Automatic extrinsic calibration for structured light camera and repetitive LiDARs

Yangtao Ge [iD], Chen Yao [iD], Zirui Wang, Bangzhen Huang, Haoran Kang, Wentao Zhang, Zhenzhong Jia and Jing Wu [iD]

Department of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen, Guangdong Province, P R China
**Corresponding author:** Jing Wu; Email: wuj@sustech.edu.cn

Abstract

The integration of camera and LiDAR technologies has the potential to significantly enhance construction robots' perception capabilities by providing complementary construction information. Structured light cameras (SLCs) are a desirable alternative as they provide comprehensive information on construction defects. However, fusing these two types of information depends largely on the sensors' relative positions, which can only be established through extrinsic calibration. This paper introduces a novel calibration algorithm considering a customized board for SLCs and repetitive LiDARs, which are designed to facilitate the automation of construction robots. The calibration board is equipped with four symmetrically distributed hemispheres, whose centers are obtained by fitting the spheres and adoption with the geometric constraints. Subsequently, the spherical centers serve as reference features to estimate the relationship between the sensors. These distinctive features enable our proposed method to only require one calibration board pose and minimize human intervention. We conducted both simulation and real-world experiments to assess the performance of our algorithm. And the results demonstrate that our method exhibits enhanced accuracy and robustness.

1. Introduction

Automatically assessing the quality of buildings is a significant concern of researchers in the field of construction robotics. This evaluation includes crack, evenness, alignment, and hollows. In the field of intelligent perception, numerous autonomous robots are utilized, including wheeled [1] and legged robots [2, 3]. However, wheeled robots are commonly utilized in construction robotics due to their superior stability during movement. By implementing sensors capable of capturing corresponding defect information, robots can assist workers in resolving these issues. Nonetheless, the detection results of individual sensors are only applicable to themselves and cannot be integrated into the robot system, resulting in the inability to accurately obtain the location and quantity of defects [4]. Fig. 1 illustrates the construction quality inspection robot we have developed, equipped with a structured light camera(SLC), a thermal camera, and a LiDAR. The SLC is capable of capturing the environmental texture and conducting 3D measurements to effectively detect cracks and flatness. The thermal camera, accurately connected with the SLC through location holes, detects hollow relying on variances in the thermal capacity of building materials. Meanwhile, LiDARs can perceive the indoor environment and generate a prior map for construction robots. Thus, it is essential to appropriately align these sensor data for the construction quality inspection robot by LiDAR-SLC extrinsic calibration.

Yangtao Ge and Chen Yao contributed equally to this paper.
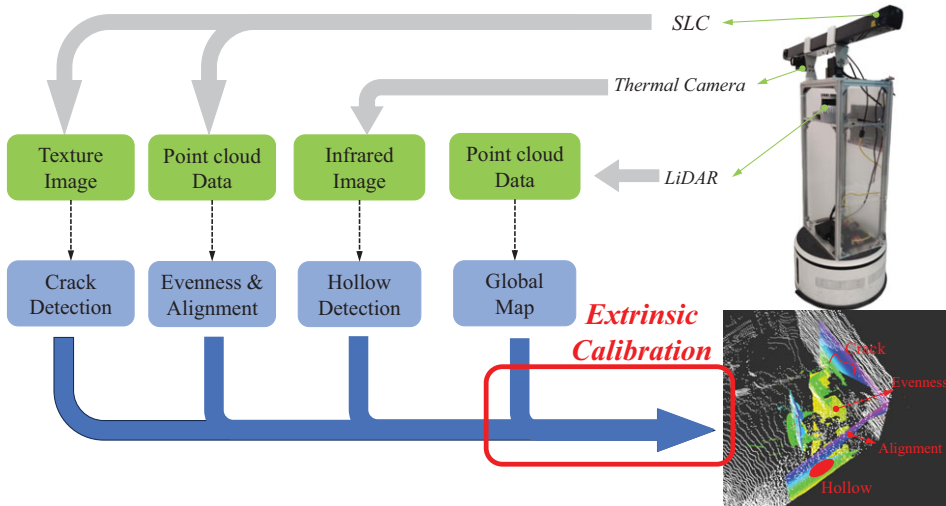
Figure 1. *The total workflow of the quality inspection robot system.*

Extrinsic calibration, which refers to the method of aligning different types of data, entails processing data obtained from individual sensors to estimate the relative position and orientation among them. Generally, the calibration procedure involves three essential steps: (1) feature extraction, (2) matching strategies, and (3) optimization methods. Depending on the method of feature extraction, calibration can be classified into two categories: target-based methods and target-less methods. Based on the process of feature extraction, calibration can be classified into two categories: target-less and target-based methods. The primary contrast between the two methods lies in the requirement of a calibration board. Features for the non-target method are gathered from the surrounding environment, whereas the calibration board-based method requires artificial setup to extract reference features.

1.1. Target-less methods

Target-less methods aim to incorporate a range of techniques to extract feature information from the environment and establish the appropriate sensor relationships. The available extrinsic parameters can be determined by projecting two types of features onto the same coordinate system and by minimizing the error between them.

Edge extraction is a popular method for extrinsic calibration due to its simplicity. It combines gradient changes of image pixels with the discontinuity or continuity of the LiDAR point cloud [5]. Zhang et al. extracted pole line characteristics from images and point clouds and determined the appropriate extrinsic parameters by developing a synthesized cost function in both horizontal and vertical directions [6]. They obtained line features from the point cloud and used an adaptive optimization method to calculate the calibration results. Additionally, researchers also explored the use of sensor intensity as a feature. For example, [7] used the statistical similarity of object surface intensities as feature information and obtained optimal extrinsic parameters for cameras and 3D LiDARs. However, the accuracy of intensity information may not be guaranteed due to environmental factors such as illumination.

Machine learning is a powerful approach for problem-solving due to its capability of handling diverse and numerous features, as well as the continuous development of computer technologies. RegNet [8] and CalibNet [9] are two prominent techniques for joint calibration of LiDARs and camera. RegNet can generate annotated data automatically and use the iterative refinement calibration method to cope with large variances. Nevertheless, this process is time-consuming, and the feature extraction-matching ability is restricted. Conversely, CalibNet incorporated a corresponding loss function into the network

to accommodate the point cloud geometry. However, the devised training strategy limited CalibNet's further development. To address these limitations, LCCNet [10] was proposed and performed exceptionally well. Additionally, features that are semantically segmented from images and point clouds can be utilized as feature points. Wang et al. [11] utilized the centroid of semantics with identical labels from image and point cloud data as reference points for the sensors; however, the efficacy of this approach heavily depends on the semantic segmentation outcomes.

1.2. Target-based methods

Target-based methods artificially define features that units such as cameras and LiDARs can recognize as reference points. These marks are utilized to associate the sensors, which subsequently transforms the calibration of extrinsic parameters into a Perspective-n-Points(PnP) [12] or an optimized problem [13].
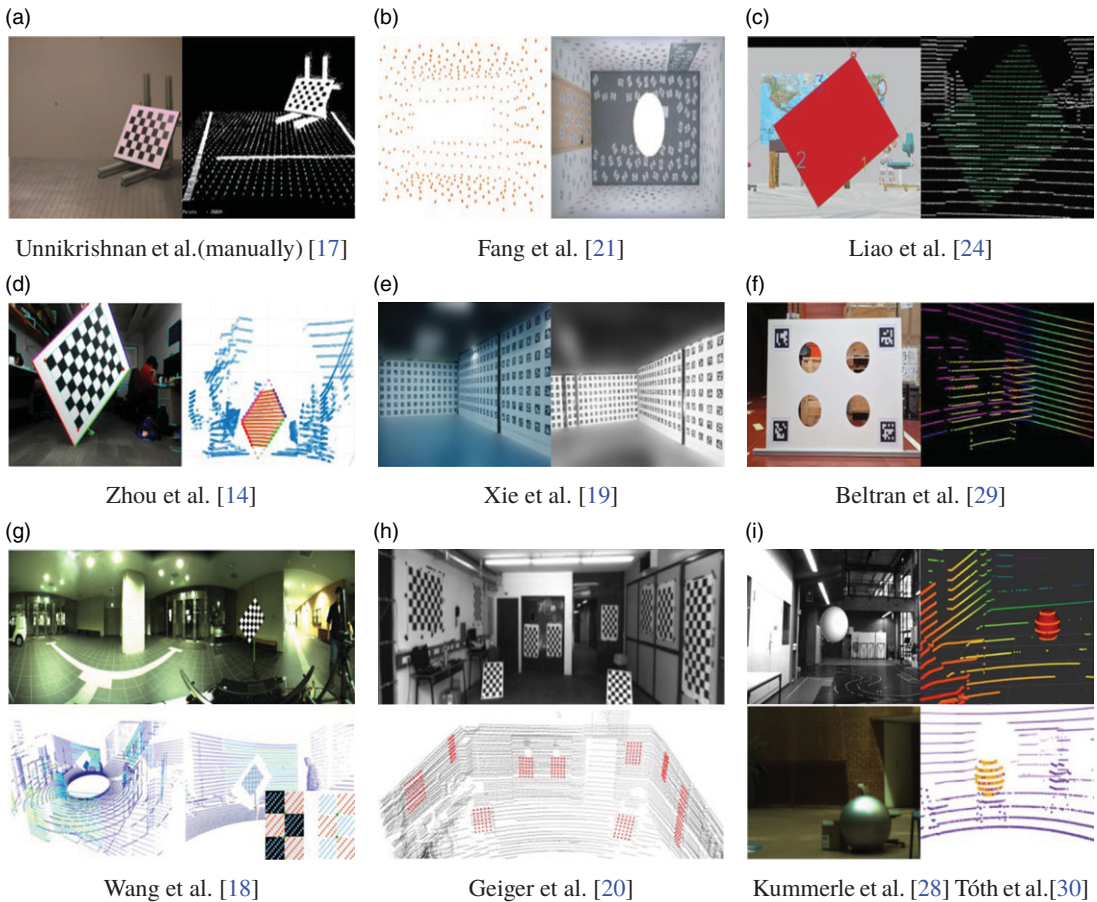
1.2.1 Single chessboard

One simple option to obtain a feasible solution is to directly use a single chessboard and employ its geometric constraints [14] or intensity [15]. Q. Zhang et al. [16] were pioneers who employed a planar checkerboard for camera and 2D LiDAR calibration, taking into account plane-line correspondences as constraints for the extrinsic parameters. However, this technique fails to achieve proper calibration accuracy because a limited number of constraints from single-frame data are insufficient for calibration, and the unstable accumulation trajectory for multi-frame data produces uncertain results. A chessboard-based calibration algorithm for cameras and 3D LiDARs is presented in Fig. 2(a). They acquire coarse parameters through plane-plane correspondences and employ point-plane constraints to enhance accuracy [17]. This method entails separate stages of data collection and processing, which requires an continuous user interface throughout the entire process. W. Wang et al. [18] utilized the correlation between the intensity of the point cloud and the color of the checkerboard to identify feature points from the detected corner points, as depicted in Fig. 2(g). However, the extrinsic calibration of the panoramic camera and 3D LiDAR sensors is unstable using this approach as the intensity is affected by factors other than color.

1.2.2 Multiple chessboards or markers

Affixing multiple chessboards or markers to an indoor setting is an extension of the calibration board technique. The placement of multiple chessboards or markers [19] within an indoor setting is an extension of the calibration board method. While these techniques require merely a single scene shot, they entail manually attaching the chessboards within the room before calibration can take place. In Fig. 2(h), multiple cameras were associated with 3D range sensors utilizing the normal vectors of multiple affixed checkerboard patterns as features, resulting in acceptable outcomes in a single shot [20]. The panoramic infrastructure, as shown in Fig. 2(b), localizes and connects sensors using the pasted marks and room corners to achieve single-shot calibration [21]. Though these methods exhibit simplicity and user-friendliness, their preparation involves significant labor and lacks flexibility. These limitations can make it challenging for dynamic systems that require frequent recalibration or when there are changes in the environment.

1.2.3 Novel calibration board

Various calibration boards with novel shapes have been proposed to generate more robust reference points, such as triangles [22], polygons [23, 24], circles [25, 26], and spheres [27, 28]. These designs provide distinguishable characteristics for various sensors. For example, a calibration method for binocular and monocular cameras [29], as well as LiDARs, was proposed using a board with four circular holes and markers, as depicted in Fig. 2(f). The appropriate arrangement of the calibration board is crucial for achieving accurate results with this method. T. Tóth et al. [30] employed spherical objects as targets and

Figure 2. *Classification of target-based methods according to different types of calibration board. Single chessboard: (a) (d) (g); Multiple chessboard or marks: (b) (e) (h); Novel shaped boards: (c) (f) (i). Top/left: 2D image; bottom/right: 3D point cloud.*

synchronized the monocular camera and LiDARs with reference points derived from the fitting of point cloud, as shown in Fig. 2(i). Nonetheless, setting up calibration scenes can be a challenging task and may not guarantee high accuracy. For simpler tasks that are not highly demanding in terms of precision, researchers may manually select feature points [17, 31]. Although feature points selected manually by these methods are robust, they are still susceptible to human error and lack complete automation.

1.3. Challenges

SLCs are an excellent alternative for quality inspection, as they can provide accurate and detailed information over a range of areas. Therefore, we can effectively utilize this specific attribute to address the extrinsic calibration problem within LiDAR-structured light camera systems. Nevertheless, current calibration techniques mainly focus on conventional cameras, and adapting them to SLCs and LiDARs can cause various issues:

1.3.1 Low-textured environments

Although environmental feature association [5, 6] is a convenient method for aligning the cameras and LiDARs, it may not be applicable for SLCs due to low-textured environments and the poor anti-interference ability of the cameras. Consequently, the implementation of existing target-less calibration
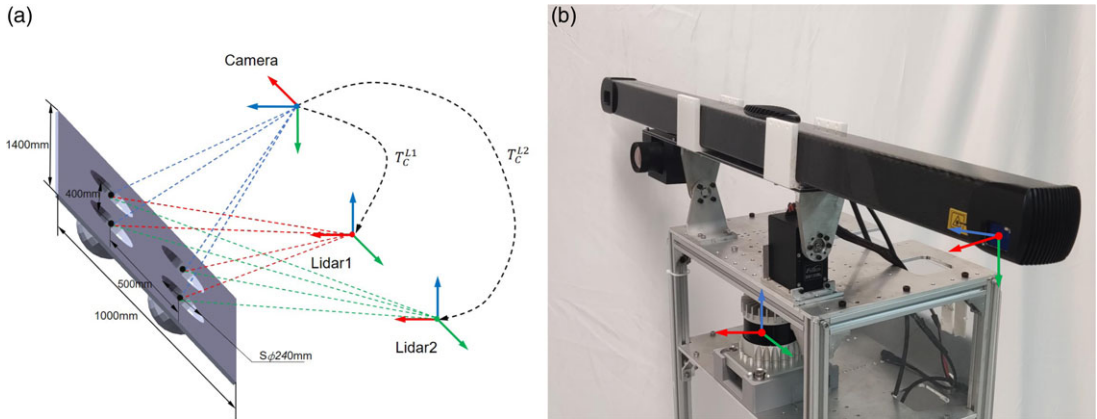
Figure 3. *(a) The customized calibration board is placed in the overlapped view of two sensors to provide additional reference points for each sensor pair. (b)The real-world SLC-LiDAR sensors.*

algorithms based on environmental features, including learning-based methods, may lead to ineffective alignment results.

1.3.2 Human intervention

To achieve accurate and efficient calibration, a single chessboard and its extended methods are insufficient due to significant human intervention, such as attaching marks [21] or setting up multiple objectives [25, 30], which renders the calibration process difficult to implement.

1.3.3 Characteristics of SLCs

SLCs are an ideal choice for building quality inspection as they provide dense point cloud information for detecting minor defects in construction. However, SLCs require relatively static conditions in order to successfully capture accurate point clouds. Therefore, employing multiple poses of calibration boards to enhance the accuracy of extrinsic calibration would be an exceedingly laborious process.

1.4. Contributions

Considering the above challenges, we propose a novel calibration method utilizing a custom hemispherical board to spatially align the LiDAR-SLC systems. The evenly distributed centers of the hemispheres on the calibration board serve as reference points for associating the two sensors, as shown in Fig. 3. Meanwhile, the reference points are adjusted to ensure a more precise joint through point-plane and point-point constraints derived from the calibration board. Instead of directly employing Iterative Closest Points (ICP) approaches, registration and optimization strategies are employed separately to estimate extrinsic parameters quickly. Specifically, our contributions can be summarized as follows:

1. We propose an automatic method for extracting feature points to calibrate SLCs and LiDARs. This method provides superior anti-interference capability as the feature points are derived by the fitted sphere centers, rather than corners or boundary points.
2. We introduce an enhanced calibration board with geometric constraints that improves the accuracy of extracting feature points. Additionally, the calibration can be completed with just a single board position, minimizing human intervention as much as possible.
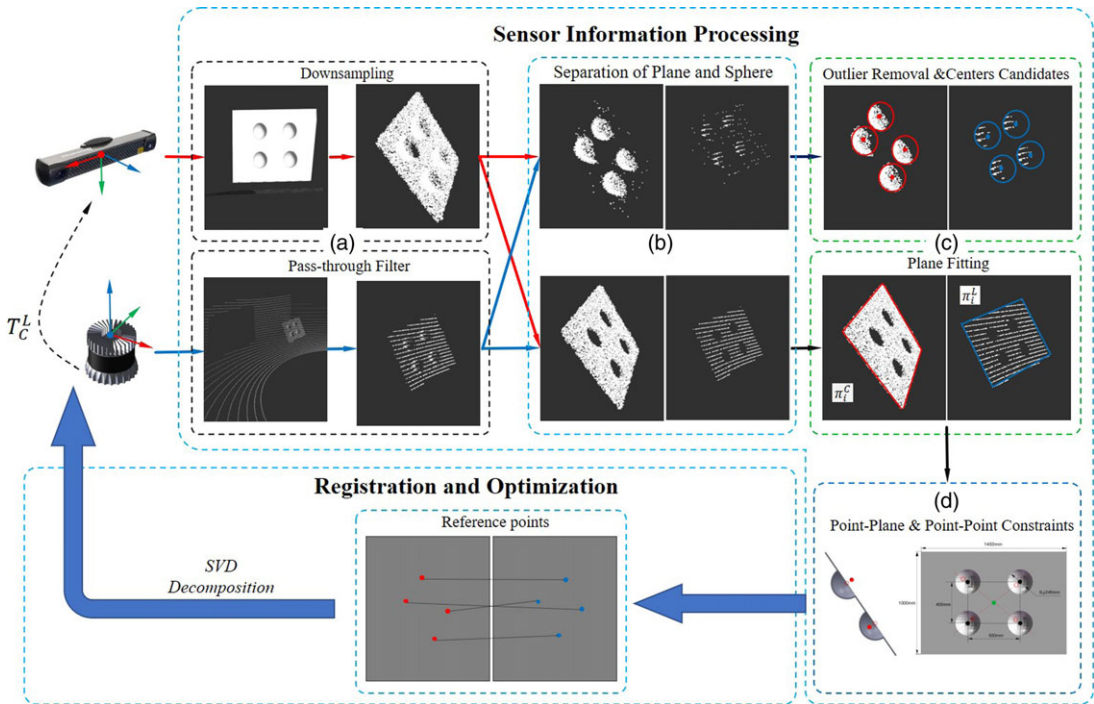
Figure 4. *The pipeline of our calibration method.*

3. We validated the advantages of proposed calibration algorithm through a comprehensive series of simulations and real experiments, being suitable for construction robotics applications.

The remainder of this manuscript is organized as follows: Section 2 describes the proposed calibration method in detail. In order to validate the accuracy and robustness of the algorithm, we conducted a set of simulations and real-world experiments in Section 3. Finally, Section 4 presents a summary of the research conducted in this paper, as well as a prospective analysis of future research.

2. Methodology

Our calibration approach comprises two main parts: (i) sensor information processing and (ii) registration and optimization. The former step involves collecting raw data from the Structured Light Cameras(SLCs) and LiDARs, and extracting the designated features through several processing steps. The latter consists of aligning the extracted reference points and performing an appropriate optimization process to determine the optimal extrinsic parameters. The pipeline of the calibration methodology is illustrated in Fig. 4: The sensor data processing can be roughly divided into four stages: downsampling and filtering, plane and spherical segmentation, outlier removal, and candidate point optimization and adjustment. These stages provide effective reference points for the subsequent optimization.

2.1. Problems formulation and assumption

The extrinsic calibration problem involves finding the relative position and orientation of a camera and a LiDAR sensor mounted on a common platform. This can be achieved by estimating a transformation matrix T_L^C, which aligns the LiDAR and camera coordinate frames. The goal of the calibration process is to minimize the distance between the corresponding points in the two frames.

The four spherical centers, derived from our custom calibration board as shown in Fig. 3(a), serve as the reference points. Our calibration board is a $1000 \times 1400mm$ rectangle with four hemispheres distributed symmetrically at a position of $400 \times 500mm$. These hemispheres, which have a diameter of 240 mm, efficiently gather information from the sensor and enable accurate fitting of the reference points. p_i^C and p_i^L are the reference points, where C and L denote the camera and LiDAR's coordinate system. The extrinsic calibration problem can be described by the following formula:

$$P^C = R_L^C P^L + t_L^C \tag{1}$$

where $T_L^C = \left[R_L^C; t_L^C \right]$, $P^C = \{p_1^C, p_2^C, \cdots\}$, $P^L = \{p_1^L, p_2^L, \cdots\}$, and R_L^C and t_L^C are the rotation and translation parameters that describe the relative pose of the LiDAR from the camera frame. SLCs capture three-dimensional geometric information, including points, shapes, surface colors, and other attributes in space. These data can be represented in three modalities: RGB-texture, Depth-map, and Point cloud. Here, high-precision point cloud information is chosen as input, eliminating the need to consider the camera's intrinsic parameters, which facilitates subsequent optimization processes.

To solve for T_L^C, we need to find the values of R_L^C and t_L^C that minimize the distance between the corresponding reference points in the two frames. This can be formulated as an optimization problem, where we minimize the sum of the squared distances between the corresponding points:

$$\min \sum_i^N \left\| p_i^L - T_L^C p_i^C \right\|^2 \tag{2}$$

Upon solving the optimization problem, the transformation matrix T_L^C can be determined, which means that the extrinsic calibration between SLCs and LiDARs is completed.

2.2. Sensor information processing

The primary objective of this step is to extract the four centroids of the predefined hemisphere and optimize the adjustment of the hemisphere centers by applying geometric constraints derived from the calibration board, thus obtaining precise reference points. This step comprises four main components: filtering and downsampling, spherical and planar segmentation, spherical fitting, and geometric constraints for candidate centers. To enhance clarity in this section, the following symbols have been defined: $P_{\{\},\{\}}^{\{\}}$ represents a point cloud cluster, where the top right corner {} denotes the corresponding camera or LiDAR coordinate system and the bottom left corner {} , {} indicates which point belongs to the spherical or planar point cloud. $[n;d]$ and $\pi^{\{\}}$ represent planar models, where n and d respectively denote the normal vector of the plane and a point on the plane. $[p, R]$ represents a spherical model, where p and R respectively represent the center and the radius of the sphere.

2.1. Filtering and downsampling

The sensors capture raw data from various sources, including the calibration board, floor, and wall. As the feature points are derived from the calibration board, we apply pass-through filtering to the original data to preserve the board P_b^L and P_b^C, as demonstrated in Fig. 4(a). It is noteworthy that the threshold of the pass-through filter should be adjusted for varying situations. Nevertheless, sparse sampling is essential for cameras because they generate a large number of high-precision point cloud. Since high-precision dense point cloud has been chosen as the camera data, sparse sampling is essential to ensure optimal performance. Additionally, the dense point cloud is uniformly divided into small cubes τ_{sample} to preserve the geometric characteristics of the calibration board. The geometric center of the cube is chosen to represent the point cloud within the small cube, which prevents errors generated by downsampling.
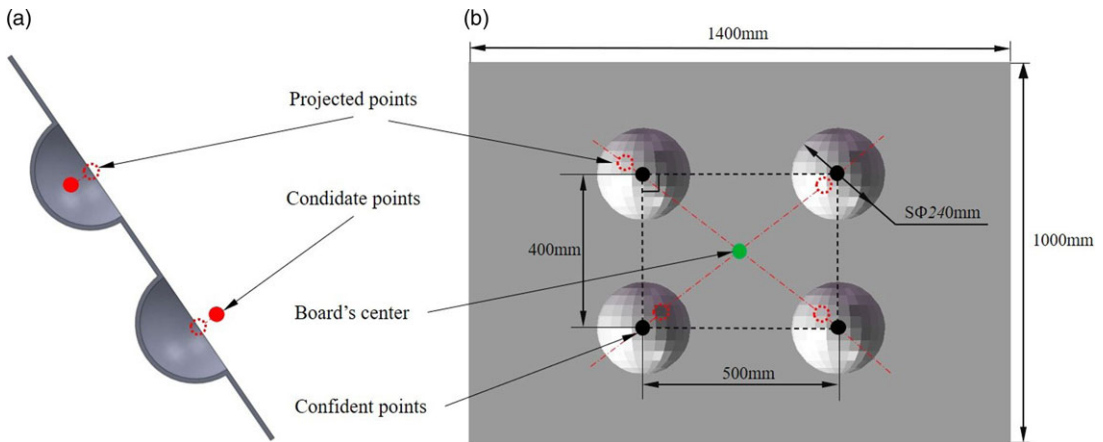
Figure 5. *Geometric constraints on calibration board. (a) The fitted centers are projected onto the plane of calibration board; (b) Adjustment of the distance from the projected points to the approximate point of the center of the rectangle.*

2.2. Spherical and planar segmentation

The calibration board's point cloud is subjected to segmentation to decompose it into two segments: a planar point cloud P_p and a spherical point cloud P_s as shown in Fig. 4(b). The planar model of the calibration board π^c and π^l, is generated through the Random Sample Consensus method(RANSAC), with the model's parameters represented by $[n_p^C; d_p^C]$ and $[n_p^L; d_p^L]$ respectively. The planar point cloud P_p comprises of points located within the threshold δ_{plane} of the model, while the spherical point cloud P_s contains all remaining points. The spherical point cloud will have some unexpected outliers due to threshold values and the presence of noise, which is detrimental to subsequent classification and spherical center fitting accuracy. Therefore, statistical filtering is employed to eliminate outliers, which are defined as points with an Euclidean distance greater than one standard deviation from the mean, to generate a clear spherical point cloud.

2.3. Spherical fitting

In order to simplify the process of fitting reference points, it is recommended that we perform Euclidean clustering on the clear spherical point cloud. Setting the Euclidean clustering threshold $\delta_{cluster,s}$ at an appropriate level will lead to the creation of four-cluster point clouds $p_{s,j} \in P_s, j \in \{1, 2, 3, 4\}$, each of which corresponds to one of the four hemispheres on the calibration board. Subsequently, the spherical models $[p_{c,j}; R_j]$ can then be obtained by separately spherical fitting with RANSAC method and a tolerable threshold δ_{sphere}. These candidate reference points are the spherical centers represented by $p_{c,j}$, as seen in Fig. 4(c).

2.4. Geometric constraints for candidate centers

The sensor-derived data invariably contain noise due to various environmental and sensor factors. Furthermore, the point clouds generated by repetitive LiDARs on a hemisphere consist of only a few lines, resulting in a lower density compared to that of SLCs. This also impacts the fitting of the candidate reference points. These two aspects can cause a substantial deviation between the final calibration result and the actual value.

Therefore, it is essential to utilize the calibration board's available characteristics to optimize the candidate spherical centers in Fig. 4(d). Since the hemispherical surfaces are located on the calibration board's plane, it is reasonable to project the candidate centers onto the plane, as illustrated in Fig. 5(a). The fitted candidate points are indicated by red solid circles, and the points projected onto the calibration

Table I. *The sensors' parameters in the experiments.*

Devices	Modality	Resolution	HFOV
Ouster HDL-64	LiDAR	64 layers, 0.2°	360°
Velodyne HDL-32	LiDAR	32 layers, 0.2°	360°
Velodyne VLP-16	LiDAR	16 layers, 0.2°	360°
PhoXi 3D Scanner XL	Structured light camera	1954 × 1459	45.57°

board are represented by dashed circles. Depicted in Fig. 5(b) is the standard geometric arrangement of the four hemispheres, which implies that the spherical centers are equidistant from the calibration board's center. We determine the center of the calibration board by taking the average of the four projected points. Subsequently, each candidate spherical center is adjusted in the direction of the board's center until it reaches the actual value. Fig. 5(b) shows the selected calibration board with a green dot representing its center and the black dots indicating updated confidence points. Finally, the black dots $\pi^c\left(p_{c,j}^C\right) \pi^l\left(p_{c,j}^L\right), j \in \{1, 2, 3, 4\}$ serve as the reference points that we require.

2.3. Registration and optimization

The second stage aims to determine the rigid body transformation T_L^C between the cameras and LiDAR coordinate systems by utilizing the reference points obtained in the previous steps. Since the above procedures rely on single-frame data, it is possible to accumulate N_{acc} frames for a single calibration board position. The sets of sphere centers acquired from the point clouds can serve as reference points between the two sensors. The loss function can be established easily by referring to the problem definition described previously:

$$\arg\min_{R,t} \sum_{i=1}^{4 \cdot N_{acc}} \sum_{j=1}^{4} \left\| \pi^l\left(p_{c,j}^L\right) - R\pi^c\left(p_{c,j}^C\right) - t \right\|_2^2 \tag{3}$$

where $R^T R = I$, the rigid body transformation from LiDARs to cameras, denoted as T_C^L, can be described by a rotation matrix $R \in \mathbb{R}^{3 \times 3}$ and a translation vector $t \in \mathbb{R}^3$. The estimation of the transformation matrix between p_i^L and p_i^C is typically achieved using the widely used ICP method. However, the customized calibration board has unique characteristics that enable us to properly sort the sphere centers based on their inclination angles from the origin of the coordinate system, ensuring that the points of p_i^L and p_i^C are appropriately associated. Once the association mentioned above is established, we can determine the optimal values for R and t via singular value decomposition of the loss function defined in Equation (3). This approach not only obviates the necessity of an initial guess and iterative optimization but also improves the efficiency of the calibration algorithm.

3. Experiments
3.1. Experimental setup
Our proposed algorithm is evaluated both on the simulation and real-world datasets. The simulated sensor suit is built on the Gazebo [32] that incorporates sensor models with actual parameters. It consists of simulated 16-beam, 32-beam, 64-beam LiDARs, and a SLC. For the real-world experiments, we conducted experiments in various environments using our mobile platform designed for building quality inspection, which is equipped with an Ouster-64 LiDAR and a Photoneo scanning camera. Table I presents the sensors utilized in our experiment along with their associated parameters.

Table II. *The setting parameters of our algorithm.*

Main steps	Parameter	Description
Filtering and Downsampling	$\tau_{sample}^{C} = 2cm$	Density of downsampling(Structured Light Camera)
Spherical and Plane Segmentation	$\delta_{plane}^{C} = 2.5cm$	Distance threshold(Structured Light Camera)
	$\delta_{plane}^{L} = 2.8cm$	Distance threshold(LiDAR)
	$N_{meanK} = 50$	The number of adjacent points
Spherical Fitting	$\delta_{cluster,s} = 12cm$	Euclidean clustering threshold
	$\delta_{sphere}^{C} = 10cm$	Distance threshold for spherical fitting()
	$\delta_{sphere}^{L} = 12cm$	Distance threshold for spherical fitting(LiDAR)
	$D_{adjust} = 16cm$	Reference distance for sphere center adjustment
Registration and Optimization	$N_{acc} = 30$	Accumulated data frames

3.2. Performance evaluation

To enhance the representation of the calibration algorithm's accuracy, we compare the acquired extrinsic parameters with the ground truth (GT). The calibration error consists of two parts, rotation error and translation error, which are specifically expressed as follows [20]:

$$e_t = \left\| t - t_g \right\|_2$$

$$e_r = \angle \left(R^{-1} R_g \right) \tag{4}$$

where t_g and R_g are the GT derived from the settings of the sensors in the simulated environment. t_g is generated by the translation vector $\left(t_x, t_y, t_z \right)^T$, while the rotation matrix R_g is represented as a combination of roll, pitch, and yaw angles $(\varphi_x, \theta_y, \phi_z)$. e_t represents the Euclidean distance between the measured value and GT, while e_r is the minimum rotation error on the three axes.

3.3. Calibration results on simulated data

We first verify the LiDAR-frame camera calibration with simulated data. In our experiments, we selected three different resolutions of LiDARs for extrinsic calibration with the SLC. The setting parameters for the main steps of our method, mentioned in the second 2, are shown in Table II.

In order to evaluate the effectiveness of our method, we conducted two types of experiments in a simulated environment: *1) single-sensor experiments* and *2) synthetic experiments*. The former was intended to examine the accuracy of the extracted reference points by changing the location of the calibration board. The second experiment provided a comprehensive evaluation of the algorithm, focusing on the accuracy and robustness of the calibration results. We also compared our method with the algorithm proposed by C. Guindel et al. [29], using ROS implementation. For the fairness of the experiment, the sensors were substituted with SLCs and LiDARs of varying resolutions to assess the applicability of the algorithm. This means that we can compare the performance of different algorithms under the same environment and sensor conditions. In simulation experiments, the GT can be easily obtained from Gazebo.

3.3.1. Single-sensor experiments

In this section, we aim to analyze the precision of the fitted spherical centers for individual sensors by varying the rotation angle of the calibration board. The relative position from LiDAR to SLC is assumed to be (0.1, 0.2, 0.5, 0, 0,0), which corresponds to $\left(t_x, t_y, t_z \right)$ and $(\varphi_x, \theta_y, \phi_z)$, respectively. Additionally, the
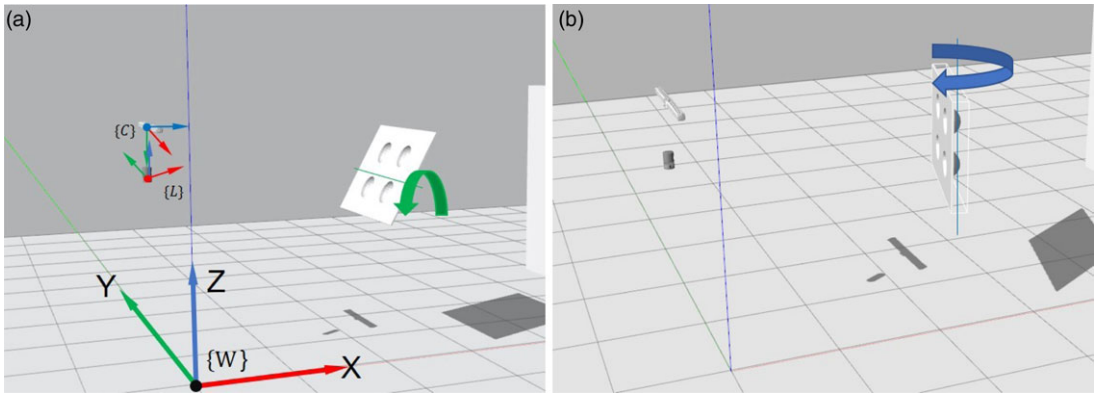
Figure 6. *(a) and (b) represent the rotation of the calibration board around the y-axis and the z-axis.*

center of the calibration board is placed at coordinates (2.2, 0, 1.8). This location is randomly selected from the overlap, because as long as the camera and LiDAR are within overlap, the results will be similar. The calibration board is tilted along the y-axis by 0 to 45 degrees and rotated along the z-axis from −45 to 45 degrees, with a 5-degree interval between each trial, as shown in Fig. 6.

Fig. 7 shows the Euclidean distance error between the fitted spherical centers and the actual ones at each corresponding angle. The proposed method provides more accurate reference points than the compared algorithm in both the camera and LiDAR's coordinate systems. It is worth noting that the compared algorithm's reference points deviate greatly when the rotation angle of the calibration board exceeds 30 degrees. This implies that the proposed algorithm can effectively find reference features regardless of the placement of the calibration board. Importantly, the camera's data yields a more stable spherical center position than that of LiDAR's at varying rotation angles. This can be attributed to the fact that the LiDAR's point cloud on the hemispheres is often sparse, consisting of only a few lines. Furthermore, rotation of the calibration board may cause some lines to become unstable and, in turn, impact the subsequent fitting of spherical centers. Fig. 8 illustrates the translation and rotation errors at each position corresponding to Fig. 7. The trends for both errors are similar, indicating that the proposed method is capable of improving the precision of the fitted spherical centers. This, in turn, leads to more accurate calibration results and significantly enhances the performance of the proposed method.

3.3.2. Synthetic experiments
Accuracy test.   The objective of this experiment is to assess the accuracy of the proposed method relative to a comparative algorithm by testing different positions. We selected SLCs and 64-layer LiDARs as sensors for their ability to generate proper point clouds for the application algorithms. We assessed the effectiveness of our approach across ten distinct relative settings between the two sensors, accounting for both translation and rotations. Table III depicts the settings of each calibration pattern. The initial position where both sensors have a clear view of the calibration board is designated as setting 1. The parameters $t_x, t_y, t_z, \varphi, \theta, \phi$ describe the GT values of the SLC with respect to the LiDARs. Settings 2 to 5 and settings 6 to 8 involve only rotation or translation between the two sensors. Complicated scenarios have also been considered in the experiment, such as settings 9 and 10, where the rigid transformations of the two sensors combined both rotations and translations.

Table III presents the quantitative experimental results of our proposed method and the comparative algorithm obtained under ideal conditions without noise. The proposed method proved to be effective across all experimental settings. The translational and rotational errors consistently remained below 1 cm and 0.1 degrees, respectively. Conversely, the comparative algorithm produced unsatisfactory results, displaying significant errors during settings 3 and 9, as well as being unable to complete calibration
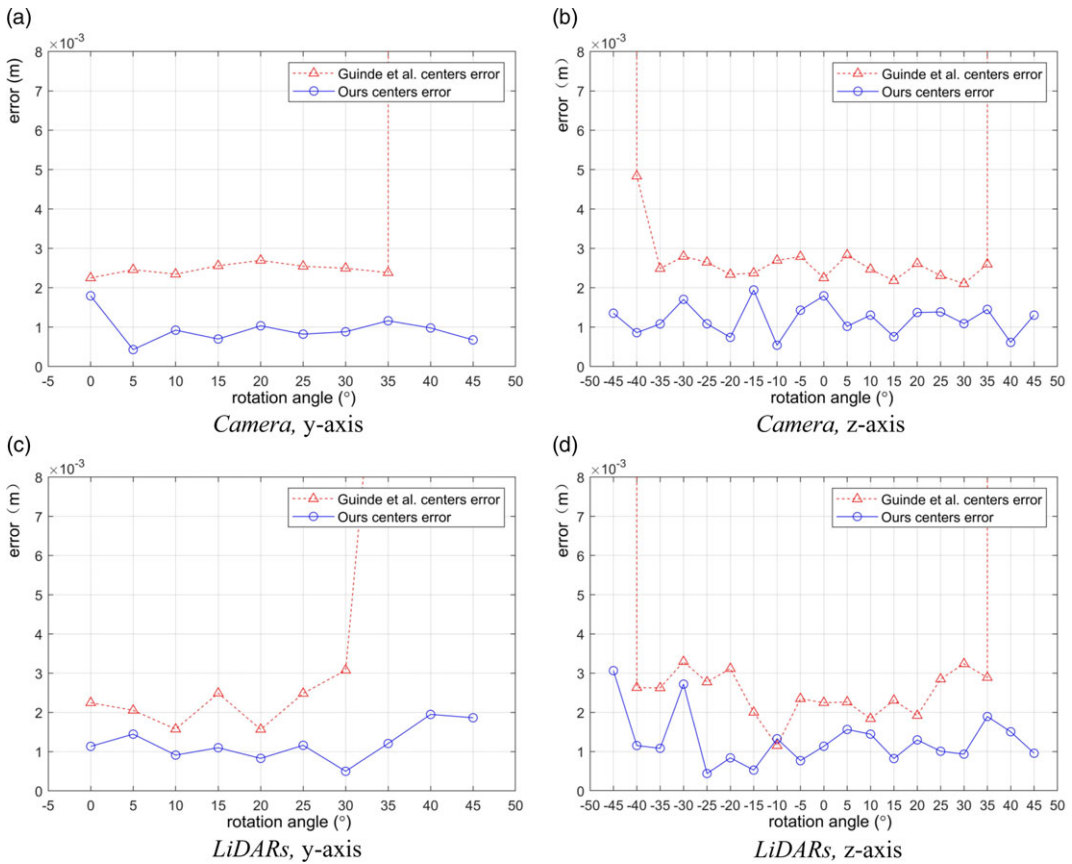
Figure 7. *Euclidean error of the centers in single frame. (a)(c) and (b)(d) demonstrate the error in the LiDARs and camera systems.*

during setting 5. These experimental results provide evidence of the superiority of our proposed method over the comparative algorithm, even in complex scenarios.

In order to present the error reduction achieved by the proposed algorithm in a more intuitive way, we validated the experimental results by conducting a reprojection experiment. Fig. 9 shows the reprojection outcomes for settings 1, 2, 6, and 9, which represent four different relative pose scenarios: the initial position, pure rotation, pure translation, and rotation plus translation. The white and red point clouds in the figure correspond to the cameras and LiDAR data. The degree of overlap between two point clouds indicates the accuracy of calibration. A higher degree of overlap corresponds to a smaller reprojection error, and therefore higher precision in the calibration results.

The results of setting 1 are displayed in Fig. 9(a)(e). The figures reveal that the proportions and shades of the two colors do not show a noticeable difference, suggesting that performance of both methods is similar to each other. Fig. 9(b)(f) reveals the outcomes of setting 2. The re-projection color of the proposed method is darker than the contrast algorithm, which indicates superior performance. The figures for setting 6, namely Fig. 9(c)(g), illustrate that the calibration board of contrast algorithm is lighter in the lower right corner, indicating a higher level of error compared to the proposed method. Fig. 9(d)(h) demonstrates the results obtained for setting 9. Specifically, the contrast algorithm calibration board shows a significant white area in the upper right corner. However, the proposed method continues to yield satisfactory performance. The re-projection results can fundamentally correspond to the errors in Table III, which provides an intuitive confirmation of the proposed algorithm's accuracy.
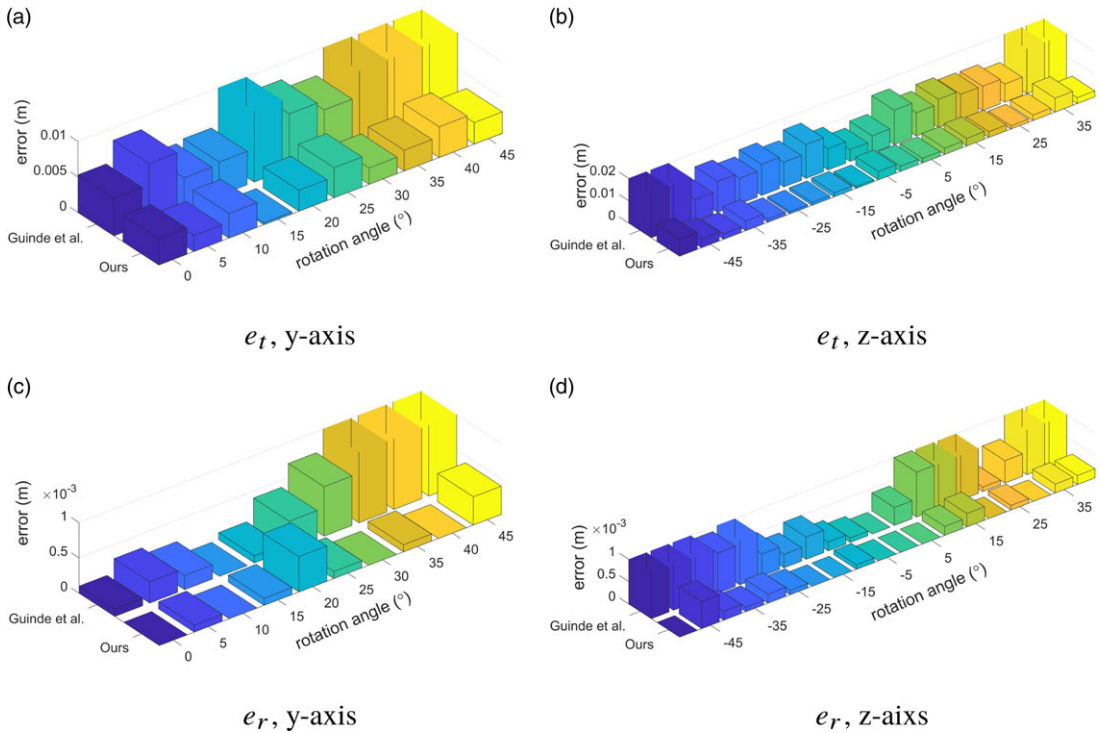
Figure 8. *The calibration result error at different rotation angles.*

Robustness test. We also evaluate the robustness of our algorithm by introducing Gaussian noise to the sensor data. The installation of these two sensors is situated in more challenging locations, specifically at the 10th setting, where these sensors assume a more complex relative pose. Additionally, we assessed the case of SLC and other different resolution LiDARs, such as VLP-16, HDL-32, and HDL-64. We simulate real-world scenarios by adding Gaussian noise $N\left(0, \delta_0^2\right)$ to the sensor measurements.

Each set of experiments was performed 20 times due to the variation in calibration results after adding noise to the sensor information. A statistical analysis of the data is shown in Fig. 10. Our experimental results demonstrate that our proposed method outperforms the comparison method in both translation and rotation errors. Specifically, the results from the proposed method are shown in the blue boxplot of Fig. 10. The mean of this method is closer to the zero baseline, with most of the data clustering around this value. In contrast, the red boxplot displays results from the compared method, with a larger spread of data and a mean further from the zero baseline. These differences indicate that our proposed method is more robust in terms of both translational and rotational errors. As the noise levels increase, the proposed algorithm consistently generates reliable output, while the comparison algorithm exhibits significant deviations.

3.4. Calibration results on real-world data

This section tests the proposed method in two real-world scenarios that represent the two most common types of scenes occurring in architectural settings, as shown in Fig. 11. The first scenario (S01) is a narrow-cluttered corner with other objects such as wall, prefabricated components, chairs, and aluminum profile racks, which affect the extraction of board information. The second scenario (S02) is a spacious and well-organized room, with little obstacles or interference. In this study, we conducted practical experiments to compare the proposed method with two other algorithms mentioned in the references [14, 29]. These particular algorithms have gained popularity in the open-source community due to their

Table III. *The sensors' parameters in the simulation experiments.*

| | | | | | | | | Experiment results | | | |
| | | | | | | | | Ours | | Guindel et al. | |
Categories	Setting	$t_x(m)$	$t_y(m)$	$t_z(m)$	$\varphi(rad)$	$\theta(rad)$	$\phi(rad)$	$e_t(m)$	$e_r(rad)$	$e_t(m)$	$e_r(rad)$
Initial	1	0.1	0.2	0.5	0	0	0	**0.0053**	**0**	0.0063	0.0001
Rotation	2	0.104	−0.051	0.535	0.912	−0.144	−0.171	**0.0014**	**0.0001**	0.0082	0.0019
	3	0.262	0.311	0.36	−0.656	0.014	−0.528	**0.0116**	**0.0007**	0.2009	0.0253
	4	0.279	0.082	0.464	0.124	0.311	−0.806	**0.0010**	**0.0002**	0.0042	0.0003
	5	0.251	0.067	0.482	0.420	0.077	0.118	**0.0045**	**0.0009**	NaN	NaN
Translation	6	−0.2	1	−0.15	0	0	0	**0.0041**	**0.0002**	0.0099	0.0003
	7	0.25	0.75	0.1	0	0	0	**0.0014**	**0.0007**	0.0110	0.0019
	8	0.1	0.22	0.55	0	0	0	**0.0009**	**0.0002**	0.0044	0.0004
Trans and Rot	9	−0.149	0.051	0.505	0.090	0.089	−0.110	**0.0007**	**0.0003**	0.1650	0.0114
	10	−0.009	0.709	0.321	−0.064	0.050	0.343	**0.0009**	**0.0002**	0.0027	0.0062
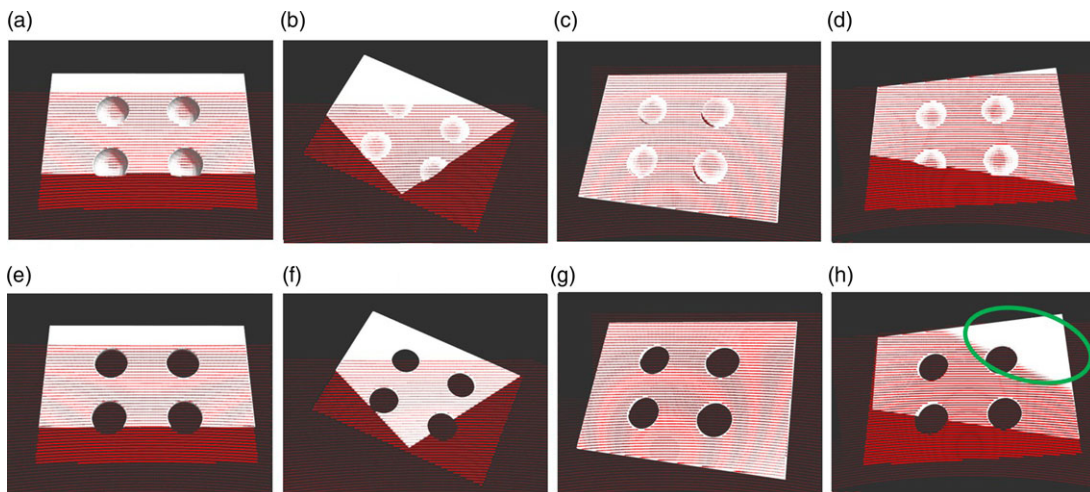
Figure 9. *The reprojection results of LiDARs to SLC are evaluated under four experimental settings. (a)(b)(c)(d) and (e)(f)(g)(h) are the reprojections from LiDARs to SLC under setting 1, 2, 6, and setting 9, respectively; (a)(b)(c)(d) are the qualitative results of proposed method and (e)(f)(g)(h) are the qualitative results of compared method.*
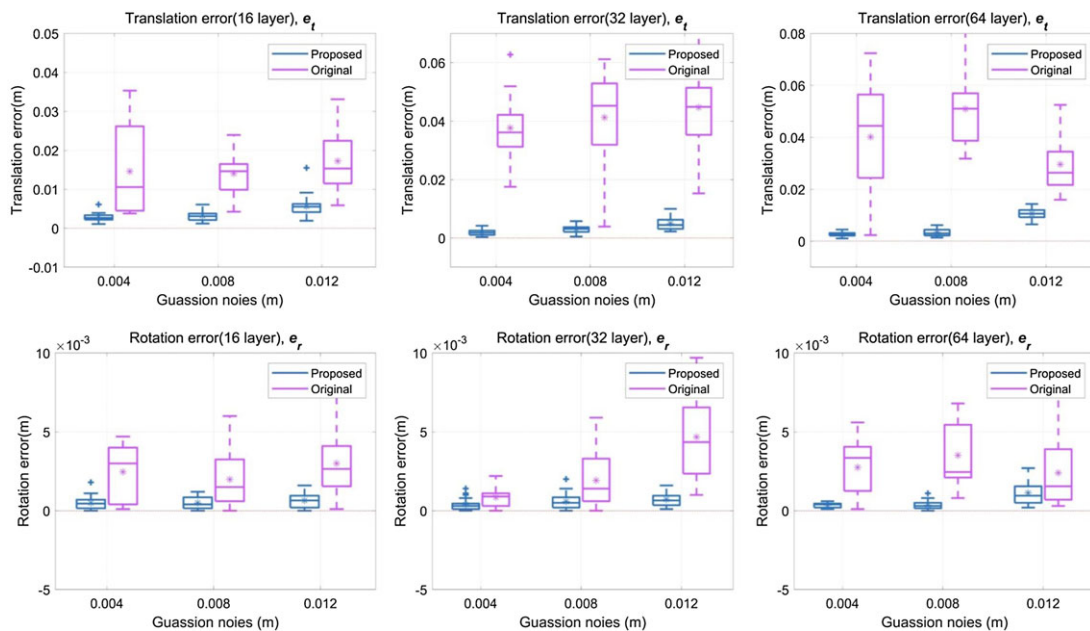
Figure 10. *The calibration results of the cameras and LiDARs are presented for various levels of noise, including 0.004m, 0.008m, and 0.012m.*

potential to effectively calibrate camera with a LiDAR. To present a comprehensive analysis of our method, we employed a combination of qualitative and quantitative approaches to examine and interpret the calibration results.

Fig. 12 and Fig. 13 show the reprojection results of the extrinsic parameters in two different scenes. To enhance the visibility of the reprojection results, we individually projected the results of calibration
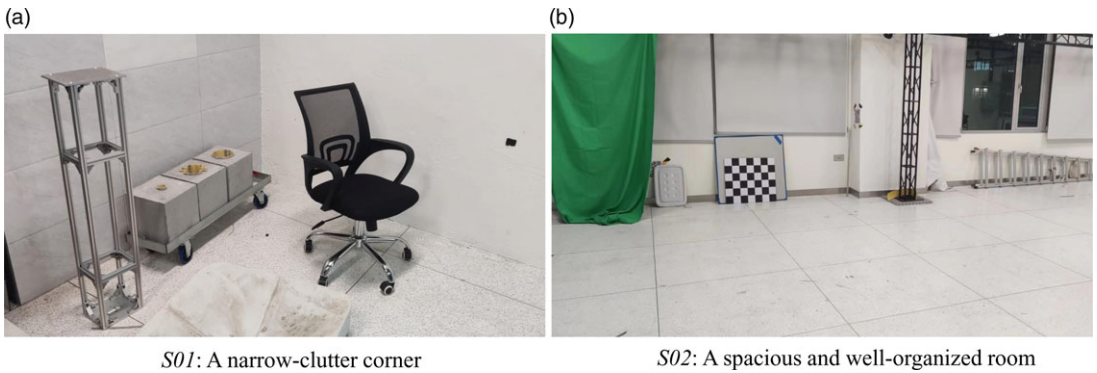
(a)

(b)

S01: A narrow-clutter corner

S02: A spacious and well-organized room

Figure 11. *The real-world experiment scenarios.*

(a)

(b)

(c)

Zhou et al.-S01 (Image)

Guindel et al.-S01 (Image)

Ours.-S01 (Image)

(d)

(e)

(f)

Zhou et al.-S01 (Cloud)

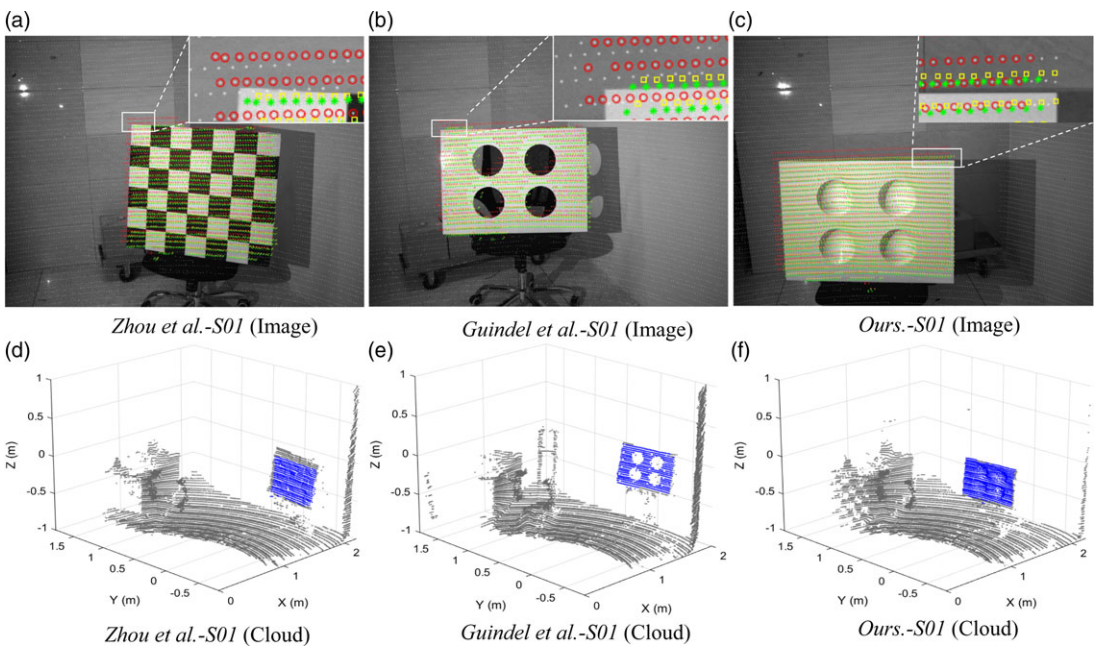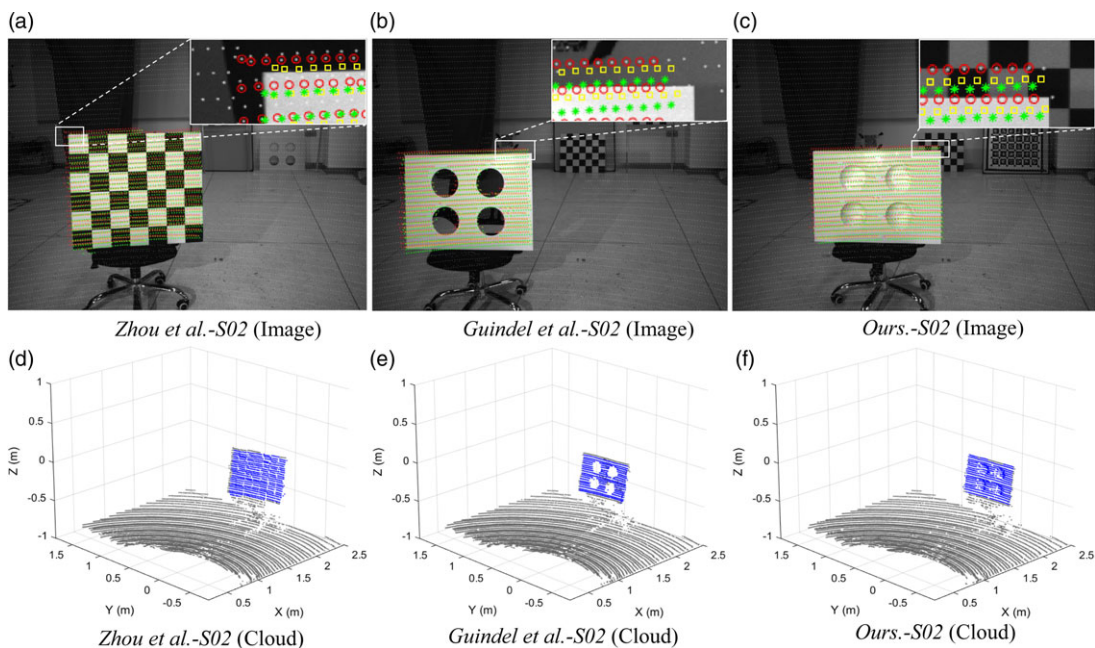*Guindel et al.-S01* (Cloud)

Ours.-S01 (Cloud)

Figure 12. *The reprojection of LiDAR points onto the image in S01: our algorithm (green), Zhou et al. (red) and Guindel et al. (yellow).*

methods onto each board. The reprojection error of the extrinsic parameters was assessed by examining the overlap between the calibration board's point cloud and the corresponding image. Notably, our algorithm's projected point cloud (colored green) demonstrates a closer alignment to the position of the calibration board in comparison to the other two methods. We consider that this is mainly caused by the discontinuity of the point cloud, resulting in the inaccuracy of directly obtained boundary feature points. In contrast, our method indirectly obtains feature points by fitting the point cloud, independent of its discontinuity. In addition, Fig. 12(d)(e)(f) and 13(d)(e)(f) illustrate the point cloud (colored blue) that falls within the calibration board in the reprojection results. A greater proportion of the board point cloud that is occupied by the blue region indicates a more accurate calibration result. The proportion of

Figure 13. *The reprojection of LiDAR points onto the image in S02: our algorithm (green), Zhou et al. (red) and Guindel et al. (yellow).*

the blue region in all three methods further emphasizes the superiority of our algorithm over the other two methods.

For qualitative experiments, as it is unfeasible to determine the exact rigid body transformation between sensors, we adopt the approach developed by Jiao [33] to compute a "pseudo-GT." This is achieved by manually selecting corresponding 3D-3D point pairs and utilizing the ICP algorithm. Table IV displays the errors of the extrinsic parameters obtained by various calibration algorithms based on the computed pseudo-GT. We randomly conducted three sequences of experiments for each scenario illustrated in Fig. 11, using a randomized approach. The results of the six sequences in the table clearly indicate that our algorithm demonstrates more accuracy and robustness in comparison to the other two contrast algorithms. It is worth noting that although Guindel's method can sometimes yield acceptable results, its outcomes are often unstable and even fail to calibrate. One possible explanation for this is the method's limited capability to consistently and effectively remove point cloud that does not belong to the calibration board, leading to the inaccurate extraction of edge features.

4. Conclusion

This paper proposes a novel approach with a customized board to calibrate extrinsics between Structured light cameras(SLCs) and LiDARs, which considers fitted sphere centers as feature points. This method can significantly reduce human intervention and utilize the geometric constraints of the calibration board to extract features accurately. The proposed method has been validated through a combination of simulation and real-world experiments, demonstrating performance well with accuracy and robustness.

However, the proposed method is limited to sensors that are capable of providing 3D geometric information and may not be compatible with ordinary cameras. In future research, we can enhance

Table IV. *The translation and rotation error in the real-world experiments.*

	Seq.	Ours		Guindel et al.		Zhou et al.	
		$e_t(m)$	$e_r(rad)$	$e_t(m)$	$e_r(rad)$	$e_t(m)$	$e_r(rad)$
Scene01	1	**0.0210**	**0.0025**	0.0298	0.0102	0.0882	0.0389
	2	**0.0235**	**0.0026**	0.1720	0.0091	0.0607	0.0098
	3	**0.0349**	**0.0162**	-	-	0.0997	0.0286
Scene02	4	**0.0098**	**0.0042**	0.0172	0.0079	0.0713	0.0218
	5	**0.0219**	**0.0158**	-	-	0.0821	0.0341
	6	**0.0386**	**0.0019**	0.1430	0.0584	0.0958	0.0397

the universality of the calibration algorithm by integrating an optimal number of QR codes onto the calibration board, thereby incorporating other sensors into the calibration framework.

Author contribution. The ideas, methodology, and experimental validation of this work were proposed and implemented by Yangtao Ge and Chen Yao. All the studies and experiments were supervised by Jing Wu. Zirui Wang provided theoretical and technical support in the calibration method and simulation experiments. Wentao Zhang and Haoran Kang raised a series of questions regarding the initial ideas, providing necessary support for further improvement. Huang provided assistance to revise the methodology and conclusion sections, as well as conducting practical experiments with "pseudo-ground truth". Zhenzhong Jia provided valuable suggestions for improving the paper's structure and offered suitable experimental scenarios and sensors for practical experiments.

Competing interests. The authors declare no conflicts of interest exist.

Ethical standards. Not applicable.

References

[1] C. Yao, G. Shi, P. Xu, S. Lyu, Z. Qiang, Z. Zhu, L. Ding and Jia Z., "STAF: Interaction-Based Design and Evaluation of Sensorized Terrain-Adaptive Foot for Legged Robot Traversing on Soft Slopes," *IEEE/ASME Trans Mechatro*, (2024). doi: 10.1109/TMECH.2024.3350183.

[2] A. Seeni, B. Schäfer and G. Hirzinger, "Robot mobility systems for planetary surface exploration–State-of-the-art and future outlook: A literature survey," *Aerosp Technol Advance* **492**, 189–208 (2010).

[3] C. Yao, F. Xue, Z. Wang, Y. Yuan, Z. Zhu, L. Ding and Z. Jia, "Wheel vision: Wheel-terrain interaction measurement and analysis using a sensorized transparent wheel on deformable terrains," *IEEE Robot Auto Lett* **8**(12), 7938–7945 (2023). doi: 10.1109/LRA.2023.3324291.

[4] T. Yu, Q. Man, Y. Wang, G. Q. Shen, J. Hong, J. Zhang and J. Zhong, "Evaluating different stakeholder impacts on the occurrence of quality defects in offsite construction projects: A Bayesian-network-based model," *J Clean Prod* **241**, 118390 (2019). doi: 10.1016/j.jclepro.2019.118390.

[5] Y. Zhu, C. Zheng, C. Yuan, X. Huang and X. Hong, "CamVox: a Low-Cost and Accurate LiDARs-Assisted Visual SLAM System," **In:** *2021 IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China (2021) pp. 5049–5055.

[6] X. Zhang, S. Zhu, S. Guo, J. Li and H. Liu, "Line-Based Automatic Extrinsic Calibration of LiDARs and Camera," **In:** *2021 IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China (2021) pp. 9347–9353.

[7] Y. Zhao, Y. Wang and Y. Tsai. 2D-image to 3D-range Registration in Urban Environments via Scene Categorization and Combination of Similarity measurements. **In:** *2016 IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm (2016) pp. 1866–1872.

[8] N. Schneider, F. Piewak, C. Stiller and U. Franke, "RegNet: Multimodal Sensor Registration Using Deep Neural Networks," **In:** *2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, USA (2017) pp. 1803–1810.

[9] G. Iyer, R. K. Ram, J. K. Murthy and K. M. Krishna, "CalibNet: Geometrically Supervised Extrinsic Calibration using 3D Spatial Transformer Networks," **In:** *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain (2018) pp. 1110–1117.

[10] X. Lv, B. Wang, Z. Dou, D. Ye and S. Wang, "LCCNet: LiDARs and Camera Self-Calibration using Cost Volume Network," **In:** *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Nashville, TN, USA (2021) pp. 2888–2895. doi: 10.1109/CVPRW53098.2021.00324

[11] W. Wang, S. Nobuhara, R. Nakamura and K. Sakurada, "SOIC: Semantic online initialization and calibration for liDARs and camera (2020).arXiv, Mar. 09.

[12] V. Lepetit, F. Moreno-Noguer and P. Fua, "EPnP: An accurate O(n) solution to the PnP problem," *Int J Comput Vision* **81**(2), 155–166 (2009).

[13] R. Kummerle, G. Grisetti, H. Strasdat, K. Konolige and W. Burgard. G 2 o: A General Framework for Graph Optimization. **In:** *2011 IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China (2011) pp. 3607–3613.

[14] L. Zhou, Z. Li and M. Kaess, Automatic Extrinsic Calibration of a Camera and a 3D LiDARs Using Line and Plane Correspondences. **In:** *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid (IEEE, 2018) pp. 5562–5569. doi: 10.1109/IROS.2018.8593660

[15] G. Koo, J. Kang, B. Jang and N. Doh, Analytic Plane Covariances Construction for Precise Planarity-based Extrinsic Calibration of Camera and LiDARs. **In:** *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France (IEEE, 2020) pp. 6042–6048. doi: 10.1109/ICRA40945.2020.9197149

[16] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," **In:** *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE Cat. No.04CH37566), Sendai, Japan, (IEEE, 2004) pp. 2301–2306. doi: 10.1109/IROS.2004.1389752.

[17] R. Unnikrishnan and M. Hebert, "Fast extrinsic calibration of a laser rangefinder to a camera," *Robotics Institute*, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-05-09 (2005).

[18] W. Wang, K. Sakurada and N. Kawaguchi, "Reflectance intensity assisted automatic and accurate extrinsic calibration of 3D liDARs and panoramic camera using a printed chessboard," *Rem Sen* **9**(8), 851 (2017).

[19] Y. Xie, R. Shao, P. Guli, B. Li and L. Wang, "Infrastructure Based Calibration of a Multi-Camera and Multi-LiDARs System Using Apriltags," **In:** *2018 IEEE Intelligent Vehicles Symposium (IV)*, Changshu (2018) pp. 605–610. doi: 10.1109/IVS.2018.8500646

[20] A. Geiger, F. Moosmann, O. Car and Schuster B., "Automatic Camera and Range Sensor Calibration Using a Single Shot," **In:** *2012 IEEE International Conference on Robotics and Automation*, Saint Paul, MN, USA (2012) pp. 3936–3943.

[21] C. Fang, S. Ding, Z. Dong, H. Li, S. Zhu and P. Tan, "Single-Shot is Enough: Panoramic Infrastructure Based Calibration of Multiple Cameras and 3D LiDARs," **In:** *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Prague, Czech Republic (2021) pp. 8890–8897.

[22] S. Debattisti, L. Mazzei and M. Panciroli, "Automated Extrinsic Laser and Camera Inter-Calibration using Triangular Targets," **In:** *2013 IEEE Intelligent Vehicles Symposium (IV)*, Gold Coast, QLD, Australia (2013) pp. 696–701.

[23] Y. Park, S. Yun, C. S. Won, K. Cho, K. Um and S. Sim, "Calibration between color camera and 3D LiDARs instruments with a polygonal planar board," *Sensors* **14**(3), 5333–5353 (2014).

[24] Q. Liao, Z. Chen, Y. Liu, Z. Wang and M. Liu, "Extrinsic Calibration of LiDARs and Camera with Polygon," **In:** *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Kuala Lumpur, Malaysia (2018) pp. 200–205. doi: 10.1109/ROBIO.2018.8665256

[25] Z. Deng, L. Xiong, D. Yin and F. Shan, "Joint calibration of dual LiDARs and camera using a circular chessboard," SAE International, Warrendale, PA, (2020). SAE Technical Paper 2020-01-0098.

[26] V. Fremont, S. A. Rodriguez F. and P. Bonnifait, "Circular targets for 3D alignment of video and LiDARs sensors," *Adv Robotics* **26**(18), 2087–2113 (2012).

[27] M. Pereira, D. Silva, V. Santos and P. Dias, "Self calibration of multiple LiDARs and cameras on autonomous vehicles," *Robot Auton Syst* **83**, 326–337 (2016).

[28] J. Kummerle, T. Kuhner and M. Lauer, "Automatic Calibration of Multiple Cameras and Depth Sensors with a Spherical Target," **In:** *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid (2018) pp. 1–8. doi: 10.1109/IROS.2018.8593955

[29] J. Beltran, C. Guindel, A. de la Escalera and F. Garcia, "Automatic extrinsic calibration method for LiDARs and camera sensor setups," *IEEE Trans Intell Transport Syst* **23**(10), 17677–17689 (2022). doi: 10.1109/TITS.2022.3155228.

[30] T. Tóth, Z. Pusztai and L. Hajder. "Automatic LiDARs-Camera Calibration of Extrinsic Parameters Using a Spherical Target," **In:** *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France (2020) pp. 8580–8586.

[31] A. Dhall, K. Chelani, V. Radhakrishnan and K. M. Krishna.LiDARs camera calibration using 3D-3D point correspondences(2017). ArXiv eprints

[32] N. Koenig and A. Howard, "Design and use Paradigms for Gazebo, an Open-Source Multi-Robot Simulator," **In:** *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sendai, Japan **3**, (2004) pp. 2149–2154.

[33] J. Jiao, F. Chen, H. Wei, J. Wu and M. Liu, "LCE-calib: Automatic LiDAR-frame/Event camera extrinsic calibration with a globally optimal solution," *IEEE/ASME Trans Mechatro* **28**(5), 2988–2999 (2023). doi: 10.1109/TMECH.2023.3259444.