Proceedings of CONF-FMCE 2025 Symposium: Semantic Communication for Media Compression and Transmission
DOI: 10.54254/2755-2721/2025.GL27341

A Review of Sensor Technology in Environmental Perception for Intelligent Driving

Junkai Lu

College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China
2021280136@email.szu.edu.cn

Abstract. This study focuses on multi-sensor fusion technology applied in environmental perception for intelligent driving systems and explores its role in enhancing the accuracy and robustness of environmental perception. Analysis shows that multi-sensor fusion can overcome the limitations of single sensors through complementary advantages. Among mainstream sensors, cameras provide 2D image details for target recognition and road condition judgment; LiDAR generates 3D point clouds for high-precision modeling, suitable for L3 and above levels; millimeter-wave radar has strong environmental adaptability but limited detail recognition. Multi-sensor fusion operates at three levels: data-level (retaining original information but requiring high computing power), feature-level (efficient but with data loss), and decision-level (fusing based on independent detection results). Combining multi-sensor fusion with deep learning (e.g., end-to-end fusion) can improve accuracy and real-time performance. Currently, intelligent driving faces challenges of low cost, high real-time performance, and high accuracy. Multi-sensor fusion is an important direction. In the future, it is necessary to optimize sensor performance (such as 4D millimeter-wave radar) and algorithms to meet the needs of complex environments.

Keywords: sensors, intelligent driving, environmental perception, multi-sensor fusion

1. Introduction

Nowadays, with the increasing complexity of motion scenarios, a single sensor gradually struggles to meet the needs of high-precision environmental detection. In the context of intelligent driving, sensors—serving as perceptual tools for vehicles and various mobile devices to perceive the surrounding environment—constitute the fundamental basis for vehicles to execute safe decision-making and achieve autonomous driving. Sensor technology is a core pillar of intelligent driving systems, known as the "sensory center" of vehicles, playing a decisive role in key links such as autonomous navigation, environmental perception, and precise decision-making execution. Each advancement in sensor technology directly promotes the improvement of vehicle safety performance and accelerates the commercialization of unmanned driving technology.

Early research on intelligent driving mainly relied on a single sensor, such as millimeter-wave radar, LiDAR, or cameras for environmental perception. This approach significantly constrains the unique advantages of various sensors and fails to fully leverage the performance potential of the perception system. In practical scenarios, when using a single LiDAR for detection, it is susceptible

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.GL27341

to weather and light conditions and difficult to identify small objects—information that camera sensors can capture. Cameras can also obtain more detailed information about objects, such as color and shape. However, a single camera is inaccurate in judging distance and speed. Therefore, projects on multi-sensor fusion have attracted attention. The fusion of sensors—including millimeter-wave radar and cameras—enables the achievement of enhanced accuracy and robustness in target detection through the mutual complementation of their respective advantages. As the "perceptual organs" of vehicles, sensors can real-time acquire environmental information around the vehicle, such as obstacles, pedestrians, and traffic signs. For example, object distance and driving speed are crucial data in intelligent driving. Tesla's early Autopilot system relied excessively on pure vision, which eventually led to multiple accidents because pure vision systems are prone to misjudging targets under extreme light conditions. After introducing fusion solutions, the system's environmental perception ability became stable. Multi-sensor fusion can effectively overcome these limitations and improve the comprehensive perception ability of the system.

Researchers have begun to explore how to efficiently fuse data from different types of sensors to build a comprehensive and accurate environmental perception system and perception methods. These methods aim to integrate the advantages of radar and cameras. The choice of fusion level for different sensors directly affects system performance. Directly superimposing data, while retaining complete original information, results in excessively high data processing complexity, which may cause calculation delays and affect the real-time performance of recognition. For example, Bird's Eye View (BEV) fusion [1], which integrates radar point clouds and image information captured by cameras to improve spatial perception ability, has a complex feature extraction network and easily causes waste of computing resources. In contrast, feature-level fusion can reduce data processing burden and improve system response speed while retaining key information. By optimizing fusion algorithms, the advantages of various sensors can be combined more effectively to ensure the stability and reliability of intelligent driving systems in different environments. For example, using deep learning algorithms for joint training of multi-source data can not only improve the accuracy of target recognition and reduce false alarm rates but also realize real-time adjustment of fusion strategies.

In addition to sensors, methods related to target detection are also developing rapidly. In recent years, with the development of deep learning, various convolutional neural network models have been proposed. Using convolutional neural networks to extract deep features from image data acquired by sensors has greatly improved the accuracy of target detection. There are mainly two types of target detection methods using deep learning: two-stage detection, with representative algorithms such as the R-CNN series, and single-stage algorithms, which are faster than two-stage detection algorithms, with representative ones including SSD and YOLO algorithms [2]. Based on these continuously developing algorithms, information collected by various sensors can play a more effective role. This paper focuses on the application and evolution of multi-sensors in the field of intelligent driving, uses the advantages of multi-sensors to make up for the limitations of single sensors, analyzes the application scenarios and development trends of various sensors in intelligent driving, aiming to support the innovation, iteration, and wide application of sensor technology.

2. Principles and characteristics of mainstream sensors in intelligent driving

2.1. Visual sensors (cameras): advantages and limitations

Visual sensors such as cameras are core devices that simulate human visual perception in driving. They provide road conditions and target recognition by capturing 2D image information. Their main

DOI: 10.54254/2755-2721/2025.GL27341

working principle is very similar to that of human eyes, both based on optical imaging, including optical sensors and image processing algorithms, and can recognize lane lines, traffic signs, pedestrians, etc. Currently, on-vehicle cameras in the market mostly use high-resolution CMOS sensors, with wide-angle and night vision functions. Combined with various on-vehicle chips, they can perform preprocessing such as noise reduction, compression, and enhancement on the extracted image data, and then use computer vision algorithms such as deep learning models to extract features from image data, such as edges, textures, and target contours, and finally realize target detection and classification.

There are many types of on-vehicle cameras in intelligent driving systems, which are divided into front-view, surround-view, rear-view, etc., each undertaking different tasks in the intelligent driving process of vehicles [3]. For example, front-view cameras are responsible for observing the front of the vehicle and can warn of potential collisions and lane departures. Because this function plays an important role in intelligent driving, the quantity and quality of front-view cameras are constantly improving. With the continuous development and update of intelligent driving systems, multi-camera design has gradually become mainstream to enhance environmental perception capabilities. This is particularly evident in Advanced Driver Assistance Systems (ADAS), where fused information from multiple cameras enables functions such as reverse imaging and holographic imaging. However, like human eyes, cameras are easily affected by harsh environments, and their environmental adaptability needs to be enhanced [4].

2.2. LiDAR: principles, advantages, and limitations

LiDAR (Light Detection and Ranging) is a technology that emits laser beams and receives reflected signals, calculates the laser propagation time or phase difference to determine the distance of the target, and then generates 3D point cloud data of the surrounding environment to realize 3D modeling and target detection of the environment. LiDAR is divided into single-line and multi-line types. Single-line LiDAR only generates a single scanning line, forming 2D point cloud data; multi-line LiDAR can emit and receive multiple laser pulses simultaneously, generating more detailed 2.5D or 3D point cloud models [5]. Here, 2.5D point cloud refers to data that contains height information but lacks complete three-dimensional structural details, as it samples height in a discrete manner rather than capturing the full contour of the target.

The invention of lasers has enabled the wide application of LiDAR technology, from aerospace and surveying to intelligent driving systems of L3 and above levels [6]. The real-time performance, high precision, and ultra-long-distance monitoring characteristics of LiDAR technology are the main reasons for its becoming an important technology in intelligent driving. Its characteristic of being unaffected by lighting conditions enables intelligent driving vehicles to accurately identify road conditions, obstacles, and other important traffic signs in various environments. However, like cameras, LiDAR is greatly affected by severe weather (such as rain, snow, and heavy fog). Laser signals are easily blocked or scattered, resulting in reduced perception accuracy. At the same time, traditional LiDAR has a high cost, which limits its large-scale application in some scenarios [7].

2.3. Millimeter-wave radar: principles, advantages, and limitations

Millimeter-wave radar emits millimeter waves (electromagnetic waves with a wavelength of 1-10 mm) and receives reflected signals, calculates the distance, speed, and azimuth of the target using the Doppler effect and time difference, thereby realizing the detection of obstacles in the surrounding environment. Compared with the previous two sensors, millimeter-wave radar has

DOI: 10.54254/2755-2721/2025.GL27341

strong environmental adaptability, can directly obtain target speed information, and performs well in long-distance obstacle detection, which can effectively make up for the shortcomings of visual sensors in complex environments. However, millimeter-wave radar also has certain limitations. Compared with LiDAR, its resolution is lower, its ability to recognize target details is limited, and it is difficult to distinguish specific categories of targets (such as pedestrians and non-motor vehicles). At the same time, it is susceptible to multipath effects, which may cause false detections in complex scenarios (such as urban building groups).

3. In-depth application of multi-sensor fusion and algorithm integration

3.1. Applications of multi-sensor fusion

Intelligent driving, especially Level L4 and above, has extremely high requirements for the accuracy of environmental perception. However, single sensors have significant limitations: visual sensors are easily affected by illumination and weather, LiDAR is costly, and millimeter-wave radar has limited height measurement capabilities. Therefore, multi-sensor fusion technology can integrate the advantages of various sensors to improve the system's ability to identify obstacles, as well as solve problems such as route planning and trajectory prediction. In intelligent driving, sensors usually used for fusion include optical sensors and non-optical sensors, and the collaboration logic between different types of sensors is different, such as visual sensors and millimeter-wave sensors. Visual sensors are good at recognizing traffic signs, lane lines, and describing pedestrians, but are affected by light, while millimeter-wave radar has strong environmental adaptability and can accurately measure speed. Fusing the two can improve the stability of vehicle detection functions such as FCW and AEB [8]. For example, in pedestrian detection, vision can provide various color and detail features, millimeter-wave radar can ensure detection capability in harsh weather, and infrared sensors can be combined when necessary to enhance detection capability at night. In addition, visual sensors and LiDAR: visual sensors provide rich color and texture information, and LiDAR generates high-precision 3D point cloud models. Their fusion can enhance environmental modeling capabilities in complex scenarios. For example, Xpeng's Xpilot system takes vision as the core, integrates LiDAR to achieve safety redundancy, and converts multi-camera data into Bird's Eye View (BEV) through XNet technology, reducing dependence on high-precision maps.

3.2. Methods of multi-sensor fusion

The essence of multi-sensor fusion is to fuse information collected by multiple different types of sensors. Since the information comes from multiple sources (i.e., multiple sensors), such information is often more reliable. In multi-sensor information fusion, it can be divided into three categories: data-level fusion, feature-level fusion, and decision-level fusion [8].

Data-level fusion refers to the direct fusion of raw data collected by different sensors. It is a low-level fusion method. After completing the raw data fusion, various algorithms need to be used for feature extraction. The advantage of this is that it reduces the loss of raw data, ensures the accuracy of target detection to a certain extent, and provides more comprehensive information. However, its disadvantage is obvious: excessive raw data often consumes huge computing power, increasing computing costs and reducing computing speed, which is contrary to the goal of real-time performance required by intelligent driving.

Feature-level fusion is a higher level than data-level fusion, belonging to the intermediate level. Its main function is to process raw data before use, that is, to extract features from raw data, fuse

DOI: 10.54254/2755-2721/2025.GL27341

them into feature vectors, and then identify target results through target recognition algorithms. Unlike data-level fusion, which directly uses raw data, feature-level fusion extracts features from raw data before fusion, which greatly reduces the amount of data to be processed, thereby improving computing efficiency. However, it also loses part of the raw data and lacks data integrity.

Decision-level fusion is a high-level fusion method. Different from the previous two fusion methods, decision-level fusion requires each sensor to collect data, extract features, and have an independent target recognition mechanism. After completing their respective detection tasks, these sensors output results separately, which are finally processed by the decision layer to form and output the final conclusion.

There are many applications of these fusion methods. For example, data matching fusion algorithms, a type of data-level fusion, use GPS/INS modules to obtain the latitude, longitude, and elevation coordinate data of intelligent vehicles, retrieve matching spatial data in OSM data, extract the number of lanes and lane width of the current road, and provide basic road information for subsequent positioning and navigation.

Since deep learning has been widely applied in the field of images [9], people have begun to apply deep neural networks to intelligent driving systems for various driving decisions, target positioning, and route planning. The adoption of end-to-end fusion approaches based on neural networks—such as target detection networks that integrate radar point clouds and camera images—can significantly enhance the accuracy and real-time responsiveness of environmental perception systems [10]. Additionally, the integration of deep learning algorithms with diverse sensor fusion methodologies constitutes one of the primary trends in the future development of intelligent driving perception technologies.

4. Conclusion

Currently, the various driving environments faced by human society are becoming increasingly complex. Although LiDAR is accurate and precise, its high cost hinders large-scale application; cameras can directly receive a large amount of information but are easily affected by the environment and light, making them difficult to function in night environments. How to achieve low cost and high real-time performance while ensuring the accuracy of target detection is a major challenge for intelligent driving systems. Therefore, the fusion of multi-sensors is a promising direction. At the same time, developing more optimized algorithms and advanced sensors remains a key direction for future research. For instance, high-resolution 4D millimeter-wave radar—when compared with conventional millimeter-wave radar—exhibits higher-dimensional measurement capabilities and enhanced resolution for target objects. On the other hand, continuous optimization of signal processing algorithms, coupled with the adoption of advanced machine learning algorithms to analyze and process raw signals, enables the extraction of more effective information from complex echo signals.

After specifically analyzing the application scenarios and limitations of various single sensors, this paper introduces the advantages and application methods of multi-sensors. By utilizing the respective advantages of each sensor to make up for each other's disadvantages, the real-time performance and accuracy of intelligent driving are improved. At the same time, it is proposed to combine current deep learning and large models and other technologies, not only to improve the quality of sensors but also to develop corresponding target detection algorithms to match the use of sensors, so as to meet the detection needs of increasingly complex environments.

DOI: 10.54254/2755-2721/2025.GL27341

References

[1] Zhou W. Research on multimodal fusion 3D object detection algorithm based on BEV perspective [D]. Beijing: Beijing Jiaotong University, 2024.
[2] Ling Y J. Research on optimization of video surveillance object detection algorithms based on YOLO, SSD and Faster R-CNN [J]. Modern Computer, 2024, 30(21): 49-53, 58. DOI: 10.3969/j.issn.1007-1423.2024.21.009.
[3] LI Y X, ZHA Y F. Applications and Development Trends of Common Field of Intelligent Driving [J]. Automotive Digest (Chinese), 2025(4): 12-22.
[4] Cheng Z M. The "eyes" of intelligent driving — present and future of vehicle-mounted camera technology (Part 1) [J]. Automobile Maintenance and Repair, 2022(8): 43-46.
[5] LI Y, IBANEZ GUZMAN J. Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems [J]. IEEE Signal Processing Magazine, 2020, 37(4): 50-61.
[6] SHI P, ZHANG Y, LI J. LiDAR-Based Place Recognition For Autonomous Driving: A Survey [J/OL]. (2023-07-29) [2024-03-13]. https: //www.researchgate.net/publication/371729036_LiDAR-Based_Place_Recognition_For_Autonomous_Driving_A_Survey.
[7] Nie G Y. Design and verification of predictive adaptive cruise control algorithm [D]. Jilin: Jilin University, 2022.
[8] Li L, Xiao S D, Li X K, Dong Q F. Design of multi-sensor fusion based intelligent vehicle positioning and navigation system [J]. Journal of Intelligent Transportation Systems, 2023, 15(2): 45-58. DOI: 10.1007/s13177-023-00321-x.
[9] Review of YOLO Based Target Detection for Autonomous Driving MAO Zhihui, ZHU Jiali, WU Xin, LI Jun College of Information and Intelligent Engineering, Zhejiang Wanli University, Ningbo, Zhejiang 315100, China.
[10] Y. Li, A. W. Yu, T. Meng, B. Caine, J. Ngiam, D. Peng, J. Shen, Y. Lu, D. Zhou, Q. V. Le, A. Yuille, and M. Tan, “DeepFusion: Lidar - Camera Deep Fusion for Multi - Modal 3D Object Detection, ” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 17182 - 17191, doi: 10.1109/CVPR52688.2022.01667.