

Visual and Radar Data Fusion: Complementary or Alternative Sensors in Intelligent Driving Systems

Junzhe Chen

University of Wisconsin-Madison, Wisconsin, USA

jchen2497@wisc.edu

Abstract. This article explores the role and advantages and disadvantages of visual and radar data fusion in intelligent driving systems. With the rapid development of intelligent driving technology, the comprehensive perception of the environment by car perception systems has become crucial, usually requiring the combination of low, medium, and high-precision sensors. Visual sensors perform excellently in acquiring high-resolution images and recognizing detailed information, while radar sensors have higher stability and penetration in harsh weather conditions. Due to the limitations of a single sensor in complex driving environments, data fusion has become a key strategy for improving the performance of perception systems. By comparing different perception schemes and application scenarios, this article analyzes the actual effects and potential advantages of visual and radar data fusion, revealing its important role in intelligent driving technology. Fusion technology can compensate for the shortcomings of various sensors, and improve the accuracy and reliability of environmental perception, but still faces challenges in data synchronization, algorithm selection, and real-time performance. This article aims to provide valuable insights for the development of future intelligent driving technology.

Keywords: Intelligent driving, radar, visual, sensor.

1. Introduction

The rapid development of intelligent driving technology is driving changes in the automotive industry, among which the core role of perception systems cannot be ignored. In order to achieve comprehensive environmental perception, autonomous vehicles usually rely on multiple sensors, classified into low-, medium-, and high-fidelity sensors [1]. These sensors each have their own unique advantages and limitations. Visual sensors excel in providing high-resolution images and are suitable for identifying detailed information such as traffic signs, lane lines and pedestrians; while radar sensors show higher stability and penetration in adverse weather conditions, and have irreplaceable advantages in detecting distance and speed.

However, the performance of a single sensor in a complex driving environment is often limited. Therefore, the fusion of visual and radar data has become an important strategy to improve the performance of perception systems. Most autonomous driving systems apply sensor fusion [2]. Data fusion technology can make up for the shortcomings of each sensor by processing the information of different sensors in a comprehensive manner, and achieve a more comprehensive and accurate understanding of the environment. Despite this, there are still many controversies and challenges in this

field, such as the synchronization of sensor data, the selection of fusion algorithms, and the real-time performance of the system.

This study aims to explore the role of visual and radar data fusion in intelligent driving systems and analyze whether they are complementary or alternative. By comparing different perception schemes and analyzing specific application scenarios, we hope to reveal the actual effects and potential advantages of various sensor fusions and provide valuable insights for the development of future intelligent driving technology.

2. Different perception systems in intelligent driving

2.1. LiDAR

This system consists of a LiDAR, a camera, and a radar. The LiDAR provides high-precision distance measurement and 3D environmental images, the camera captures detailed visual information such as traffic signs, lane markings, and pedestrians, and the radar detects the speed and position of objects. The characteristics of Plan A include high-precision 3D environment perception and rich visual information, improving the detection ability of dynamic objects through multimodal fusion. Its main advantages lie in accurate environmental recognition and excellent performance in nighttime and low-light environments, but there are also issues of high cost and complex data processing requirements.

2.2. Millimeter wave radar

This system relies on millimeter wave radar and cameras, possibly supplemented by ultrasonic sensors. Millimeter wave radar provides accurate speed measurement and distance perception that is not affected by light, while cameras provide environmental images to assist in identifying traffic signs and lane information. The characteristic of this scheme lies in the stability of the radar and its adaptability to harsh weather conditions, while the camera supplements visual information [3]. Its advantages lie in its high cost-effectiveness and strong environmental adaptability, but due to the low resolution of millimeter wave radar, it may not be able to accurately recognize small objects, and the effectiveness of image recognition depends on the image processing capability of the camera.

2.3. Ultrasonic sensors

This system mainly includes cameras and ultrasonic sensors, which may be supplemented with a small amount of radar. The camera provides the main visual information and performs environment recognition through deep learning algorithms, while ultrasonic sensors are used for close-range obstacle detection, such as parking assistance systems. The characteristic of this solution is to rely on low-cost cameras and ultrasonic sensors, but it lacks accuracy and environmental adaptability. Its advantages lie in lower hardware costs and advances in image processing technology, but it performs poorly in light changes and harsh weather conditions, with lower resolution and measurement accuracy.

2.4. Comparison

A comparison of schemes shows that LiDAR performs the best in accuracy, but due to high costs and complex data processing requirements, it may not be suitable for all application scenarios. The millimeter wave radar has significant advantages in cost-effectiveness and stability and is suitable for various environments, but its resolution and image recognition capabilities are limited. The ultrasonic sensors have the lowest cost and are suitable for applications with limited budgets, but it has significant shortcomings in environmental adaptability and accuracy. After comprehensive consideration, the advantages and disadvantages of each solution should be reasonably balanced according to specific application scenarios and requirements to achieve the best performance of the intelligent driving system.

3. Different sensors in intelligent driving

The selection of various sensors is crucial in intelligent driving systems, as each sensor plays a unique role in environmental perception and has its own specific advantages and disadvantages. Monocular

cameras are one of the most common visual sensors that capture two-dimensional images of the environment through a lens and image sensor. It has low cost, compact design, and easy integration, but is highly sensitive to environmental conditions such as lighting and weather changes, and has limitations in depth perception, requiring algorithm supplementation [4]. The binocular camera system uses two cameras to simulate human eyes and provides accurate three-dimensional depth information through disparity calculation. This system can achieve strong depth perception and stereoscopic vision, improving the accuracy of object detection and scene understanding. However, binocular cameras are expensive and require processing complex image data to calculate depth information. Conventional radar sensors such as millimeter wave radar measure the distance and velocity of objects by emitting electromagnetic waves and receiving their reflected waves. After radar radiates electromagnetic waves, it gathers the scattered waves of the target through the receiving antenna, and then performs a series of signal processing to obtain target information [5]. Radar sensors are not affected by lighting and weather conditions and can operate stably in various environments. However, their resolution is low, making it difficult to detect small objects or details of the environment, and may be subject to interference from other radar systems. Lidar generates a three-dimensional environmental map by exciting a laser beam and measuring reflection time, providing extremely high accuracy and detailed point cloud data. This makes LiDAR excellent in environmental modeling and obstacle detection, but its equipment cost is expensive and requires a powerful computing platform to process large amounts of data. To sum up, the application of these sensors in intelligent driving has its own advantages and challenges. Reasonable selection and combination of different sensors can significantly improve the perception ability and overall performance of the auto drive system.

4. Complex scene analysis in intelligent driving

It's of vital importance to analyze complex situations to ensure the safety of automatic driving systems. Here are 2 scenarios which should be attached the most importance to due to their commonality.

4.1. The complex traffic environment at urban intersections

In this scenario, the intelligent driving system must handle a large amount of traffic information, including oncoming vehicles, pedestrians, bicycles, as well as various traffic signs and signals. The characteristics of urban intersections are high-density traffic flow and complex intersection paths, which require the system to have high-level real-time perception and decision-making capabilities. The main difficulties include how to accurately identify and predict the behavior of traffic participants from different directions, especially in busy intersections where the state of traffic flow, pedestrian flow, and signal lights changes rapidly and various unexpected situations often occur. This requires the system to not only have strong data fusion capabilities but also to have fast response and processing capabilities.

4.2. Lane changes and overtaking operations on highways

The characteristic of highway scenarios is that vehicles travel at high speeds, and lane markings are usually clear, but complex lane changes and overtaking requirements pose different challenges to the system. The difficulty lies in how to accurately evaluate the speed and position of surrounding vehicles at high speeds, as well as how to safely perform overtaking operations when making lane changes. This requires intelligent driving systems to have extremely high detection accuracy and decision-making speed to ensure that they can quickly and accurately handle tasks such as lane changes, distance adjustments, and obstacle avoidance during high-speed driving. At the same time, driver expectations and compliance with traffic regulations need to be considered to ensure the safety and reliability of the system.

5. Optimization of intelligent driving perception solution

In intelligent driving systems, optimizing the perception scheme is the key to achieving safe and efficient autonomous driving. Perception systems rely on various sensors and algorithms to perceive the environment, identify obstacles, predict the behavior of traffic participants, and make decisions and

controls. In order to improve the performance of intelligent driving systems, comprehensive optimization must be carried out from the aspects of fusion technology, machine learning, large models, sensor hardware progress, and computing power.

5.1. Fusion

The theoretical basis of vision and radar data fusion lies in the comprehensive utilization of the advantages of different sensors to make up for the shortcomings of a single sensor. Visual sensors (such as cameras) can capture high-resolution two-dimensional images, providing rich color and texture information that is crucial for identifying traffic signs, lane markings, pedestrians, and more. However, the main limitation of visual data is that it is highly sensitive to lighting conditions and weather changes, such as at night or in adverse weather conditions, where camera performance may significantly degrade. Radar sensors (such as millimeter wave radar) measure the distance and velocity of objects by emitting electromagnetic waves and receiving their reflected signals. The main advantage of radar lies in its strong environmental adaptability, which enables stable operation under various lighting and weather conditions. However, the resolution of radar is low, making it difficult to provide detailed environmental information, especially in identifying small objects or environmental details.

The theoretical basis for integrating visual and radar data lies in utilizing the stability of radar and the high resolution of vision, and combining the advantages of both through data fusion algorithms to improve the accuracy and reliability of overall environmental perception. This fusion usually combines data from different sensors into a comprehensive environment model through algorithms to help the system better identify and predict various elements in the environment.

5.1.1. Specific fusion algorithm. Specific fusion algorithms include data-level fusion, feature-level fusion, and decision-level fusion. Data level fusion refers to the direct combination of raw data from different sensors. Common algorithms include Kalman filter and particle filter. Kalman filter is a recursive algorithm that estimates the state of a target by weighted averaging sensor data, suitable for handling noisy data in linear systems [6]. Particle filtering is suitable for systems with nonlinear and non-Gaussian noise, and approximates the state distribution of the target by generating a large number of particle samples. Feature level fusion is the process of fusing extracted features from sensor data after initial processing. For example, by extracting edge features of objects through visual sensors and combining them with distance information provided by radar sensors, more accurate target recognition and tracking can be achieved. Common algorithms include support vector machines (SVM) and fusion networks in deep learning, which can process data in high-dimensional feature spaces and improve the accuracy of object detection. Decision-level fusion is the fusion of the judgment results made by each sensor separately. Common methods include weighted voting and confidence fusion, which combine the judgment results of various sensors to make the final decision. For example, in pedestrian detection, visual sensors may recognize multiple potential pedestrians, while radar sensors provide distance information for these targets. Through decision level fusion, the accuracy and reliability of the final recognition can be improved.

5.1.2. Implementation results and challenges. The implementation effect of fusion technology significantly improves the environmental perception ability of intelligent driving systems. By integrating vision and radar, the system can maintain high detection accuracy under various lighting and weather conditions, improving the reliability of target recognition and scene understanding. However, achieving these effects faces some challenges.

Firstly, the data acquisition frequency and format of different sensors may be different. How to accurately synchronize them together is an important issue. This requires efficient time synchronization mechanisms and data preprocessing algorithms to ensure the accuracy of the fused data.

Although data fusion algorithms can significantly improve system performance, their computational complexity is high, requiring powerful computing resources to process real-time data.

The errors and noise of the sensor itself may also affect the fusion effect, so it is necessary to design effective filtering and denoising algorithms to improve the quality of the fused data.

5.2. Machine learning / large models

5.2.1. Application of machine learning in data fusion. The application of machine learning in data fusion is mainly reflected in automatically learning and extracting important features from data through algorithms, thereby improving the perception ability of the system. Traditional data fusion methods rely on rules and models, while machine learning can automatically discover patterns and patterns in data by training on large amounts of historical data. For example, a model called “FIERY” enables the camera to get rid of GPS and predict safe directions only by modeling the inherent randomness of the future from camera driving data [7]. Statistics have shown that their model outperforms previous prediction baselines on the NuScenes and Lyft datasets.

5.2.2. The role of deep learning models. The role of deep learning models in intelligent driving perception systems is particularly prominent. Deep learning, especially convolutional neural networks (CNNs), performs well in processing image data. By training deep neural networks, the system can automatically learn complex feature representations from a large amount of annotated data, achieving excellent performance in tasks such as object detection, lane recognition, and traffic sign recognition. For example, the YOLO (You Only Look Once) series models can detect and classify objects in images in real-time, and can supplement and validate radar data during processing, thereby improving overall recognition accuracy.

5.2.3. Case study on performance improvement. In practical applications, there are numerous cases of performance improvement for deep learning models. For example, Tesla's auto-drive system uses the deep convolutional network to process the data from the camera, realizing efficient pedestrian detection and obstacle recognition [8]. In addition, Baidu's Apollo project achieves high-precision environmental perception and autonomous driving decision-making by combining deep learning and sensor fusion technology [9]. The application of deep learning models can significantly improve the detection accuracy and processing speed of perception systems, but it also requires a large amount of computing resources and training data to support.

5.3. Sensors

5.3.1. Progress in sensor hardware. The progress of sensor hardware is an important aspect of optimizing intelligent driving perception solutions. In recent years, sensors such as LiDAR, millimeter wave radar, and cameras have significantly improved in terms of performance and cost. The resolution and detection range of laser radar continue to improve, and the new generation of laser radar can provide more detailed three-dimensional environmental data. The frequency band and detection capability of millimeter wave radar has also been enhanced, which can work stably under high-speed driving and adverse weather conditions. The image sensor technology of cameras is also constantly advancing, such as improving image quality by enhancing low light performance and high dynamic range.

5.3.2. Improved performance analysis. The improved sensor hardware can significantly enhance the perception capability of intelligent driving systems. High-resolution LiDAR provides more accurate environmental models, enhanced millimeter wave radar can maintain stable performance in high-speed and complex environments, and advanced camera technology improves image quality in low light and high-contrast scenes. These hardware advancements enable the system to more accurately perceive the surrounding environment, improving safety and driving experience. However, these improvements also bring about issues of increased costs and system integration complexity, which need to be comprehensively considered to achieve optimal system performance.

5.4. Computing power

5.4.1. Calculation requirements for data processing. The data processing and computing requirements of intelligent driving systems are very high. The amount of data collected by sensors is enormous, including high-resolution images, radar point clouds, and three-dimensional point cloud data from LiDAR [10]. These data require real-time processing to achieve accurate object detection, obstacle recognition, and path planning. The computing requirements mainly include data preprocessing, feature extraction, fusion processing, and decision making, each step requiring powerful computing resources to ensure real-time response and high accuracy.

5.4.2. Application of hardware acceleration technology. In order to meet high computing demands, hardware acceleration technology is widely used in intelligent driving systems. Graphics Processing Units (GPUs) and Application Specific Integrated Circuits (ASICs) are the main acceleration technologies that can significantly improve data processing speed. GPUs are capable of processing large amounts of data in parallel, making them highly suitable for training and reasoning deep learning models. ASIC is an integrated circuit optimized for specific applications, which can provide advantages in power consumption and computational efficiency. For example, NVIDIA's Drive PX platform and Intel's Mobileye series chips have been widely used in intelligent driving systems to support efficient data processing and real-time decision-making.

5.4.3. Computing power optimization strategy. Computing power optimization strategies include algorithm optimization, hardware upgrades, and distributed computing. Algorithm optimization reduces computational complexity and resources by improving data processing and model inference algorithms.

6. Conclusion

In intelligent driving systems, the fusion of visual and radar data demonstrates the complementary advantages of sensor technology, rather than simply replacing each other. Visual sensors and radars each have unique advantages: visual sensors provide high-resolution image information, suitable for recognizing complex environmental details and traffic signs, but their performance is greatly affected by lighting and weather conditions; radar sensors have stable distance measurement and speed detection capabilities, and can maintain stable operation even in adverse weather conditions. However, the resolution of radar is low, making it difficult to provide detailed environmental information. Therefore, it is difficult for a single sensor to meet the overall perception needs of the auto drive system in various environments.

By integrating visual and radar data, intelligent driving systems can comprehensively utilize the advantages of both and compensate for their respective shortcomings. Specifically, the stable distance information provided by radar can effectively supplement the limitations of visual sensors in low light or complex weather conditions, while visual sensors can enhance the ability to recognize environmental details in radar data. This multimodal fusion technology not only improves the accuracy and reliability of environmental perception, but also optimizes data processing and decision-making processes.

In addition, technology evolution and emerging technologies such as deep learning and edge computing further promote the development of vision and radar fusion technology. These technologies can improve data processing speed and accuracy, supporting more complex perception tasks. Although there are still challenges such as data synchronization, processing complexity, and sensor errors in the fusion process, the significant performance improvement and system reliability it brings make the complementary fusion of vision and radar the best choice in intelligent driving systems.

In summary, the fusion of visual and radar data is not only an effective integration of their respective advantages, but also a step towards higher precision and safety in intelligent driving technology. The complementarity of sensors determines that they play a complementary role in the auto drive system, rather than simply replacing each other. This fusion strategy will continue to drive the development of intelligent driving technology, bringing a more intelligent and safe driving experience.

References

- [1] Schlager, B., Muckenhuber, S., Schmidt, S., Holzer, H., Rott, R., Maier, F. M., ... & Ruebsam, J. (2020). State-of-the-art sensor models for virtual testing of advanced driver assistance systems/autonomous driving functions. *SAE International Journal of Connected and Automated Vehicles*, 3(12-03-03-0018), 233-261.
- [2] Yeong, D. J., Velasco-Hernandez, G., Barry, J., & Walsh, J. (2021). Sensor and sensor fusion technology in autonomous vehicles: A review. *Sensors*, 21(6), 2140.
- [3] Zhang, Z., Wang, X., Huang, D., Fang, X., Zhou, M., & Zhang, Y. (2021). MRPT: Millimeter-wave radar-based pedestrian trajectory tracking for autonomous urban driving. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-17.
- [4] Martins, P. F., Costelha, H., Bento, L. C., & Neves, C. (2020, April). Monocular camera calibration for autonomous driving—a comparative study. In *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)* (pp. 306-311). IEEE.
- [5] Wang, Z., Wu, Y., & Niu, Q. (2019). Multi-sensor fusion in automated driving: A survey. *Ieee Access*, 8, 2847-2868.
- [6] Khodarahmi, M., & Maihami, V. (2023). A review on Kalman filter models. *Archives of Computational Methods in Engineering*, 30(1), 727-747.
- [7] Hu, A., Murez, Z., Mohan, N., Dudas, S., Hawke, J., Badrinarayanan, V., ... & Kendall, A. (2021). Fiery: Future instance prediction in bird's-eye view from surround monocular cameras. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 15273-15282).
- [8] Mit, R., Zangvil, Y., & Katalan, D. (2020, September). Analyzing tesla's level 2 autonomous driving system under different gnss spoofing scenarios and implementing connected services for authentication and reliability of gnss data. In *Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020)* (pp. 621-646).
- [9] Zhang, Y., & Wu, T. (2021). Will Baidu's "All in AI" Strategy Bring It Back to the High-Speed Growth Train?. *Journal of Applied Business and Economics*, 23(4).
- [10] Li, Y., & Ibanez-Guzman, J. (2020). Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. *IEEE Signal Processing Magazine*, 37(4), 50-61.