Theory and Practice of Science and Technology
2025, VOL. 6, NO. 6, 137-140
DOI: 10.47297/taposatWSP2633-456927.20250606

# Pure Vision vs Lidar: Why Cameras Are the Future of Autonomous Driving

Zihao Wang

Veritas Collegiate Academy of Beijing, Beijing, 065603, China

**ABSTRACT**
The current perception systems of autonomous driving technology can be divided into two paths. One is the pure vision camera solution, where Tesla is the most representative. The other is the multi-sensor fusion solution led by Waymo using lidar. These two approaches have significant differences in terms of principle, cost, safety reliability, and future market potential. This article argues that the pure vision camera solution has more advantages, whether from the cost advantage or from the potential for algorithm evolution. This article will prove the author's viewpoint through a comparison of these four points: the technical principle, cost, performance comparison, and case analysis - pure vision camera solutions have greater potential in the future. With the maturity of deep learning models, in the future, pure vision autonomous driving technology will completely replace lidar technology and can thereby promote the full commercialization layout of autonomous driving technology.

**KEYWORDS**
Pure vision; Lidar; Comparison; Safety; Performance; Cost; Anti-attack; Measure accuracy

## 1 Introduction

The perception system of autonomous driving affects major car manufacturers in terms of technology, safety, and commercialization. First, the perception system of autonomous driving is like its eyes and ears. Autonomous vehicles need a sensitive sensory system to analyze obstacles, pedestrians, and vehicles on the road in real time, thereby planning the route. A better perception system can help the autonomous system better understand the road and provide important information for decision-making. When the perception system fails, it is like a person's eyes being covered, and the autonomous vehicle cannot change the driving route in time or make incorrect decisions due to different situations on the road. Moreover, the perception system is directly related to the safety of the vehicle. An effective perception system can make timely predictions by observing the information on the road to prevent possible accidents. According to statistics on road traffic accidents, the vast majority of accidents are related to the perception and decision-making errors of the driver. Therefore, a good perception system is very important. When the perception system fails, such as a significant delay between the perception system and the decision-making system, the safety of the vehicle will be greatly reduced. In handling extreme weather conditions, a good perception system is even more important. For example, when visibility is reduced due to heavy fog, if information cannot be accurately captured, the probability of traffic accidents will significantly increase. Supporters of the pure vision camera solution at present believe that if AI can perfectly simulate the visual ability of human drivers who rely on both eyes to drive, additional sensors will become redundant. The implementation costs of pure vision and lidar solutions also have significant differences. A precise lidar costs up to several thousand dollars, greatly delaying the commercialization of autonomous driving technology. While the cost of cameras is only a few dollars, it is more conducive to the full popularization of autonomous driving technology and has a higher ability to obtain information, lower cost, and is more conducive to the continuous optimization and faster algorithm iteration of AI models. However, the lidar-led solution has a higher cost for data annotation due to the scarcity of data. Therefore, from the perspective of full commercialization, the pure vision camera-led solution has greater development potential. Comprehensive from the perspectives of technology, safety, and commercialization, the combination of vision cameras and AI will ultimately completely replace the lidar solution.

## 2 Technical comparison

### 2.1 Principle comparison

The vision perception system of pure vision cameras uses multiple high-definition cameras to achieve 360-degree unobstructed perception of the surrounding environment. The pure vision perception system mostly relies on precise information capture capabilities, large data accumulation, and powerful computing power. The most advanced high-definition cameras can capture high frame rate and high resolution images in real time, effectively obtaining more detailed road conditions. At the same time, the integration of AI enables the entire autonomous driving system to analyze the surrounding environment more quickly and accurately. (Li Zihao, 2024) Currently, Tesla dominates the autonomous driving technology based on pure visual perception. Many models, such as Model 3, are equipped with a total of 8 high-definition cameras, which not only have a 360-degree unobstructed view but also have the ability to detect up to 250 meters away. In the United States and Europe, Model 3, a self-driving vehicle relying on pure visual perception, has a higher safety level compared to other vehicles relying on radar. The focus on continuously improving the performance of the pure visual perception system lies in the cameras. Currently, the market mainly uses two types of cameras: single-lens cameras and dual-lens cameras. A single-lens camera refers to the use of a single camera to recognize and capture the surrounding environment and road conditions (Zhang Jingyi, 2024). When capturing spatial data, single-lens cameras will have a certain degree of distortion and cannot fully reproduce the road conditions (Liu Hongli, 2024). Dual-lens cameras simultaneously use two cameras for information capture, with one as the reference and the other for comparison, and then combine the information from the two perspectives. Compared to single-lens cameras, dual-lens cameras have higher stability, more precise information acquisition capabilities, and lower errors in close-range situations, typically around 5% (Ni Wentao, 2021). After the road conditions are identified and collected, they will be transmitted to the system, where algorithms such as YOLOv4-tiny and DeepSORT will convert the surrounding environment from a 2D form to 3D for monitoring and tracking. At the same time, this algorithm can also be combined with intelligent algorithms, such as target warning or automatic obstacle avoidance (Guo Pengyu, 2023).

The laser radar perception system uses the simulation localization and mapping (SLAM) technology, using the fusion of multiple sensors to collect real-time data and draw environmental maps based on its own positioning. Usually, automotive manufacturers use positioning methods such as global positioning navigation systems, positioning, and wheel speed measurement. During the process of fusing multiple sensors, different types of SLAM sensors are often used, mainly including mechanical rotating laser radar, solid-state laser radar, and semi-solid-state laser radar (Yang Chao, 2024). Laser radar is often considered to have higher resolution, accuracy, and anti-interference capabilities. Represented by Google's Waymo technology, LiDAR sensors can quickly emit laser pulses. When the laser signal reaches an obstacle, it will reflect back to the sensor, and through the calculation of the return time, the distance to the obstacle can be accurately calculated (Jiang Jiarui, 2019). Using the real-time calculated data for modeling, after integrating intelligent analysis algorithms, it can accurately determine the road conditions like bat echolocation and make correct driving decisions.

## 2.2 Cost comparison

The consensus in the industry is that high-precision laser radar is expensive and has a low production volume due to the need for precise optical components. However, cameras have been updated for a long time, so they have a lower cost. Moreover, Musk mentioned in an interview in 2023 that "Laser radar is an expensive and unnecessary crutch, and the marginal cost of cameras is approaching zero." In the process of fusing multiple sensors, different types of SLAM sensors are often used, mainly including mechanical rotating laser radar, solid-state laser radar, and semi-solid-state laser radar (Yang Chao, 2024). Laser radar is often considered to have higher resolution, accuracy, and anti-interference capabilities. Represented by Google's Waymo technology, LiDAR sensors can quickly emit laser pulses. When the laser signal reaches an obstacle, it will reflect back to the sensor, and through the calculation of the return time, the distance to the obstacle can be accurately calculated (Jiang Jiarui, 2019). Using real-time calculated data for modeling, after integrating intelligent analysis algorithms, it can accurately determine the road conditions and make correct driving decisions like bat echolocation.

## 2.3 Performance comparison

### 2.3.1 Reliability in Extreme WeatherConditions

Reliability in Extreme Weather Conditions In extreme weather conditions, the bottleneck of the pure visual solution lies in (1) low light, (2) rain, snow, and fog, and (3) strong reflection. Due to the reflective nature of the camera, it may lead to missed detections of targets, such as obstacles on the road or the position and size of pedestrians, which cannot be accurately determined in low-light nights. In the 2024 investigation, the night visibility of the pure visual solution for autonomous driving technology was only 30 meters, far lower than the 200 meters visibility of lidar. Strong rainfall, snowfall can cover the camera, causing blurring or obstruction, and in rainy weather, visibility is only 50 meters. While heavy fog reduces the contrast of the image, increasing the difficulty of post-processing data analysis. When encountering backlight water surfaces or roads with strong reflection, the camera may overexpose, resulting in the inability to accurately identify lane lines (Wei Hao, 2025). Lidar and multi-fusion sensors perform better in extreme weather conditions, although they are also affected by rain, snow, and fog, but because the laser wavelength is 1550nm, it has better penetration and stronger information acquisition capabilities. However, raindrops and snowflakes can also cause scattering noise, and algorithms need to be integrated to reduce interference (Wu Xiaoyu, 2020).

### 2.3.2 Anti-attack capability

The attacks that the lidar solution may receive can be classified as LiDAR-Adv attacks and reflection object deception. LiDAR-Adv attacks may cause the lidar to misidentify obstacles in front as empty space, resulting in the inability to detect the vehicle ahead, leading to collisions. Reflection object deception occurs because specific angles of the laser are reflected, causing the system to misjudge and resulting in unnecessary emergency braking, causing traffic accidents. The attacks that the pure visual solution may receive can be classified as sticker attacks and projection attacks. Sticker attacks refer to the situation where, under the influence of specific symbols, the decision system cannot accurately distinguish traffic signs from stickers, and may mistakenly identify stickers as traffic signal signs, thereby causing accidents such as sudden stops. Projection attacks involve using a high-power projector to project lane lines on the ground, thereby luring vehicles to deviate (Zheng Junjie, 2022).

From the perspective of anti-attack, there are two solutions that both have four defense methods: (1) hardware layer defense, (2) data layer defense, (3) algorithm layer defense. The hardware-level defense of lidar can be achieved by encrypting the laser encoding to prevent signal forgery, while the visual solution can directly use infrared cameras. Data layer defense, for lidar, can use noise filtering to combat scattering noise, while the pure visual solution requires training models and performing image physical consistency detection. Algorithm layer defense, for lidar, can be achieved through the fusion of multiple sensor algorithms, while the pure visual solution requires the use of multi-model integration. Considering the defense from three different perspectives, the defense deployment of the pure visual solution has lower costs and greater development potential. By prioritizing the training of models to prevent physical attacks such as sticker attacks, a higher security level can be achieved (Zhao Yuan Yuan, 2020).

### 2.3.3 Measurement accuracy

Lidar calculates the distance between the object and the vehicle in real time by collecting the time of laser return from different positions, achieving centimeter-level accuracy. While the pure visual solution uses indirect estimation, relying on the features of the image and the algorithm model of deep learning for matching, with an accuracy of meters (Cheng Xuexiao, 2024). Lidar is active emission, thus being less affected by weather conditions, and using the laser wavelength of 1550nm, it has stronger penetration and anti-interference capabilities. The pure visual solution relies on environmental light, and its measurement accuracy is easily affected by weather conditions, with a visibility distance of only 30 meters at night. In dynamic scenarios, lidar, due to its high frame rate advantage, can more accurately capture fast-moving information, such as pedestrians suddenly crossing the road or vehicles suddenly merging. However, the pure visual solution tends to produce blurry images in certain scenarios. However, when combined with Occupancy Networks, it can enhance the ability to model dynamic objects (Wei Hao, 2025). Additionally, multiple articles have stated that the pure visual solution combined with enhancement algorithms can also ensure low errors (Cheng Xuesha, 2024), which meets the safety requirements of general passenger vehicles.

## 3 Conclusion

Although lidar outperforms the pure visual solution in terms of extreme weather conditions, anti-attack capabilities, and measurement accuracy, due to the stronger algorithm evolution potential of cameras and their similar perception logic to human drivers, the integration of deep learning models and algorithms can significantly narrow the performance gap between the two solutions. Since camera manufacturing costs are low, and compared to lidar, there are almost no usage and maintenance costs, and the market share of cameras is high. Therefore, considering cost and performance, in the future when AI algorithms are powerful enough, the pure visual solution will completely replace the lidar solution and become the mainstream choice for autonomous driving.

## About the Author

Zihao Wang, 3554860950@qq.com.

## References

[1] Wei, H., & Cao, Y. (2025). A brief analysis of safety issues in purely visual autonomous driving. Times Auto, (4), 153-156.

[2] Li, Z. H. (2024). "Pure vision" vs. "integrated perception": Who will lead the future of autonomous driving? Automobile Manufacturing, (4), 12-15.

[3] Xu, Y. N., Zhu, X. C., Ma, Z. X., et al. (2014). Three-dimensional road reconstruction in front of the vehicle based on a monocular camera. Automotive Engineering, (2), 48-52.

[4] Zhang, J. Y., Chen, X. D., & Wang, H. (2024). Research on the slow start problem of hill following based on a monocular camera by Bertoli. Times

Auto, (10), 169-171.

[5] Liu, H. L., Wang, Y. L., Shao, L., et al. (2024). Research on monocular vision vehicle ranging based on the lower edge of the detection box. Computer Science, 51(z1), 415-420.

[6] Ni, W. T., Zhong, D. S., & Tang, X. Y. (2021). Implementation of a ranging system based on binocular cameras. Wireless Internet Technology, 18 (2), 31-32.

[7] Guo, P. Y. (2023). Vehicle detection, tracking, and ranging based on binocular vision. Intelligent Computer and Applications, 13(4), 147-150, 157.

[8] Zhang, G., Yang, C., Wang, W. D., et al. (2024). A review of autonomous driving simultaneous localization and mapping methods based on LiDAR. Journal of Automotive Engineering, 14(1), 1-13.

[9] Jiang, J. R. (2019). Analysis of the core sensor system in unmanned driving. Times Auto, (2), 145-146.

[10] Zheng, J. J. (2022). Perception attack and defense of autonomous driving based on LiDAR (Doctoral dissertation). Zhejiang University of Technology, Zhejiang.

[11] Wu, X. Y., Zhang, X., & Wang, W. Z. (2020). Working principle, performance comparison, and safet analysis of LiDAR for intelligent connected vehicles. Information Security and Communication Privacy, (9), 92-98.

[12] Zhao, Y. Y., & Zeng, Y. T. (2024). Research and application of pure vision BEV fusion in autonomous driving. Computer Applications Digest, 40 (14), 107-108, 111.

[13] Cheng, X. X. (2024). Research and application of high-precision vision detection technology for new energy vehicle autonomous driving. Times Auto, (5), 104-106.

[14] Hou, J. (2011). Research on lane detection and forward vehicle ranging technology based on monocular vision (Doctoral dissertation). Tianjin Polytechnic University, Tianjin. https://doi.org/10.7666/d.y2059493.

[15] Cheng, X. X. (2024). Research and application of high-precision vision detection technology for new energy vehicle autonomous driving. Times Auto, (5), 104-106.

[16] Hyundai Motor Company, & Kia Corporation. (2022, November 8). Method and device for tracking object using LiDAR sensor and recording medium: CN202210468670.0 [Patent].