The Application of Multi-sensor Fusion Technology in Intelligent Vehicles

Abstract: At present, the rapid development of the new energy vehicle industry has also promoted the intelligence of automobiles. Meanwhile, multi-sensor fusion technology has become the core support for intelligent vehicles. By integrating data from various sensors such as lidar, millimeter-wave radar, cameras, and ultrasonic radar, the shortcomings of a single sensor can be effectively compensated for, thereby enhancing the accuracy, reliability, and robustness of environmental perception. This article will explore multi-sensor fusion technology and analyze its specific applications in key links such as environmental perception, target recognition, and path planning of intelligent vehicles. Analyze the technical challenges it faces, such as data synchronization, heterogeneous data processing, and optimization of fusion algorithms. In addition, based on the current development status and trends of multi-sensor fusion technology, this paper will also look forward to the future development direction of intelligent vehicles, aiming to provide theoretical references for the progress of intelligent vehicle technology.

Key words: Multi-sensor fusion;Intelligent car;Environmental perception; Perceptual fusion

# Introduction

Against the backdrop of the global automotive intelligent transformation, multi-sensor fusion technology is also gradually coming into the public eye. Due to the certain limitations of a single sensor, it cannot meet the requirements under all-scenario usage conditions. However, multi-sensor fusion technology can precisely make up for its own deficiencies [15]. By integrating the advantages of different sensors such as cameras, millimeter-wave radars, and lidars [1], it effectively addresses the shortcomings of individual sensors in terms of environmental adaptability and target recognition accuracy. It has become a key technology to break through the environmental perception deficiencies of intelligent vehicles [16].

At present, multi-sensor fusion technology has been applied in some scenarios, but it still faces many challenges in its development process: there are deviations in spatio-temporal calibration, the computing power demand for data processing is extremely high, and the redundancy mechanism in extreme scenarios needs to be optimized, etc [17]. In-depth research on this technology is of profound value. It not only improves the theoretical system of environmental perception for intelligent vehicles but also effectively enhances the reliability of vehicle decision-making in complex scenarios, providing strong support for highly autonomous driving. In the long run, this technology will help build a safe and efficient future mobility ecosystem. This article will conduct a systematic exploration from three dimensions: key integration technologies, typical application cases, and future development trends.

# Key technologies of multi-sensor fusion

## Time and Space synchronization

### Time Synchronization

In the field of autonomous driving, sensors not only need to achieve precise synchronization of the spatial coordinate system, but also the synchronization of the temporal coordinate system is equally crucial. Time synchronization refers to the fact that a unified host provides a reference time for each sensor, and each sensor then adds a timestamp to the independently collected data based on its own calibrated time, thereby ensuring that the timestamps of all sensors remain consistent [18]. However, the sampling frequencies of different sensors vary. For instance, liDAR typically has a frequency of 10Hz, while cameras usually have a frequency of 30Hz. Moreover, there is a certain delay in data transmission between sensors. Therefore, it is necessary to determine the nearest neighbor frame by looking for adjacent timestamps. However, if the two timestamps differ significantly and the sensor or obstacle is in motion, a large synchronization error is likely to occur. The function of a unified clock is precisely to synchronize the timestamps of different sensors.

If the sensor supports hardware triggering, the GPS timestamp can be used as the reference for hardware triggering. At this point, the timestamp in the sensor output data is the global timestamp (i.e., GPS timestamp), rather than the timestamp of the sensor itself. Of course, due to the different sampling frequencies of sensors and the delay in data transmission, significant synchronization errors may still occur when there is a large gap between the two timestamps and the sensor or obstacle is in motion. To address this issue, in addition to using hard synchronous triggering to reduce errors caused by finding timestamps, the time difference can also be narrowed by adjusting the natural frequency of the sensor (for example, setting the camera frequency to 20Hz).

### Spatial Synchronization

Currently, spatial synchronization of sensors is a key step in data fusion. Among them, the registration methods of lidar and cameras are currently mainly divided into the following categories:

One type is to first extract feature sets such as points and lines, and then solve the registration parameters of the two through the matching of laser point clouds and image features [19]. Another type is to first calculate the mutual information loss function between the laser point cloud and the camera image, and then obtain the relative pose parameters with the help of an optimization algorithm. This method is more suitable for scenarios where the relative pose movement of the sensor on a straight road is relatively small. In addition, there is another method that utilizes the historical frame information of multiple cameras or a single camera to apply the SFM algorithm to the camera. Judging from the current configuration of autonomous driving, this method is also quite applicable.

## Data Preprocessing

As a core step in data processing, the role of filtering technology is particularly crucial. Among them, the application of Kalman filtering is based on the premise of linear systems and Gaussian noise. It makes the optimal estimation of sensor data through the recursive process of "predictive - update" by means of the system state equation and measurement equation, and can efficiently filter out Gaussian noise. Therefore, it is widely used in the estimation of target position and speed by radar. Particle filtering is more suitable for nonlinear and non-Gaussian systems: it generates a particle set through random sampling to approximate the posterior probability distribution, and can handle sensor data in complex environments. For example, in dynamic target tracking, it can better adapt to the irregular movement of the target.

After filtering, the denoising technology will further purify the original data. When processing image data, median filtering can eliminate impulse noise, and Gaussian filtering can remove Gaussian noise. For LiDAR point cloud data, methods such as statistical filtering and voxel filtering can eliminate outliers, reduce redundant information, and improve the quality of the point cloud.

Feature extraction is an important means of mining core information from data. In image processing, the Canny operator and the Sobel operator can extract edge features, which are then used to identify road boundaries and traffic sign contours [2]; When dealing with radar point clouds, Euclidean clustering and DBSCAN clustering algorithms can divide the point clouds into different target clusters, achieving the detection and classification of obstacles. These technologies not only reduce the data dimension but also extract the most valuable information for target recognition and environmental perception. While enhancing the efficiency of data processing, they provide high-quality basic data for subsequent multi-sensor data fusion and decision-making.

## Fusion Algorithm

In the multi-sensor fusion system of intelligent vehicles, early fusion, late fusion and hybrid fusion are three strategies with their own characteristics. The differences between them will have a significant impact on perception accuracy and system efficiency [20]. The specific situations are as follows:

The core of early integration is to complete information integration at the source of data. It will perform cross-modal matching between the point cloud data generated by the lidar and the image pixels captured by the camera, and then, with the help of coordinate transformation, time synchronization and other technologies [3], unify the data from different sensors into the same spatiotemporal coordinate system. This strategy can preserve the integrity and details of the original data to the greatest extent, enabling the system to carry out feature extraction and target recognition based on the fused comprehensive data. This is very helpful for improving the perception accuracy in complex scenarios, such as identifying targets that are partially occluded. However, due to the need to process massive amounts of raw data in real time, it has extremely high requirements for the computing power, storage bandwidth and data transmission rate of the computing platform. Meanwhile, its algorithm has a relatively high complexity, and even minor errors during data alignment may be magnified, thereby affecting the final perception effect [4].

Late-stage fusion adopts a different approach: first, each sensor is allowed to independently complete data processing and target detection tasks. For instance, the camera outputs the bounding box and category information of the target through deep learning algorithms, while the millimeter-wave radar provides the distance and speed parameters of the target. Subsequently, the system conducts decision-level fusion based on these detection results and makes comprehensive judgments on the detection results of different sensors through algorithms such as voting mechanisms and Bayesian reasoning [5]. This approach has a relatively low computational load, a relatively simple algorithm implementation, and each sensor module is independent of one another, which is convenient for maintenance and upgrade.

The goal of the hybrid fusion strategy is to take into account the advantages of both early fusion and late fusion, achieving mutual complementarity. Such systems typically adopt a hierarchical architecture: Firstly, early fusion is carried out on sensors with similar characteristics and closely related data processing logics, such as fusing the point cloud data of millimeter-wave radar and lidar at the data layer, taking advantage of their complementarity in distance detection to generate more accurate three-dimensional target information. On this basis, the fused point cloud data and camera image data are re-fused at the feature layer or decision-making layer [6]. Through the complementarity of multimodal features, the recognition ability of target categories and postures is enhanced. This phased fusion approach can not only alleviate the processing pressure of raw data to a certain extent, but also enhance the robustness and accuracy of the perception system through multi-level information interaction. It is currently the mainstream development direction that takes into account both performance and efficiency. Of course, it also faces challenges such as complex fusion strategy design and high difficulty in parameter optimization.

# Typical application scenarios in intelligent vehicles

## Environmental Perception and Target Detection

To achieve autonomous driving, environmental perception and target detection are the most fundamental links in the core functional modules of intelligent vehicles. To break through the performance limitations of a single sensor, multi-sensor collaborative perception is the core technology. Different types of sensors vary in working principles and characteristics, each playing an irreplaceable role in environmental perception: Cameras, with the help of computer vision algorithms, analyze features such as texture, shape, and color in images, and can accurately identify the categories of objects such as pedestrians, vehicles, and traffic signs. Lidar, through high-frequency laser scanning, can generate high-precision three-dimensional point cloud data in real time, accurately capturing the position, size and contour information of the target. Millimeter-wave radar, based on the principle of electromagnetic wave reflection, can stably obtain the speed, distance and azimuth parameters of the target even under complex weather and lighting conditions. The organic integration of these three types of sensors can form a complete perception chain from target attribute recognition to spatial dynamic monitoring, effectively enhancing the tracking accuracy and prediction ability of dynamic targets.

From the perspective of industry practice, Tesla's Autopilot and Waymo's self-driving cars respectively represent different technical routes of multi-sensor fusion. Tesla Autopilot adopts a vision-driven solution, building a surround visual perception system through multiple cameras and combining the detection capability of millimeter-wave radar for distant targets to achieve real-time recognition and tracking of lane lines, traffic signs, and obstacles. Its deep learning algorithm conducts multi-frame analysis on the images collected by the camera and integrates the distance data from millimeter-wave radar to enhance the spatial positioning ability of the target. It can demonstrate efficient perception performance in both complex urban roads and high-speed scenarios. Waymo's self-driving cars adopt a multi-combination strategy of lidar, cameras, and millimeter-wave radar. They build a three-dimensional environment model with lidar as the core, providing centimeter-level precision map information for the self-driving cars. The camera is responsible for supplementing detailed texture information and assisting in completing target classification [7]; Millimeter-wave radar focuses on the speed monitoring of dynamic targets. This multimodal fusion solution enables Waymo's driverless vehicles to accurately perceive various targets and track them for a long time in complex urban traffic environments, thereby providing solid data support for autonomous driving decisions.

## Positioning and Map Construction

In the field of localization and mapping (SLAM) for intelligent vehicles, if precise positioning and environmental modeling are to be achieved, multi-sensor SLAM technology is the core support. A single sensor has obvious limitations in SLAM tasks, while multi-sensor fusion can integrate the advantages of each sensor, significantly improving system performance.

The working principle of lidar is to emit laser beams to detect the environment. This feature enables it to quickly collect massive amounts of three-dimensional spatial point cloud data. Whether it is tall buildings, trees by the roadside, or undulating road surfaces, these point cloud data can precisely outline their geometric contours, thereby constructing high-precision three-dimensional maps [8]. However, although such 3D maps can provide a basic spatial framework for smart cars, their recognition is relatively low due to the lack of color and texture information in point cloud data. At this point, the role of the camera comes into play. The image data collected by the camera contains rich texture and semantic information, such as the color of road signs and the appearance features of buildings. By fusing and matching the camera images with the lidar point cloud data, it can not only add visual textures to the 3D map, improving its readability, but also help smart cars more accurately identify key elements in the environment and enhance the reliability of positioning.

The core function of the Inertial Navigation System (INS) is to compensate for the position deviation of the vehicle during dynamic movement [9]. It measures the angular velocity and acceleration of the vehicle in real time with the help of a gyroscope and an accelerometer, and then calculates the vehicle's attitude and position changes based on these data. For instance, in environments where satellite signals are easily blocked, such as tunnels and sections of canyons with numerous high-rise buildings, GPS and other satellite positioning systems may experience signal loss or a decline in accuracy. However, INS can continuously provide vehicle movement information through its own inertial measurement, ensuring the continuity of positioning [10]. When INS combines the data from liDAR and cameras, it can also correct the calculated position, ultimately achieving centimeter-level real-time positioning and map updates in complex urban roads, providing reliable location references for the path planning and autonomous driving decisions of intelligent vehicles [11].

## Path Planning and Decision Control

In the path planning and decision-making control module of intelligent vehicles, to achieve safe and efficient autonomous driving, integrating data-driven decision-making is the core logic. The traffic scenarios in which intelligent vehicles operate are complex and ever-changing, and the information from a single sensor is difficult to support reliable decision-making. The environmental information integrated after multi-sensor fusion precisely provides a comprehensive and dynamic reference basis for path planning and decision control [12].

In actual driving, sensor fusion data can accurately capture potential risks and dynamic changes in the environment. For instance, when traffic congestion is detected ahead, the system will integrate the target position, speed and density information provided by liDAR, cameras and millimeter-wave radar, and re-evaluate the optimal route through a path planning algorithm, choosing a road with less traffic to detour. If a pedestrian suddenly crosses the road, the real-time data from multi-sensor fusion can quickly determine the pedestrian's movement trajectory. The decision control system will immediately trigger emergency braking or adjust the driving path, thereby avoiding a collision accident.

When encountering extreme weather, the advantages of multi-sensor fusion will be more obvious. Take rainy and foggy weather as an example: The camera will have blurred images due to reduced visibility and be unable to accurately identify the target. At this time, millimeter-wave radar, with its strong penetration and less influence by weather, becomes the main sensor for obstacle detection, continuously monitoring the distance and speed of the target in front of the vehicle. The camera is responsible for assisting in identifying key visual information such as traffic lights and lane lines. Although the image quality has declined, it can still extract faint color and contour features through deep learning algorithms. The two work in coordination, enabling the autonomous driving system to promptly detect potential obstacles while ensuring it complies with traffic rules and drives safely in harsh environments. Not only that, the system will also adjust parameters such as vehicle speed and following distance in real time based on sensor fusion data, further enhancing driving safety and comfort.

# Technical Challenges and Development Trends

## Main Challenges

In the development process of multi-sensor fusion technology for intelligent vehicles, it mainly faces three core challenges: computing power, data processing and cost adaptation.

The limitations of computing power and power consumption are the primary challenges. When intelligent vehicles are in motion, they need to process high-resolution images, massive liDAR point clouds and other data in real time, which poses extremely high demands on the computing power of on-board chips. In the scenario of autonomous driving alone, tens of gigabytes of data may be generated every second, which requires the chip to have powerful parallel computing capabilities. However, the vehicle space is limited and the range needs to be taken into account, so the on-board chip must strike a balance between performance and power consumption. Take the NVIDIA Orin chip as an example. Although it can provide up to 254TOPS of computing power and support the operation of complex fusion algorithms, it also faces the challenges of heat dissipation and power consumption control.

The heterogeneity of multi-source data also brings prominent problems. Data collected by different sensors vary in format, dimension, time and frequency, etc. : cameras output two-dimensional images, lidars generate three-dimensional point clouds, and millimeter-wave radars provide structured data such as target distance and speed. These multi-source data are not only difficult to directly integrate in terms of expression form, but also pose challenges in time synchronization and alignment of spatial coordinate systems. If these problems are not effectively solved, the fusion results may be biased or even incorrect. Therefore, developing efficient feature alignment algorithms and data preprocessing techniques to achieve precise fusion of multi-source data has become an urgent problem to be solved.

Cost and mass production compatibility are important factors restricting the large-scale application of this technology. At present, the cost of high-performance sensors such as lidar remains high: the unit price of mechanical lidar once reached tens of thousands of dollars, and even solid-state lidar, which has more advantages in performance and cost, its price has become an obstacle for car manufacturers to mass-produce. In addition, issues such as the size of sensor hardware and installation compatibility also need to be addressed. For instance, some lidars are too large in size to be integrated into the exterior design of regular vehicle models. It can be seen from this that reducing the cost of sensors while ensuring stable performance is the key to the large-scale application of this technology.

## Development Trends

With the iterative development of intelligent vehicle technology, the core directions of multi-sensor fusion technology are lightweighting and low cost, end-to-end deep learning fusion, and vehicle-road coordination (V2X) fusion, which is reshaping the application landscape [13].

Lightweighting and low cost are reflected in the innovation of hardware. Among them, solid-state lidar has no mechanical rotating parts, reducing costs and size, and accelerating mass production. The 4D millimeter-wave radar has added azimuth and pitch Angle detection, achieving three-dimensional positioning, reducing the reliance on lidar, and also lowering deployment costs. The new type of sensor, through functional integration and performance optimization, reduces hardware and improves accuracy, making the fusion solution more economical and practical.

End-to-end deep learning fusion, such as the Transformer architecture efficiently handling the association of multi-source heterogeneous data; BEV technology converts multi-view data into a unified top view. Occupancy Network voxelization builds a three-dimensional space occupancy model. These architectures directly extract features from raw data and output decisions, eliminating traditional complex steps and enhancing the real-time performance and accuracy of the system.

Vehicle-to-everything (V2X) integration expands the boundary of perception, and vehicles interact with roadside units, other vehicles, etc. in real time through V2X [14]. Roadside equipment can transmit information about obstacles in blind spots, and satellites, in combination with traffic data, plan the optimal path. This mode breaks through the limitations of on-board sensors, achieving beyond-visual-range perception, enhancing the reliability of decision-making in complex scenarios, and facilitating the construction of an intelligent transportation system.

# Conclusion

Multi-sensor fusion technology is the core driving force for intelligent vehicles to upgrade from assisted driving to highly automated, and it is also a key technology for achieving L2+ to L4 level autonomous driving. By complementing the advantages and disadvantages of different sensors, the vehicle's perception ability in complex environments is significantly enhanced. In practical applications, sensors such as cameras, lidars, and millimeter-wave radars work in coordination, effectively compensating for the deficiencies of a single sensor in environmental adaptability and information integrity. In the future, as hardware costs decline, algorithms continue to be optimized, and the vehicle-road coordination system is constantly improved, the integration technology will develop in a more efficient, intelligent and safe direction, promoting the large-scale commercialization of autonomous driving.

The future multi-sensor fusion technology will achieve breakthroughs in multiple dimensions. In terms of hardware, new types of sensors such as solid-state lidar and 4D millimeter-wave radar will gradually become popular and their costs will continue to decline. This will drive the development of in-vehicle perception systems towards lightweight and integration. In terms of algorithms, end-to-end deep learning models will continue to evolve. For instance, the optimization of Transformer, BEV, and Occupancy Network architectures will simplify the data processing flow. In terms of ecology, the improvement of the vehicle-to-everything (V2X) system enables vehicles to integrate information from roadside units, satellite navigation, and other traffic participants, thereby breaking through the physical limitations of on-board sensors and achieving beyond-visual-range perception and collaborative decision-making. The coordinated development of these technologies will drive the multi-sensor fusion technology to upgrade towards greater efficiency, intelligence and safety, reshape the future travel ecosystem, and lay a solid foundation for building a safe, efficient and sustainable intelligent transportation system.

**COMPETING INTERESTS DISCLAIMER:**

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

# References:

1. Zhang Jun, Yang Zichen Research on Data Fusion in Multi-Sensor Data Acquisition System [J]. Sensors and Microsystems,2014,33(03):52-54+57.DOI:10.13873/ J.1000-97872014.03.035.
2. Yang Tao, Li Xiaoxiao. The research development of machine vision technology in modern agricultural production [J]. Journal of Chinese agricultural mechanization, and 2021 (03) : 171-181. The DOI: 10.13733 / j.j CAM. Issn 2095-5553.2021.03.024.
3. Zhang Shuliang, Tan Xiangquan, Wu Qingwen. Research on Positioning of Indoor Mobile Robots Based on Multi-Sensor Fusion Technology [J]. Sensors and Microsystems,2021,40(08):53-56.DOI:10.13873/J.1000-9787(2021)08-0053-04.
4. Deng X ,Shi X ,Wang H , et al.An Indoor Fire Detection Method Based on Multi-Sensor Fusion and a Lightweight Convolutional Neural Network[J].Sensors,2023,23(24):
5. Benedict M ,Hamid A S ,Hamid B .A Critical Review of Deep Learning-Based Multi-Sensor Fusion Techniques[J].Sensors,2022,22(23):9364-9364.
6. Yixiao F ,Xiangyu T ,Tiemin L , et al.Measurement of mobile manipulator chassis pose change caused by suspension deformation and end-effector accuracy improvement based on multi-sensor fusion[J].Robotics and Autonomous Systems,2023,170.
7. Wang Hai, Xu Yansong, CAI Yingfeng, et al Review of Multi-Object Detection Technology for Intelligent Vehicles Based on Multi-Sensor Fusion [J]. Journal of Automotive Safety and Energy Conservation,2021,12(04):440-455.
8. Zhang Yi, Du Fanyu, Luo Yuan, et al. A SLAM map creation method integrating Laser and depth Vision Sensor [J]. Journal of Computer Application Research,2016,33(10):2970-2972+3006.
9. Wang Tingting, CAI Zhihao, Wang Yingxun. Unmanned aerial vehicle (uav) indoor visual/ins integrated navigation method [J]. Journal of Beijing university of aeronautics and astronautics, 2018, 44 (01) : 176-186. The DOI: 10.13700 / j.b h. 1001-5965.2016.0965.
10. Liang Z ,Wang K ,Zhang J , et al.An Underwater Multisensor Fusion Simultaneous Localization and Mapping System Based on Image Enhancement[J].Journal of Marine Science and Engineering,2024,12(7):1170-1170.
11. He G ,He Q .Study on the indoor mobile robot localization based on multi-sensor fusion[J].Journal of Physics: Conference Series,2024,2803(1):012041-012041.
12. Shirshahi A ,Moshiri B ,Shoorehdeli A M .Intelligent fault diagnosis based on similarity analysis using generative model and multi-sensor fusion in industrial processes[J].Process Safety and Environmental Protection,2025,197107097-107097.
13. Editorial Department of China Journal of Highway and Transportation China automotive engineering research review, 2023 [J]. China journal of highway and transport, 2023, 4 (11) : 1-192 DOI: 10.19721 / j.carol carroll nki. 1001-7372.2023.11.001.
14. Ma Xiaoyi. Thoughts on the Role of Vehicle-to-Everything (V2X) Technology in Assisting Autonomous Driving of Vehicles [J].Commercial Vehicles 2022, (01): 71-73+76.
15. Zhang, Y., Zhang, B., Shen, C., Liu, H., Huang, J., Tian, K., & Tang, Z. (2024). Review of the field environmental sensing methods based on multi-sensor information fusion technology. *International Journal of Agricultural and Biological Engineering*, *17*(2), 1-13.
16. De Farias, C. M., Pirmez, L., Fortino, G., & Guerrieri, A. (2019). A multi-sensor data fusion technique using data correlations among multiple applications. *Future generation computer systems*, *92*, 109-118.
17. Liu, Z., Xiao, G., Liu, H., & Wei, H. (2022). Multi-sensor measurement and data fusion. *IEEE Instrumentation & Measurement Magazine*, *25*(1), 28-36.
18. Senel, N., Kefferpütz, K., Doycheva, K., & Elger, G. (2023). Multi-sensor data fusion for real-time multi-object tracking. *Processes*, *11*(2), 501.
19. Yang, B., Li, J., & Zeng, T. (2025). A review of environmental perception technology based on multi-sensor information fusion in autonomous driving. *World Electric Vehicle Journal*, *16*(1), 20.
20. Zhou, X., & Peng, T. (2020). Application of multi-sensor fuzzy information fusion algorithm in industrial safety monitoring system. *Safety science*, *122*, 104531.