Research on Comparison of LiDAR and Camera in Autonomous Driving

Peide Wang
College of Engineering, Computer Engineering, Auburn University, Main Campus, Auburn, Alabama, AL 36849 US
*Corresponding author’s e-mail: pzw0020@auburn.edu

Abstract. With the improvement of vehicles automation, autonomous vehicles become one of the research hotspots. Key technologies of autonomous vehicles mainly include perception, decision-making, and control. Among them, the environmental perception system, which can convert the physical world’s information collection into digital signals, is the basis of the hardware architecture of autonomous vehicles. At present, there are two major schools in the field of environmental perception: camera which is dominated by computer vision and LiDAR. This paper analyzes and compares the two majors schools in the field of environmental perception and concludes that multi-sensor fusion is the solution for future autonomous driving.

1. Introduction
Autonomous driving is an advantageous industry in the field of intelligent travel. Among them, the perception module plays a vital role in the automatic driving system. The key technologies of autonomous vehicles mainly include the perception of the road environment, the planning of the driving path, the intelligent decision-making of vehicle motion behavior and, the realization of adaptive motion control of the vehicle. At present, the immaturity of the development of environmental perception technology is still the main reason hindering the improvement of the overall performance of autonomous vehicles, and it is also the biggest obstacle to the large-scale commercialization of autonomous vehicles.

The sensors currently used in the perception module of autonomous driving mainly include cameras, millimeter-wave radar, ultrasonic radar, and LiDAR. The camera has the advantages of high resolution, fast speed, rich information transmission, and low cost; LiDAR has the advantages of accurate 3D perception and is rich in information [1].

A variety of sensors are used in autonomous vehicles, and different types of sensors complement each other in function to improve the safety factor of the autonomous driving system. To take advantage of different sensors, fusion technology plays a key role.

This paper uses the method of literary analysis to analyze and compare the advantages and disadvantages of cameras and LiDAR, and discuss the impact of the fusion of the two sensors in the field of environmental perception.

2. Basic Principles of Autonomous System
As shown in Figure 1, the latest revision of SAE J3016, formulated by SAE International, provides a taxonomy that defines six levels of driving automation in detail, from level 0 to level 5 [2]:
However, there are still some gaps in the application of autonomous vehicles to practical and complex road traffic scenarios. Therefore, the environment perception system is an important part of the autonomous vehicle. The main task is to identify and classify road obstacles, traffic signs, signal lights, pedestrians and vehicles, etc., able to analyze and judge the location of the vehicle environment[3]. The Environment perception system is one of the key technologies of the autonomous vehicle. It is very important for the accurate understanding of traffic scene semantics and the corresponding behavior decisions of vehicles, and it is an important guarantee for driving safety and intelligence.

The autonomous system includes the decision layer, perception layer, control layer, as well as the support of high-precision maps and the Internet of Vehicles. Among them, various hardware sensors in the perception layer capture the position information of the vehicle and external environment information, and the existing on-board sensors include LiDAR, camera, millimeter-wave radar, etc. Based on the input information from the perception layer, the decision layer models the environment, forms a global understanding, makes decisions and judgments, and sends out the signal instructions for vehicle execution. Finally, the control layer converts the signal of the decision layer into the action behaviour of the vehicle [4].
2.1. LiDAR
The LiDAR can sense the surrounding environment information in real time and form high-definition 3D graphics. LiDAR works by emitting pulse modulated light, observing the time difference between the reflected light and the emitted light to calculate distance, and by scanning or synchronously measuring multiple beams of light to obtain angle information [5]. It has the advantages of fast response, long detection distance and high accuracy, so it is the sensor solution with the highest measurement accuracy among the known environmental measurement solutions.

During the course of the on-board LiDAR system, noise is inevitably included in the point cloud data obtained by the system due to factors such as acceleration, deceleration, change of driving direction, and so on. At present, most on-board LiDAR systems can provide the intensity information of the reflected pulse, which reflects the energy reflected from the target surface and is a reflection of the properties of different objects. But when the LiDAR systems, weather conditions or other specific conditions are different, the intensity information reflected by the laser point cloud will also be very different. Moreover, the on-board LiDAR point cloud data is a kind of blind data. The point cloud data is used as the spatial coordinate information of the surface of the object, and there is no information about the category to which the object belongs. This leads to great difficulties in the automatic recognition and feature extraction of objects such as buildings and human beings in the point cloud data.

LiDAR is too expensive to be used in mass production. There are two potential trends in the market: one is to reduce the cost of performance by producing LiDAR that can only accomplish a specific task with a limited detection distance; another is the use of solid-state LiDAR to reduce costs.

2.2. Camera
Cameras are a lower-cost solution for applications where both radar and cameras can be used. The installation position of the camera in front, side, rear, and built-in. Mainly used for forwarding collision warning system, lane departure warning system, traffic sign recognition system, parking assistant system, blind spot monitoring system.

Cameras work by illuminating an array of photosensitive cells with light from the outside, creating a photoelectric effect that generates a certain electrical charge. The row selection unit selects the corresponding row pixel unit, and the image signals in the row pixel unit are transmitted to the corresponding analog signal processing unit and digital-to-analog conversion circuit through the signal bus of their respective columns, which are converted into digital image signals and output [6]. Cameras are usually divided into the monocular cameras and binocular camera. In general, the perspective of the monocular vehicle camera is from 50° to 60°, and the visual distance is from 100m to 200m. The binocular camera can simulate human visual imaging for 3D imaging, compare the different image signals obtained by two cameras, identify the object more reliably, and obtain the distance and speed information of the object through the algorithm.

Compare with LiDAR, the cameras can use natural light during the day. It can recognize the color of cars and traffic lights and recognize objects far away with higher resolution and lower costs in full-light conditions. However, it also has some disadvantages. For example, it is easy to be affected by rain and snow weather and illumination, the change of illumination has a great impact on its recognition accuracy, and the current camera technology is difficult to identify distant objects in static images.
2.3. Comparison

| Performance aspect                  | Human | AV Radar | Lidar | Camera | CV DSRC | CAV CV+A* |
|-------------------------------------|-------|---------|-------|--------|--------|-----------|
| Object detection                    | Good  | Good    | Good  | Fair   | n/a    | Good      |
| Object classification               | Good  | Poor    | Fair  | Good   | n/a    | Good      |
| Distance estimation                 | Fair  | Good    | Fair  | Good   | Good   | Good      |
| Edge detection                      | Good  | Poor    | Good  | Good   | n/a    | Good      |
| Lane tracking                       | Good  | Poor    | Good  | n/a    | Good   |           |
| Visibility range                    | Good  | Good    | Fair  | Good   | Good   | Good      |
| Poor weather performance            | Fair  | Good    | Fair  | Poor   | Good   | Good      |
| Dark or low illumination performance| Poor  | Good    | Good  | Fair   | n/a    | Good      |
| Ability to communicate with other traffic and infrastructure | Poor | n/a | n/a | n/a | Good | Good |

Figure 3 Comparison of sensing capabilities of human drivers and sensors in highly automated vehicles

As can be seen from Figure 3, different sensors have different principles and functions, which can play their advantages in different use scenarios and are difficult to replace each other.[7] For example, if LiDAR is mounted on the roof of the car, then the underside of the vehicle is a circular area that cannot be measured. In the future, it is impossible to realize intelligent driving and autonomous driving only by relying on a single sensor. An important trend is to combine different sensors so that they can complement each other and ultimately improve accuracy. Today's driverless vehicles often get information by combining data from multiple sensors, such as LiDAR, cameras, and millimeter-wave radar.

3. Comparison between computer vision and LiDAR

There are currently two schools of thought for perceptual solutions in autonomous driving technology: computer vision (camera) school and LiDAR school. The computer vision school thinks LiDAR is unnecessary and too expensive; LiDAR advocates believe that the computer vision algorithm is insufficient in data form and precision. The following will analyze the two schools respectively.

3.1. Computer Vision

Humans can use their brains to estimate distances. They also have some gifts, such as stereoscopic vision, but only for estimating medium distances. Another gift is motion parallax. The movement direction and speed of objects seen when the line of sight moves laterally in the field of view are different, which also provides information for judging the distance. So, the human brain is perfectly up to the task, and in fact, it is possible to estimate distance while driving with one eye open. Now, researchers are trying to use neural networks to build machine learning techniques to estimate the distance from an image.

A technical paper from Cornell University proposes a new approach to closing the performance gap between pure computer vision technology architectures and LiDAR [8]. By changing the 3D information presentation form of the target detection system of the stereo camera, this paper converts the image-based stereo vision data into 3D point clouds similar to those generated by LiDAR, and then converts the data into the final view format. Using a relatively inexpensive camera on either side of the windshield, the new method has brought the performance of the camera close to that of LiDAR for object detection. On the KITTI standard dataset, the detection accuracy of the proposed method is improved from 22% to 74% in the range of 30m.

The study suggests that it is possible to use stereoscopic cameras in autonomous vehicles, which could significantly reduce costs and improve safety. But it will be a long time before cameras replace LiDAR.
3.2. LiDAR
Compared with the camera, LiDAR has a powerful three-dimensional spatial resolution ability. At present, in the field of autonomous driving and assisted driving, LiDAR is one of the core sensors to realize environmental perception. Among the sensors used for road information detection, LiDAR has certain advantages compared with millimeter-wave radar in detection range and accuracy. Its most important characteristic is the accuracy of distance measurement and high angular resolution. But high cost is almost the biggest barrier to large-scale use of on-board LiDAR. In addition, the detection beam of on-board LiDAR is easily affected by atmospheric absorption, scattering and refraction [9].

While micro-scanning LiDAR is currently the dominant type in the automotive industry, the solid-state LiDAR offers more reliable performance and compact size, which is critical for integrating equipment into the exterior of a vehicle. In contrast, although the field of view of solid-state LiDAR is narrower, manufacturers can also choose to install several solid-state LiDAR at the same time to expand the field of view by means of relatively lower costs [10].

3.3. LiDAR and Computer Vision
The core decision of the computer vision school is to bet that artificial intelligence technology will enable vehicles to achieve the same perception as the human eyes and the same thinking ability as human brain through visual perception and cognitive decisions made by optical cameras before LiDAR gets down to near the price of millimeter-wave radar. However, under the premise that existing artificial intelligence cannot reach human intelligence, a sensor must be used to make up for the limited range perception of passive optical imaging and solve the problem of uncertain detection direction of the millimeter-wave radar. By comparison, LiDAR can better solve the problem of distance measurement, and improve the safety of autonomous vehicles on the road under the existing technical conditions. This cannot be achieved by computer vision.

Although LiDAR has the characteristics of high measurement accuracy, rich information collection and good real-time performance, it is still unable to complete all the tasks of target detection and recognition, environmental modeling and so on in the complex and changeable autonomous driving environment. Therefore, the information fusion with multi-sensor is very important.

Many sensors are mounted on the same vehicle and use the same system to collect and process data. In order to standardize them, it is necessary to unify the coordinate system and clock for these sensors. The purpose is to make the same target appear at the same world coordinate of different types of sensors at the same time [8]. A Unified clocks are timestamps that synchronize different sensors. Unified coordinate system has two steps, one is motion compensation, the other is sensor calibration. Taking camera and LiDAR as an example, LiDAR data contains explicit 3D position information. Through the calibration parameters and the internal parameters of the camera itself, multi-sensor fusion can project 3D points on the image, mark pixels of the image with depth information and help the perception system based on image segmentation or practice deep learning model.

Autonomous driving is by no means a single field of technology, but the result of the integration of many technologies. Only by constructing the perception system of multi-source information fusion and making full use of multi-source information fusion, can reliable and effective surrounding environment information be established during the driving process of autonomous driving vehicles.

With the continuous development of technology, using deep learning network models to train LiDAR point cloud data and visual information, fusion of different modal data characteristics, and real-time monitoring and recognition of targets in the surrounding environment of autonomous vehicles have achieved a series of results [8].

With the continuous optimization of the price and precision of LiDAR and the continuous improvement of computer vision technology, the two complement each other and make up for the shortcomings of each other, which is the only solution to realize automatic driving in the future.
4. Conclusion

Autonomous vehicles are the development trend in the future, and environment perception system is one of the important hardware architecture foundations of autonomous vehicles. Among many environmental perception sensors, LiDAR has the highest measurement accuracy but high cost. The measurement accuracy of the camera is low, but the cost is low. LiDAR is developing towards miniaturization and integration, while the camera is the absolute mainstream vision sensor.

In order to ensure safe driving and the widespread popularization of autonomous vehicles, using effective multi-sensor information fusion technology and integrating multi-sensor information such as LiDAR and cameras can improve the accuracy of environmental detection and identification of autonomous vehicles in complex environments.

However, the use of multiple sensors will greatly increase the amount of information that needs to be processed, and there may even be conflicting information. So, it is very critical to ensure that the system quickly processes data, filter useless and wrong information, and makes timely and correct decisions. It can be seen from the above analysis that multi-sensor fusion is not difficult to achieve at the hardware level, but it is more complex in algorithm. Multi-sensor fusion software and hardware are difficult to separate, so it will be a difficult task in the future.

References

[1] X. Li, C. Liu, J. Li, M. Baghdadi, and Y. Liu, “A Multi-Sensor Environmental Perception System for an Automatic Electric Shovel Platform,” Sensors, vol. 21, no. 13, p. 4355, Jun. 2021.

[2] “SAE J3016 automated-driving graphic.” [Online]. Available: https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic. [Accessed: 14-Jul-2021].

[3] N. M. Tahir, U. I. Bature, M. A. Baba, K. A. Abubakar, and S. M. Yarima, “Image Recognition Based Autonomous Driving: A Deep Learning Approach,” Int. J. Eng. Manuf., vol. 10, no. 6, pp. 11–19, Dec. 2020.

[4] Y. W. S. P. Luo, “Research on Autonomous Driving Perception Test Based on Adversarial Examples.” 2020.

[5] A. Carteni, “The acceptability value of autonomous vehicles: A quantitative analysis of the willingness to pay for shared autonomous vehicles (SAVs) mobility services,” Transp. Res. Interdiscip. Perspect., vol. 8, 2020.

[6] K. Geng, G. Dong, G. Yin, and J. Hu, “Deep dual-modal traffic objects instance segmentation method using camera and lidar data for autonomous driving,” Remote Sens., vol. 12, no. 20, pp. 1–22, Oct. 2020.

[7] Y. Wang, W. L. Chao, Di. Garg, B. Hariharan, M. Campbell, and K. Q. Weinberger, “Pseudo-lidar from visual depth estimation: Bridging the gap in 3D object detection for autonomous driving,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2019-June, pp. 8437–8445, 2019.

[8] Y. Li et al., “Deep Learning for LiDAR Point Clouds in Autonomous Driving: A Review,” IEEE Trans. Neural Networks Learn. Syst., pp. 1–21, 2020.

[9] W. Wei, B. Shirinzadeh, R. Nowell, M. Ghafarian, M. M. A. Ammar, and T. Shen, “Enhancing solid state lidar mapping with a 2d spinning lidar in urban scenario slam on ground vehicles,” Sensors, vol. 21, no. 5, pp. 1–18, 2021.

[10] H. Bae, G. Lee, J. Yang, G. Shin, G. Choi, and Y. Lim, “Estimation of the closest in-path vehicle by low-channel lidar and camera sensor fusion for autonomous vehicles,” Sensors, vol. 21, no. 9, 2021.