A Multi-sensor Information Fusion Method for Autonomous Vehicle Perception System

Peng Mei[1,2(✉)], Hamid Reza Karimi[2], Fei Ma[3], Shichun Yang[1], and Cong Huang[2,4]

[1] School of Transportation Science and Engineering, Beihang University, Beijing, China
maple@buaa.edu.cn
[2] Department of Mechanical Engineering, Politecnico di Milano, Milan, Italy
[3] Autonomous Driving Center, SAIC Motor Passenger Vehicle Co., Shanghai, China
[4] School of Transportation and Civil Engineering, Nantong University, Nantong, China

Abstract. Within the context of the environmental perception of autonomous vehicles (AVs), this paper establishes a sensor model based on the experimental sensor fusion of lidar and monocular cameras. The sensor fusion algorithm can map three-dimensional space coordinate points to a two-dimensional plane based on both space synchronization and time synchronization. The YOLO target recognition and density clustering algorithms obtain the data fusion containing the obstacles' visual information and depth information. Furthermore, the experimental results show the high accuracy of the proposed sensor data fusion algorithm.

Keywords: Autonomous driving · Sensor data fusion · Lidar and Monocular camera

1 Introduction

Autonomous driving has recently become a hot topic in the automotive field, while low-level autonomous vehicles have gradually entered people's daily lives [1–3]. Based on the driving environment, the vehicle perception system can provide real-time information for the AVs.

As a vital part of AVs, the environment perception system can provide information about the driving environment for the AVs. The driving environment is usually quite complex, and it is hard to be predicted. Thus, AVs need to focus on the changes in the surrounding environment. It is challenging to complete all perception tasks with a specific sensor [4]. Generally, the sensors used in AVs include onboard cameras, lidars, radars, and millimeter-wave radars [5]. Compared to other sensors, lidar has more obstacle detection and tracking advantages. According to the running time of the laser, we can get the relative distance between the AVs and the obstacle. Due to the high energy density and frequency, the laser beams can easily capture the contour details of obstacles in the environment. Compared with lidar, millimeter-wave radar has no advantage in capturing detailed information. However, its anti-interference ability is

S. Paiva et al. (Eds.): SmartCity360° 2021, LNICST 442, pp. 633–646, 2022.
https://doi.org/10.1007/978-3-031-06371-8_40

strong, and it is often used for sensing road conditions. The camera can express the surrounding environment intuitively, but it is hard to obtain accurate depth information. Radar is cheap and has a long service life. Because of the low resolution, radar has significant limitations in acquiring dynamic information [6].

As the sensor combination of lidar and camera can well identify surrounding vehicles and other obstacles, AVs sensing systems have begun to combine lidar and camera in recent years [7]. The lidar receives three-dimensional space coordinate information, while the camera obtains two-dimensional plane coordinate information. The key is to establish a mapping relationship between three-dimensional space points and a two-dimensional plane to fuse these sensors' data.

Some scholars have done related research on the multi-sensor fusion algorithm of AVs. Elhousni et al. [8] believe that the material of the calibration target will affect the reflection intensity of the laser signal, which will affect the final calibration result. Therefore, they use a glass calibration target, and the surface is smooth and has high reflectivity. By fitting the contour of the calibration target, its edge line equation is obtained. Due to the penetrability of the laser beam is not considered, there is a gap between the calibration result and the expected value. Vasconcelos et al. [9] used a single-line lidar instead of a multi-line lidar case to reduce the error. Specifically, a single-line lidar is used to scan the entire calibration target area. With the small number of laser beams, the amount of data is less than that of the multi-line lidar. Therefore, this method can obtain a more accurate checkerboard corner detection. To simplify the calibration process, Zewei et al. [10] achieved the external parameter calibration of the camera by using Zhang's calibration method. However, the obtained rotation matrix ignored the relative distance between the camera and lidar coordinate systems. The experimental approach has shown some limitations; when the radar's relative position and the camera are far, the calibration result's deviation error will be obvious. On this basis, Park et al. [11] optimized the calibration process and obtained a better fusion effect. However, this method only achieves fusion from the data level, which is not optimized for specific goals. Based on the independence of two planes, Li et al. [12] described the corresponding relationship between laser point cloud coordinates and image coordinates. Finally, they obtained the rotation and translation relationship.

In summary, as two indispensable sensors for AVs, camera and lidar have their advantages and disadvantages, respectively. It is necessary for the perception system to fuse these two sensors' data and then obtain environmental information. The experimental results show that the accuracy of the data fusion method reaches up to 98.42%.

2 Sensor Model

2.1 Camera Observation Model

Through the principle of optical imaging, the vehicle camera maps the coordinate information of the three-dimensional space to the two-dimensional plane, thereby realizing the imaging of the object [13]. Typically, the specific position of the camera in space is uncertain, so it is necessary to establish a spatial coordinate reference system to describe this mapping relationship. Thus, this article introduces four coordinate system references, as shown in Fig. 1.

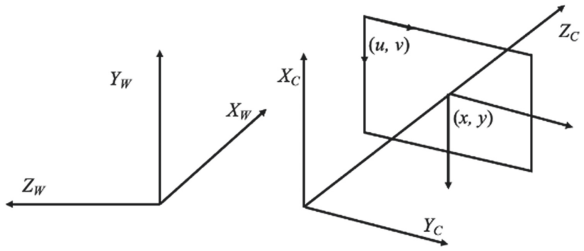


Fig. 1. Coordinate system diagram.

The coordinate points are represented by (X_w, Y_w, Z_w) in the world coordinate system, it can describe the proper object position in three-dimensional space. In the camera coordinate system, the coordinate points are represented by (X_c, Y_c, Z_c). Generally, a specific coordinate axis of the camera coordinate system coincides with the camera's optical axis. The coordinate point in the image coordinate system is represented by (x, y), and the coordinate points in the pixel coordinate system are represented by (u, v).

The target under each coordinate system can be scaled, rotated, and translated to realize the transformation of the research object. The rotation transformation can be achieved by using a rotation vector through a three-dimensional vector, it can be shown as follows,

$$R = [R_1, R_2, R_3] \tag{1}$$

where R means the rotation matrix.

The calculation formula can be shown as follows,

$$t = \left[t_x, t_y, t_z\right]^T \tag{2}$$

where t denotes the translation vector.

Through the rotation matrix R and the translation vector t, the conversion relationship between the world coordinate system and the specific coordinate points in the camera coordinate system can be described as follows [12],

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + t = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{3}$$

In order to simplify the correspondence between the camera coordinate system and the image coordinate system, it is assumed that the X_c axis and Y_c axis of the camera coordinate system are corresponding to the x-axis and y-axis of the image coordinate system, respectively. Therefore, the corresponding relationship between the two coordinate systems can be described as follows,

$$\begin{bmatrix} x \\ y \end{bmatrix} = f \begin{bmatrix} \frac{X_c}{Z_c} \\ \frac{Y_c}{Z_c} \end{bmatrix} \tag{4}$$

where f is the focal length of the camera.

The transformation between the pixel coordinate system and the image coordinate system is mainly obtained through the scaling of the corresponding coordinate axis. Assuming that any pixel point coordinates are the corresponding image coordinate point, it is scaled by α times and β times in the x-axis and y-axis respectively. The origin of the pixel coordinate system is located at the top left corner of the image, and the image coordinate system will produce c_x and c_y offsets in the x-axis and y-axis, respectively. Based on the above analysis, the corresponding relationship between the pixel coordinates and the image coordinates can be obtained as follows,

$$u = \alpha x + c_x, v = \beta y + c_y \tag{5}$$

where c_x and c_y are the offsets parameters.

In summary, the coordinate points in a certain world coordinate system can be mapped to the pixel coordinate system by formula (6).

$$\begin{aligned} z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} &= K \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_x \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \\ &= \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_x \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \right) \end{aligned} \tag{6}$$

There are usually two types of distortion, including barrel distortion and pincushion distortion, as shown in Fig. 2. The characteristic of barrel distortion is that as the distance between the imaging point and the optical axis increases, the image magnification decreases. As the distance between the imaging point and the optical axis increases, the image magnification increases, and it is called pincushion distortion. The processing of imaging distortion is generally achieved by calibrating the camera's internal parameter matrix.

(a) Normal (b) Barrel distortion (c) pincushion distortion

Fig. 2. Imaging distortion.

2.2 Lidar Point Cloud Processing

This paper proposes straight-through filtering to preprocess the original point cloud data. The specific method is given as follows. First, define the coordinate system of the point

cloud data. As shown in Fig. 3, assume that the laser radar installation position is the origin of the coordinates, the lateral direction of the vehicle is the x-axis, the longitudinal direction of the vehicle is the y-axis, and the vertical direction of the vehicle body is the z-axis. According to the above coordinate system, the straight-through filtering method limits the range of point cloud data. Among them, $x \in [-10,10]$, $y \in [-10,30]$, $z \in [-1,2.5]$. This range can ensure the obstacles are detected in the surrounding environment, and it can also filter out point cloud data that is not related to environmental perception. The point cloud data preprocessing is shown in Fig. 4.

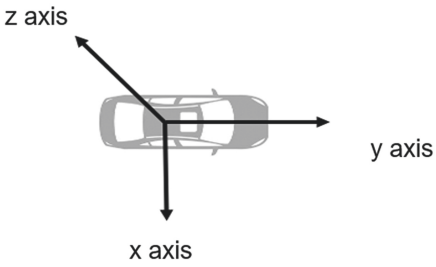
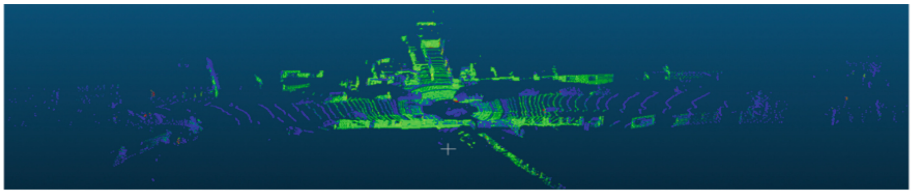
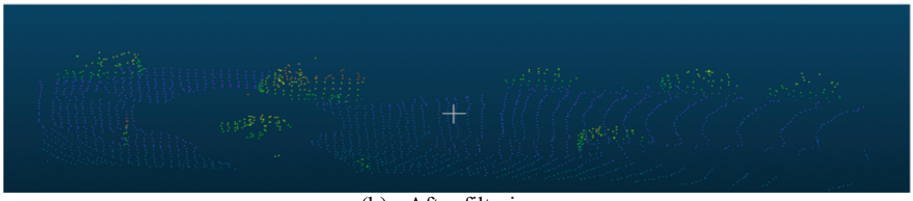


Fig. 3. Point cloud data coordinate system.



(a) Before filtering.



(b) After filtering.

Fig. 4. Data point cloud preprocessing.

2.3 Sensor Parameters

The sensors used in this article are Beike Tianhui mechanical Lidar and LI-USB30-AR023ZWDRB monocular camera to verify the sensor fusion algorithm. The technical parameters of the sensor are shown in Tables 1 and 2.

Table 1. Lidar technical parameters.

Type	R-Fans 32
Detection distance	200 m
Measurement accuracy	<2 cm
Horizontal angular	<0.1°
Working frequency	15 Hz
Vertical field of view	32°
Vertical angular	1°
Operating voltage	24VDC
Operating temperature	(−20,55)°C
Size	113 mm(D) × 70 mm(H)
Power consumption	<15 W

Table 2. Camera technical parameters.

Type	LI-USB30
Resolution	1920 × 1080
Frame rate	30 fps
Size	30 mm(D) × 30 mm(H)
Operating voltage	5VDC
Operating temperature	(−20,80)°C
Power consumption	<5 W

3 Sensor Data Synchronization Mapping

3.1 Spatial Synchronization

Sensor spatial synchronization is mainly used to realize the unification of lidar coordinates and camera pixel coordinates. Generally, the mapping relationship between the three-dimensional space coordinate system and the two-dimensional image coordinate system is obtained by calibrating the camera's external parameters. The two-dimensional pixel point be expressed as follows,

$$m = [u, v]^T \tag{7}$$

where m means pixel projection point.

Its corresponding three-dimensional space point can be expressed as follows,

$$M = [X, Y, Z]^T \tag{8}$$

where M denotes three-dimensional space point.

Perform homogeneous coordinate transformation of the above two vectors, ones get,

$$\tilde{m} = [u, v, 1]^T \tag{9}$$

$$\tilde{M} = [X, Y, Z, 1]^T \tag{10}$$

According to the imaging principle of the camera, the corresponding relationship between M and m can be shown as follows,

$$\tilde{m} = \frac{1}{s}\mathrm{A}[R\,t]^T\tilde{M} \tag{11}$$

where s stands for the scale factor, and A means the internal parameters of the camera. $[R, t]$ is to transform the world coordinate system to the camera coordinate system for rotation matrix and translation vector, which is the external parameters of the camera.

Define A as follows,

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12}$$

where (u_0, v_0) are the principal coordinate points, α and β are scaling factors, which contains the focal length information of the camera, γ is used to describe the skewness of the two image coordinate axes and the pixel coordinate axis.

It is assumed that the plane of the calibration target coincides with the base plane of the world coordinate system, that is $Z = 0$. Denote the i-th column in the rotation matrix R as r_i, which can be obtained by formula (11):

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A\begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix}\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = A\begin{bmatrix} r_1 & r_2 & t \end{bmatrix}\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{13}$$

Then the relationship between the point M and its corresponding pixel m can be defined by a homography matrix H, we can obtain the following equation,

$$s\tilde{m} = H\tilde{M} \tag{14}$$

$$H = A\begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \tag{15}$$

From the Eq. (13), we can get the following equations,

$$u = \frac{h_{11}X + h_{12}Y + h_{13}}{h_{31}X + h_{32}Y + h_{33}} \tag{16}$$

$$v = \frac{h_{21}X + h_{22}Y + h_{23}}{h_{31}X + h_{32}Y + h_{33}} \tag{17}$$

From the Eq. (15), we can get the following equation,

$$\begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = \lambda A\begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \tag{18}$$

where λ is any non-zero constant. According to the nature of the rotation matrix, r_1 and r_2 are orthogonal vectors, then we can get:

$$h_1^T A^{-T} A^{-1} h_2 = 0 \tag{19}$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \tag{20}$$

Define the matrix B as follows,

$$B = A^{-T} A^{-1} = [B_{i,j}]_{3*3} \ (i, j \ = \ 1, 2, 3) \tag{21}$$

Since B is a symmetric matrix, set the vector b as follows,

$$b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \tag{22}$$

The i-th column of the homography matrix H can be expressed as,

$$h_i = [h_{i1}, h_{i2}, h_{i3}]^T \tag{23}$$

$$h_i^T B h_j = v_{ij}^T b \tag{24}$$

where $v_{ij} = \left[h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}\right]$.

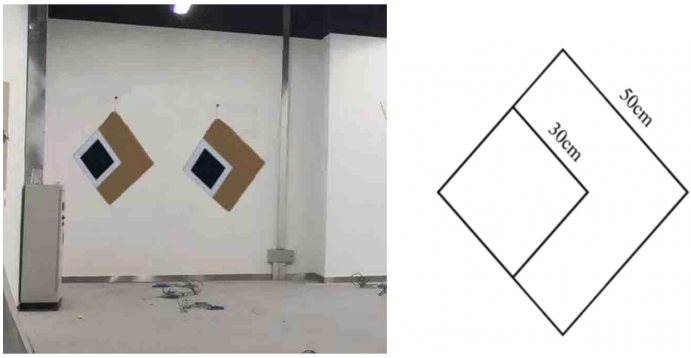
From (23) and (24), we can get the following formula,

$$\begin{bmatrix} v_{11}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \tag{25}$$

$$vb = 0 \tag{26}$$

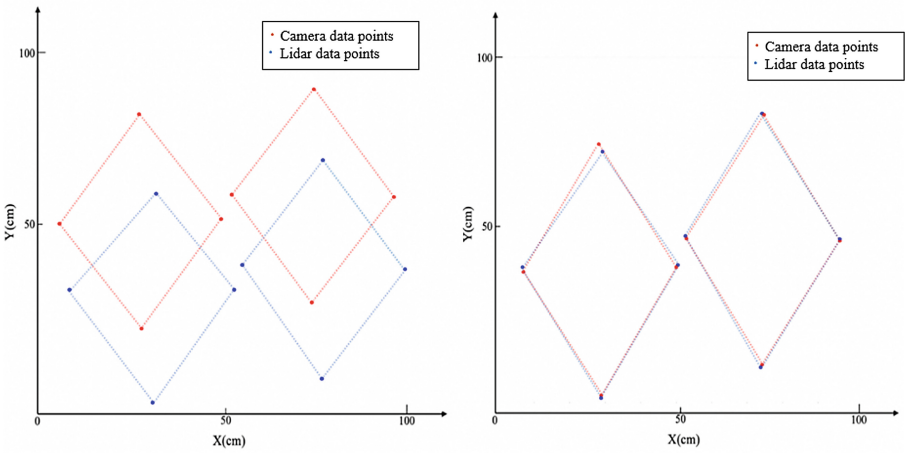
Based on the calibration target to collect images at different positions, the matrix B and homography matrix H can be obtained. The homography matrix H contains the camera's internal and external parameters. To achieve spatial synchronization, Cholesky decomposition is adopted to obtain the camera's internal and external parameters.

According to the spatial synchronization method, the lidar and the camera are spatially synchronized. As shown in Fig. 5, the square paper calibration target is used, and it is hung vertically about 3 m in front of the sensor. Figure 6 is a comparison diagram of the sensors used before and after spatial synchronization. The red color represents the contour of the object in the pixel coordinate system. After coordinating conversion, the blue point represents the object contour obtained by the laser radar, and then output to the pixel coordinate system. Among them, Fig. 6(a) indicates the fusion effect without spatial synchronization, Fig. 6(b) shows the effect after synchronization. It can be seen that after space synchronization, the objects in the two coordinate systems can achieve the effect of overlapping.



(a) Actual calibration target (b) The size of calibration target

Fig. 5. Calibration target parameters.



(a) Before space synchronization (b) After space synchronization

Fig. 6. Spatial synchronization effect comparison. (Color figure online)

3.2 Time Synchronization

Due to the different operating frequencies of lidar and vehicle camera, their respective observation times are also inconsistent. If we use non-synchronized sensor data to perform data fusion directly, it will affect the accuracy of the fusion result. Therefore, time synchronization is a prerequisite for the integration of lidar and camera data. In the environment of the robot operating system, a certain moment is taken as a benchmark. Find the time stamp of the lidar point cloud data and camera image data at this moment, and then keep the data with the same time stamp.

Based on both space synchronization and time synchronization, the synchronous mapping of lidar and camera data can be achieved. The laser radar point cloud data is output to the camera image data, and the processing effect is shown in Fig. 7. It can be

seen that the sensor can not only work under a unified time and space reference, but also can identify obstacle information in the surrounding environment.

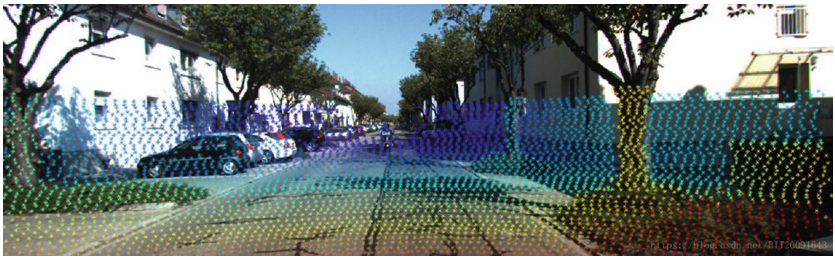


Fig. 7. Synchronous mapping effect.

4 Target Recognition and Depth Information Association

4.1 Camera Target Detection

Target detection refers to the ability to accurately identify the type and location of an object through a target recognition algorithm. YOLO target detection treats the target detection problem as a regression problem and then directly returns the position and category of the Bounding Box in the output layer [1, 14, 15]. Its advantages are fast recognition speed and high accuracy, which can be used in real-time systems. This article uses the YOLO algorithm to achieve target recognition.

4.2 In-Depth Information Association

Above-mentioned shows that the sensor data was synchronized in time and space to complete the vehicle target recognition. However, the integration of data is not related to each other, so this data fusion needs to be processed.

To solve this problem, this paper is based on the DBSCAN (Density-based spatial clustering algorithm with noise) algorithm to achieve sensor depth information association [16–19]. This algorithm is a commonly used density clustering algorithm. The main idea is to divide the dense points in the area into the same cluster. It can realize clusters of any shape without inputting category information in advance. Therefore, it can be used for the preliminary fusion data to realize deep information association.

4.3 Experimental Verification

This paper uses the yellow and blue cones, a diameter of 20 cm and a height of 20 cm, to verify the target recognition algorithm and depth information association, as shown in Fig. 8.

First, a training data set is generated by labeling 1000 pictures containing cones, and the marked feature information is the color of the cones. Then train the Darknet53 neural

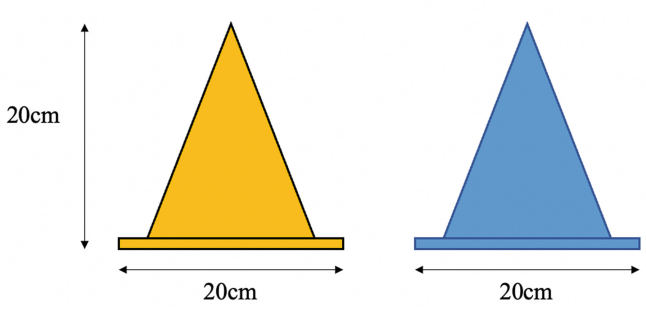


Fig. 8. The size of the cone.

network under the YOLO target recognition framework. The camera can accurately identify the cone information, while mark the shape and color of the cone in the picture. Secondly, the image data and the lidar point cloud data are synchronized in time and space, simultaneously, to obtain preliminary fusion data. Then, we can find a cone point group in the pixel coordinate system. By calculating the sum of the coordinate variances of each point from the remaining points, we can obtain a point with the smallest variance sum. In this process, we regard the coordinates of this point as the cone coordinate reference point. Finally, the integration of data with cone color information and depth information is obtained.

The experiment process is shown in Fig. 9. The two-color cones are arranged as shown in the figure, and the initial position is the observed position 1.

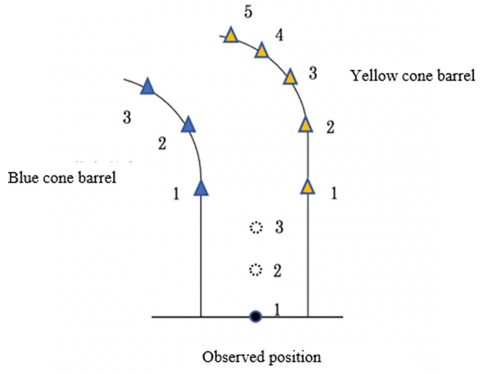


Fig. 9. A schematic diagram of fusion effect verification experiment.

The data fusion of the sensor moving from position 1 to position 3 is recorded, and 100 frames of fusion data of three observation points are respectively taken to obtain 300 sets of data. Visualize one group of data, as shown in Fig. 10.

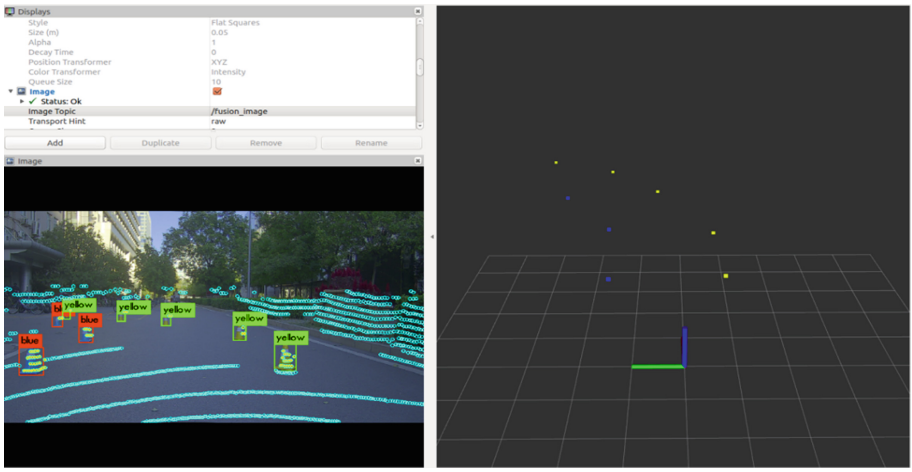


Fig. 10. Verify experiment visualization.

4.4 Results and Discussion

(1) Target recognition error discussion

Above all, each data group contains three blue cones and five yellow cones, with 2400 recognition results. The data obtained from the statistical experiment is shown in Table 3, and the table is called the confusion matrix. The sum of each row represents the true sample size of the category. Each column of the confusion matrix represents an accurate classification of the cones. The result shows its accuracy rate reaches 98.42%.

Table 3. Statistics of experimental results.

n = 2400	Recognized as blue	Recognized as yellow	Total
Actually blue	887	13	900
Actually yellow	25	1475	1500
Accuracy	(887 + 1475)/2400 = 98.42%		

(2) Depth information error discussion

Since the depth information in the fusion algorithm used in this paper comes directly from the measurement data of the lidar, the distance error of the fusion data is consistent with the measurement error of the lidar. Moreover, the distance error is less than 2 cm. Therefore, the data fusion method proposed in this paper has high accuracy.

5 Conclusion

Aiming to improve the environment perception of AVs, this paper combines the sensor data of lidar and monocular cameras. This paper proposes a sensor fusion algorithm that combines space synchronization and time synchronization simultaneously to solve the problem of different sensor time and space coordinate systems. For the spatial synchronization of the lidar and the camera, the external parameter matrix of the camera is calibrated through a checkerboard grid, which realizes the mapping of three-dimensional space coordinate points to a two-dimensional plane. Based on YOLO's target recognition algorithm, the recognition of obstacles is completed. In view of the density clustering algorithm, the target depth information association is realized. Finally, the fusion data containing the visual information and depth information of obstacles is obtained. Through the experiment result, we can verify the effect of the sensor data fusion algorithm. In the following work, we will compare the accuracy of different algorithms and apply them to AVs with these two sensors.

References

1. Kiran, B.R., Sobh, I., Talpaert, V., et al.: Deep reinforcement learning for autonomous driving: a survey. IEEE Trans. Intell. Transp. Syst. (2021)
2. Karimi, H.R., Lu, Y.: Guidance and control methodologies for marine vehicles: a survey. Control. Eng. Pract. **111**, 104785 (2021)
3. Mei, P., Karimi, H.R., Yang, S., et al.: An adaptive fuzzy sliding-mode control for regenerative braking system of electric vehicles. Int. J. Adapt. Control Signal Process. **36**(2), 391–410 (2022)
4. Thuruthel, T.G., Shih, B., Laschi, C., et al.: Soft robot perception using embedded soft sensors and recurrent neural networks. Sci. Robot. **4**(26), eaav1488 (2019)
5. Zang, S., Ding, M., Smith, D., et al.: The impact of adverse weather conditions on autonomous vehicles: how rain, snow, fog, and hail affect the performance of a self-driving car. IEEE Veh. Technol. Mag. **14**(2), 103–111 (2019)
6. Bilik, I., Longman, O., Villeval, S., et al.: The rise of radar for autonomous vehicles: signal processing solutions and future research directions. IEEE Signal Process. Mag. **36**(5), 20–31 (2019)
7. Roriz, R., Cabral, J., Gomes, T.: Automotive LiDAR technology: a survey. IEEE Trans. Intell. Transp. Syst. 1–16 (2021)
8. Elhousni, M., Lyu, Y., Zhang, Z., et al.: Automatic building and labeling of HD maps with deep learning. Proc. AAAI Conf. Artif. Intell. **34**(08), 13255–13260 (2020)
9. Vasconcelos, F., Barretoj, P., Nunes, U.: A minimal solution for the extrinsic calibration of a camera and a laser-rangefinder. IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2097–2107 (2012)
10. Liu, Z., Lu, D., Qian, W., et al.: Extrinsic calibration of a single-point laser rangefinder and single camera. Opt. Quantum Electron. **51**(6), 1–13 (2019)
11. Park, S., Chung, M.: Extrinsic calibration between a 3D laser scanner and a camera using PCA method. In: 2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), pp. 527–528 (2012)
12. Li, N., Hu, Z., Zhao, B.: Flexible extrinsic calibration of a camera and a two-dimensional laser rangefinder with a folding pattern. Appl. Opt. **55**(9), 2270–2280 (2016)

13. Yu, H., Tseng, H.E., Langari, R.: A human-like game theory-based controller for automatic lane changing. Transp. Res. Part C Emerg. Technol. **88**, 140–158 (2018)
14. Santos, H., Pereira, G.V., Budde, M., Lopes, S.F., Nikolic, P.: Science and Technologies for Smart Cities: 5th EAI International Summit, SmartCity360, Braga, Portugal, 4–6 December 2019, Proceedings. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-51005-3
15. Huang, C., Karimi, H.R.: Non-fragile H∞ control for LPV-based CACC systems subject to denial-of-service attacks. IET Control Theory Appl. **15**(9), 1246–1256 (2021)
16. Guo, F., Hao, K., Xia, M., Zhao, L., Wang, L., Liu, Q.: Detection of insulator defects based on YOLO V3. In: Han, S., Ye, L., Meng, W. (eds.) AICON 2019. LNICSSITE, vol. 287, pp. 291–299. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22971-9_25
17. Huang, C., Mei, P., Wang, J.: Event-triggering robust fusion estimation for a class of multi-rate systems subject to censored observations. ISA Trans. **110**, 28–38 (2021)
18. Cai, Z., Wang, J., He, K.: Adaptive density-based spatial clustering for massive data analysis. IEEE Access **8**, 23346–23358 (2020)
19. Lu, Y., Karimi, H.R.: Recursive fusion estimation for mobile robot localization under multiple energy harvesting sensors. IET Control Theory Appl. **16**, 20–30 (2021)