural information and local detail features of the target, and improve the discriminative and robust feature expression. Then, a 3DOR method combining feature descriptors and ISS FPD algorithm is proposed to achieve efficient and accurate target recognition in complex scenes.

2.1. 3D OFE Method Based on GASD and PDSH Descriptors

In 3DOR tasks, feature extraction is a crucial step in determining recognition performance. Traditional methods often rely on global features or single type local features, making it difficult to balance robustness and discriminability in complex scenarios. Especially in LiDAR PC, traditional feature descriptors often perform poorly due to factors including viewpoint changes, occlusion, noise, and uneven point density distribution [25]. Therefore, this paper constructs a 3D OFE method that combines GASD descriptor and PDSH descriptor, aiming to simultaneously capture the global structural information and local detail features of the target, and improve the robustness of the recognition system.

In PC feature description, Point Feature Histogram (PFH) and Fast PFH (FPFH) can fully express the feature information of LiDAR PCs, providing discriminative information for subsequent tasks like PC registration and target recognition [5]. This method calculates the feature angle by describing the geometric relationship between the query point and its neighboring points, as shown in Equation (1).

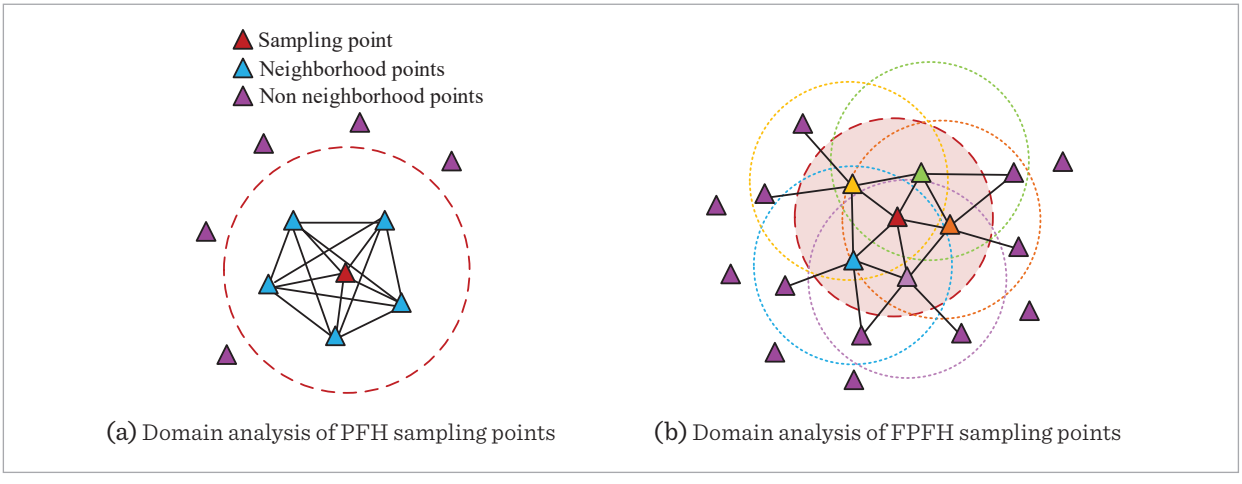
$$\alpha = \arctan\left(\frac{\mathbf{w}\cdot\mathbf{n}_i}{\mathbf{u}\cdot\mathbf{n}_i}\right). \tag{1}$$

In Equation (1), α is the characteristic angle. $\mathbf{u}$ and $\mathbf{w}$ are the basis vectors of the local Coordinate System (CS) at the query point. $\mathbf{n}_i$ is the normal vector of the i-th point in the neighborhood. The domain analysis of sampling points for PFH and FPFH is shown in Figure 1.

Figure 1(a) shows domain analysis of PFH sampling points, which calculates the relationships between all neighboring point pairs using a fully connected method. Figure 1(b) shows the sampling point domain diagram of FPFH, which simplifies the calculation strategy to improve efficiency and reflects the different strategies of two descriptors in processing the local geometric structure of PCs. In global feature extraction, this study uses GASD descriptors to process PC captured from different perspectives to ensure that the extracted global features are not affected by transformations. In local feature extraction, to ensure the invariance of local surface translation and rotation of PCs, this study constructs a local reference framework to achieve rotation and 3D PC projection on its coordinate plane. The GASD descriptor aligns PC collected from different perspectives to extract their spatial distribution features, thereby achieving effective description of the global structure of the target [34]. The calculation of its center of gravity coordinates and centralization processing is shown in Equation (2).

$$\begin{cases} \mu = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} p_i \\ p_i' = p_i - \mu \end{cases}. \tag{2}$$

Figure 1
Domain analysis of sampling points for PFH and FPFH.



(a) Domain analysis of PFH sampling points

(b) Domain analysis of FPFH sampling points

In Equation (2), μ is the centroid coordinate of the PC. N is the gross number of points in the PC. p_i means the coordinate of the i-th point. p_i' is the new p_i after centralization processing. To accurately describe the global spatial distribution of PCs, this article captures the distribution differences of PCs in various directions by calculating the Covariance Matrix (CM) of the centralized PC, as shown in Equation (3).

$$C = \frac{1}{N}\sum_{i=1}^{N}(p_i')(p_i')^T. \tag{3}$$

In Equation (3), C is the CM of the PC. $(p_i')(p_i')^T$ is the outer product of column vectors and row vectors. After establishing a global CS, it is necessary to project the PC onto the CS and quantify its spatial distribution. The 3D space is segmented into uniform voxel grids, and the number of points within each voxel is counted to form a feature histogram that describes the global spatial distribution. The formula of GASD feature vectors is given by Equation (4) [3].

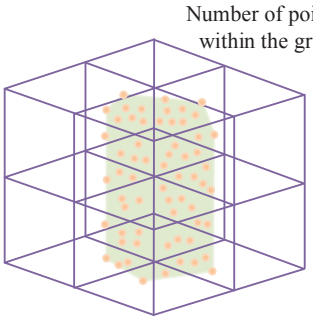
$$H_{\text{GASD}}(k) = \sum_{i=1}^{N} I(p_i' \in V_k). \tag{4}$$

In Equation (4), $H_{\text{GASD}}(k)$ is the value of the k-th histogram bucket of the GASD feature vector. V_k is the k-th voxel grid. $I(p_i' \in V_k)$ is the indicator function. Based on this global distribution, local features are further extracted. To enhance the discriminative power of local features, this study proposes a PDSH descriptor, whose core is to construct a local reference framework and calculate the distribution histograms of projection points on three coordinate planes. By calculating the CM of the neighborhood and capturing the distribution characteristics of the neighborhood point set, the calculation of the local CM is shown in Equation (5).

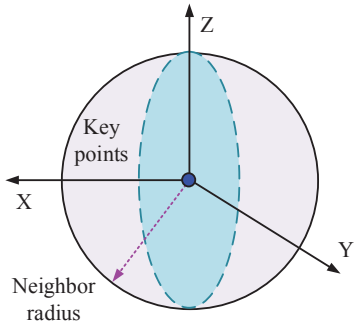
$$\begin{cases} N(p) = \left\{ q_j \middle| \left\| q_j - p \right\|_2 \le r \right\} \\ C_{\text{local}} = \dfrac{1}{m}\sum_{j=1}^{m}(q_j - \mu_Q)(q_j - \mu_Q)^T \end{cases}. \tag{5}$$

In Equation (5), $N(p)$ means a set of local neighborhood points. p is the feature point currently being processed. q_j is the j-th point in the neighborhood.

Figure 2
Schematic diagram of GASD features and local reference framework.



(a) Schematic diagram of GASD features



(b) Schematic diagram of local reference framework

r is a pre-set neighborhood radius. C_{local} is the local neighborhood CM of point p. m and μ_Q denote the number of midpoints and centroid coordinate of the local neighborhood. The GASD features and local reference framework are shown in Figure 2.

Figure 2(a) is a schematic diagram of GASD features, which quantifies the spatial distribution of PCs through voxel grids. The points within each grid are counted and used to construct a feature histogram that describes the global structure, reflecting the whole distribution features of the PC in 3D space. Figure 2(b) determines the principal direction by calculating the CM of the local neighborhood, establishes a local CS with rotational invariance, and provides a stable reference benchmark for subsequent local feature projection and statistics. Furthermore, the inherent consistency between the proposed multi-scale feature fusion and radar weak target recognition mechanisms has been established, showing that weak geometric features are effectively amplified through comple-

mentary global-local representation [29]. After constructing a local reference framework, this projection method not only reduces computational complexity but also preserves sufficient geometric information by statistically analyzing the distribution information of projection points [2]. To quantify the distribution characteristics of projection points, it is necessary to divide each projection plane into regular grid regions [9]. The calculation of the distribution histogram of the projection points is shown in Equation (6).

$$\begin{cases} q_j^{xy} = (q_j \cdot u_1, q_j \cdot u_2) \\ H^{xy}(b) = \sum_{j=1}^{m} I(q_j^{xy} \in G_b) \end{cases}. \tag{6}$$

In Equation (6), q_j^{xy} is the 2D coordinate of the neighboring point projected onto the plane. u_1 and u_2 are the basis vectors of the local reference frame. $H^{xy}(b)$ is the value of the b-th bin of the histogram on the XY plane. G_b is the b-th grid region on the XY plane. Finally, the distribution histograms of the three projection planes are concatenated into a complete feature vector, as shown in Equation (7).

$$\begin{cases} H_{\text{PSDH}} = \left[H^{xy}, H^{yz}, H^{zx}\right] \\ \hat{H}_{\text{PSDH}}(e) = H_{\text{PSDH}}(e) / \sum_j H_{\text{PSDH}}(j) \end{cases}. \tag{7}$$

In Equation (7), H_{PSDH} is the final PDSH feature vector. H^{xy} is the distribution histogram on the XY plane. H^{yz} and H^{zx} are distribution histograms on the YZ and ZX planes. $\hat{H}_{\text{PSDH}}(e)$ and $H_{\text{PSDH}}(e)$ are the value of the th element of the normalized and original PDSH feature vectors. $\sum_j H_{\text{PSDH}}(j)$ is the sum of all elements of the original feature vector. e is the index of the feature vector. From an information-theoretic perspective, the concatenated triple-plane histogram has been analyzed. The joint entropy is increased compared to a single plane due to complementary spatial information. Feature dimension (192 bins) has been selected to balance entropy maximization and computational cost, avoiding the curse of dimensionality where marginal discriminative gain diminishes rapidly. Thus, near-optimal discriminative ability per dimension is achieved. To improve the adaptability of descriptors to scale changes, this study scales the projection coordinates of neighboring points according to the search radius [10]. This study uses scale normalization to process PCs with different sampling densities, as given by Equation (8).

$$(x_n', y_n') = \left(\frac{p' \cdot u_1}{r'}, \frac{p' \cdot u_2}{r'} \right). \tag{8}$$

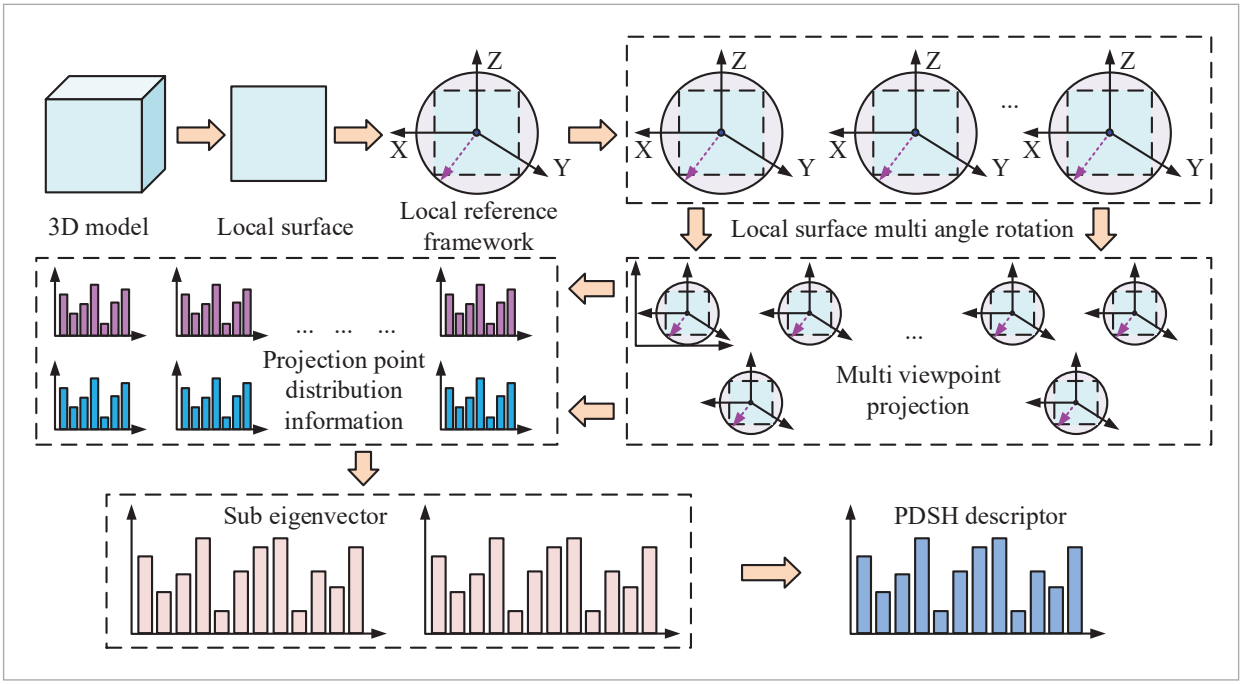
In Equation (8), (x_n', y_n') is the normalized projection coordinate of neighboring point p' in constructing a local reference frame. r' is the search radius. Compared to Rotational Projection Statistics (RoPS), three substantial improvements have been made in the proposed PDSH descriptor. First, a scale-normalized projection strategy based on the adaptive search radius has been introduced, by which improved robustness to varying PC densities is achieved. Second, a lightweight histogram concatenation scheme with fixed bin allocation has been designeds. Third, the combination with GASD global descriptors has been established, by which discriminability under occlusion and noise is strengthened without increasing local computational overhead. The construction process of PDSH descriptor is shown in Figure 3.

In Figure 3, the construction of the PDSH descriptor begins with the extraction of the local surface of the 3D model, and rotation invariance is achieved by establishing a local reference frame. Subsequently, multi-view projection is performed to project neighboring points onto three coordinate planes, and the distribution information of the projected points is statistically analyzed. These distribution histograms are ultimately concatenated to form a discriminative PDSH descriptor.

2.2. 3DOR method Integrating Feature Descriptors and ISS FPD Algorithm

After extracting 3D target features through GASD and PDSH descriptors in this study, further efficient and accurate recognition is needed. Traditional methods rely on a single type of feature or global information, making it difficult to balance recognition accuracy and robustness in complex scenarios. The ISS algorithm can stably extract feature points with strong invariance to rotation, translation, and noise by analyzing the intrinsic characteristics of 3D shapes in the local neighborhood of the PC, to improve the accuracy of 3DOR [33]. Therefore, this study proposes a 3DOR

Figure 3
The construction process of PDSH descriptor.



method that integrates feature descriptors with ISS FPD algorithm, aiming to improve recognition robustness and efficiency in complex scenes by combining the advantages of local and global features.

The ISS algorithm selects points with significant changes in different directions as feature points by calculating the CM eigenvalues of each point in its local neighborhood [1]. The algorithm first establishes a local CS for the points in the PC, searches for points with radius r' centered on point p_g. The weight assigned to each neighboring point is computed by Equation (9).

$$w_{gh} = \frac{1}{\left\| p_g - p_h \right\|}, \left| p_g - p_h \right| < r' . \tag{9}$$

In Equation (9), w_{gh} is the weight of point p_g and point p_h. Then, the CM of the points is calculated to capture the trend of local surface changes, as shown in equation (10).

$$\mathrm{cov}(p_g) = \frac{\sum_{|p_g - p_h| < r'} w_{gh} (p_g - p_h)(p_g - p_h)^T}{\sum_{|p_g - p_h| < r'} w_{gh}} . \tag{10}$$

In Equation (10), p_g is the CM of the point. The characteristic value $\lambda_1 \geq \lambda_2 \geq \lambda_3$ corresponds to the degree of change in the three main directions, as shown in Equation (11).

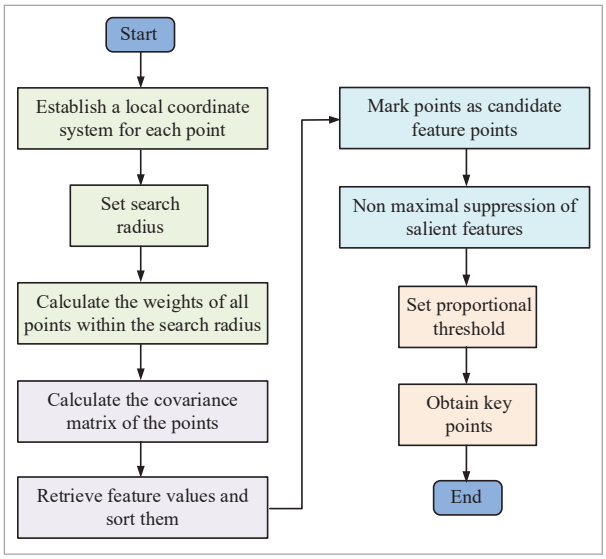
$$\begin{cases} (\lambda_2 / \lambda_1) \leq \tau_{12} \\ (\lambda_3 / \lambda_2) \leq \tau_{23} \end{cases} . \tag{11}$$

In Equation (11), λ_1, λ_2, and λ_3 are eigenvalues. τ_{12} and τ_{23} are preset thresholds.

After keypoint detection, local descriptors are constructed. The flowchart of ISS FPD is shown in Figure 4.

In Figure 4, the ISS starts by establishing a local CS for each point, constructing a neighborhood by setting a search radius, and calculating the weights and covariance matrices of points within the neighborhood. Then, it decomposes and sorts the eigenvalues to obtain the characteristic values that represent the local shape changes. To extract candidate target regions from complex scenes, this study improves RGSA and proposes an adaptive RGSA [28]. This algorithm can better handle the problems of uneven PC density and noise interference by dynamically adjusting the growth threshold and termination conditions.

Figure 4
Schematic diagram of ISS FPD process.



The core lies in combining geometric distance and normal vector differences to achieve more accurate region partitioning, as shown in Equation (12).

$$\begin{cases} \left\| q_j - s \right\|_2 < d_{\max} \\ \angle(n_s, n_{q_j}) < \theta_{\max} \end{cases}. \tag{12}$$

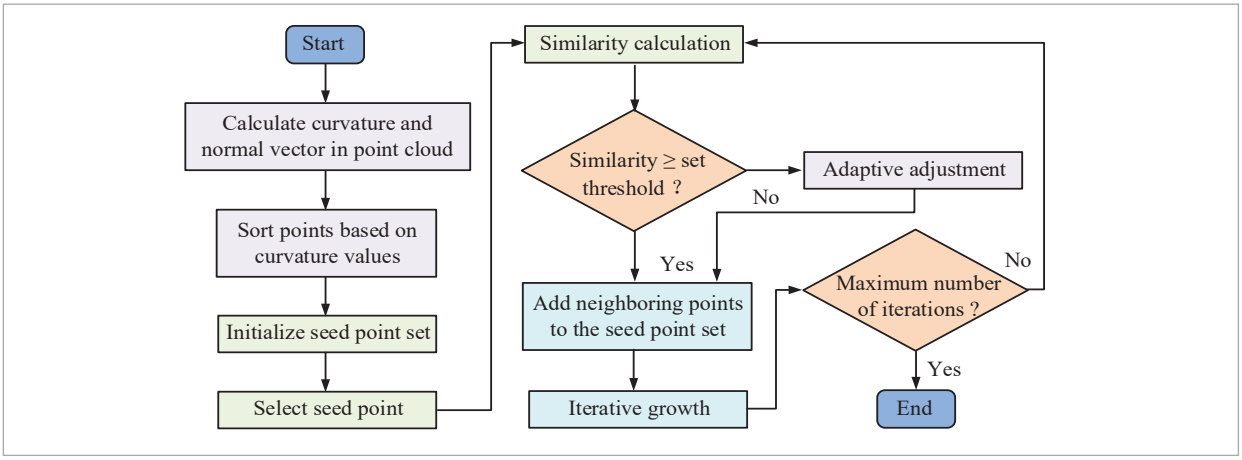
In Equation (12), s is the starting point of the current growth region. $d_{\max}$ is the maximum distance threshold. n_s and n_{q_j} are the normal vectors of the starting and neighboring points in the growth region. $\angle(n_s, n_{q_j})$ is the angle between normal n_s and n_{q_j}. $\theta_{\max}$ is the maximum angle threshold. The termination conditions for regional growth are shown in Equation (13).

$$\begin{cases} |R_i| < N_{\min} \\ \max\limits_{q \in R_i} \left\| a - \mu_R \right\|_2 > \sigma_{\max} \end{cases}. \tag{13}$$

In Equation (13), $|R_i|$ is the number of points contained in the current region R_i. $N_{\min}$ is the minimum point threshold. a is any point in region R_i. μ_R is the center of gravity of region R_i. $\sigma_{\max}$ is the maximum dispersion threshold. This mechanism demonstrates how segmentation robustness can be enhanced against noise and outliers, providing a theoretical foundation for the adaptive threshold adjustment strategy in the proposed method [29]. The improved RGSA process is shown in Figure 5.

In Figure 5, the improved RGSA first calculates the curvature and normal vector of each point in PC, and then sorts the points based on their curvature values. The core steps are similarity calculation and threshold judgment, dynamically controlling the growth process through adaptive adjustment strategies. The iterative growth process will continuously merge neighboring points that meet similarity criteria into the region until the maximum iteration termination condition is met. After obtaining the segmented candidate regions and their feature descriptors, the target identity needs to be determined through fea-

Figure 5
Flow diagram of improved RGSA.



ture matching. This study uses chi square distance to evaluate the differences between two descriptors. The smaller the value, the more similar the two features are, and the higher the likelihood of matching [15]. Then, the Nearest Neighbor Distance Ratio is used to filter the matches. The similarity between two feature histograms is measured by Equation (14).

$$\begin{cases} D_{\chi^2}(H_1,H_2)=\sum_{b=1}^{B}\dfrac{(H_1(b)-H_2(b))^2}{H_1(b)+H_2(b)} \\ \text{NNDR}=\dfrac{D_1}{D_2}<\gamma \end{cases}. \quad (14)$$

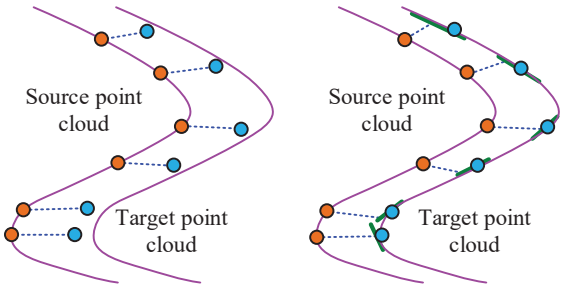
In Equation (14), $D_{\chi^2}(H_1,H_2)$ is the chi square distance between two histograms H_1 and H_2. b is the bin index of the histogram. B is the total number of bins in the histogram. $H_1(b)$ and $H_2(b)$ are the values of histogram H_1 and H_2 at bin b. D_1 and D_2 are both the chi square distance between the current feature point and its nearest neighbor feature point. γ is the threshold for the NNDR.

To eliminate potential redundancy, this study implemented a cascaded fusion strategy to decouple global and local features in the distance metric space [31, 11]. The system no longer uses joint vectors, but instead utilizes weighted chi square distances, as shown in Equation (15).

$$D_{\text{total}}=\alpha\cdot D_{\text{GASD}}+(1-\alpha)\cdot D_{\text{PDSH}}. \quad (15)$$

In Equation (15), D_{total} represents the total weighted distance. α represents the weight coefficient. D_{GASD} and D_{PDSH} respectively represent the global feature distance and local feature distance. Among them, α is dynamically set based on local point density to prioritize global features in sparse areas. This mechanism ensures that GASD constrains the search manifold, while PDSH achieves precise matching, theoretically eliminating information overlap and optimizing the computational process. For candidate targets that have been preliminarily matched successfully, it is necessary to further improve the estimation accuracy through refined registration. The ICP algorithm optimizes the rotation and translation parameters iteratively to minimize the distance error between the corresponding points of the source PC and the target PC [8]. The ICP algorithm includes two methods: point-to-point (P2P) and point-to-plane (P2Pl). The P2P ICP algorithm solves the optimal transformation by minimizing the Euclidean distance between corresponding points, which is simple to calculate but easily affected by noise and outliers. The P2Pl ICP algorithm achieves registration by minimizing the distance from the source point to the plane where the target point is located, which has better robustness to noise, as displayed in Figure 6.

Figure 6
Schematic diagram of ICP algorithm based on P2P and P2Pl.



(a) Schematic diagram of point to point ICP algorithm

(b) Schematic diagram of ICP algorithm based on point to plane

Figure 6(a) shows the P2P ICP algorithm, which iteratively optimizes the rotation and translation parameters to minimize the overall distance error, and achieves PC alignment. Figure 6(b) shows the P2Pl ICP algorithm, which optimizes the transformation parameters by minimizing the vertical distance from the source point to the tangent plane where the target point is located, reducing the influence of noise and outliers. The error function of P2P ICP is to minimize the sum of squared Euclidean distances between corresponding points. The error function of P2Pl ICP is to minimize the distance from the source point to the target point tangent plane, as shown in Equation (16) [22].

$$\begin{cases} E_{\text{point}}(R_o,t)=\sum_{i=1}^{N}\left\|R_o\cdot p_i+t-q_i\right\| \\ E_{\text{plane}}(R_o,t)=\sum_{i=1}^{N}\left[(R_o\cdot p_i+t-q_i)\cdot n_i\right]^2 \end{cases}. \quad (16)$$

In Equation (16), $E_{\text{point}}(R_o,t)$ and $E_{\text{plane}}(R_o,t)$ are the error function of P2P ICP and P2Pl ICP. $R_o \cdot p_i + t$ is the coordinate obtained by applying rotation R_o and translation t to point p_i. n_i is the normal vector at point q_i. To characterize the impact of matching inaccuracy, registration error is modeled as a function of the initial corresponding set. The total propagation error is shown in Equation (17).

$$\boldsymbol{E}(\boldsymbol{R},\boldsymbol{t}) = \sum_{(p_i,q_i)} w_i \left\| \boldsymbol{R} p_i + t - q_i \right\|^2 + \eta(C_{\text{false}}). \tag{17}$$

In Equation (17), $\boldsymbol{E}(\boldsymbol{R},\boldsymbol{t})$ represents the total registration error. $\boldsymbol{R}$ and $\boldsymbol{t}$ represent the rotation matrix and translation vector, respectively. p_i and q_i respectively represent the source PC and the target PC. w_i represents the weight coefficient. $\eta(C_{\text{false}})$ represents the error disturbance term. C_{false} represents a false corresponding set. False correspondences introduce non-Gaussian noise into the objective function, shifting the global minimum and causing the ICP algorithm to converge toward suboptimal local minima. This results in a persistent rotational bias and translation drift. Theoretically, if the outlier ratio exceeds the algorithm's breakdown point, the quadratic convergence is disrupted, leading to increased iterations or complete divergence, thereby significantly degrading the final registration accuracy and structural fidelity.

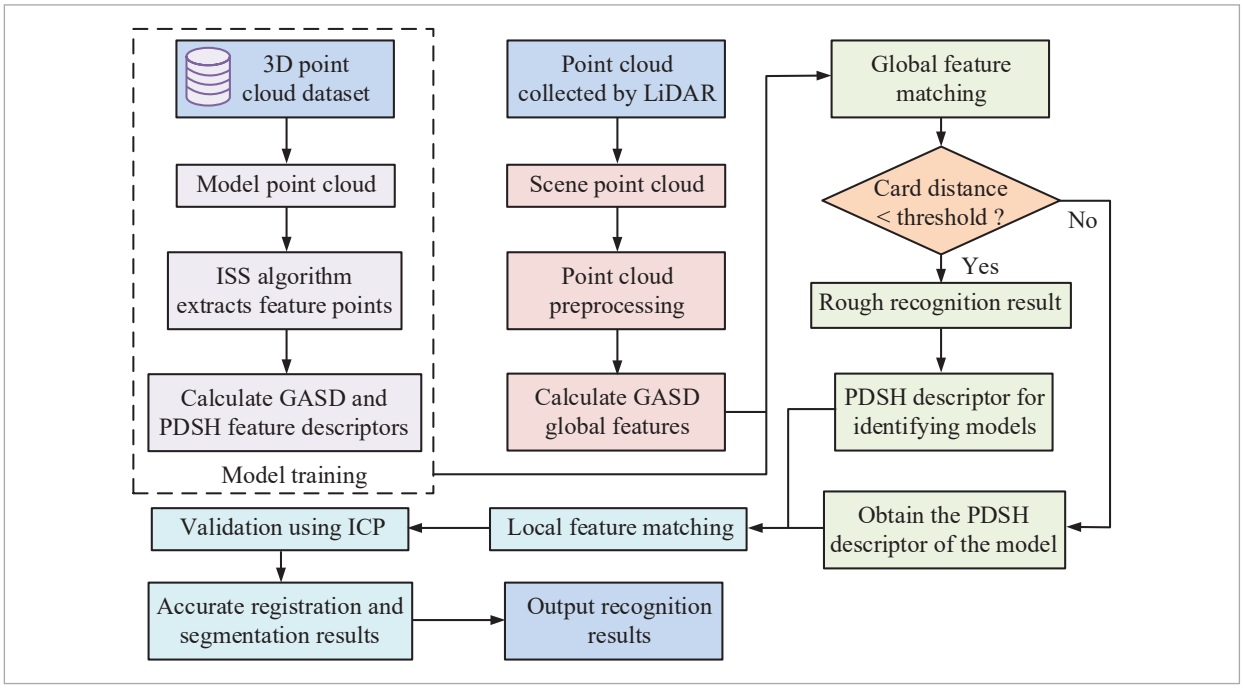
The final recognition result is output through refined registration and segmentation. The flowchart of LiDAR 3DOR is shown in Figure 7.

In Figure 7, the process has two main lines: model training and scene recognition. The ISS keypoint detector provides geometrically stable centers for constructing the PDSH local descriptor, while the GASD global descriptor guides the adaptive initialization of the improved RGSA by aligning scene and model centroids. Furthermore, the feature matching residuals from PDSH are fed back as a weight term into the ICP fine registration error function, enabling iterative cross-refinement between descriptor similarity and spatial alignment. This closed-loop coupling ensures that global structure, local geometry, and registration quality are jointly optimized rather than processed serially.

The pseudocode of the proposed method is as follows:

Figure 7

Schematic diagram of the process of 3DOR using LiDAR.



Algorithm: LiDAR 3DOR using PDSH descriptors and ISS algorithmInput: Scene point cloud P_scene, Model point cloud P_modelOutput: Recognition result (matched object identity)

// Global feature extraction
Compute centroid μ of P_scene and P_model using Eq. (2)
Compute covariance matrix C using Eq. (3)
Construct GASD feature vectors H_GASD for both clouds using Eq. (4)
// ISS keypoint detection
For each point p in P_scene and P_model:
    Compute weights w_gh using Eq. (9)
    Compute covariance matrix cov(p_g) using Eq. (10)
    Compute eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ using Eq. (11)
    Select keypoints where $\lambda_2/\lambda_1 \leq \tau_{12}$ and $\lambda_3/\lambda_2 \leq \tau_{23}$
Apply non-maximum suppression to obtain final keypoints K
// PDSH local descriptor construction
For each keypoint k in K:
    Extract local neighborhood N(k) within radius r using Eq. (5)
    Build local reference frame (u_1, u_2, u_3)
    Project neighborhood points onto XY, YZ, ZX planes using Eq. (6)
    Compute distribution histograms Hˆxy, Hˆyz, Hˆzx
    Concatenate histograms to form PDSH vector H_PSDH using Eq. (7)
// Feature matching
Compute chi-square distance $D_\chi^2(H_1, H_2)$ between descriptors using Eq. (14)
Filter matches using NNDR with threshold γ
Compute total weighted distance D_total = $\alpha \cdot$ D_GASD + (1-α)$\cdot$ D_PDSH using Eq. (15)
// Adaptive region growing segmentation
Sort points by curvature in ascending order
For each unvisited seed point:
    Initialize an empty region R
    Repeat (region growing):
        Add neighboring points that satisfy:
            distance < d_max and angle difference < θ_max // from Eq. (12)
    Until no more points can be added
    If R < N_min then:
        Discard region R as too small
    Else:
        If dispersion of R > σ_max then:
            Discard region R as excessively dispersed
    Else:
        Keep region R as a valid segment
// ICP fine registration
Initialize transformation T_0 from best matching candidate
Repeat until convergence:
    Find correspondences between source and target
    Minimize error function E_point(R,t) or E_plane(R,t) using Eq. (16)
    Update transformation R, t
Output recognition result with minimized registration error

3. Results and Analysis

The performance of the constructed 3D OFE method was validated, and the recognition results of the method were analyzed to verify its robustness and effectiveness in different scenarios.

3.1. Performance Validation of 3D OFE Methods

To verify the performance of the proposed descriptor-based 3D OFE method, this study conducted experimental evaluations on two common PC sets, KITTI and Bologna 3D Retrieval (B3R). Among them, the KITTI dataset was collected from real road environments, including multiple targets such as vehicles, pedestrians, bicycles, etc. The PC density is uneven, there are occlusions and noise, and it is suitable for evaluating the robustness of descriptors in complex scenes. The B3R dataset contains various 3D object models with different levels of complexity, suitable for testing the discriminative ability of descriptors under different geometric structures. The reference/modl data consisted of complete 3D object models extracted from each dataset. The test data comprised independent scene PC frames containing the same target classes. Ten-fold cross-validation was performed solely for statistical evaluation, without any training. No frame-level data leakage occurred, as PCs from the same LiDAR sweep or adjacent frames were assigned to the same fold. For reproducibility, ten-fold cross-validation was performed on both public datasets to ensure statistical reliability. Specifically, each dataset was randomly divided into 10 equal-sized subsets; nine subsets were used for feature extraction and matching validation, and the remaining one subset was used for testing. This process was repeated ten times, and the average performance metrics were reported. For the self-collected UrbanRoad and TunnelLite datasets, the same ten-fold cross-validation protocol was adopted. The experiment compared the proposed descriptor with current mainstream PC feature descriptors, including FPFH, RoPS, Signature of Histograms of OrienTations (SHOT), and Normal Aligned Radial Features (NARF). For noise robustness evaluation, Gaussian noise with zero mean and standard deviation σ (in meters) was added to the PC coordinates. The noise deviation values tested ranged from 0.2 m to 0.8 m. Table 1 lists specific environmental configurations.

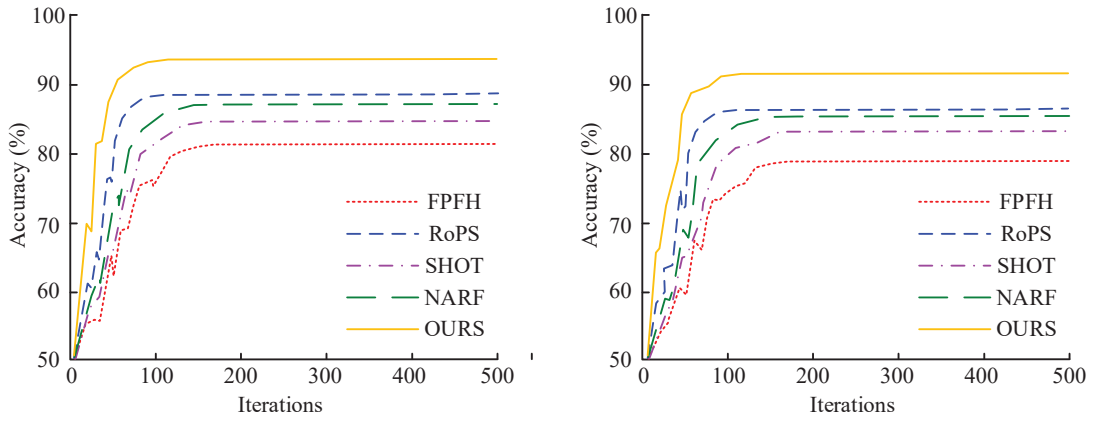
This study first compared the feature extraction accuracy of various descriptors in the KITTI and B3R datasets to validate the effectiveness of the proposed method. Iteration number refers to the number of repeated independent random trials, not ICP or matching iterations. The reported accuracy at each iteration represents the cumulative average over all completed trials. In Figure 8(a), in the KITTI dataset, when the iteration number was 50, the feature extraction accuracies of FPFH, RoPS, SHOT, and NARF were 68.24%, 85.15%, 73.28%, and 74.36%, and the accuracy of the proposed descriptor was 91.12%. When iterating to 200, the accuracy of FPFH, RoPS, SHOT, NARF, and the proposed descriptor was 82.31%, 88.72%, 85.64%, 87.36%, and 93.68%. In Figure 8(b), in the B3R dataset, when the number of iterations was 200, the feature extraction accuracies of the five descriptors were 78.52%, 86.35%, 83.72%, 85.05%, and 91.46%. The results showed that the proposed descriptor could more

Table 1
Experimental environment and parameter settings.

Environment		Parameters	
Processor	Intel Core i9-12900K	Neighbor radius	0.05 m
Memory	64 GB DDR4	Voxel grid size	0.02 m
Graphics card	NVIDIA GeForce RTX 4090	Number of projected grids	8 × 8
Operating system	Ubuntu 20.04 LTS	Feature dimension (GASD)	512
Programming language	Python 3.8	Feature dimension (PDSH)	192
PC library	Open3D 0.15	ISS threshold	0.75

Figure 8

Comparison results of feature extraction accuracy for different descriptors.

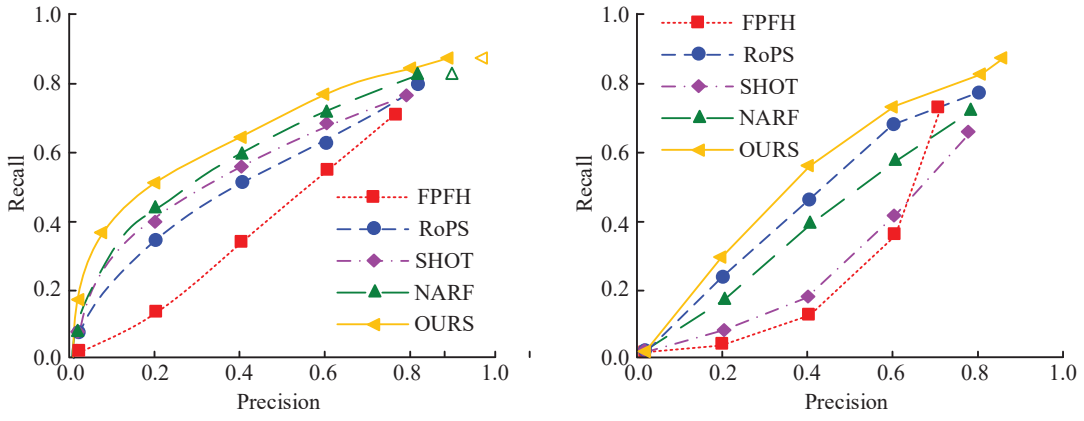


(a) The accuracy of feature extraction for different descriptors on the KITTI dataset

(b) The accuracy of feature extraction for different descriptors on the B3R dataset

Figure 9

Recall@1-Precision performance description on KITTI and B3 datasets.



(a) Recall@1-Precision performance description on KITTI dataset

(b) Recall@1-Precision performance description on B3R dataset

effectively capture the global structure and local detail features of PCs, and had stronger discriminative ability and robustness in complex scenes and different geometric structures.

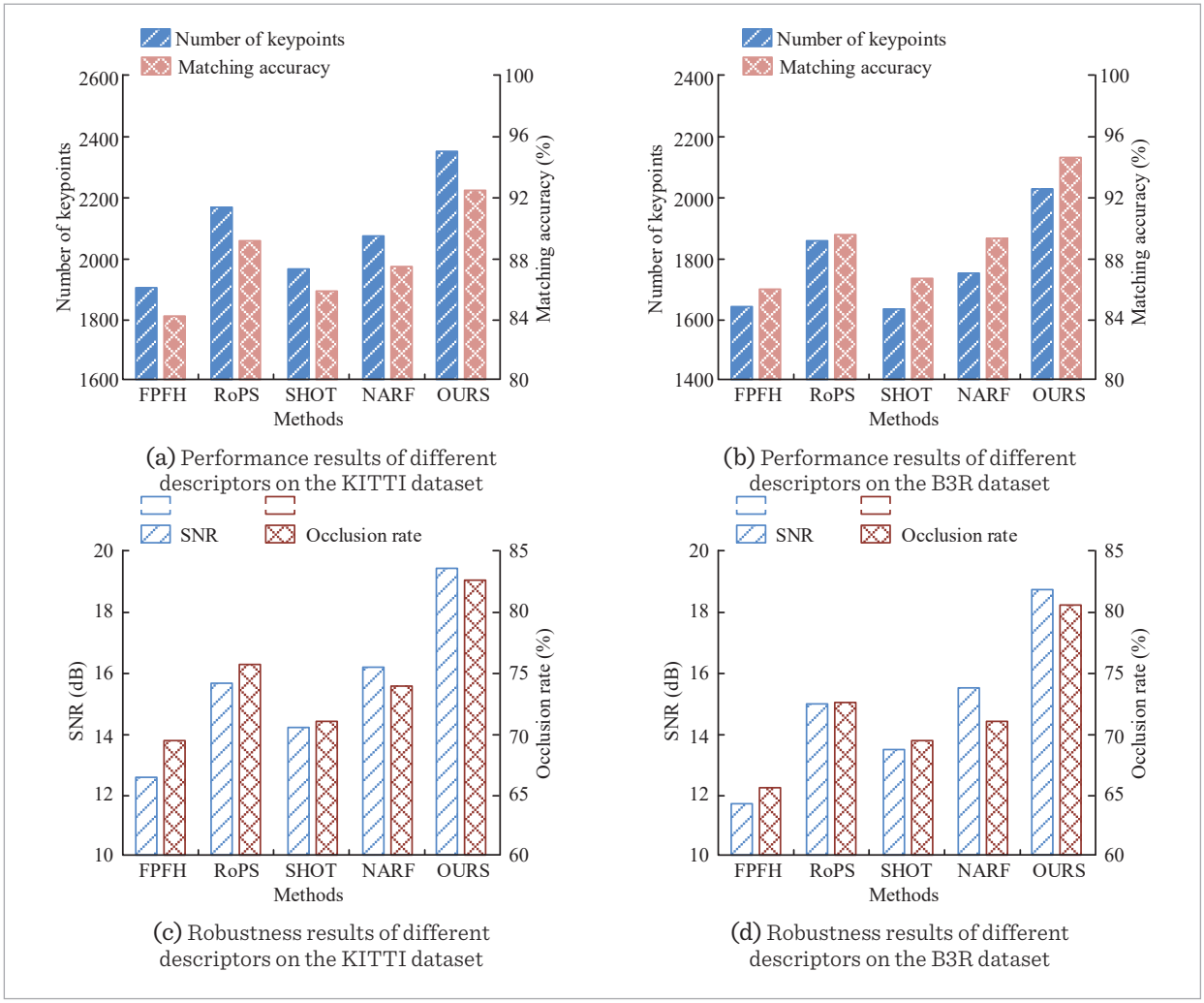
To investigate the performance of different descriptors, this study tested their descriptive properties using the Recall@1-Precision curve in two datasets, as shown in Figure 9. In Figure 9(a), the feature extraction accuracies of FPFH, RoPS, SHOT, and NARF in the KITTI dataset were 0.78, 0.83, 0.79, and 0.82, with recall rates of 0.71, 0.77, 0.75, and 0.79. The precision and recall of the proposed descriptor were 0.91 and 0.84. In Figure 9(b), the feature extraction accuracies of the five descriptors in the B3R dataset were 0.69, 0.81, 0.79, 0.80, and 0.88, with recall rates of 0.71, 0.79, 0.67, 0.70, and 0.86. The proposed descriptor could achieve more accurate feature matching while maintaining a high recall rate, indicating

that it had stronger target recognition ability while maintaining a lower false detection rate, verifying its superior descriptive performance and stability.

This study further analyzes the number of feature points extracted by different descriptors and the accuracy of feature matching in two datasets, and tests the robustness of feature extraction for different descriptors, including noise robustness and occlusion robustness. Among them, noise robustness is evaluated using Signal-to-Noise Ratio (SNR), while occlusion robustness is evaluated using the acceptable occlusion rate. For noise robustness evaluation, two types of synthetic noise were added to the PCs: Gaussian noise (zero mean, standard deviation ranging from 0.01 to 0.05 m) and outlier noise (randomly distributed points with a proportion of 1% to 5%). For occlusion experiments, structured occlusion was simulated by removing points within randomly generated spherical volumes (radius 0.1 to 0.5 m) placed at random positions on the target surface, mimicking object occlusion in real scenes. The occlusion rate was defined as the ratio of removed points to total points in the target PC. In Figure 10(a), the amount of extracted feature points for FPFH, RoPS, SHOT, NARF, and the proposed descriptor in the KITTI dataset were 1912, 2167, 1956, 2085, and 2374, with feature matching accuracies of 84.05%, 89.62%, 86.21%, 87.35%, and 92.42%. In Figure 10(b),

Figure 10

The performance and robustness results of different descriptors.



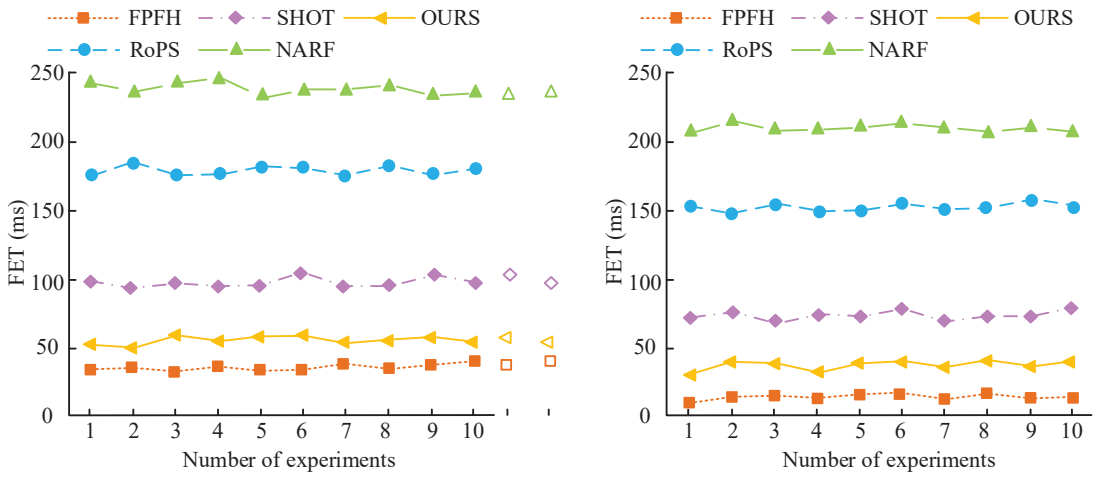
(a) Performance results of different descriptors on the KITTI dataset

(b) Performance results of different descriptors on the B3R dataset

(c) Robustness results of different descriptors on the KITTI dataset

(d) Robustness results of different descriptors on the B3R dataset

Figure 11
FET for different descriptors on two datasets.



(a) FET for different descriptors on the KITTI dataset

(b) FET for different descriptors on the B3R dataset

in the B3R dataset, the number of extracted feature points for the five descriptors was 1638, 1885, 1626, 1754, and 2023, with feature matching accuracies of 85.92%, 89.67%, 85.32%, 89.69%, and 94.28%. In Figure 10(c), the SNRs of FPFH, RoPS, SHOT, and NARF in the KITTI dataset were 12.52 dB, 15.80 dB, 14.25 dB, and 16.13 dB, with occlusion rates of 68.35%, 75.26%, 71.64%, and 73.82%. The SNR and occlusion rate of the proposed descriptor were 19.46 dB and 82.75%. In Figure 10(d), the SNRs of the five descriptors in the B3R dataset were 11.82 dB, 14.91 dB, 13.58 dB, 15.32 dB, and 18.76 dB, with occlusion rates of 65.42%, 72.63%, 69.24%, 71.55%, and 80.90%. It was found that the proposed descriptor could effectively suppress noise interference and had stronger tolerance for occlusion by fusing global and local features, and could maintain stable feature extraction performance in complex environments.

To validate the superiority of the research method, this study analyzed the efficiency of feature extraction based on different descriptors in two datasets, as shown in Figure 11. In Figure 11(a), the average Feature Extraction Time (FET) of FPFH, RoPS, SHOT, and NARF descriptors in the KITTI dataset was 38.62 ms, 175.26 ms, 99.87 ms, and 242.63 ms. The average FET of the proposed descriptor was 52.37 ms, which was 70.12%, 47.56%, and 78.42% shorter than that of RoPS, SHOT, and NARF descriptors. In Figure 11(b), the average FET of the five descriptors in the B3R dataset was 16.72 ms, 151.37 ms, 75.86 ms, 225.67 ms, and 38.25 ms. The results showed that although FPFH achieved faster feature extraction than the proposed descriptor, the proposed method offered a superior accuracy-robustness-efficiency trade-off.

The study further compared and analyzed the proposed method with recent deep learning based 3D recognition methods, including PointNet++, Learning Local Geometric Descriptors from RGB-D Reconstructions (3DMatch), Joint Learning of Dense Detection and Description of 3D Local Features (D3Feat). The results of Full Convolutional Geometric Features (FCGF) are shown in Table 2. In Table 2, on the KITTI dataset, the proposed method improved feature matching accuracy by 2.69 percentage points, SNR by 3.22 dB, and occlusion rate by 5.90 percentage points compared to D3Feat. On the B3R dataset, the proposed method also demonstrated significant advantages, with a feature matching accuracy of 94.28%, which was 7.33 percentage points higher than D3Feat. This superior performance was attributed to the effective fusion of global GASD features and local PDSH features, robust ISS keypoint detection, and improved RGSA.

Table 2
Comprehensive performance comparison between the proposed method and recent deep learning-based 3D recognition methods.

Datasets	Methods	Feature matching accuracy (%)	Number of feature points	SNR (dB)	Occlusion rate (%)	Feature extraction time (ms)
KITTI	PointNet++	87.32	2156	14.58	71.24	156.81
	3DMatch	85.94	1987	13.89	69.73	128.50
	D3Feat	89.73	2294	16.24	76.85	185.62
	FCGF	88.69	2210	15.62	74.93	142.35
	OURS	92.42	2374	19.46	82.75	52.37
B3R	PointNet++	84.56	1842	13.72	68.53	132.44
	3DMatch	82.38	1653	13.15	66.81	109.32
	D3Feat	86.95	1968	15.48	73.42	162.71
	FCGF	85.77	1908	14.86	71.68	121.57
	OURS	94.28	2023	18.76	80.90	38.25

3.2. Analysis of Recognition Results of 3DOR Method

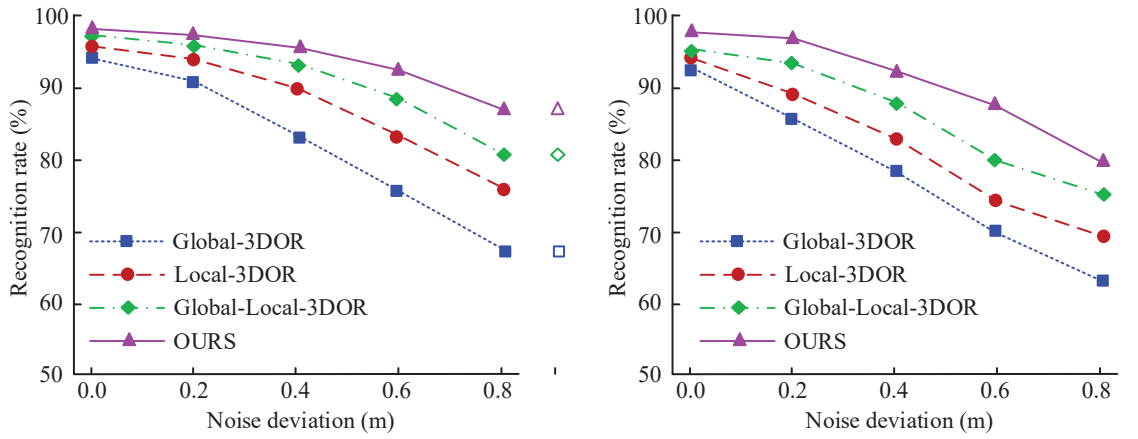
To evaluate the recognition performance of the designed 3DOR method, this study constructed two real-world PC datasets: UrbanRoad and TunnelLite. The UrbanRoad dataset utilized Robosense RS-LDAR-32 line LiDAR to collect PC data from multiple scenes including urban main roads, intersections, overpasses, etc. This dataset focuses on challenging cases such as dense traffic flow, severe occlusion, and distant small targets, and has been accurately manually annotated. The TunnelLite dataset collects scenes including long tunnels, underground garages, nighttime city roads, and foggy weather, totaling 8200 frames of PC. The experiment compared the proposed method with existing popular methods, including Global Feature-based 3DOR (Global-3DOR), Local Feature-based 3DOR (Local-3DOR) and Global and Local Feature Fusion-based 3DOR (Global-Local-3DOR).

This study compared the recognition rates of various 3DOR methods under different noise biases in two real-world PC sets to assess the robustness and stability of the research method under noise interference, as shown in Figure 12. In Figure 12(a), in the UrbanRoad dataset, when the noise deviation was 0.2 m, the recognition rates of Global-3DOR, Local-3DOR, Global-Local-3DOR, and the proposed method were 91.27%, 94.68%, 96.52%, and 97.36%. When the noise deviation increased to 0.8 m, the recognition rates of the four 3DOR methods were 67.63%, 75.24%, 81.26%, and 89.86%. In Figure 12(b), in the TunnelLite dataset, when the noise deviation was 0.8 m, the recognition rates of the four 3DOR methods were 63.35%, 69.59%, 75.68%, and 80.72%. The proposed method could maintain stable target recognition performance in high noise environments and had strong practicality and adaptability.

To evaluate the performance and generalization ability of the proposed method at different PC resolutions, this study compared and analyzed the recognition rates at different resolutions in two real-world PC sets, as exhibited in Figure 13. In Figure 13(a), when the downsampling ratio is 1/2, the recognition rates of Global-3DOR, Local-3DOR, Global-Local-3DOR, and the proposed method in the UrbanRoad dataset were 85.62%, 89.27%, 93.66%, and 97.62%. When the downsampling ratio was 1/16, the recognition rates of the four methods were 45.32%, 58.59%, 78.12%, and 88.76%. In Figure 13(b), when the downsampling ratio was 1/16, the recognition rates of the four 3DOR methods in the TunnelLite dataset were 28.69%, 48.62%, 66.32%, and 80.78%. The proposed method combined multi-scale fea-

Figure 12

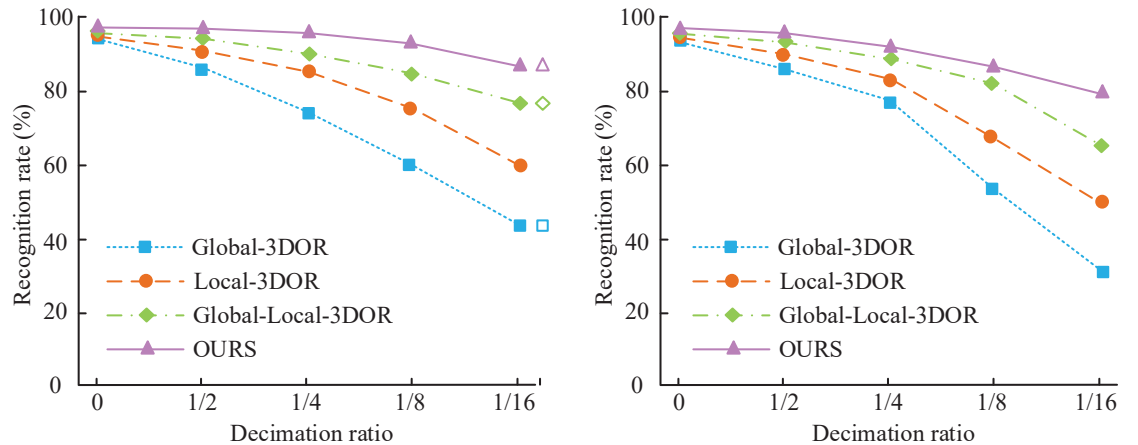
Recognition rates of 3DOR methods under various noise biases.



(a) Recognition rates on the UrbanRoad dataset (b) Recognition rates on the TunnelLite dataset

Figure 13

Recognition rates of different 3DOR methods at different resolutions.



(a) Recognition rates on the UrbanRoad dataset (b) Recognition rates on the TunnelLite dataset

ture description with ISS keypoint detection, which could effectively overcome the impact of PC density changes and maintain stable recognition performance in low resolution scenes.

To evaluate the contribution of each module to the overall recognition performance, ablation experiments were conducted on the UrbanRoad dataset, and the results are shown in Table 3. In Table 3, each module contributed to the overall recognition performance. When PDSH was used alone to describe subunits, the recognition rate reached 81.23%, the feature extraction time was only 72.5 ms, and the efficiency was better than GASD. When using ISS alone, improving RGSA or ICP, the recognition rate improvement was limited and the time consumption increased. After combining GASD with PDSH, the recognition rate increased to 86.57% and the feature extraction time decreased to 68.2 ms, indi-

cating a significant synergistic effect between global and local features. The complete method integrates all modules, achieving a recognition rate of 89.86%, a feature matching accuracy of 92.4%, and a feature extraction time of 116.6 ms. The results showed that the full proposed method (116.6 ms) was slower than its individual components and the combination GASD+PDSH. However, it achieved the highest recognition rate.

To investigate the computational efficiency of the proposed method in real-world environments, this study compared the computation time and memory usage of different object recognition methods in two real-world PC sets, as shown in Figure 14. In Figure 14(a), the average computation time of Global-3DOR, Local-3DOR, Global-Local-3DOR, and the research method in the UrbanRoad dataset was 103.28 ms, 289.69 ms, 416.85 ms, and 232.74 ms. Compared with Local-3DOR and Global-Local-3DOR, the computation time of the proposed method was reduced by 19.66% and 44.17%. The average computation time of the four comparative methods in the TunnelLite

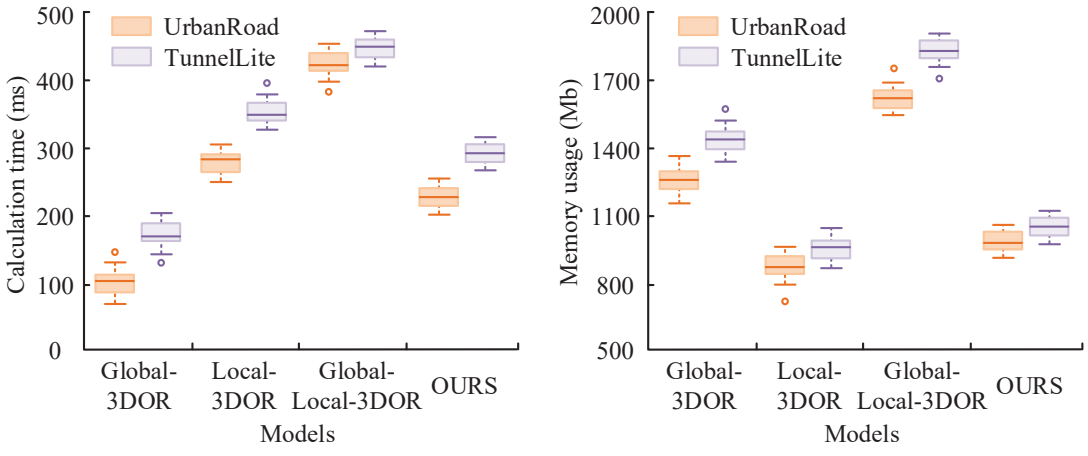
Table 3

Results of ablation experiment.

Configuration	Recognition rate (%)	Feature matching accuracy (%)	Feature extraction time (ms)
Baseline	72.34	85.6	103.2
+ GASD only	78.56	87.2	98.7
+ PDSH only	81.23	88.5	72.5
+ ISS only	75.89	86.1	115.8
+ Improved RGSA only	76.42	86.4	118.4
+ ICP (P2Pl) only	74.15	85.2	126.3
GASD + PDSH	86.57	90.6	68.2
ISS + Improved RGSA + ICP	84.32	89.3	142.6
Full proposed method	89.86	92.4	116.6

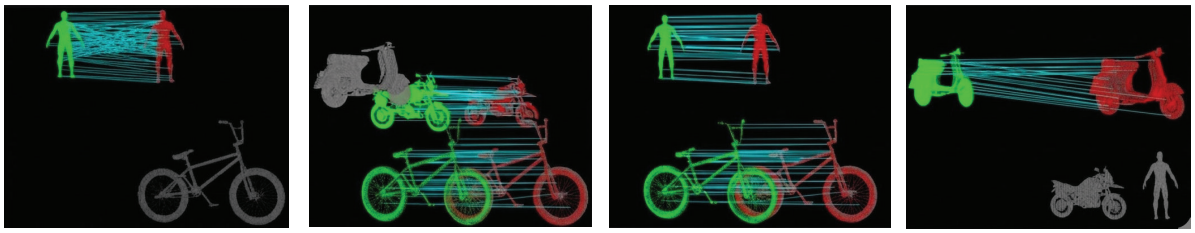
Figure 14

Comparison of computational efficiency of different 3DOR methods.



(a) Calculation time on two datasets (b) Memory usage on two datasets

Figure 15
Feature activation heatmaps of the proposed method.



(a) (b) (c) (d)

dataset was 174.36 ms, 342.12 ms, 452.69 ms, and 297.35 ms. In Figure 14(b), the memory consumption of the four 3DOR methods in the UrbanRoad dataset was 1267.23 Mb, 889.62 Mb, 1619.27 Mb, and 992.36 Mb. The memory consumption of Global-3DOR, Local-3DOR, and Global-Local-3DOR in the TunnelLite dataset was 1436.58 Mb, 945.62 Mb, and 1825.91 Mb, and the proposed method was 1047.35 Mb. It could ensure recognition accuracy while having higher computational efficiency, suitable for real-time 3DOR applications.

To qualitatively evaluate the focusing ability of the proposed method on the discriminative geometric structure of PCs, the feature activation visualization heatmap of the proposed laser radar 3D target recognition was analyzed, and the results are shown in Figure 15. In Figure 15, the proposed method could effectively focus on the discriminative geometric structure of PC targets. On different targets such as vehicles and pedestrians, the activation area of the heatmap was mainly concentrated on geometric features such as contour edges, local protrusions, and depressions, while the activation was weaker on flat areas. Activation refers to geometric saliency, defined as the normalized response of the proposed feature descriptor-keypoint detector combination. For each point, activation was computed by weighting the PDSH histogram entropy with the ratio of ISS eigenvalues, indicating local shape distinctiveness. Higher activation values corresponded to regions with significant geometric variations, such as edges, corners, and surface discontinuities. This saliency measure did not rely on neural network features but was derived directly from geometric invariants. The heatmap visualized this saliency, where red represents high activation (highly discriminative structures) and green represents low activation (flat or uniform regions). This indicated that the combination of PDSH descriptors and ISS keypoint detection could accurately capture the local geometric information of the target, enhance the discriminative ability of feature expression, and provide a reliable feature basis for subsequent high-precision target recognition.

4. Discussion and Interpretation

This study proposed a LiDAR 3DOR method that integrates PDSH and ISS algorithms, and validated its performance on multiple datasets. The superior feature extraction accuracy of the proposed descriptor (93.68% on KITTI and 91.46% on B3R) has been attributed to three design features: the multi-plane projection strategy of the PDSH descriptor, which encodes complementary spatial information from three orthogonal views and increases joint entropy without causing the curse of dimensionality; the scale-normalized projection, which provides robustness to varying PC densities; and the integration of GASD global features, which compensates for the limited receptive field of local descriptors.

The hypothesis stated in Section 1 has been quantitatively verified by the ablation study in Table 3: local features (PDSH) contribute an 8.89 percentage point improvement, global features (GASD) contributed a 6.22 percentage point improvement, and their combination yielded a synergistic gain of 14.23 percentage points, confirming the necessity of fusing global and local features. Noise robustness (SNR of 19.46 dB) has been causally linked to the eigenvalue thresholds of the ISS keypoint detector and the histogram aggregation strategy of PDSH. Occlusion robustness

(82.75%) has been attributed to the multi-view projection strategy, where complementary information remains available from the other two orthogonal planes when one projection plane is occluded.

The necessity of each module in the serial pipeline has also been validated by the ablation study in Table 3. The baseline achieved only 72.34% on UrbanRoad. Incrementally adding modules monotonically improved performance to 89.86%, confirming that each module addressed a distinct challenge and their serial integration followed a theoretically sound coarse-to-fine paradigm. As shown in Table 2, the proposed handcrafted descriptor outperformed deep learning methods (PointNet++, 3DMatch, D3Feat, FCGF) because its geometric design principles provided inherent robustness without requiring large-scale training data. The feature extraction time (52.37 ms) was substantially lower than that of D3Feat because no GPU-intensive computation was required.

Compared with reference [30], this paper only verified the robustness under Gaussian noise and did not involve complex interference types such as structural disturbances. In addition, reference [30] focuses on ultra high speed weak target recognition, while the data in this paper comes from low-speed scenarios. The potential of the proposed method for transferring to high-speed scenarios remains to be explored.

Despite the superior performance of the proposed method, several limitations should be acknowledged. First, the method relies on normal vector estimation quality, which tends to fail in sparse PCs or regions with sharp geometric features. Second, the computational advantage diminishes when processing large-scale scenes. Third, robustness under extreme weather conditions has not been validated. Fourth, the framework assumes rigid objects and cannot handle non-rigid targets. Future work will address these issues through adaptive techniques and multimodal fusion.

5. Summary

This study constructed a recognition method that integrated PDSH descriptors and ISS algorithm to deal with insufficient robustness and low computational efficiency in LiDAR 3DOR. Research showed that the research method outperformed existing methods on multiple datasets, especially maintaining stable recognition performance in noisy, occluded, and low resolution environments. Future research will combine multimodal data to further enhance the robustness and generalization ability of the system in complex environments.

Acknowledgement

The research is supported by: the National Natural Science Foundation of China under Grant, Remote Sensing Monitoring and Analysis of Cropland Abandonment Based on Biased Depth Feature Fusion, No.: 42271282.

References

1. Cattaneo, D., Vaghi, M., Valada, A. LCDNet: Deep Loop Closure Detection and Point Cloud Registration for LiDAR SLAM. IEEE Transactions on Robotics, 2022, 38(4), 2074-2093. https://doi.org/10.1109/TRO.2022.3150683
2. Chiang, K. W., Srinara, S., Tsai, S., Lin, C. X., Tsai, M. L. High-Definition-Map-Based LiDAR Localization Through Dynamic Time-Synchronized Normal Distribution Transform Scan Matching. IEEE Transactions on Vehicular Technology, 2023, 72(6), 7011-7023. https://doi.org/10.1109/TVT.2023.3237275
3. Cho, Y., Kim, G., Lee, S., Ryu, J. H. OpenStreetMap-Based LiDAR Global Localization in Urban Environment Without a Prior LiDAR Map. IEEE Robotics and Automation Letters, 2022, 7(2), 4999-5006. https://doi.org/10.1109/LRA.2022.3152476
4. Cui, Y., Zhang, Y., Dong, J., Sun, H., Chen, X., Zhu, F. Link3D: Linear Keypoints Representation for 3D LiDAR Point Cloud. IEEE Robotics and Automation Letters, 2024, 9(3), 2128-2135. https://doi.org/10.1109/LRA.2024.3354550
5. Guadagnino, T., Chen, X., Sodano, M., Behley, J., Grisetti, G., Stachniss, C. Fast Sparse LiDAR Odometry Using Self-Supervised Feature Selection on Intensity Images. IEEE Robotics and Automation Letters, 2022, 7(3), 7597-7604. https://doi.org/10.1109/LRA.2022.3184454

6. Han, X., Feng, Z., Sun, S., Xiao, G. 3D Point Cloud Descriptors: State-of-the-Art. Artificial Intelligence Review, 2023, 56(10), 12033-12083. https://doi.org/10.1007/s10462-023-10486-4

7. Katkade, S. N., Bagal, V. C., Manza, R. R., Yannawar, P. L. Advances in Real-Time Object Detection and Information Retrieval: A Review. Artificial Intelligence Applications, 2023, 1(3), 123-128. https://doi.org/10.47852/bonviewAIA3202456

8. Kuçak, R. A., Erol, S., Erol, B. The Strip Adjustment of Mobile LiDAR Point Clouds Using Iterative Closest Point (ICP) Algorithm. Arabian Journal of Geosciences, 2022, 15(11), 1017-1029. https://doi.org/10.1007/s12517-022-10303-2

9. Li, H. T., Todd, Z., Bielski, N., Carroll, F. 3D LiDAR Point-Cloud Projection Operator and Transfer Machine Learning for Effective Road Surface Features Detection and Segmentation. The Visual Computer, 2022, 38(5), 1759-1774. https://doi.org/10.1007/s00371-021-02103-8

10. Li, J., Huang, X., Zhan, J. High-Precision Motion Detection and Tracking Based on Point Cloud Registration and Radius Search. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(6), 6322-6335. https://doi.org/10.1109/TITS.2023.3250209

11. Li, T., Zhang, Z., Zhu, M., Cui, Z., Wei, D. Combining Transformer Global and Local Feature Extraction for Object Detection. Complex & Intelligent Systems, 2024, 10(4), 4897-4920. https://doi.org/10.1007/s40747-024-01409-z

12. Lin, C., Tian, D., Duan, X., Zhou, J., Zhao, D., Cao, D. CL3D: Camera-LiDAR 3D Object Detection with Point Feature Enhancement and Point-Guided Fusion. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10), 18040-18050. https://doi.org/10.1109/TITS.2022.3154537

13. Liu, S., Luo, F., Li, Q., Liu, X., Hu, L. AWEDD: A Descriptor Simultaneously Encoding Multiscale Extrinsic and Intrinsic Shape Features. The Visual Computer, 2024, 40(4), 2537-2554. https://doi.org/10.1007/s00371-023-02935-6

14. Meng, Z., Xia, X., Xu, R., Liu, W., Ma, J. HYDRO-3D: Hybrid Object Detection and Tracking for Cooperative Perception Using 3D LiDAR. IEEE Transactions on Intelligent Vehicles, 2023, 8(8), 4069-4080. https://doi.org/10.1109/TIV.2023.3282567

15. Na, K. I., Park, B. Real-Time 3D Multi-Pedestrian Detection and Tracking Using 3D LiDAR Point Cloud for Mobile Robot. ETRI Journal, 2023, 45(5), 836-846. https://doi.org/10.4218/etrij.2023-0116

16. Nagar, R. Robust Extrinsic Symmetry Estimation in 3D Point Clouds. The Visual Computer, 2025, 41(1), 115-128. https://doi.org/10.1007/s00371-024-03313-6

17. Poiesi, F., Boscaini, D. Learning General and Distinctive 3D Local Deep Descriptors for Point Cloud Registration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3), 3979-3985. https://doi.org/10.1109/TPAMI.2022.3175371

18. Rani, S., Lakhwani, K., Kumar, S. Three-Dimensional Objects Recognition and Pattern Recognition Technique; Related Challenges: A Review. Multimedia Tools and Applications, 2022, 81(12), 17303-17346. https://doi.org/10.1007/s11042-022-12412-2

19. Romero-Gonzalez, C., García-Varea, I., Martinez-Gomez, J. Shape Binary Patterns: An Efficient Local Descriptor and Keypoint Detector for Point Clouds. Multimedia Tools and Applications, 2022, 81(3), 3577-3601. https://doi.org/10.1007/s11042-021-11586-5

20. Song, Z., Liu, L., Jia, F., Luo, Y., Jia, C., Zhang, G., Wang, L. Robustness-Aware 3D Object Detection in Autonomous Driving: A Review and Outlook. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(11), 15407-15436. https://doi.org/10.1109/TITS.2024.3439557

21. Usmani, K., O'Connor, T., Wani, P., Javidi, B. 3D Object Detection Through Fog and Occlusion: Passive Integral Imaging vs Active (LiDAR) Sensing. Optics Express, 2022, 31(1), 479-491. https://doi.org/10.1364/OE.478125

22. Vizzo, I., Guadagnino, T., Mersch, B., Wiesmann, L., Behley, J., Stachniss, C. KISS-ICP: In Defense of Point-to-Point ICP: Simple, Accurate, and Robust Registration If Done the Right Way. IEEE Robotics and Automation Letters, 2023, 8(2), 1029-1036. https://doi.org/10.1109/LRA.2023.3236571

23. Vodisch, N., Unal, O., Li, K., Van Gool, L., Dai, D. End-to-End Optimization of LiDAR Beam Configuration for 3D Object Detection and Localization. IEEE Robotics and Automation Letters, 2022, 7(2), 2242-2249. https://doi.org/10.1109/LRA.2022.3142738

24. Wang, H., Liu, Y., Hu, Q., Wang, B., Chen, J., Dong, Z., Yang, B. RoReg: Pairwise Point Cloud Registration with Oriented Descriptors and Local Rotations. IEEE Transactions on Pattern Analysis and Machine

Intelligence, 2023, 45(8), 10376-10393. https://doi.org/10.1109/TPAMI.2023.3244951

25. Wen, L., Jo, K. H. Fast LiDAR R-CNN: Residual Relation-Aware Region Proposal Networks for Multiclass 3-D Object Detection. IEEE Sensors Journal, 2022, 22(12), 12323-12331. https://doi.org/10.1109/JSEN.2022.3172446
26. Xiong, W., Liu, J., Huang, T., Han, Q. L., Xia, Y., Zhu, B. LXL: LiDAR Excluded Lean 3D Object Detection with 4D Imaging Radar and Camera Fusion. IEEE Transactions on Intelligent Vehicles, 2023, 9(1), 79-92. https://doi.org/10.1109/TIV.2023.3321240
27. Xu, Q., Zhong, Y., Neumann, U. Behind the Curtain: Learning Occluded Shapes for 3D Object Detection. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(3), 2893-2901. https://doi.org/10.1609/aaai.v36i3.20194
28. Xu, Z. H., Guo, G., Sun, Q. C., Wang, Q., Zhang, G. D., Ye, R. Q. Structural Plane Recognition from Three-Dimensional Laser Scanning Points Using an Improved Region-Growing Algorithm Based on the Robust Randomized Hough Transform. Journal of Mountain Science, 2023, 20(11), 3376-3391. https://doi.org/10.1007/s11629-023-7914-z
29. Xue, B., Zheng, Q., Li, Z., Wang, J., Mu, C., Yang, J., Li, X. ISAR Weak Feature Enhancement with Perturbation Defense Using Hybrid Clustering Oversegmentation. IEEE Transactions on Aerospace and Electronic Systems, 2024, 60(5), 6256-6274. https://doi.org/10.1109/TAES.2024.3408139
30. Xue, B., Zheng, Q., Li, Z., Wang, J., Mu, C., Yang, J., Li, X. Perturbation Defense Ultra High-Speed Weak Target Recognition. Engineering Applications of Artificial Intelligence, 2024, 138, 109420. https://doi.org/10.1016/j.engappai.2024.109420
31. Yang, J., Zhang, J., Cai, Z., Fang, D. Novel 3D Local Feature Descriptor of Point Clouds Based on Spatial Voxel Homogenization for Feature Matching. Visual Computing for Industry, Biomedicine, and Art, 2023, 6(1), 18-39. https://doi.org/10.1186/s42492-023-00145-4
32. Zhang, Y., Han, H., Han, X., Fu, J. Point Cloud Registration Method Based on Improved Covariance Matrix Descriptor. Journal of System Simulation, 2023, 35(5), 979-986. DOI: https://doi.org/10.16182/j.issn1004731xjos.22-0010
33. Zhao, Y., Song, N., Bao, W. Hyperspectral Image Classification Based on Adaptive Spectral Feature Decoupling with Global Local Feature Fusion Network. Earth Science Informatics, 2024, 17(5), 4619-4637. https://doi.org/10.1007/s12145-024-01415-2
34. Zhou, B., Li, C., Chen, S., Xie, D., Yu, M., Li, Q. ASL-SLAM: A LiDAR SLAM with Activity Semantics-Based Loop Closure. IEEE Sensors Journal, 2023, 23(12), 13499-13510. https://doi.org/10.1109/JSEN.2023.3270871
35. Zhou, Z., Huang, D., Liu, Z. Point Cloud Registration Algorithm Based on 3D Shape Context Features. Journal of Applied Optics, 2023, 44(2), 330-336. https://doi.org/10.5768/JAO202344.0202005

This article is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 (CC BY 4.0) License (http://creativecommons.org/licenses/by/4.0/).