A Modified Online Intelligent Method to Calibrate Radar and Camera Sensors for Data Fusing

Guobin Xu¹, Dongdong Li², Xiangyang Chen², Qiankun Li² and Jun Yin²

¹Zhejiang University, Hangzhou, Zhejiang, China
²Zhejiang Dahua Technology Co., Ltd. Hangzhou, Zhejiang, China
Email: aiguozhefei@126.com

Abstract. This paper introduces a modified online intelligent calibration method to calibrate radar and camera sensors in traffic scenarios. Radar and camera sensor calibration is to construct a one-to-one mapping between the radar coordinate system and the image pixel coordinate system. Based on the currently widely used artificial intelligence method (deep neural networks), camera can accurately recognize the target type. For the research and application of radar sensor and camera sensor fusion, the most important thing is to synchronize the sensor sampling time of the radar and the camera and space calibration. This paper emphasizes on spatial calibration. At present, the common calibration methods are the calibration method based on the internal and external parameters of the camera, the four-point calibration method, and the neural network calibration method. However, these three calibration methods have significant advantages and disadvantages. In the field of transportation, the disadvantages are even more prominent. This paper proposes a modified online intelligent calibration method for the application field of radar and camera sensors in traffic scenarios. We have proved through simulation experiments that this method is less time-consuming and manual operation while ensuring high-precision calibration performance.

Keywords. Multiple sensors; calibration; data fusing.

1. Introduction

In recent years, due to the rapid development of artificial intelligence technology [1, 2], intelligent monitoring systems have emerged and are widely used in various fields of social production. As the core technology of the intelligent monitoring system, the object detection and tracking algorithm has attracted the attention of more and more experts and scholars, and has become a hot research field [3, 4]. Intelligent monitoring equipment has higher detection accuracy and detection efficiency, and can replace humans in some dangerous scenarios. For example, more and more cities choose to use intelligent cameras or radars in the traffic field to assist traffic police in the detection of illegal activities.

In the field of intelligent transportation, the industry is concerned about a core issue: who did what and when? That is, sufficient evidence must be provided to indicate that someone did a specific thing at a specific time. “When” is the issue of time, which is easier to solve. The issue “Who” can be solved by intelligent cameras. The key is how to solve the issue “Doing what”. Some scholars in the industry are studying to solve the third problem solely relying on intelligent cameras with “Action Detection” [5, 6], the effect is not satisfactory, and it cannot be put into actual products in a short period of time. Fortunately, the issue “Doing what” can be solved by radar. Radar can obtain information such as the object’s position, speed and RCS (Radar Cross Section) by measuring the
electromagnetic waves reflected by the object [7]. Through this information, it can be judged whether the vehicle has illegal behaviors such as speeding, underspeeding, and illegal lane change. This is the essential advantage of radar and camera fusion.

The advantages of radar and camera fusion are significant, but there are some problems to be overcome, such as the synchronization of the sampling time of radar and camera sensors, the spatial calibration between the radar and camera coordinate system, the identity fusion and state estimation of the object, redundant attribute data processing. The first problem to be solved is the spatial calibration, which refers to the construction of an one-to-one mapping between the radar coordinate system and the camera pixel coordinates. Based on this, the corresponding pixel coordinates can be obtained by any radar coordinate point, or the radar coordinate point can be obtained by any pixel coordinate.

At present, there are three common methods of radar and camera calibration: the calibration method based on the internal and external parameters of the camera, the calibration method based on four points, and the online intelligent calibration method [8]. The calibration method based on the camera’s internal and external parameters refers to obtaining the camera’s internal parameters and installation information (three-dimensional coordinate position and installation attitude angle), as well as radar installation information. Based on the camera imaging principle, we can get the conversion relationship between the radar coordinate system and the pixel. The advantage of this method is strong adaptability; but its shortcomings are also obvious. It is necessary to know the internal parameters of each camera and obtain high-precision equipment installation information through the sensor. As we all know, even cameras of the same model and batch have significant differences in internal parameters; obtaining high-precision equipment installation information will undoubtedly increase equipment construction, measurement difficulty and equipment cost. Therefore, the calibration method based on the internal and external parameters of the camera is not suitable for batch products. The four-point calibration method has the advantage of high calibration accuracy; but the disadvantage is that the four key points in the radar coordinate system need to measure in advance. It is usually not easy to meet this requirement, such as in dangerous areas such as highways. Moreover, every time the device moves, it needs to be re-calibrated manually. The online intelligent calibration method is a calibration method based on multi-layer neural network proposed in 2020. It is a kind of intelligent calibration and self-correction without manual intervention. This is a major advancement in the field of radar and camera calibration, but it also has a significant shortcoming: it takes a long time. Because of online intelligent calibration, a large amount of data needs to collect and store to train the network. If the amount of data is insufficient or the data quality is not high, the network will converge slowly. That is, online intelligent calibration is a calibration method is a continuously optimizing method. During the long calibration process, the equipment cannot carry out normal business (business that depends on the calibration result).

This paper proposes a new calibration method—a modified online intelligent calibration method, which can effectively reduce the calibration complexity, reduce human intervention, and the calibration speed is fast. The calibration accuracy is high.

2. Introduction for R-C Spatial Calibration

Figure 1 shows the appearance of an intelligent radar camera monitoring device. The upper part is the camera, and the radar is installed under the camera. Radar sensor and camera sensor detect and track targets at the same time, and the actual detection results are shown in figure 2. The figure on the left is the result of camera monitoring, and the figure on the right is the result of radar monitoring. In the camera figure, the AI algorithm automatically detects the vehicle target and uses a box to mark the target vehicle. The current location and historical movement trajectory of each target can be displayed in the radar figure.
The core of multi-sensor fusion detection is the space calibration and time synchronization between different types of sensors [9]. Compared with the calibration of lidar and camera, the sampling data of millimeter wave radar is more sparse, so the calibration method of lidar is not suitable for the calibration of millimeter wave radar and camera [10]. Most of the radar-camera equipment in traffic scenes is installed at a specific location above the road. The special environment prevents the use of checkerboard calibration boards and other calibration equipment [11].

Assuming that there is a target point in the real world, the target can be expressed as $X = [x, y]^T$ in the radar coordinate system, and $U = [u, v]^T$ in the camera coordinate system. Therefore, the relationship between the real target point in the radar coordinate system and the camera coordinate system can be obtained as:

$$U = F(X)$$

$$X = G(U)$$
\[ F = G^{-1} \]  
(3)

\[ G = F^{-1} \]  
(4)

where \( F \) and \( G \) are the mapping relationships that need to be solved for space calibration. In other words, the space calibration of the radar and the camera is to find a coordinate system mapping function, which can correctly convert the radar coordinates and the camera coordinates.

3. Modified Online Intelligent Calibration Method

3.1. Modified Calibration Method

The four-point calibration method and the online intelligent calibration method give us a great inspiration. The advantage of the four-point calibration method is that the amount of data is small and the mathematical principle is clear; the disadvantage is that in the field of radar application. There is great difficulty in obtaining the precise coordinate positions of the four points in the radar coordinate system. The online intelligent calibration method is a process in which enough data is obtained at a certain time cost, and the data is trained online to gradually improve the calibration accuracy. From the perspective of the number of solving parameters, the four-point calibration method has eight unknown variables to be solved; the online intelligent calibration method has thousands of unknown variables to be solved. This also determines the amount of data required for the latter, the time-consuming solution, and the amount of calculation far greater than the former.

The modified online intelligent calibration method absorbs the advantages of both the four-point calibration method and the online calibration method, while avoiding the disadvantages of both, that is, the use of online data collection and online solution of the eight parameters in the four-point calibration method. This not only avoids the situation of a large number of unknown variables to be solved, but also avoids the situation of manually participating in the configuration of the midpoint coordinates of the radar coordinate system.

For the modified online intelligent calibration method, the key is to identify the ideal environment. The ideal environment means that the vehicles are very sparse, that is, there is at most one target in each lane in the overlapping area of radar and video surveillance. The goal is to avoid target association/matching ambiguity during the online calibration process.

When the ideal conditions have met and data have been collected on any two lanes, the mapping relationship between the radar coordinate system and the image coordinate system can be dissolved online. Assuming that the collected data radar video target coordinate information is:

\[ t_1, BoxPt_1, CartPt_1 \]
\[ t_2, BoxPt_2, CartPt_2 \]
\[ \ldots \]
\[ t_i, BoxPt_i, CartPt_i \]
\[ \ldots \]
\[ t_n, BoxPt_n, CartPt_n \]

where, \( t_i \) is time, \( BoxPt_i = (u_i, v_i) = U_i \) is the position of a certain point of the camera target frame, usually the center point, \( CartPt_i = (x_i, y_i) = X_i \) is a certain radar target point. Construct the following homogeneous equation:

\[
\begin{bmatrix}
  x_i \\
  y_i \\
  1
\end{bmatrix}
= 
\begin{bmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{bmatrix}
\begin{bmatrix}
  u_i \\
  v_i \\
  1
\end{bmatrix}
\]

(5)
Further transformation, we can get:

\[ \begin{align*}
\dot{x}_i &= a_{1i} \cdot u_i + a_{12} \cdot v_i + a_{13} \\
\dot{y}_i &= a_{2i} \cdot u_i + a_{22} \cdot v_i + a_{23} \\
1 &= a_{31} \cdot u_i + a_{32} \cdot v_i + a_{33}
\end{align*} \tag{6, 7, 8} \]

Regarding equation (8), regard it as a constraint, then we get the mathematically equivalently equations:

\[ \begin{align*}
\dot{x}_i &= a_{1i} \cdot u_i + a_{12} \cdot v_i + a_{13} \\
\dot{y}_i &= a_{2i} \cdot u_i + a_{22} \cdot v_i + a_{23} \\
1 &= a_{3i} \cdot u_i + a_{32} \cdot v_i + a_{33} \\
\end{align*} \tag{9, 10} \]

From the analysis on the right side of equations (9) and (10), the solution is concerned with the proportional value, which can be made:

\[ \begin{align*}
\dot{x}_i &= b_{1i} \cdot u_i + b_{12} \cdot v_i + b_{13} \\
\dot{y}_i &= b_{2i} \cdot u_i + b_{22} \cdot v_i + b_{23} \\
1 &= b_{3i} \cdot u_i + b_{32} \cdot v_i + 1
\end{align*} \tag{11, 12} \]

Equations (11) and (12) are the radar and camera calibration functions. Then the coefficient matrix can be constructed as follows:

\[
\begin{bmatrix}
u_i & v_i & 1 & 0 & 0 & 0 & -x_i u_i & -y_i v_i \\
0 & 0 & 0 & u_i & v_i & 1 & -y_i u_i & -y_i v_i \\
\end{bmatrix}
W = \begin{bmatrix}
x_i \\
y_i
\end{bmatrix}
\tag{13}
\]

where, \( W = [b_{11}, b_{12}, b_{13}, b_{21}, b_{22}, b_{23}, b_{31}, b_{32}]^T \). Based on equation (13), there are eight unknown quantity demand solutions. In theory, only four pairs of radar video calibration points are enough to solve the unknown variables. However, the radar video measurement in the actual scene has errors after all, and a large number of data point pairs have been obtained. Therefore, we construct the following measurement equations:

\[
\begin{bmatrix}
u_i & v_i & 1 & 0 & 0 & 0 & -x_i u_i & -x_i v_i \\
0 & 0 & 0 & u_i & v_i & 1 & -y_i u_i & -y_i v_i \\
u_2 & v_2 & 1 & 0 & 0 & 0 & -x_2 u_2 & -x_2 v_2 \\
0 & 0 & 0 & u_i & v_i & 1 & -y_2 u_2 & -y_2 v_2 \\
\vdots
\end{bmatrix}
W = \begin{bmatrix}
x_i \\
y_i \\
x_2 \\
\vdots
\end{bmatrix}
\tag{14}
\]

Simply write equation (14) as:

\[
AW = b
\tag{15}
\]

Get:
\[ W = (A^T A)^{-1} A^T b \] (16)

This is the result of the coefficient solution, and the calibration function of the radar and camera coordinate system can be obtained by putting into equations (11) and (12).

3.2. Flow Diagram
We summarize the major steps of the modified online intelligent calibration method, whose flow diagram is shown in figure 3.

![Flow diagram](image)

**Figure 3.** Flow diagram.

Step 1: Initializing. We should initialize traffic road information, and data length threshold of equation (14), and so on. Skip to step 2.

Step 2: Recognize environment. The purpose of this step is to judge when we begin to collect data. That is to say, ideal environment makes unambiguous pairs to ensure the correctness of collected data set. At this step, the target tracking algorithm of radar sensor can provide the target trajectory information real-time, and the artificial intelligent target detecting algorithm of camera sensor can provide the target information real-time. Skip to step 3.

Step 3: Is ideal environment. If current environment is ideal environment, Skip to step 4 to collect data. Otherwise, skip to step 2, wait for next scan, and recognize environment. With the actual experience and business requirements, most of environment is not ideal environment, but we have enough time to collect data.

Step 4: Collect data. Attaching time stamp for radar and camera data and selecting one point in the target area of the camera image. Skip to step 5.

Step 5: Time synchronizing. Based on the time stamp, we get the paired radar and camera data by time synchronizing process. Skip to step 6.
Step 6: Quality and quantity judge. In order to get better solution, we should judge the quality and quantity of data set. It is required that the data is sufficient and the distribution is balanced. If the setting is satisfied, skip to step 7.

Step 7: Is satisfying. If quality and quantity has been satisfied for the collected data, skip to step 8. Otherwise, skip to step 2 (waiting for next scan).

Step 8: Solving equations. We can solve equations (14), (11), (12), then obtain calibration relationship. Skip to step 9.

Step 9: Over.

4. Simulation and Analyses
In order to verify the performance of our new method, we provide some simulations and analyses in this section. We evaluate the performance of different methods by MSE

\[ MSE = \frac{1}{m} \sum_{i=1}^{m} [(x_i^p - x_i^r)^2 + (y_i^p - y_i^r)^2] \] (17)

where \( x_i^p \) and \( y_i^p \) is the predict value of the responding calibration method at time \( i \), \( x_i^r \) and \( y_i^r \) is the real measurement. Three methods are compared: our new method (modified online intelligent calibration method), online intelligent calibration method and modified Zhang calibration method. The result is shown figure 4.

![Simulation Result](image-url)

**Figure 4.** Simulation result.

If we set the right external and intrinsic parameters for the camera sensor and installation information for the radar, the errors of these three methods are similar in value, and the performance is good, as shown in figure 4a. In the second phase of the experiment, we added errors to the measurement of the gyroscope and altimeter. The error of the Zhang method is greater than that of the AI method and the modified AI method. At the same time, the modified AI method has the smallest
error. The result is shown in Figure 4b. In the third stage of the experiment, we moved the device, which is equivalent to modifying the parameters of the space calibration. In the initial stage of the three methods, the error increased, but after a period of time, the AI method and the modified AI can self-correct the parameters, so the error will decrease, and the adjustment speed of the modified AI will be faster, and Zhang method does not have the ability to self-calibrate, even if it runs for a long time, it cannot effectively reduce the system error. The result is shown in the figure 4c. Then we specially focus on rate of convergence between AI method and modified AI method, as shown in the figure 4d. The X axis denotes precision, and Y axis denotes time-consuming, so we find that the modified AI method has a faster convergence speed than the AI method.

5. Conclusions
In the spatial calibration of radar-camera, traditional methods have obvious shortcomings: long calibration time, poor robustness, insufficient anti-distortion ability, and insufficient self-correction ability. The method proposed in this paper further simplifies the manual operation process in the calibration process, can complete the radar-camera calibration more quickly, and has a certain self-correction ability to cope with changes in the environment. The theoretical analysis and simulation experiment prove that the method proposed in this paper improves the calibration performance of the radar-camera.

References
[1] Krizhevsky A, Sutskever I and Hinton G 2012 Imagenet classification with deep convolutional neural networks [C]. Proceedings of the Advances in Neural Information Processing Systems pp 1097-1105.
[2] Ouyang W, Zeng X and Wang X 2015 Single-Pedestrian Detection Aided by Two-Pedestrian Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence 37 (9) 1875-1889.
[3] Ricardo C and Aycard O 2016 Multiple sensor fusion and classification for moving object detection and tracking IEEE Transactions on Intelligent Transportation Systems 17 (2) 525-534.
[4] Ren S Q, He K M, Girshick R and Sun J 2017 Faster R-CNN: Towards real-time object detection with region proposal networks IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (6) 1137-1149.
[5] Liu X C, Liu W and Mei T 2018 Progressive and multimodal vehicle re-identification for large-scale urban surveillance IEEE Transactions on Multimedia 20 (3) 645-658.
[6] Liu M, Li D D, Li Q K and Liu W 2020 An online intelligent method to calibrate radar and camera sensors for data fusing 2nd International Conference on Artificial Intelligence and Computer Science pp 1-9.
[7] Wang D H, Cai Z Y, Zeng J Q and Zhang G Z 2020 Review of traffic data collection research on urban traffic control Journal of Transportation Systems Engineering and Information Technology 20 (03) 95-102.
[8] Ji S W, Xu W and Yang M 2013 3D convolutional neural networks for human action recognition IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (1) 221-231.
[9] Yang J D, Chen L S and Cheng S H 2021 Detecting the calibration failure of multiple sensors Intelligent Computer and Applications 11 (02) 101-104+108.
[10] Liu J Y, Tang X, Jia X H, Yang D and Li T J 2019 High efficient extrinsic parameter calibration method of 3D LiDAR-camera system Chinese Journal of Scientific Instrument 40 (11) 64-72.
[11] Niu P J and Liu L 2019 Joint calibration method of millimeter wave radar and vision sensor based on neural network Journal of Tiangong University 38 (05) 64-69.