


A multi-sensor fusion and object tracking algorithm for self-driving vehicles

Proc IMechE Part D:
J Automobile Engineering
2019, Vol. 233(9) 2293–2300
© IMechE 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0954407019867492
journals.sagepub.com/home/pid


Chunlei Yi¹, Kunfan Zhang and Nengling Peng

Abstract

Vehicles need to detect threats on the road, anticipate emerging dangerous driving situations and take proactive actions for collision avoidance. Therefore, the study on methods of target detection and recognition are of practical value to a self-driving system. However, single sensor has its weakness, such as poor weather adaptability with lidar and camera. In this article, we propose a novel spatial calibration method based on multi-sensor systems, and the approach utilizes rotation and translation of the coordinate system. The validity of the proposed spatial calibration method is tested through comparisons with the data calibrated. In addition, a multi-sensor fusion and object tracking algorithm based on target level to detect and recognize targets is tested. Sensors contain lidar, radar and camera. The multi-sensor fusion and object tracking algorithm takes advantages of various sensors such as target location from lidar, target velocity from radar and target type from camera. Besides, multi-sensor fusion and object tracking algorithm can achieve information redundancy and increase environmental adaptability. Compared with the results of single sensor, this new approach is verified to have the accuracy of location, velocity and recognition by real data.

Keywords

Self-driving vehicles, spatial calibration, multi-sensor fusion, object tracking

Date received: 5 February 2019; accepted: 17 June 2019

Introduction

Increasing road traffic safety and at the same time reducing the number of fatal car accidents is one of the most challenging future tasks for both car manufacturers and research institutions worldwide. Besides intelligent roadside infrastructures, advanced traffic routing and information services, considerable effort is to spend on enhancing the intelligence of individual vehicles within the traffic flow.

Intelligent vehicles are equipped with a variety of environmental perception sensors to detect vehicles and pedestrians. The quality of context perception by a set of sensors is of utmost importance for the so-called Advanced Driver Assistance Systems (ADAS) which rely on the sensor data. Intelligent vehicles utilize sensor systems, control systems, information processing systems to perceive the surrounding environment and possible threats around the vehicle either actively or passively. This significantly enhances the car's ability to anticipate dangerous driving situations and to act early and effectively in order to avoid a collision or at least mitigate the accident severity by proactive activation of

adequate protection means. The technologies for intelligent vehicle have made remarkable progress in recent years.^{1–4} Among them, the environmental perception is very important to detect and recognize targets. Many efforts have been devoted to this field.^{5–13}

For the self-driving vehicles, environmental perception system is more important than ADAS. Different environmental perception sensors, such as lidar, millimeter-wave radar and camera, can be used for environmental perception system. The performance for diverse sensors is different because of their different working principles. Single sensor systems often have undesired weaknesses that suggest the use of multi-sensor systems. Multi-sensor information fusion can make full use of the advantages of each sensor. Multi-

Autonomous Driving R&D Department, Zhengzhou Yutong Bus Co., Ltd., Zhengzhou, People's Republic of China

Corresponding author:

Nengling Peng, Autonomous Driving R&D Department, Zhengzhou Yutong Bus Co., Ltd., Zhengzhou 450000, People's Republic of China.
Email: pengnl@yutong.com

sensor fusion algorithm for environment perception of self-driving vehicles is regarded as a promising instrument to obtain dependable vehicular context information.⁵ It will play a critical role in target detection, tracking and recognition.

The Navlab group at Carnegie Mellon University proposed a high-level fusion approach for object tracking using cameras and lidars.¹⁴ Another interesting effort was the work of Monteiro et al.¹⁵ where they used lidar and camera for obstacle detection and tracking. Since then, an approach of fusing lidar measurements with camera outputs has gained popularity for vehicle and pedestrian tracking. However, most of the teams mainly research target tracking methods by fusing sensor measurements in different modalities.¹⁵⁻²²

In this article, a multi-sensor fusion and object tracking algorithm (MFOTA) is proposed, which mainly focuses on the accuracy of target information such as target location, velocity and type. This algorithm inherits the advantages of lidar, radar, camera and Kalman filter. The MFOTA can provide more accurate target location compared with radar and camera, more accurate target velocity compared with lidar, radar and camera, more accurate target type compared with lidar and radar. Besides, the MFOTA has also higher weather robustness compared with single sensor.

This article includes four sections that are spatial calibration method for multi-sensor systems, the proposed MFOTA, performance comparison and analysis and concluding remarks.

Spatial calibration method for multi-sensor systems

The calibration among multiple sensors contains time and space. In our self-driving vehicles, the timestamps of all the measurements from multiple sensors are accurately known with respect to a single clock. However, the frame rate among multiple sensors is different. Thus, the time gap of the target measured by different sensors exists. In practice, we utilize the way of velocity integral to bridge the time gap. In this article, we just focus the spatial calibration method among lidar, radar and camera. The purpose of space calibration is to minimize installation errors as far as possible for a variety of sensors.

For different kinds of sensors, spatial calibration is necessary to realize information fusion in the same region. Location accuracy for lidar is very high. Thus, other sensors need to be calibrated to the target location measured by lidar. Figure 1 shows the installation structure diagram and the field of view (FOV) of various sensors on the self-driving vehicle. The FOV of 77GHz millimeter-wave radar (MWR) contains three angles with 120°, 80° and 18° in Figure 1. The FOVs of lidar and camera are 110° and 50°, respectively. Lidar, MWR and camera are mounted on the front of the vehicle. The FOV of various sensors overlaps in the

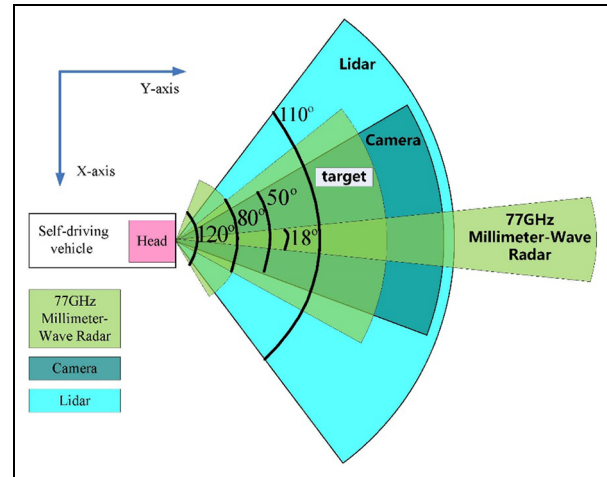


Figure 1. Installation structure diagram and the FOV of various sensors.

space. The target is represented by the rectangle box. The longitudinal axis of the vehicle is Y-axis. The lateral axis of the vehicle is X-axis. All sensors can cover at least 100 m range in Y-axis.

In order to realize the spatial calibration of various sensors, collection of target location information from various sensors in the FOV is required first. The different targets (pedestrians) measured are shown for the three sensors in Figure 2. The symbols of blue stars, red circles and red squares represent the targets of lidar, radar and camera, respectively.

The method of spatial calibration is to project the targets of radar or camera to the targets of lidar. We assume that the installation position of the lidar has no deviation and is in the origin of coordinates. The theory of calibration references equations from (1) to (9)

$$x_r = x_{r0} + \Delta x \quad (1)$$

$$y_r = y_{r0} + \Delta y \quad (2)$$

where x_{r0} and y_{r0} are the x-axis and y-axis ordinates of the target measured by radar, respectively; Δx and Δy are the corresponding installation position offsets between lidar and radar, respectively; and x_r and y_r are the results of coordinate transformation for radar.

In equations (1) and (2), Δx and Δy are known. The target location calibrated (x_{lr} and y_{lr}) can be obtained as

$$x_{lr} = R_{lr} \cos(\theta_{lr}) \quad (3)$$

$$y_{lr} = R_{lr} \sin(\theta_{lr}) \quad (4)$$

with

$$R_{lr} = \beta R_r \quad (5)$$

$$R_r = \sqrt{x_r^2 + y_r^2} \quad (6)$$

$$\theta_{lr} = \theta_r + \Delta\theta \quad (7)$$

$$\theta_r = \begin{cases} \arctan(y_r/x_r), & x_r > 0 \\ \pi/2 - \arctan(x_r/y_r), & x_r \leq 0 \end{cases} \quad (8)$$

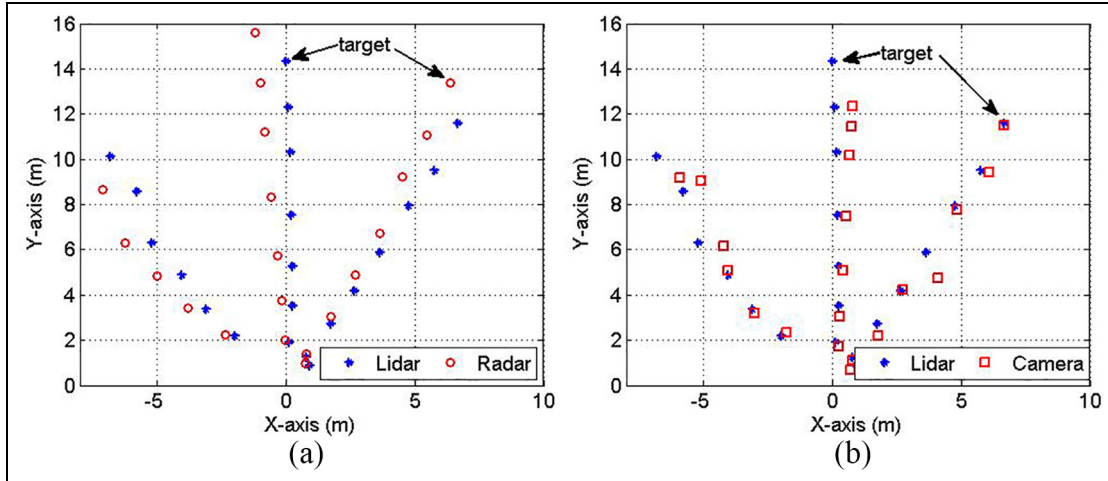


Figure 2. Target results from different sensors. (a) The targets from lidar and radar and (b) the targets from lidar and camera.

where R_{lr} is the target radial distance calibrated, θ_{lr} is the angle from x-axis to target (counterclockwise is positive), β is ratio coefficient, $\Delta\theta$ is the angle from the target calibrated to the target of lidar, x_r and y_r are the x-axis and y-axis ordinates measured by radar, respectively, and x_{lr} and y_{lr} are the x-axis and y-axis ordinates calibrated, respectively.

From equations (3) to (8), the only two parameters β and $\Delta\theta$ are unknown. They can be calculated by

$$\{\beta, \Delta\theta\} = \arg \min \left\{ \sum_{i=0}^{n-1} \left[(x_{lr}(i) - x_l(i))^2 + (y_{lr}(i) - y_l(i))^2 \right] \right\} \quad (9)$$

where $x_{lr}(i)$ is the x-axis location for the i th target, $y_{lr}(i)$ is the y-axis location for the i th target, and n is the number of samples utilized in Figure 2.

Utilizing the data in Figure 2, the results of β and $\Delta\theta$ calculated by equation (9) are shown in Table 1. The corresponding results of spatial calibration are shown in Figure 3. It can be seen from Figure 3 that the targets calibrated from radar and camera can match those of lidar better. The validity of the proposed spatial calibration method is tested in Figures 2 and 3.

MFOTA

After spatial calibration, the targets from multiple sensors can be correlated efficiently in space. The next step is to realize multi-sensor information fusion. The purpose of multi-sensor fusion is to ensure that all the information of the targets detected is more accurate compared with any of other sensors. The framework of information fusion is shown in Figure 4. The fusion framework mainly contains four parts that are measured targets preprocessing for lidar, association for measured targets, information update for associated targets and Kalman filter at velocity. Because target location measured by lidar is accurate, so it is

Table 1. The results of β and $\Delta\theta$ for different sensor data.

Sensors	Parameters	
	β	$\Delta\theta$
Radar	1.1	$-\pi/36$
Camera	0.95	$\pi/60$

unnecessary to use Kalman filtering at position. Now that Kalman filtering is well known, it will not be repeated in this article.

Measured targets preprocessing for lidar

For the state information of target measured, we only focus on location, velocity and type in this article. Compared with radar and camera, lidar can measure target location more accurately. The velocity of the target measured by lidar is calculated by the rate of change of location. However, because of the existence of location noise, sometimes the rate of change of location is large. Thus, target information pre-processing for lidar is necessary. Based on Figure 1, we assume that v_x is the x-axis velocity and v_y is the y-axis velocity for a given target. The velocity of the given target can be estimated by

$$\bar{v}_x = \frac{1}{N} \sum_{i=1}^N v_x(i) \quad (10)$$

$$\bar{v}_y = \frac{1}{N} \sum_{i=1}^N v_y(i) \quad (11)$$

where \bar{v}_x is the average x-axis velocity for a given target, \bar{v}_y is the average y-axis velocity for a given target, $v_x(i)$ is the x-axis velocity of the same target for the i th frame, $v_y(i)$ is the y-axis velocity of the same target for the i th frame, and N is the number of the latest frames observed continuously.

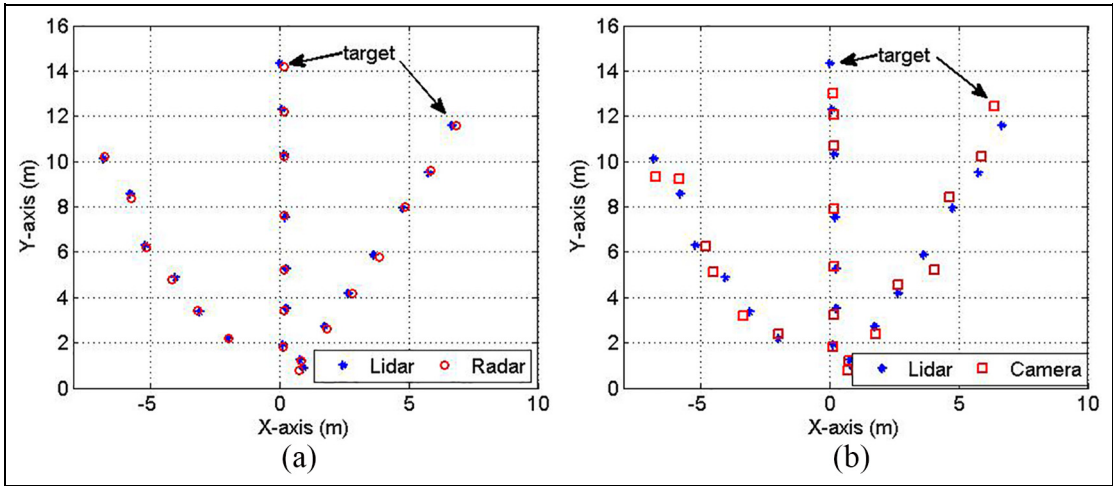


Figure 3. Target calibration results from different sensors. (a) The targets from lidar and radar calibrated and (b) the targets from lidar and camera calibrated.

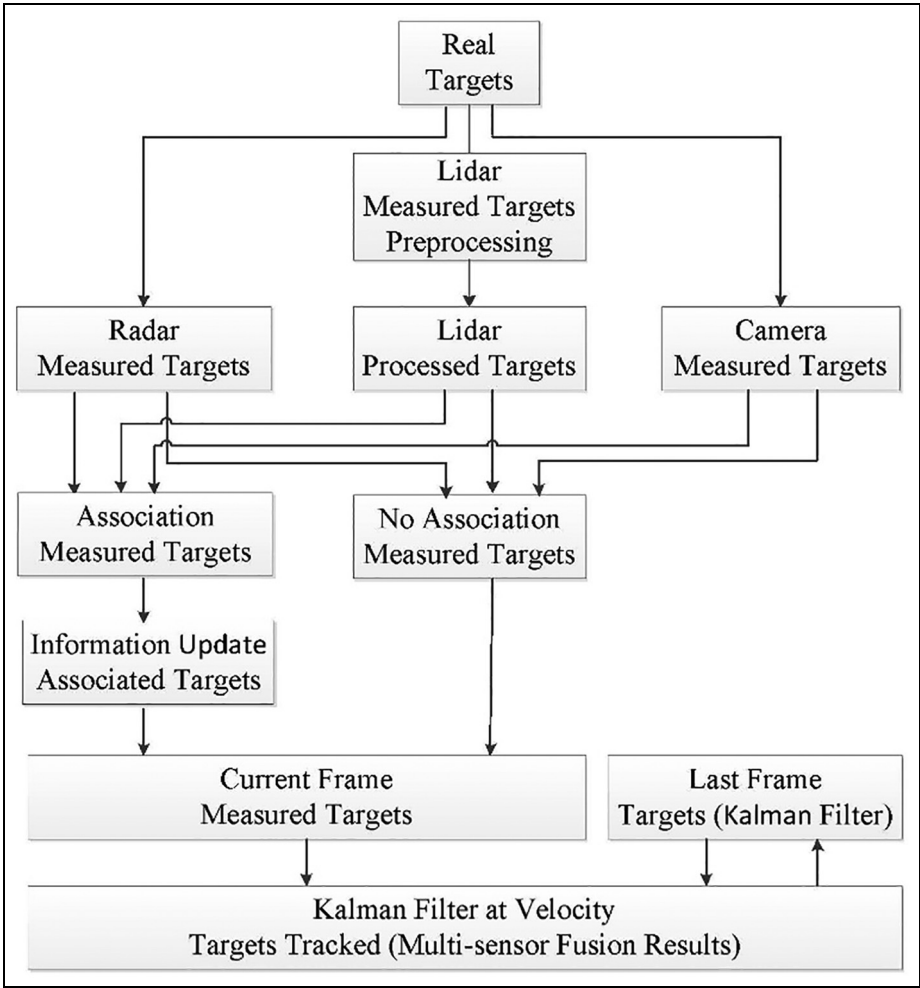


Figure 4. The framework of information fusion.

Association for measured targets

The purpose of multi-sensor information fusion is to utilize the advantages of diverse sensors. In order to do this, the most important part is the association. Take

the target measured by lidar and radar for example, the association logic is shown in Figure 5. The association logic between other sensors is the same as that of lidar and radar.

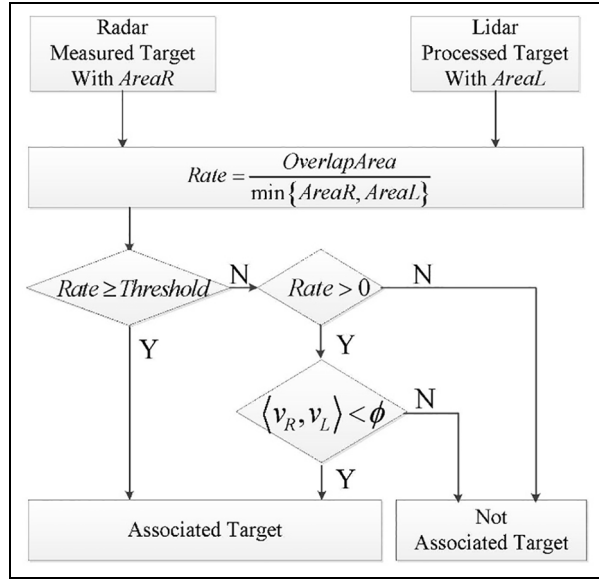


Figure 5. The association logic.

In Figure 5, $AreaL$ and $AreaR$ are the areas of the target measured by lidar and radar, respectively. The $AreaL$ and $AreaR$ both consider the measurement standard deviation of the target shape provided by sensors. $OverlapArea$ is the overlap area between the target of lidar and that of radar. $Rate$ is the ratio of $OverlapArea$ to the minimum of $AreaL$ and $AreaR$. The target velocities from lidar and radar are v_L and v_R . The angle between v_L and v_R is ϕ . The threshold of rate is *threshold*. The outputs are “Associated Target” and “Not Associated Target.”

Information update for associated targets

For those associated targets, the information needs to be updated utilizing the three kinds of sensors. Assuming that all the targets consist of four points (rectangle), they can be calculated by

$$\begin{cases} p_{xf} = \alpha_{xl}p_{xl} + \alpha_{xr}p_{xr} + \alpha_{xc}p_{xc} \\ p_{yf} = \alpha_{yl}p_{yl} + \alpha_{yr}p_{yr} + \alpha_{yc}p_{yc} \\ w_f = \kappa_l w_l + \kappa_r w_r + \kappa_c w_c \\ l_f = \gamma_l l_l + \gamma_r l_r + \gamma_c l_c \\ v_{xf} = \lambda_{xl}v_{xl} + \lambda_{xr}v_{xr} + \lambda_{xc}v_{xc} \\ v_{yf} = \lambda_{yl}v_{yl} + \lambda_{yr}v_{yr} + \lambda_{yc}v_{yc} \end{cases} \quad (12)$$

where p is the center of the target bottom, w is the width of the target, l is the length of the target, v is the velocity of the target, the subscript x is x-axis, the subscript f is fusion, the subscript y is y-axis, the subscript l is lidar, the subscript r is radar, the subscript c is camera, α is the weight of the target location, κ is the weight of the target width, γ is the weight of the target length, and λ is the weight of the target velocity.

The selection of weight depends on the accuracy of information supplied by sensors. For example, lidar can provide more accurate target location and shape

Table 2. Assignment logic for weights in equation (12).

Weights	Sensors fused		
	Lidar and radar	Lidar and camera	Radar and camera
α_{xl}	1	1	0
α_{xr}	0	0	1
α_{xc}	0	0	0
α_{yl}	1	1	0
α_{yr}	0	0	1
α_{yc}	0	0	0
κ_l	1	1	0
κ_r	0	0	0
κ_c	0	0	1
γ_l	1	1	0
γ_r	0	0	0
γ_c	0	0	1
λ_{xl}	1	1	0
λ_{xr}	0	0	1
λ_{xc}	0	0	0
λ_{yl}	0	1	0
λ_{yr}	1	0	1
λ_{yc}	0	0	0

compared with other two sensors, radar can provide more accurate target velocity, and camera can provide more accurate target type realized by deep learning. Thus, the type of the target associated is provided by camera.

Experiments, comparison and analysis

According to the characteristics of lidar, radar and camera, any one of them cannot provide all the information including target's location, target's velocity and target's type. Thus, the three kinds of sensors need to be fused for self-driving vehicles.

In this part, the value of N is set to 5 in equations (10) and (11). The parameters *Threshold* and ϕ in Figure 2 are 0.2 and $\pi/4$, respectively. Based on the way of information update described above, the assignment logic in this article is shown in Table 2. In this article, the accuracy of target type cannot be compared, and can only be elaborated by the theory of target recognition.

In order to prove the feasibility of multi-sensor fusion, vehicle and pedestrian objectives are used to compare the performance of multiple sensors and single sensor. The two scenarios are shown in Figure 6. The pedestrian and vehicle objectives are used to verify the location and velocity accuracy obtained by MFOTA, respectively.

Pedestrian

Here, pedestrian objectives are used to verify the location accuracy obtained by MFOTA. In Figure 6, the pedestrian changes his location, then records the distances measured by the three sensors and fusion. The

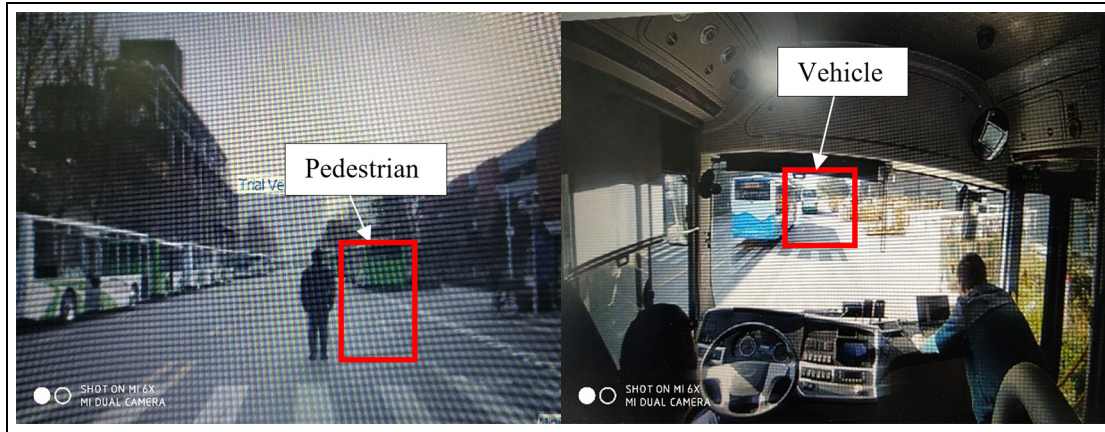


Figure 6. The two scenarios including pedestrian and vehicle.

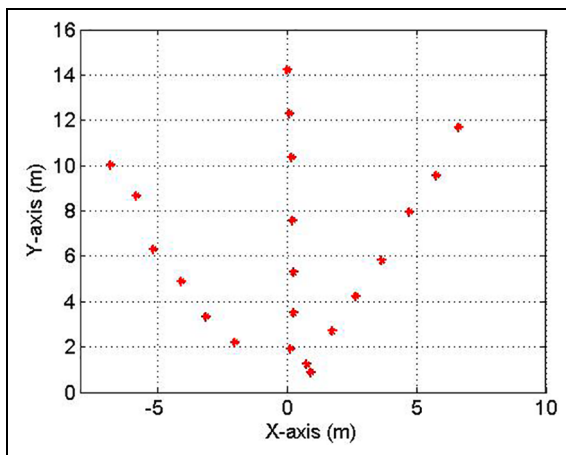


Figure 7. The real locations of pedestrian.

spatial calibration of the three sensors is completed. The real locations of the pedestrian are measured by a tapeline, and shown in Figure 7.

The computation method of position error can be obtained by

$$error_{pf} = \frac{1}{n_0} \sum_{j=1}^{n_0} \left[(p_{xf}(j) - p_x(j))^2 + (p_{yf}(j) - p_y(j))^2 \right] \quad (13)$$

$$error_{pl} = \frac{1}{n_0} \sum_{j=1}^{n_0} \left[(p_{xl}(j) - p_x(j))^2 + (p_{yl}(j) - p_y(j))^2 \right] \quad (14)$$

$$error_{pr} = \frac{1}{n_0} \sum_{j=1}^{n_0} \left[(p_{xr}(j) - p_x(j))^2 + (p_{yr}(j) - p_y(j))^2 \right] \quad (15)$$

$$error_{pc} = \frac{1}{n_0} \sum_{j=1}^{n_0} \left[(p_{xc}(j) - p_x(j))^2 + (p_{yc}(j) - p_y(j))^2 \right] \quad (16)$$

Table 3. Position errors from sensors based on (13)~(16).

Sensor errors	$error_{pf}$	$error_{pl}$	$error_{pr}$	$error_{pc}$
Value (m ²)	9.4×10^{-4}	9.4×10^{-4}	0.0179	0.2603

where $error_p$ is the error of target position, j is the index of sample, n_0 is the number of sample and other signs are the same as equation (12). The corresponding errors from sensors are shown in Table 3. The parameter n_0 is 21 in Table 3.

It can be seen from Table 2, the position accuracy for fusion and lidar is highest compared with other two sensors. The position accuracy of radar exceeds that of camera.

Vehicle

Here, vehicle objective is used to verify the velocity accuracy obtained by MFOTA. In Figure 6, the vehicle keeps its velocity, then records the velocities measured by the three sensors and fusion. The real velocities of the vehicle are obtained by its position module, and shown in Figure 8.

The computation method of velocity error can be obtained by

$$error_{vf} = \frac{1}{n_1} \sum_{k=1}^{n_1} \left[(v_{xf}(k) - v_x(k))^2 + (v_{yf}(k) - v_y(k))^2 \right] \quad (17)$$

$$error_{vl} = \frac{1}{n_1} \sum_{k=1}^{n_1} \left[(v_{xl}(k) - v_x(k))^2 + (v_{yl}(k) - v_y(k))^2 \right] \quad (18)$$

$$error_{vr} = \frac{1}{n_1} \sum_{k=1}^{n_1} \left[(v_{xr}(k) - v_x(k))^2 + (v_{yr}(k) - v_y(k))^2 \right] \quad (19)$$

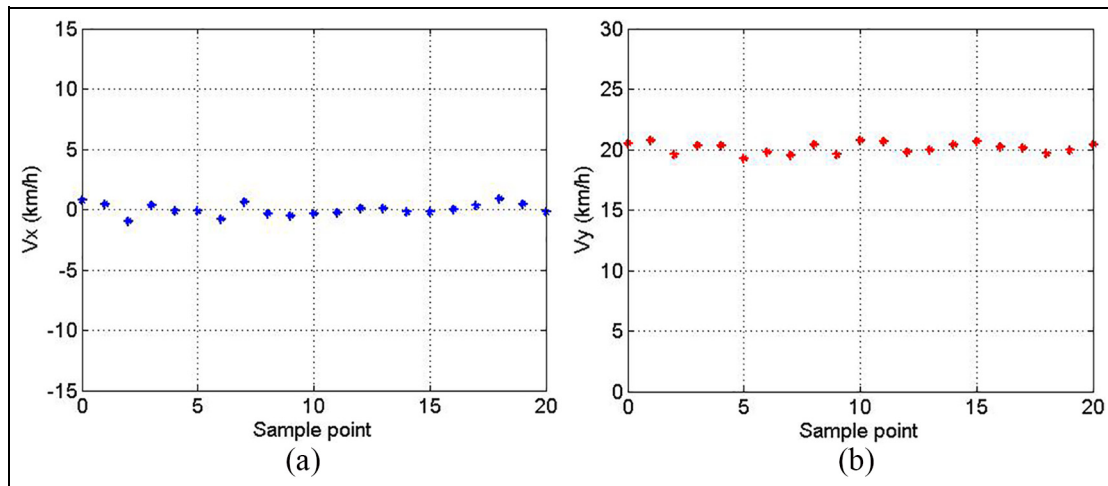


Figure 8. The real velocities of vehicle.

Table 4. Velocity errors from sensors based on (17)–(20).

Sensor errors	$error_{vf}$	$error_{vl}$	$error_{vr}$	$error_{vc}$
Value (m^2/s^2)	0.0732	1.9164	0.7764	9.3807

$$error_{vc} = \frac{1}{n_1} \sum_{k=1}^{n_1} \left[(v_{xc}(k) - v_x(k))^2 + (v_{yc}(k) - v_y(k))^2 \right] \quad (20)$$

where $error_v$ is the error of target velocity, k is the index of sample, n_1 is the number of sample, and other signs are the same as equation (12). The corresponding errors for velocity from other sensors are shown in Table 4. The parameter n_1 is 21 in Table 4.

In Table 4, the velocity accuracy for fusion is highest compared with other three sensors. The performance of radar is better than those of lidar and camera. And the performance of velocity for camera is the worst.

Performance analysis

It can be seen from Table 3, the position accuracy for fusion and lidar is highest compared with other two sensors. That is because fusion output for target position is the same as lidar. And the position accuracy for lidar is higher than radar and camera.

In Table 4, the velocity accuracy for fusion is highest compared with other three sensors. That is because fusion output for target velocity inherits the advantages of lidar, radar and Kalman filter. The accuracy of v_x of lidar is higher than that of radar, and the accuracy of v_y of radar is higher than that of lidar. The velocity of fusion is the result of Kalman filter at v_x of lidar and v_y of radar. For the vehicle objective, the v_y is much higher than v_x , then $error_{vr}$ is smaller than $error_{vl}$. Because of the limitation of camera characteristics, the v_x and v_y

Table 5. Observation intervals.

Sensors and MFOTA	Observation intervals (ms)
Lidar	100
Radar	80
Camera	60
MFOTA	100

MFOTA: multi-sensor fusion and object tracking algorithm.

for camera are worst compared with fusion and other two sensors.

Considering practical use, it is important to clarify the observation intervals from sensors and MFOTA shown in Table 5. The observation interval of MFOTA is no longer than that of lidar, which can maintain acceptable real-time performance in a low-velocity scenario.

Conclusion

In this article, a novel spatial calibration method based on multiple sensors was proposed by utilizing rotation and translation of the coordinate system, and its validity is tested through comparisons with the data calibrated. Moreover, MFOTA is proposed to increase environmental adaptability and provide accurate information of target such as position, velocity and type for self-driving vehicles. MFOTA inherits the advantages of lidar, radar, camera and Kalman filter. The real data are used to test the performance of MFOTA that is compared in both position and velocity dimensions with that of single sensor. MFOTA can not only provide position accuracy like lidar, but type accuracy like camera. MFOTA has also accurate lateral velocity and accurate longitudinal velocity utilizing Kalman filter. Besides, the observation interval of MFOTA is no longer than that of lidar, which can meet acceptable

real-time requirement in low-velocity scenario. In general, MFOTA shows good application prospects from the performance and observation interval.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The project is supported by the National Key Research and Development Program of China (Project No. 2018YFB0105003).

ORCID iD

Chunlei Yi  <https://orcid.org/0000-0003-0564-8781>

References

1. Kelly A, Stentz A, Amidi O, et al. Toward reliable off road autonomous vehicles operating in challenging environments. *Int J Robot Res* 2006; 25: 449–483.
2. Darms M, Rybski PE, Baker C, et al. Obstacle detection tracking for the urban challenge. In: *Proceedings of the 12th IEEE conference on intelligent transportation systems*, St. Louis, MO, 3–7 October 2009. New York: IEEE.
3. Petrovskaya A and Thrun S. Model based vehicle detection and tracking for autonomous urban driving. *Autonom Robot* 2009; 26: 123–139.
4. Yang D, Jiang K, Zhao D, et al. Intelligent connected vehicles: current status future perspectives. *Sci China* 2018; 61: 20–45.
5. Wang F, Mirchandani PB and Wang Z. The VISTA project and its applications. *IEEE Intell Transp Syst* 2002; 17: 72–75.
6. Wang F, Wang X, Li L, et al. Creating a digital-vehicle proving ground. *IEEE Intell Transp Syst* 2003; 18: 12–15.
7. Li L, Song J, Wang F, et al. IVS 05: new developments and research trends for intelligent vehicles. *IEEE Intell Transp Syst* 2005; 20: 10–14.
8. Yim Y and Oh S. Three-feature based automatic lane detection algorithm (TFALDA) for autonomous driving. *IEEE Trans Intell Transp Syst* 2003; 4: 219–225.
9. Jeong S, Kim CS, Yoon KS, et al. Real-time lane detection for autonomous vehicle. In: *Proceedings of the IEEE international symposium on industrial electronics*, vol. 3, Pusan, South Korea, 12–16 June 2001, pp.1466–1471. New York: IEEE.
10. Huang S, Chen C-J, Hsiao P-Y, et al. On-board vision system for lane recognition and front-vehicle detection to enhance driver's awareness. In: *Proceedings of the IEEE international conference on robotics and automation*, vol. 3, New Orleans, LA, 26 April–1 May 2004, pp.2456–2461. New York: IEEE.
11. McCall J and Trivedi M. An integrated, robust approach to lane marking detection and lane tracking. In: *Proceedings of the IEEE intelligent vehicles symposium*, Parma, 14–17 June 2004, pp.533–537. New York: IEEE.
12. Chen S, Huang L, Bai J, et al. Multi-Sensor information fusion algorithm with central level architecture for intelligent vehicle environmental perception system. SAE technical paper 2016-01-1894, 2016.
13. Zhao G, Xiao X, Yuan J, et al. Fusion of 3D-LIDAR and camera data for scene parsing. *J Vis Commun Imag Rep* 2014; 25: 165–183.
14. Aufrere R, Gowdy J, Mertz C, et al. Perception for collision avoidance and autonomous driving. *Mechatronics* 2003; 13: 1149–1161.
15. Monteiro G, Premevida C, Peixoto P, et al. Tracking and classification of dynamic obstacles using laser range finder and vision, 2006, https://www.academia.edu/18267582/Tracking_and_Classification_of_Dynamic_Obstacles_Using_Laser_Range_Finder_and_Vision
16. Miller I, Campbell ME and Fujishima H. Team cornell's skynet: robust perception and planning in an urban environment. *J Field Robot* 2008; 25: 493–527.
17. Leonard J, How J, Teller S, et al. A perception-driven autonomous urban vehicle. *J Field Robot* 2008; 25: 727–774.
18. Mahlich M, Schweiger R, Ritter W, et al. Sensorfusion using spatio-temporal aligned video and lidar for improved vehicle detection. In: *Proceedings 2006 IEEE intelligent vehicles symposium*, Tokyo, Japan, 13–15 June 2006. New York: IEEE.
19. Montemerlo M, Becker J, Bhat S, et al. Junior: the stanford entry in the urban challenge. *J Field Robot* 2008; 25: 569–597.
20. Premevida C, Ludwig O and Nunes U. Lidar and vision-based pedestrian detection system. *J Field Robot* 2009; 26: 696–711.
21. Spinello L and Siegwart R. Human detection using multi-modal and multidimensional features. In: *Proceedings of the 2008 IEEE international conference on robotics and automation*, Pasadena, CA, 19–23 May 2008. New York: IEEE.
22. Spinello L and Siegwart R. A Multi-sensor fusion system for moving object detection and tracking in urban driving environments. In: *Proceedings of the 2014 IEEE international conference on robotics & automation (ICRA)*, Hong Kong, China, 31 May–7 June 2014. New York: IEEE.