Journal of Network Intelligence
Taiwan Ubiquitous Information

Volume **10**, Number **3**, August **2025**

Vehicle Coordinate System Calibration Based on 3D Vision Sensors and Body Feature Points

Zhi-Gang Ding[1], Zhi-Ming Dong[2,*], Ji-Shi Zheng[2], Yan-Lu Lv[1], Ling-Hua Kong[1]

[1]School of Mechanical and Automotive Engineering, Fujian University of Technology, Fuzhou 350118, China
dzgsydy@fjut.edu.cn lv163163@126.com 15392030898@163.com

[2]School of Transportation, Fujian University of Technology, Fuzhou 350118, China
1040824830@qq.com zhengjishi@fjut.edu.cn

*Corresponding author: Zhi-Ming Dong

Received December 9, 2023, revised March 24, 2024, accepted September 6, 2024.

Abstract. *Image fuzzy enhancement is a research hotspot in the field of image processing, which aims to recover enhanced beginning clear images from degraded images. Based on the research of traditional particle swarm optimization algorithm and fuzzy enhancement algorithm, an image fuzzy enhancement method based on membrane computing particle swarm algorithm is proposed. Firstly, in order to make full use of the sparse characteristics of the clear image, the coefficient decomposition under wavelet domain and tightly supported wavelet domain is performed on the image respectively. Then, a joint optimisation model is constructed using L1 parametric constraints to achieve pretzel noise cancellation. Next, the MMH-PSO algorithm is designed by improving the particle swarm algorithm using membrane computing and Metropolis-Hastings sampling. Based on the simulated annealing algorithm temperature drop process, Metropolis-Hastings sampling is used to add randomness to the particle swarm algorithm so that it has the ability to jump out of the local optimum. The use of membrane computing enhances the parallelism of the particle swarm algorithm and can reduce the time complexity in solving complex problems. Finally, MMH-PSO is used to simultaneously search out the magnitude of the two fuzzy parameters in the traditional fuzzy enhancement algorithm in order to improve the accuracy of the algorithm. The experimental results show that the proposed algorithm has better SSIM values than the traditional fuzzy enhancement algorithm, which effectively improves the image quality and makes the image edge information more abundant.*

Keywords: Image enhancement; particle swarm algorithm; membrane computing; simulated annealing algorithm; pretzel noise

1. **Introduction.** Advanced Driving Assistance Systems (ADAS), as the core products of assisted driving, refer to the use of various sensors in the vehicle to sense the environment, collect and analysis data, and assist the driver in performing driving tasks, thereby enhancing driving comfort and safety. With the increasing maturity of ADAS technology, the market size has expanded. Many mature technology applications are becoming more common in the entry-level passenger car segment [1, 2, 3]. However, in the production process, due to the installation process, vehicle individual differences and other reasons, the installation angle of the sensor on the vehicle will be errors, so the positioning and calibration of these sensors is an important part of the vehicle can be normally off the line. As shown in Figure 1, the existing equipment platform requires alignment of the vehicle coordinate system with the workstation coordinate system before sensor calibration due

to the limitations of the rigid target mounting platform. To achieve this, the vehicle to be tested is driven onto a vehicle alignment platform, and mechanical contact is made with alignment blocks on the platform, pushing the vehicle into the specified coordinate system and aligning it accordingly [4]. The existing equipment suffers from drawbacks such as high costs, insufficient flexibility, large footprint, and poor scalability.

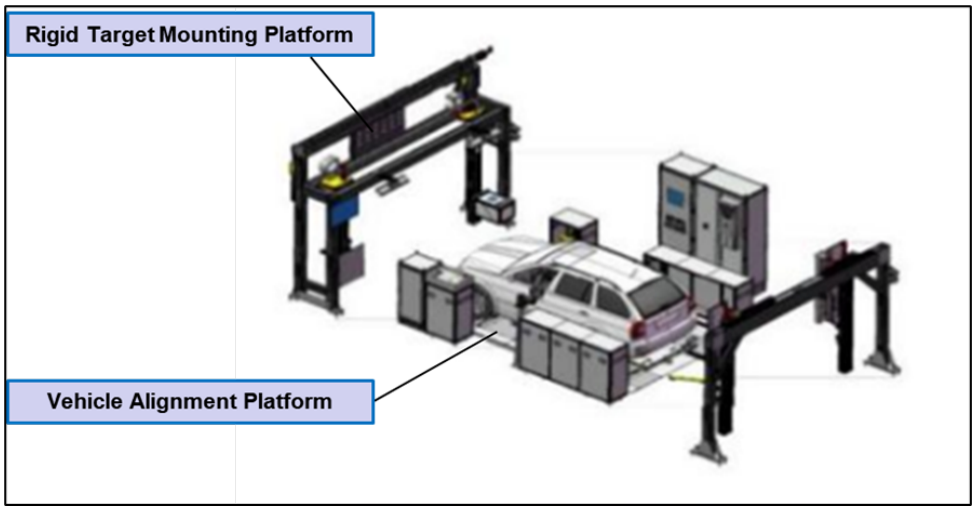


Figure 1. Existing ADAS testing equipment

This article proposes a design for a non-contact, high-precision solution based on machine vision to address the current problems. The aim is to accurately locate the position of vehicles offline in the manufacturing plant.

1.1. **Related work.** In the automotive field, the application of machine vision for global or local pose estimation of vehicles is widely used. By performing pose estimation on vehicles in different scenarios, various applications can be achieved. In urban roads, vision is usually used to locate vehicles on the road, and one of the scenarios is used in vehicle monitoring at urban intersections or parking lots. Li et al. [5] proposed a method to accurately estimate the pose of all vehicles and the ground plane in car park scenes, then achieved good results on a car park dataset. Zhang et al. [6] addressed the global vehicle pose estimation problem at intersections by inferring the 6-DoF pose of vehicles based on vehicle images and 3D information. Another scenario is autonomous driving, where vehicles themselves use on-board cameras to perceive the environment and estimate the pose and 3D centroids of surrounding vehicle targets. Currently, the main methods being studied by researchers in this scenario include direct regression methods based on deep learning. Wu et al. [7] proposed 6D-VNet, which added two branches to the object detection network and directly estimated the pose of vehicle targets in RGB images through end-to-end training. Ke et al. [8] presented an end-to-end multitask network, GSNet, that can simultaneously estimate the pose of vehicles and reconstruct dense 3D shapes. There is also an approach based on the previous 3D model of the vehicle. After extracting the features of the target vehicle, this method matches them with the 3D model of the vehicle to determine the pose of the vehicle. Barowski et al. [9] performed multi-class segmentation of different parts of the vehicle, converting the segmented multi-class masks into a set of variable keypoints and establishing a 2D-3D correspondence. Chabot et al. [10] proposed Deep MANTA, which uses convolutional neural networks to detect and localize vehicle components in autonomous driving scenarios, establish a 2D to 3D point correspondence, and solve the PnP problem to obtain the target pose. Lyu et al. [11]

introduced Mono6D, which matches the vehicle in a monocular image by searching for the best-fitting 3D model, extracting vehicle features, and performing vehicle localization. As is evident, solving for the poses of the vehicle using a 3D model can be advantageous. However, this type of method is mostly used for real-time detection of traffic flow, and the target detection of the car in the global scene is also needed before the estimation of the vehicle's pose, so the complexity of the algorithm is higher and more complicated than the static detection scenario of the vehicle coming off the line in the vehicle factory. While the automatic charging scenario of electric vehicles is similar to the scenario described in this paper. In the case of electric vehicle self-charging [12], vision is usually used to localize the vehicle to be charged and transmit the vehicle position information to the charging robot through hand-eye calibration [13]. Pan et al. [14] designed a visual localization system for electric vehicle charging ports, where the charging port image is captured by a monocular camera and the charging port is identified and localized by CNN and circular features. However, the single camera used for visual localization has a small field of view, lacks depth information, and has poor localization accuracy, while binocular vision-based target localization can recover the 3D information of the test object and improve the localization accuracy [15]. Yao [16] developed a binocular vision-based charging hole identification and localization system for electric vehicles. The method realized binocular system calibration and hand-eye calibration by Halcon software, and then combined with stereo matching, 3D positioning and coordinate transformation in binocular vision to complete the 3D position estimation of the charging hole, and sent to the robot controller to realize the robot automatically charging the electric vehicle. As can be seen, the use of vision to locate the vehicle from the line of the vehicle factory can also provide a basis for the future use of collaborative robot hand-held target instead of the rigid target mounting platform for the inspection of sensor functions.

1.2. **Innovation and contribution.** The installation position and angle of vehicle sensors are designed with reference to the vehicle coordinate system. On this basis, this paper attempts to adopt a 3D vision sensor-based for off-line calibration of the vehicle coordinate system. The method acquires images of the exposed body-in-white region at the rear of the vehicle with a known 3D model of the vehicle, extracts the key points of the body-in-white region in the RGB image with the help of image processing, recovers their 3D coordinates under the camera coordinate system with the help of a 3D sensor, and eventually realizes the calibration of the vehicle coordinate system through coordinate transformation.

2. 3D vision system design.

2.1. **Principles of binocular vision.** This paper represents a preliminary study within an industrial project that aims to calculate the coordinate positions of parked vehicles by 3D vision sensors. In this paper, the 3D vision sensors used mainly include binocular cameras and 3D depth cameras, and these sensors are used to collect the vehicle's body surface data and recover the 3D coordinates of the body feature points under the camera coordinate system. Binocular vision systems typically use two cameras to simulate the human eye viewing the test object from different viewpoints. The depth information of the object is reconstructed according to the principle of disparity, allowing its 3D information to be recovered [17]. Parallel binocular vision systems are commonly used in such setups, where two cameras with identical internal parameters are arranged in a parallel configuration. Ideally, the plane formed by the two cameras should be parallel to the X-axis direction of one of the cameras, with a baseline distance denoted as b. Let the pixel coordinates of the observation point P in the left and right camera images be

(x_l, y_l) and (x_r, y_r), respectively, and let $D = x_l - x_r$ be the disparity. With the baseline distance, disparity value, and focal length, the coordinates of point P in the camera coordinate system can be calculated. See Figure 2 for details. O_l and O_r are the origin of the left and right camera coordinate systems, b is the baseline distance, f is the focal length, and h is the depth value of the observation point.

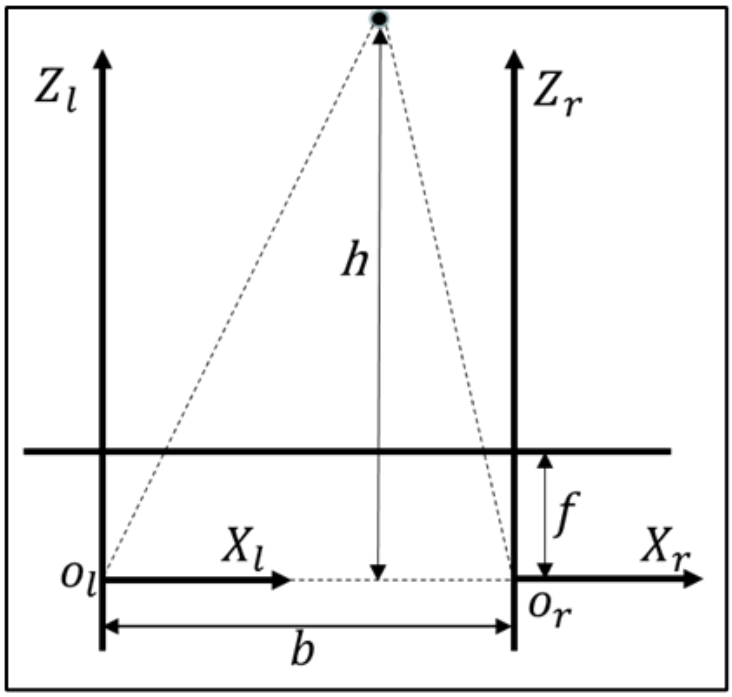


Figure 2. Schematic of binocular vision

2.2. **Camera calibration.** In machine vision applications, camera calibration is a key step in determining the 3D geometric position of a point on the surface of a spatial object in relation to its corresponding point in the image. The accuracy of calibration results directly affects the accuracy of the results generated by the camera work [18].

Camera calibration involves 4 coordinate systems: the pixel coordinate system, image coordinate system, camera coordinate system, and world coordinate system, where the pixel coordinate system is a 2D coordinate system with the left vertex of the image as the coordinate origin, and the center point coordinates are (u_0, v_0), and the unit is pixels; the image coordinate system is in the same plane as the pixel coordinate system, and the center point of the image is the coordinate origin in 2D as the coordinate system, and the unit is mm; the camera coordinate system is a 3D coordinate system with the optical center O_C of the camera as the origin, its x-axis and y-axis direction is the same as the direction of the image coordinate system, z-axis from the camera optical center, along the camera optical axis to the shooting direction positive extension; the world coordinate system refers to the coordinate system that does not move relative to the ground, also called absolute coordinate system, is the benchmark coordinate system to determine the position of the camera, object.

The mutual conversion between the coordinate systems is performed by the internal and external parameters obtained from the camera calibration. As shown in Figure 3, let the coordinates of point P be (u, v) in the pixel coordinate system, (x, y) in the image coordinate system, (X_C, Y_C, Z_C) in the camera coordinate system, and (X_W, Y_W, Z_W) in the world coordinate system. Then, the transformation relations from the pixel coordinate

system to the image coordinate system, from the image coordinate system to the camera coordinate system, and from the camera coordinate system to the world coordinate system are shown in Equations (1) to (3), respectively:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/S_x & 0 & u_0 \\ 0 & 1/S_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z_C} \left[\begin{array}{ccc|c} f_x & 0 & 0 & 0 \\ 0 & f_y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3\times3} & T_{3\times1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \tag{3}$$

Combining the above equations gives the following relationship for the transformation of point P from the world coordinate system to the pixel coordinate system:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = k \left[\begin{array}{ccc|c} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \begin{bmatrix} R_{3\times3} & T_{3\times1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \tag{4}$$

Where $k = 1/Z_C$ is the scale factor, $f_x = f/S_x$, $f_y = f/S_y$, f is the camera focal length, S_x, S_y is the physical size of a single pixel point; R is the rotation matrix, and T is the translation vector.

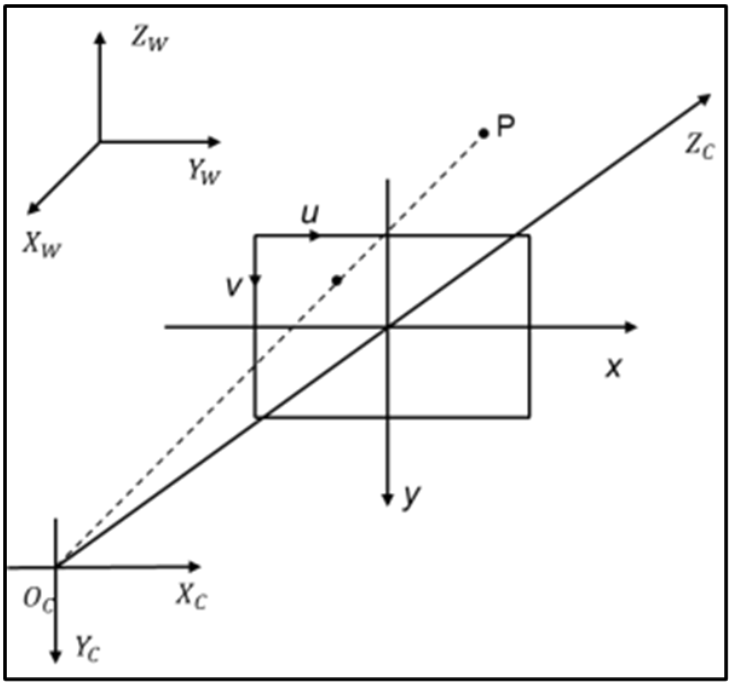


Figure 3. Schematic of the four coordinate systems included in the camera calibration

2.3. **3D coordinate calculation.** As seen in Figure 4, this study employs the ZED2i binocular camera with a self-calibration feature. Its self-calibration function after extensive and rigorous multi-step factory calibration, can re-estimate, refine the calibration and read the internal and external parameters of the camera, eliminating the process of manual calibration [19]. When using the binocular camera to photograph the vehicle' body surface, not only RGB images but also depth maps can be acquired, which are images with the distance from the camera to the points in the vehicle' body surface as pixel values, as show in Figure 5.

Figure 4. The ZED2i binocular camera used in this paper

Figure 5. Vehicle body RGB image and corresponding depth map acquired with ZED2i

This work focuses on extracting the corner points of the exposed vehicle body contour lines and the center point of the fuel cap as feature points from the RGB image of the vehicle' body surface. For the corner points of the body contour lines, the Canny edge detector is used to extract the body contour lines, and then the Harris corner detection algorithm [20] is applied to extract the corner points on the contour lines. For the center point of the fuel cap, the fuel cap contour is obtained, and an ellipse fitting algorithm [21] is used to determine the center point of the fuel cap (see Figure 6).

After extracting the pixel coordinates of the feature point, the depth map can be used to determine the depth value Z_C of the extracted feature point on vehicle's body surface. When this value is combined with the internal camera calibration parameters, the 3D coordinates of the body feature point P_i in the camera coordinate system can then be calculated as follows:

$$\begin{cases} X_C = \dfrac{(u_i - u_0)}{f_x} Z_C \\ Y_C = \dfrac{(v_i - v_0)}{f_y} Z_C \end{cases} \tag{5}$$

Where (u_i, v_i) is the coordinate of the body feature point P_i in the pixel coordinate system.

As shown in Figure 7, this paper uses ZhiSensor's 3D depth camera, whose model is D132. It adopts the surface array dynamic grating stripe structure light combined with infrared laser light source to realize high-precision data acquisition. Using its SDK (Software Development Kit), the collected point cloud data can be rearranged according

to the RGB image, so that the 2D pixel points are aligned with the 3D point cloud. Combined with the above feature point extraction method, the 3D coordinates of the feature points in the camera coordinate system can be computed.

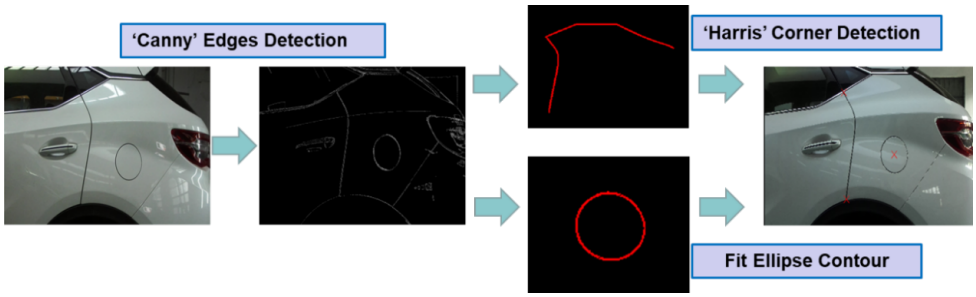
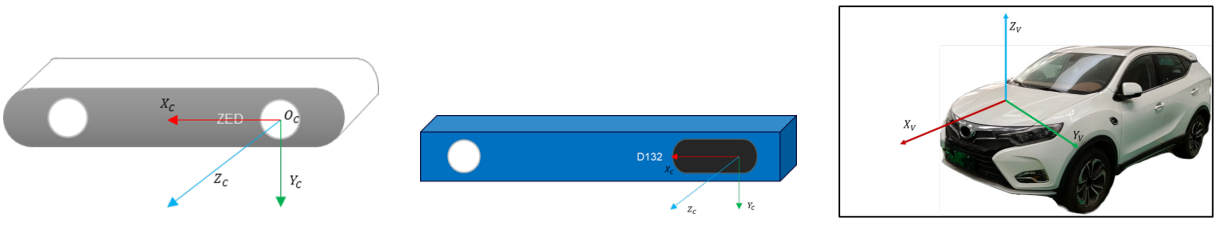


Figure 6. Flowchart of feature point extraction from RGB image of vehicle body

Figure 7. D132 depth camera used in this paper

3. Coordinate transformation.

3.1. **Definition of coordinate systems.** The camera coordinate system and vehicle coordinates are the key coordinate systems in the vehicle localization approach mentioned in this paper. Figure 8 (a) and (b) displays the camera coordinate system $O_C X_C Y_C Z_C$. The coordinate origin O_C is located at the focus center of the main lens, the X_C-axis is parallel to the right lens, the Y_C-axis is perpendicular to the camera mounting plane and points to the ground, and the Z_C-axis is perpendicular to the $X_C Y_C$ plane and points to the photographed object. The vehicle coordinate system [22] $O_V X_V Y_V Z_V$ is shown in Figure 8 (c). The coordinate origin O_V is the center of mass of the vehicle, the X_V-axis points to the forward direction of the vehicle parallel to the ground, the Y_V-axis points to the left side of the vehicle parallel to the ground, and the Z_V-axis points to the top of the vehicle perpendicular to the $X_V Y_V$ plane.



(a) The camera coordinate system of ZED2i binocular camera

(b) The camera coordinate system of D132 depth camera

(c) The vehicle coordinate system

Figure 8. Orientation of the axes of the relevant coordinate system

After the vehicle enters the designated parking zone, the vehicle's body pictures are collected according to the method mentioned above. The coordinates (X_C, Y_C, Z_C) of

the feature point P_i on the vehicle's body in the camera coordinate system are obtained. When combined with its coordinates (X_V, Y_V, Z_V) in the vehicle coordinate system, we can represent the transformation model of the two 3D coordinate systems by the rigid transformation shown below:

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} = \lambda R \begin{bmatrix} X_V \\ Y_V \\ Z_V \end{bmatrix} + \lambda \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \tag{6}$$

Where, λ is the scale factor, R is the rotation matrix in the coordinate transformation, and (T_X, T_Y, T_Z) is the translation vector.

3.2. **Rodriguez matrix.** A 3D coordinate transformation model based on the Rodriguez matrix is applied in this paper. The merits of this approach include its simplicity, ease calculation and excellent accuracy. It solves the 7 parameters in the coordinate system transformation simply by using 3 pairs of known point coordinates in the camera coordinate system and the vehicle coordinate system [23].

As shown in Equation (7), considering the rotation matrix R as a Rodriguez matrix consisting of an antisymmetric matrix S:

$$S = \begin{bmatrix} 0 & -c & -b \\ c & 0 & -a \\ b & a & 0 \end{bmatrix}$$
$$R = (I + S)(I - S)^{-1} \tag{7}$$

where I is the 3-order identity matrix and the three parameters in S are independent of each other. As a result, the transformation of the 3D coordinate system includes seven parameters: three for translation, three for antisymmetric matrix, and one for scale. The scale factor can be determined among them by figuring out the ratio of the corresponding edge lengths from two common locations in two coordinate systems, or it can be calculated and then taken as the mean value of all the ratios of corresponding edge lengths.

$$\lambda = \frac{\sqrt{(X_{Ci} - X_{Cj})^2 + (Y_{Ci} - Y_{Cj})^2 + (Z_{Ci} - Z_{Cj})^2}}{\sqrt{(X_{Vi} - X_{Vj})^2 + (Y_{Vi} - Y_{Vj})^2 + (Z_{Vi} - Z_{Vj})^2}} \tag{8}$$

The common point coordinates (X_C, Y_C, Z_C) and (X_V, Y_V, Z_V) in the camera coordinate system and the vehicle coordinate system are subtracted according to Equation (6) and combined with Equation (7) to obtain Equation (9), which is further transformed by bringing in S and I to obtain Equation (10):

$$(I - S) \begin{bmatrix} \Delta X_{Cij} \\ \Delta Y_{Cij} \\ \Delta Z_{Cij} \end{bmatrix} = \lambda (I + S) \begin{bmatrix} \Delta X_{Vij} \\ \Delta Y_{Vij} \\ \Delta Z_{Vij} \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} 0 & \Delta Z_{Cij} + \lambda \Delta Z_{Vij} & \Delta Y_{Cij} + \lambda \Delta Y_{Vij} \\ \Delta Z_{Cij} + \lambda \Delta Z_{Vij} & 0 & -\Delta X_{Cij} - \lambda \Delta X_{Vij} \\ -\Delta Y_{Cij} - \lambda \Delta Y_{Vij} & -\Delta X_{Cij} - \lambda \Delta X_{Vij} & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \lambda \Delta X_{Vij} - \Delta X_{Cij} \\ \lambda \Delta Y_{Vij} - \Delta Y_{Cij} \\ \lambda \Delta Z_{Vij} - \Delta Z_{Cij} \end{bmatrix} \tag{10}$$

The rotation matrix parameters a, b, and c can be found by the least squares method by combining the 3 common points two by two and forming 3 independent equations according to Equation (10). After finding the scale factor and the rotation matrix R, the translation vector can be calculated according to Equation (6) [24].

3.3. **Transformation of Euler angles.** In addition to the rotation matrix, three Euler angles can be used to describe a vehicle's pose in 3D space: the roll angle α represents the angle of vehicle rotation around the x-axis of the vehicle coordinate system, the pitch angle β represents the angle of vehicle rotation around the y-axis, and the yaw angle γ represents the angle of vehicle rotation around the z-axis. With $R_x(\alpha)$, $R_y(\beta)$, $R_z(\gamma)$ denoting the rotation matrix with rotation angles α, β and γ around the x, y and z axes respectively. After determining the rotation order, the rotation matrix R is related to the Euler angles as follows:

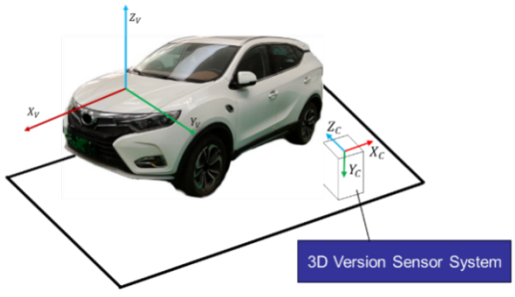
$$R = R_z(\gamma)R_y(\beta)R_x(\alpha) \tag{11}$$

Let the resulting rotation matrix be $R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$.

Equation (12) demonstrates how to calculate the Euler angles α, β and γ by resolving the inverse tangent function [25]. The position of the vehicle coordinate system in the camera coordinate system is determined by the solved Euler angle as well as the translation vector.

$$\begin{cases} \alpha = \text{atan2}(R_{21}, R_{11}) \\ \beta = \text{atan2}(-R_{31}, \sqrt{R_{32}^2 + R_{33}^2}) \\ \gamma = \text{atan2}(R_{32}, R_{33}) \end{cases} \tag{12}$$

4. **Experimental result.** The overall structure of the 3D vision sensors-based vehicle coordinate system calibration method consists of 3D vision sensors, the host computer as well as the experimental vehicle. The experimental vehicle is a well-known brand SUV. The schematic diagram of the vehicle coordinate system calibration scheme and the experimental scenario are shown in Figure 9. In the indoor environment, the simulated vehicle off-line vehicle is parked in the designated area, the 3D camera is fixed at the designated position to ensure that the images collected by the camera cover the whole vehicle's body area, extract the feature points in the camera coordinate system. Then the rotation matrix and translation vector are solved and transformed into Euler angles according to the method above to calibrate the vehicle coordinate system of the off-line vehicle.



(a) Schematic of vehicle coordinate system calibration

(b) Experimental scenarios

Figure 9. Vehicle coordinate system calibration scheme

According to the actual off-line vehicle parking scenario, the parking condition of the vehicle each time is not unique, and there may be yawing in different directions. In this

paper, rotation experiments are conducted on the vehicle at different distances from the vehicle, which is used to simulate the change of yaw angle when the off-line vehicle is parked. In the experiment, because the car is larger and not convenient to rotate, the binocular camera and 3D depth camera are placed on the turntable, and the camera is driven to rotate by turning the turntable to simulate the swing of the car's front end in two directions. When the camera rotates to the left, the car is simulated to rotate to the right, and when the camera rotates to the right, the car is simulated to rotate to the left. Taking 0° on the turntable as the reference, the car was rotated by 1° to the left and right directions in turn, to calculate the change of the vehicle's attitude angle after each camera rotation, and each angle was collected several times to analyze its repeatability accuracy. The vehicle' body images collected at a certain distance are shown in Figure 10, while the calculated results are shown in Tables 1 and 2 and Figure. 11, respectively.

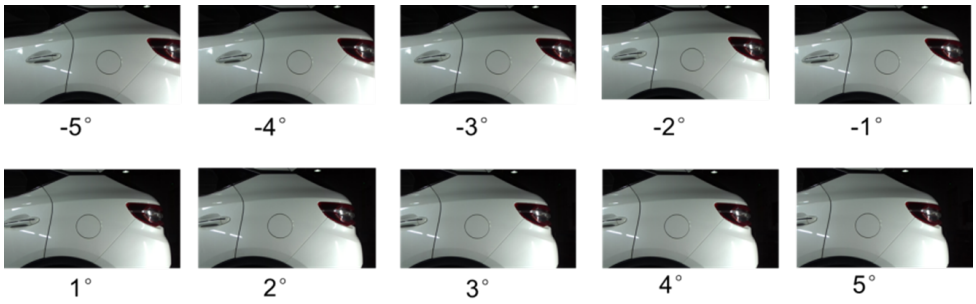
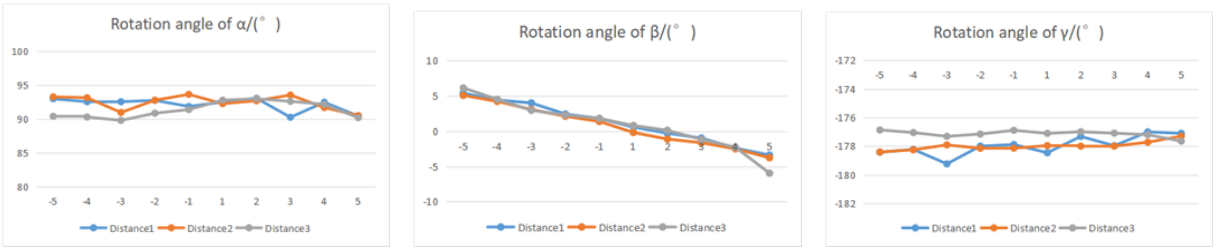


Figure 10. Raw images of the vehicle body at different angles

Table 1. Calculation results of car attitude angle change under binocular camera

Distance1	-5°	-4°	-3°	-2°	-1°	1°	2°	3°	4°	5°
α (°)	93.022	92.579	92.585	92.783	91.887	92.525	93.091	90.286	92.507	90.466
β (°)	5.387	4.417	3.996	2.431	1.796	0.582	-0.296	-0.983	-2.417	-3.387
γ (°)	-178.4	-178.2	-179.2	-178.0	-177.9	-178.5	-177.3	-178.0	-177.0	-177.1
Distance2	-5°	-4°	-3°	-2°	-1°	1°	2°	3°	4°	5°
α (°)	93.303	93.175	90.998	92.803	93.669	92.291	92.721	93.573	91.717	90.545
β (°)	5.075	4.223	3.079	2.111	1.374	-0.189	-1.120	-1.644	-2.503	-3.775
γ (°)	-178.4	-178.3	-177.9	-178.2	-178.1	-178.0	-178.0	-178.0	-177.7	-177.3
Distance3	-5°	-4°	-3°	-2°	-1°	1°	2°	3°	4°	5°
α (°)	90.433	90.328	89.828	90.890	91.419	92.813	93.024	92.639	92.194	90.228
β (°)	6.130	4.514	2.959	2.2903	1.801	0.800	0.116	-1.176	-2.285	-5.954
γ (°)	-176.9	-177.2	-177.3	-177.1	-176.9	-177.1	-177.0	-177.1	-177.2	-177.7

From Table1 and Fig. 11, it can be seen that when the car is turned to the left and right directions, its rotation angle around the Z and X axes does not change much, but the rotation angle around the Y axis changes more obviously. When the car turns to the left in turn, the attitude angle decreases. When the car turns to the right in turn, the attitude angle increases, which is consistent with the relationship between the camera coordinate system and the car coordinate system position given in Figure 8. Analyzing Table 2, we can get that: the RMS of x-translation is maximum 13.216mm, minimum 0.574mm, and average 5.865mm. The RMS of y-translation is maximum 3.950mm, minimum 0.317mm, and average 1.656mm. The RMS of z-translation is maximum 16.850mm, minimum 2.034mm. The RMS of x-rotation is maximum 1.954°, the minimum is 0.095°,



(a) Change curve of α (b) Change curve of β (c) Change curve of γ

Figure 11. Vehicle attitude angle change curve with turntable rotation

Table 2. Repeatability accuracy analysis under binocular camera

Distance1	-5°	-4°	-3°	-2°	-1°	1°	2°	3°	4°	5°
T_X(mm)	3.433	6.165	6.046	5.045	5.850	4.421	5.110	7.840	1.205	7.589
T_Y(mm)	0.666	0.834	0.947	1.439	1.439	0.373	1.514	0.756	1.851	0.880
T_Z(mm)	2.034	14.401	14.821	10.721	12.088	11.451	6.765	16.761	14.984	7.231
α (°)	0.095	0.652	0.663	0.474	0.533	0.533	0.295	1.195	1.256	0.915
β (°)	0.031	0.306	0.244	0.151	0.231	0.236	0.203	0.431	0.014	0.273
γ (°)	0.054	0.079	0.050	0.054	0.057	0.067	0.059	0.096	0.018	0.124
Distance2	-5°	-4°	-3°	-2°	-1°	1°	2°	3°	4°	5°
T_X(mm)	2.833	8.456	2.850	1.747	2.916	0.574	2.835	7.762	5.652	3.228
T_Y(mm)	0.325	1.041	2.216	3.335	1.214	0.317	2.105	2.929	1.380	1.658
T_Z(mm)	10.156	11.683	12.236	9.036	14.988	8.291	7.507	10.190	3.743	15.919
α (°)	0.440	0.520	0.510	1.236	0.845	1.037	0.326	0.405	0.183	0.679
β (°)	0.146	0.310	0.082	0.079	0.036	0.077	0.100	0.376	0.238	0.187
γ (°)	0.016	0.060	0.105	0.061	0.047	0.341	0.035	0.188	0.058	0.165
Distance3	-5°	-4°	-3°	-2°	-1°	1°	2°	3°	4°	5°
T_X(mm)	5.813	2.971	7.916	10.026	11.762	4.200	13.216	8.688	9.089	6.183
T_Y(mm)	1.702	0.659	1.031	1.466	3.319	2.255	1.922	3.950	2.866	3.289
T_Z(mm)	11.743	7.754	15.471	16.850	2.937	5.634	13.255	9.486	12.253	11.965
α (°)	0.359	1.072	1.251	1.331	1.954	0.206	1.232	1.488	0.491	0.488
β (°)	0.237	0.127	0.318	0.380	0.366	0.228	0.564	0.323	1.251	0.247
γ (°)	0.295	0.010	0.097	0.143	0.090	0.043	0.103	0.059	0.173	0.073

Table 3. Calculation results of car attitude angle change under 3D depth camera

Distance1	-5°	-4°	-3°	-2°	-1°	1°	2°	3°	4°	5°
α (°)	88.410	87.361	87.546	87.620	86.831	87.515	87.670	87.602	87.472	87.730
β (°)	8.498	7.913	6.936	6.150	5.014	2.644	1.834	0.573	-0.374	-1.487
γ (°)	-179.9	-179.9	-179.9	-180.0	-179.8	-180.2	-180.4	-180.5	-180.5	-180.5

the average is 0.755°. The RMS of y-rotation is maximum 1.251°, minimum 0.0142°, the average is 0.260°. The RMS of z-rotation is maximum 0.341°, minimum 0.010°, the average is 0.094°. Analyzing Table 3, it can be obtained that when using the 3D depth camera, with the change of the turntable angle, the attitude angle not only has the same trend with the real angle change, but also each angle change is more accurate than the results measured by the binocular camera. In other words, it is closer to 1°. In summary, the method described in this paper provides a new approach for vehicle sensor calibration in

the vehicle industry by enabling high-precision and non-contact vehicle spatial positioning using a depth camera.

5. **Conclusion.** This article aims to address the shortcomings of existing ADAS sensor calibration equipment in automotive manufacturing plants by proposing a non-contact, depth camera-based method for calibrating the coordinate system of offline vehicles. This method involves using image processing algorithms to extract body surface feature points from RGB images and calculating the three-dimensional coordinates of these points using the depth information captured by the depth camera. By combining the vehicle's 3D model and utilizing the Rodrigues matrix, the coordinate system of the vehicle is calibrated. Through the analysis of the measured vehicle poses using this method, the results indicate that not only can this method accurately reflect the trend of changes in attitude angles, but it also exhibits high repeatability precision. In addition to the existing binocular camera, this study also utilizes a higher-precision 3D depth camera. The experimental results indicate that using a higher-precision depth camera is beneficial for calibrating the vehicle coordinate system. The method proposed in this paper provides a novel design approach for vehicle sensor calibration. In future work, improvements can be made based on the 3D depth camera, such as extracting feature points directly from the collected point cloud data. This can eliminate the influence of different vehicle body colors in practical applications, making the method more versatile.

Acknowledgment. This research work is supported by the University-Enterprise Cooperation Project of Fujian Province (GY-H-21211) between Xiamen Richen Technology Co. and Fujian University of Technology.

REFERENCES

[1] T.-Y. Wu, X. Guo, L. Yang, Q. Meng, and C.-M. Chen, "A lightweight authenticated key agreement protocol using fog nodes in social internet of vehicles," *Mobile Information Systems*, vol. 2021, pp. 1–14, 2021.

[2] A. Ziebinski, R. Cupek, D. Grzechca, and L. Chruszczyk, "Review of advanced driver assistance systems (adas)," in *AIP Conference Proceedings*. AIP Publishing, 2017, No. 1.

[3] I. J. Xique, W. Buller, Z. B. Fard, E. Dennis, and B. Hart, "Evaluating complementary strengths and weaknesses of adas sensors," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2018, pp. 1–5.

[4] C. Hajdu and I. Lakatos, "Calibration measurements and computational models of sensors used in autonomous vehicles," *Periodica Polytechnica Transportation Engineering*, vol. 51, no. 3, pp. 230–241, 2023.

[5] P. Li, W. Qiu, M. Peven, G. D. Hager, and A. L. Yuille, "Car pose in context: Accurate pose estimation with ground plane constraints," *arXiv preprint arXiv:1912.04363*, 2019.

[6] S. Zhang, C. Wang, Z. He, Q. Li, X. Lin, X. Li, J. Zhang, C. Yang, and J. Li, "Vehicle global 6-dof pose estimation under traffic surveillance camera," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 114–128, 2020.

[7] D. Wu, Z. Zhuang, C. Xiang, W. Zou, and X. Li, "6d-vnet: End-to-end 6-dof vehicle pose estimation from monocular rgb images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2019, pp. 1–10.

[8] L. Ke, S. Li, Y. Sun, Y.-W. Tai, and C.-K. Tang, "Gsnet: Joint vehicle pose and shape reconstruction with geometrical and scene-aware supervision," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XV 16*. Springer International Publishing, 2020, pp. 515–532.

[9] T. Barowski, M. Szczot, and S. Houben, "6dof vehicle pose estimation using segmentation-based part correspondences," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 573–580.

[10] F. Chabot, M. Chaouch, J. Rabarisoa, C. Teuliere, and T. Chateau, "Deep manta: A coarse-to-fine many-task network for joint 2d and 3d vehicle analysis from monocular image," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 2040–2049.

[11] Y. Lyu, R. Royen, and A. Munteanu, "Mono6d: Monocular vehicle 6d pose estimation with 3d priors," in *2022 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2022, pp. 2187–2191.

[12] C.-M. Chen, Q. Miao, F. Khan, G. Srivastava, and S. Kumari, "Sustainable secure communication in consumer-centric electric vehicle charging in industry 5.0 environments," *IEEE Transactions on Consumer Electronics*, pp. 1–1, 2023.

[13] J. Jiang, X. Luo, Q. Luo, L. Qiao, and M. Li, "An overview of hand-eye calibration," *The International Journal of Advanced Manufacturing Technology*, vol. 119, no. 1, pp. 77–97, 2022.

[14] M. Pan, C. Sun, J. Liu, and Y. Wang, "Automatic recognition and location system for electric vehicle charging port in complex environment," *IET Image Processing*, vol. 14, no. 10, pp. 2263–2272, 2020.

[15] W.-P. Ma, W.-X. Li, and P.-X. Cao, "Binocular vision object positioning method for robots based on coarse-fine stereo matching," *International Journal of Automation and Computing*, vol. 17, no. 4, pp. 562–571, 2020.

[16] A. Yao, "Design of recognition and positioning system of electric vehicle charging hole based on binocular vision," M.E.dissertation, Zhejiang University of Technology, 2020.

[17] Z. Xu, H. Zheng, K. Sun, Y. Lu, T. Liu, and L. Yuan, "Research on position estimation method of intelligent forklift based on binocular vision," in *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*. IEEE, 2020, pp. 604–607.

[18] H. Wu, W. Hao, Y. Sun, Z. Wang, and Y. Huang, "Calibration error study of binocular stereo camera," in *Journal of Physics: Conference Series*. IOP Publishing, 2022, No. 1, p. 012058.

[19] V. Tadic, A. Toth, Z. Vizvari, M. Klincsik, Z. Sari, P. Sarcevic, J. Sarosi, and I. Biro, "Perspectives of realsense and zed depth sensors for robotic vision applications," *Machines*, vol. 10, no. 3, p. 183, 2022.

[20] P. Ram and S. Padmavathi, "Analysis of harris corner detection for color images," in *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*. IEEE, 2016, pp. 405–410.

[21] M. Liao, Y.-Q. Zhao, X.-H. Li, P.-S. Dai, X.-W. Xu, J.-K. Zhang, and B.-J. Zou, "Automatic segmentation for cell images based on bottleneck detection and ellipse fitting," *Neurocomputing*, vol. 173, pp. 615–622, 2016.

[22] Y. Lu, S. Kourian, C. Salvaggio, C. Xu, and G. Lu, "Single image 3d vehicle pose estimation for augmented reality," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2019, pp. 1–5.

[23] H. Zeng, "A 3d coordinate transformation algorithm," in *2010 The 2nd Conference on Environmental Science and Information Application Technology*. IEEE, 2010, pp. 195–198.

[24] F. Yang, H. Dai, and H. Xing, "Least squares based on rodrigues matrix and its application in similar material model of mining," in *2015 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE, 2015, pp. 1686–1690.

[25] J. Cashbaugh and C. Kitts, "Automatic calculation of a transformation matrix between two frames," *IEEE Access*, vol. 6, pp. 9614–9622, 2018.