ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003

RANSAC-based Point Cloud Calibration Algorithm Accounting for Point Quantity and Distribution

Kai Qiao

School of Environment and Spatial Information, China University of Mining and Technology, Xuzhou, China

Abstract

To tackle the challenge of inadequate accuracy in UAV-LiDAR point cloud data, this research introduces a RANSAC calibration algorithm that takes into account both the quantity and spatial distribution of points. Simulation experiments were carried out across four experimental sites to ascertain the optimal point placement strategy under diverse terrain conditions. Initially, control points were established within the experimental zones, and both control point data and point cloud data were collected. Subsequently, calibration strategies incorporating different point counts and distributions were devised. Ultimately, the precision of each calibration approach was evaluated. The experimental findings reveal that: (1) calibration accuracy decreases as terrain complexity increases within the experimental areas; (2) a greater number of points correlates with higher calibration accuracy; (3) given an equal number of points, a diamond-shaped point distribution yields the highest calibration accuracy.

Keywords

UAV-LiDAR; Point cloud; RANSAC.

1. INTRODUCTION

UAV-LiDAR enables efficient, large-scale monitoring and swiftly captures three-dimensional surface information with centimeter-level measurement accuracy. Furthermore, its flight parameters, detection cycles, and spatial resolution can be tailored to specific requirements, offering a high degree of flexibility. Nevertheless, it is vulnerable to weather conditions, potentially introducing errors into data collection. Consequently, point cloud calibration is essential to enhance accuracy. Chen introduced the Iterative Closest Point (ICP) algorithm, which has since served as the cornerstone of fine registration by iteratively resolving point correspondences and rigid transformations. Subsequently, Point-to-Plane ICP, Plane-to-Plane ICP, and Generalized ICP (GICP) have progressively enhanced their robustness against noise and local optima[1]. Notably, GICP has emerged as a prevailing mainstream approach by integrating various ICP losses through probabilistic modeling[2]. Magnusson proposed the Normal Distributions Transform (NDT), which represents point clouds as Gaussian distributions and facilitates registration through probabilistic matching, eliminating the need for explicit point correspondences and rendering it well-suited for sparse and unordered point clouds[3]. Myronenko introduced Consensus Point Drift (CPD), which conceptualizes registration as probability density alignment and exhibits increased robustness against low overlap and outliers[4]. Zhou proposed Fast Global Registration (FGR), which efficiently filters outliers using RANSAC and graph optimization to achieve global registration without initialization, thereby becoming a core algorithm in Open3D[5]. Koide presented VGICP, which synergizes voxelization and GICP, reduces computational load by partitioning the point cloud into voxel grids, and accelerates registration speed by 3-5 times while preserving accuracy[6]. The aforementioned

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003

methods still exhibit potential for enhancement in terms of accuracy and computational efficiency. To further improve the precision of UAV-LiDAR point cloud calibration, this paper introduces a RANSAC-based point cloud calibration algorithm that takes into account both the quantity and distribution of points.

2. AN IMPROVED RANSAC ALGORITHM

The RANSAC algorithm demonstrates remarkable robustness against noise points and outliers, enabling it to effectively process point clouds containing a substantial number of outliers. Fundamentally, it identifies the rigid transformation matrix between point clouds through a process of random sampling and consistency verification. Specifically, RANSAC initially randomly selects three point pairs from the input dataset, computes their transformation matrix, and subsequently assesses consistency by calculating the distance between the transformed point set and the target point set. If the quantity of consistent points satisfies a predetermined threshold, the model is deemed valid. To augment the robustness and stability of the registration process and enhance matching accuracy, this study increases the number of corresponding point pairs and further takes into account the influence of point distribution on calibration outcomes, adopting six calibration schemes: 3-point triangular, 4-point rhomboidal, 4-point rectangular, 6-point rhomboidal, 6-point rectangular, and 7-point uniformly distributed configurations. The specific procedural steps of the RANSAC algorithm are as follows: (1) feature point matching; (2) selection of point pairs; (3) computation of the transformation matrix; and (4) iterative computation to yield the optimal solution.

Upon the completion of steps (1) and (2), initiate the calculation of the transformation matrix. Initially, compute the centroids of the point pairs that are engaged in the registration process between the point cloud subject to registration and the reference point cloud:

$$C_p = \frac{1}{4}\sum_{i=1}^{4} p_i \tag{1}$$

$$C_q = \frac{1}{4}\sum_{i=1}^{4} q_i \tag{2}$$

Here, p_i represents the point cloud to be registered, while q_i denotes the reference point cloud.

Subtract the coordinates of each point by its corresponding centroid to derive the decentralized point sets P' and Q', and subsequently compute the covariance matrix H:

$$H = \sum_{i=1}^{4} (p_i' q_i'^T) \tag{3}$$

Employ Singular Value Decomposition (SVD) to compute the rotation matrix R and the translation matrix T.:

$$U, S, V = \mathrm{SVD}(H) \tag{4}$$

$$R = VU^T \tag{5}$$

$$T = C_q - RC_p \tag{6}$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003

Apply the transformation matrix [R, T] to transform the source point cloud, and then use a K-Dtree to search for the nearest neighbor points to the target point cloud. Calculate the distance between each transformed source point and the target point, and find a set of consistent points with a distance less than the set threshold ε:

$$||Rp_i + T - q_i|| < \varepsilon \tag{7}$$

Repeat the above process and find the optimal transformation matrix through continuous iterative sampling.

Subsequently, ICP registration is performed. The ICP algorithm is essentially an optimal matching technique utilizing the principle of least squares. Its core idea lies in determining the optimal rotation and translation matrices to achieve precise correspondence between source point cloud and target point cloud data points. The algorithm continuously optimizes these transformation parameters through an iterative process until it meets the convergence criteria, which usually means that the change in the objective function value is below a preset threshold or the maximum number of iterations has been reached.

3. SIMULATION EXPERIMENT

3.1. Experimental area

Addressing the issues of the original UAV-LiDAR point cloud accuracy being insufficient to meet the monitoring requirements of mining areas and the high cost of deploying GNSS continuous observation stations, this section designs and conducts simulation experiments in four typical areas. The experiments aim to investigate the calibration accuracy of UAV-LiDAR point clouds under different numbers and distributions of control points, and to obtain the optimal number and distribution of control points for millimeter-level accuracy. This provides a reference for the deployment of GNSS continuous observation stations in mining areas. The point cloud data is obtained through UAV-LiDAR scanning, and the ground control points are obtained through GNSS static measurement. The four typical areas are divided based on their terrain characteristics as follows: Experimental Area A: a flat and open area with small undulations, relatively sparse vegetation, no tall trees, and good visibility within the area; Experimental Area B: a flat dense forest area with minor topographic undulations, tall trees, low-lying grass, and relatively poor visibility within the area; Experimental Area C: a mountainous open area with undulations, low-lying grass, no tall trees, and good visibility within the area; Experimental Area D: a mountainous dense forest area with significant undulations, dense vegetation, tall trees, and poor visibility within the area. In Experimental Area A, 12 control points are designed and 11 are actually scanned; in Experimental Area B, 18 control points are designed and 18 are actually scanned; in Experimental Area C, 17 control points are designed and 17 are actually scanned; in Experimental Area D, 20 control points are designed and 19 are actually scanned.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003

(a) Experimental Area A (b) Experimental Area B

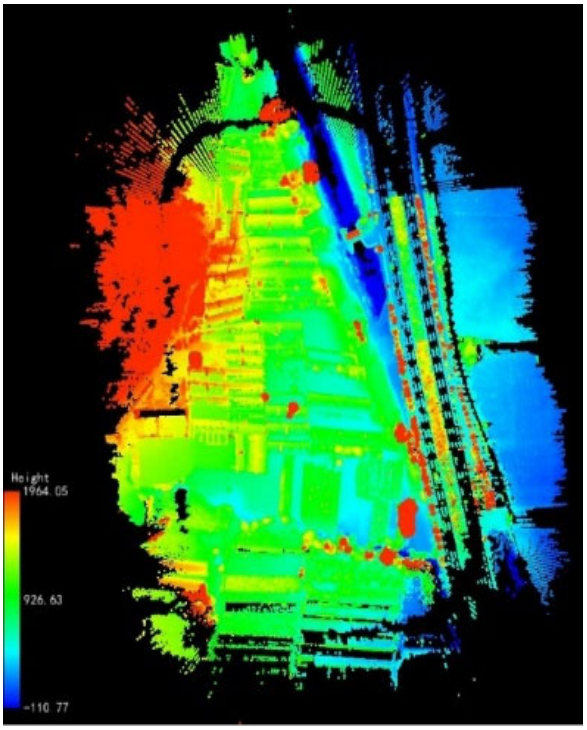
(c) Experimental Area C (d) Experimental Area D

Figure 1. 4 experimental zones

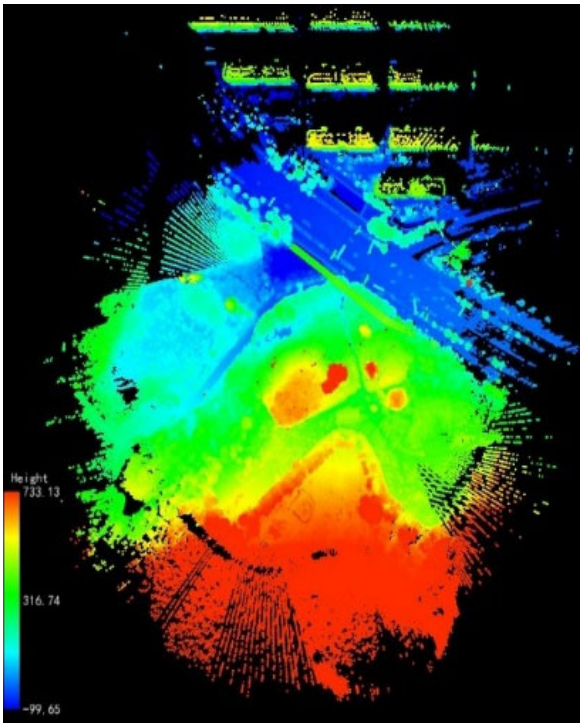
3.2. Data processing

The UAV-LiDAR point cloud data is processed through IE processing and data fusion software, while the control point data is processed through Huace CHO software. The UAV-LiDAR point cloud processing results are shown in Figure 2, and the locations of the control points are shown in Figure 3.

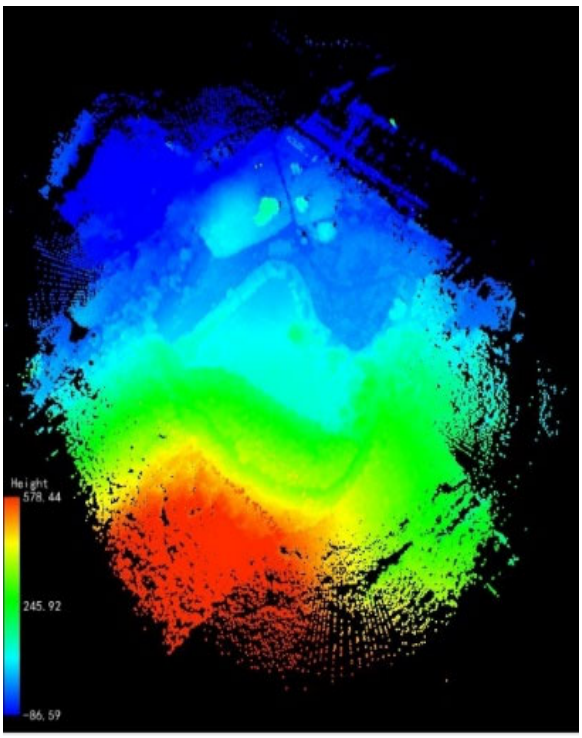
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003



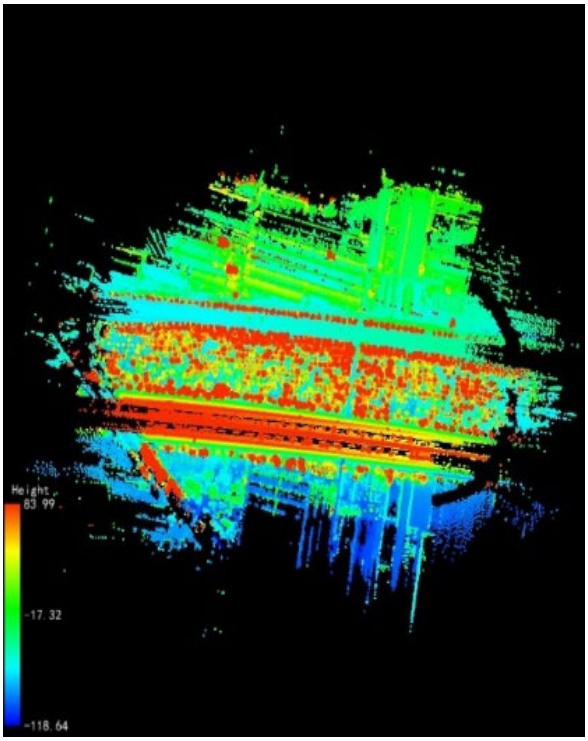
(a) Experimental Area A



(b) Experimental Area B



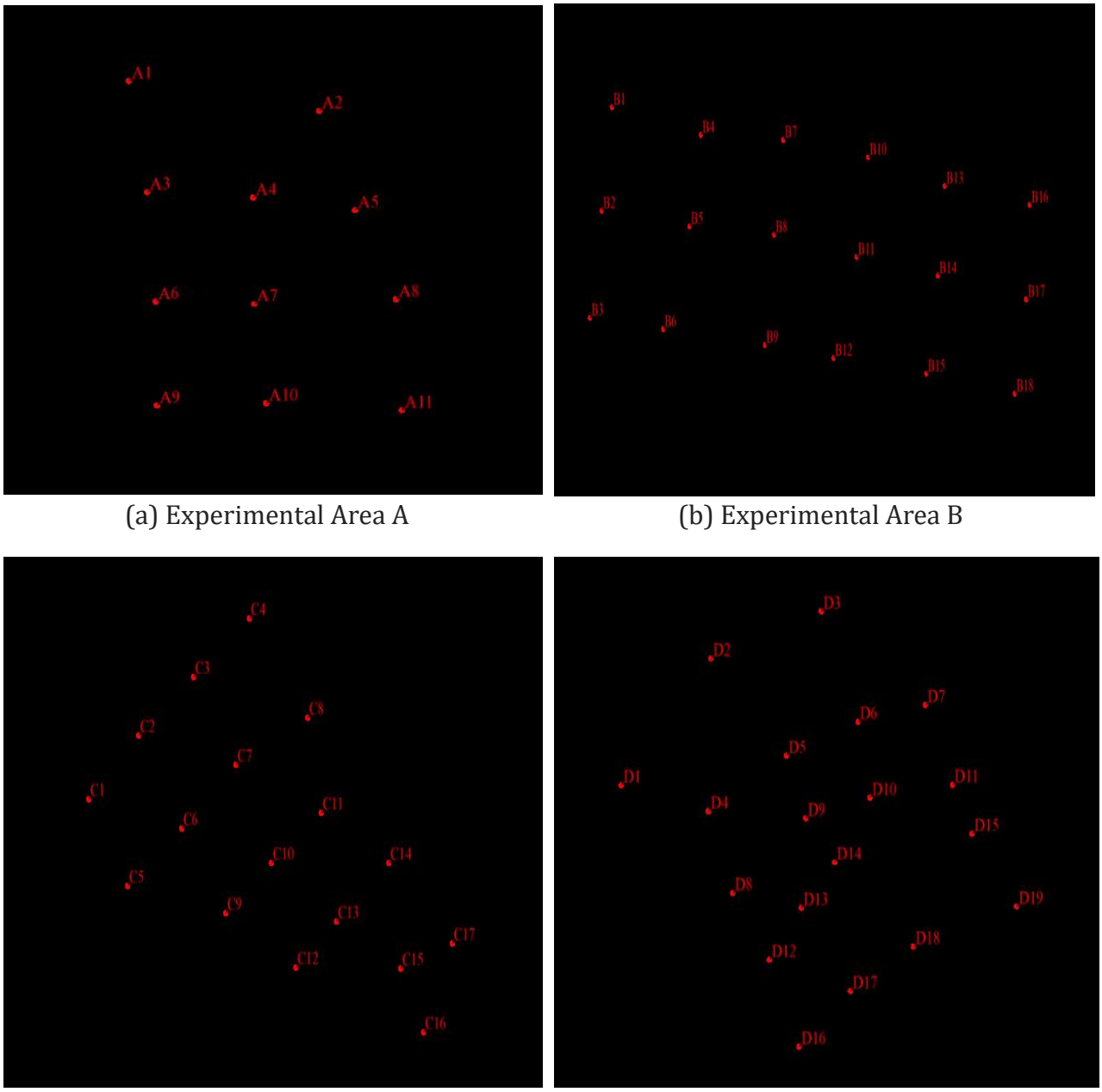
(c) Experimental Area C



(d) Experimental Area D

Figure 2. UAV-LiDAR point cloud processing results

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003



(a) Experimental Area A (b) Experimental Area B

(c) Experimental Area C (d) Experimental Area D

Figure 3. Distribution of control points in 4 experimental areas

The traditional RANSAC algorithm only utilizes three control points, making it difficult to meet the required calibration accuracy. In this experiment, we increased the number of control points and further investigated the impact of control point distribution on calibration accuracy. The experimental scheme is presented in Table 1.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003

Table 1. Number and distribution of calibration points involved

(a) Experimental Area A		
Point	Number of points	Distribution of points
1,2,3,5,6,8,9,10,11	9	Uniform distribution
2,4,6,7,8,10	6	Diamond
1,2,3,5,9,11	6	Rectangle
4,6,8,10	4	Diamond
1,2,9,11	4	Rectangle
1,5,9	3	Triangle
(b) Experimental Area B		
Point	Number of points	Distribution of points
1,2,3,7,8,9,13,14,15	9	Uniform distribution
2,5,7,9,14,17	6	Diamond
1,3,7,9,16,18	6	Rectangle
2,7,9,17	4	Diamond
1,3,16,18	4	Rectangle
1,9,16	3	Triangle
(c) Experimental Area C		
Point	Number of points	Distribution of points
1,2,4,5,7,12,14,15,17	9	Uniform distribution
3,7,9,11,13,16	6	Diamond
1,4,9,11,16,17	6	Rectangle
3,9,11,16	4	Diamond
1,4,16,17	4	Rectangle
1,8,16	3	Triangle
(d) Experimental Area D		
Point	Number of points	Distribution of points
1,2,3,8,10,11,16,18,19	9	Uniform distribution
2,8,9,10,11,18	6	Diamond
1,3,8,11,16,19	6	Rectangle
2,8,11,18	4	Diamond
1,3,16,19	4	Rectangle
1,11,16	3	Triangle

3.3. Result analysis

After the point cloud calibration is completed, the errors of each scheme are calculated, as shown in Table 2 and Figure 4.

Table 2. Errors in each experimental protocol

Experimental protocol	Number of points	Distribution of points	Error in experimental area A (mm)	Error in experimental area B (mm)	Error in experimental area C (mm)	Error in experimental area D (mm)
1,2,3,8,10,11,16,18,19	9	Uniform distribution	5.2	5.3	5.5	5.3
2,8,9,10,11,18	6	Diamond	6.6	7.1	6.8	8.8
1,3,8,11,16,19	6	Rectangle	10.9	9.1	10.3	12.4
2,8,11,18	4	Diamond	20.2	25.0	21.0	28.6
1,3,16,19	4	Rectangle	24.5	30.7	27.5	33.3
1,11,16	3	Triangle	55.9	54.4	58.3	68.2

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003

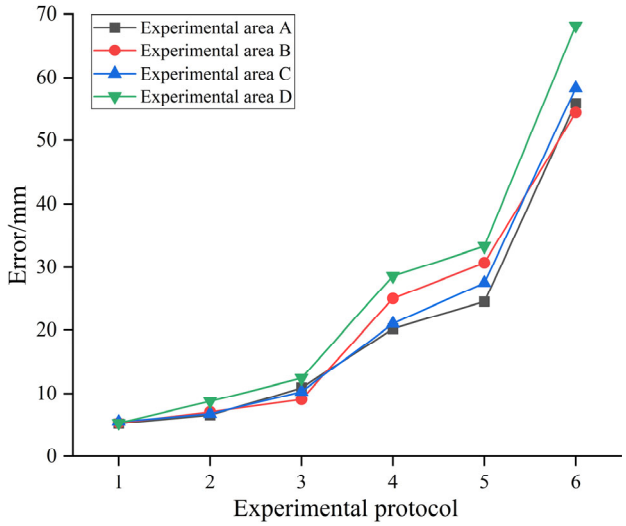


Figure 4. Errors in each experimental protocol

As shown in Table 2 and Figure 4, overall, under the same conditions of point quantity and distribution, the mean square error generally follows the pattern of Experimental Area D > Experimental Area C > Experimental Area B > Experimental Area A. This indicates that Experimental Area D has the worst calibration accuracy, while Experimental Area A is relatively optimal. Considering the terrain conditions of each experimental area, Experimental Area A is a flat and open area with small topographic relief and a vast open surface, low vegetation coverage, and relatively good monitoring conditions; Experimental Area D is a mountainous dense forest area with significant topographic relief and a dense surface vegetation, high vegetation coverage, and relatively poor monitoring conditions. With changes in the schemes, the error gap between different areas is further magnified. Comparing Experimental Area A and Experimental Area D, the accuracy gaps for schemes 1 to 6 are 0.1mm, 2.2mm, 1.5mm, 8.4mm, 8.8mm, and 12.3mm, respectively. This indicates that the requirements for control point layout are more stringent in complex environments, and the calibration accuracy is relatively lower.

Based on the calculation results, the number of points is the core factor affecting the mean square error, and the calibration accuracy decreases as the number of participating points decreases. When the number of points decreases from 6 to 4, and then to 3, the error does not increase linearly but exhibits a step-like jump, indicating that when the number of points falls below a certain threshold, the monitoring accuracy deteriorates sharply. Specifically, when the number of participating points decreases to 4 or fewer, the calibration accuracy drops to the centimeter level; with the same number of points, different point distribution patterns have a significant impact on accuracy. Among them, uniform distribution is optimal, indicating that uniformly distributing control points can effectively cover the monitoring area and is the most accurate arrangement method. The calibration accuracy of a diamond distribution is superior to that of a rectangular distribution, and when the number of points is 6, the calibration accuracy of the diamond distribution reaches the millimeter level. The reason for this is that the diagonal crossing structure of the diamond distribution provides more spatial geometric constraints in various directions, which helps to improve the stability and accuracy of the adjustment calculation and reduce rotation and translation errors in point cloud registration; moreover, the symmetrical structure of the diamond ensures more balanced coverage of the control points over the survey area, avoiding the error accumulation problem that may occur in the long side direction of rectangles, and ensuring the consistency of overall point cloud accuracy. The calibration accuracy of a triangular distribution is the worst, indicating that in the

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202604_12(4).0003

case of very few points, this simple geometric distribution cannot provide sufficient constraints, resulting in a significant decrease in accuracy.

4. CONCLUSION

Experimental area D, representing a complex environment, is the most sensitive to the deployment scheme and exhibits the worst accuracy. The 9-point uniform distribution is the scheme with the highest accuracy. A decrease in the number of points leads to a sharp decline in accuracy, and the 3-point scheme no longer has practical value. When the number of points is limited, the 6-point diamond distribution outperforms the rectangular distribution, achieving calibration accuracy down to the millimeter level. The diamond distribution is superior to the rectangular distribution in terms of geometric strength and error control. Moreover, in irregular or undulating terrain, the diamond distribution can better conform to the terrain contour by adjusting the vertex positions, avoiding local accuracy degradation caused by uneven distribution of control points. The diamond distribution is suitable for point cloud calibration in high-precision, complex scenes.

ACKNOWLEDGMENTS

The authors sincerely thanks all those who have provided assistance during the research and thesis writing process. Special thanks go to the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

[1] Chen Y, Medioni G. Object modelling by registration of multiple range images[J]. Image and vision computing, 1992, 10(3): 145-155.

[2] Segal A, Haehnel D, Thrun S. Generalized-icp[C]//Robotics: science and systems. 2009, 2(4): 435.

[3] Magnusson M, Lilienthal A, Duckett T. Scan registration for autonomous mining vehicles using 3D-NDT[J]. Journal of Field Robotics, 2007, 24(10): 803-827.

[4] Myronenko A, Song X. Point set registration: Coherent point drift[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 32(12): 2262-2275.

[5] Zhou Q Y, Park J, Koltun V. Fast global registration[C]//European conference on computer vision. Cham: Springer International Publishing, 2016: 766-782.

[6] Koide K, Yokozuka M, Oishi S, et al. Voxelized GICP for fast and accurate 3D point cloud registration[C]//2021 IEEE international conference on robotics and automation (ICRA). IEEE, 2021: 11054-11059.