Applications of Visual Perception in Autonomous Driving

Zhiyu Xi*

College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing, 211800, China

* Corresponding Author Email: 202221110063@njtech.edu.cn

Abstract. This paper focuses on the application of visual perception in autonomous driving. The types of cameras in visual perception systems are introduced, including charge-coupled device (CCD), and Complementary metal oxide semiconductor (CMOS) image sensors, and the LiDAR system with vision sensors. The applications of visual perception in autonomous driving are discussed, such as environmental perception (including road recognition, lane line detection, and traffic sign and signal light recognition) and vehicle positioning. However, visual perception technology faces challenges, such as the effects of light changes and interference from harsh weather on visual sensors. To address these challenges, solutions include the fusion of vision and other sensors, as well as progress in hardware and software technology. By optimizing and improving visual sensors in terms of precision, cost, power consumption, stability, and data perception and processing ability, the performance of autonomous driving systems can be enhanced. The multi-sensor fusion method utilizes the advantages of different sensors to obtain more comprehensive environmental information and improve the reliability and adaptability of the vehicle system. This paper can help autonomous driving researchers understand the application of visual sensors in autonomous driving more quickly.

Keywords: Visual perception, Autonomous driving, Environmental perception.

1. Introduction

With the advancement of society, automobiles have become an integral part of people's daily lives. In 2018 alone, the cumulative domestic automobile sales in China surpassed 30 million units, and the total number of automobiles in possession exceeded 320 million [1]. As the number of automobiles in possession continues to grow, the frequency of traffic accidents is also on the rise. A series of traffic safety issues caused by automobiles pose a serious threat to people's lives and property safety. Currently, many automobile enterprises both at home and abroad have developed various safety auxiliary systems, such as fatigue warning systems, emergency braking systems, and assisted driving systems. The integration of these auxiliary systems has enhanced the safety of driving. For instance, with the addition of assisted driving, the average fatigue level of truck drivers is lower than that in manual driving, and drivers in assisted driving maintain better task response capabilities during long-distance journeys [2]. Although vehicles are currently highly intelligent, people still hope to develop vehicles with fully autonomous driving capabilities to improve the safety of travel. To achieve full autonomous driving, autonomous vehicles need to continuously perceive the surrounding environment through all the sensors on the vehicle to acquire information, and then plan a safe and reliable driving route through information processing. At present, driverless cars can obtain environmental information through lidar, millimeter-wave radar, and cameras [3]. A typical autonomous driving system can generally be divided into several subsystems such as environmental perception, path planning, behavior decision-making, and motion control [4]. Among these subsystems, environmental perception, as the initial link for autonomous driving vehicles to interact with the external environment, its performance in terms of accuracy, speed, and stability has a significant impact on the overall working performance of autonomous driving [4] . Compared with radar, visual sensors possess advantages such as a smaller size, lower energy consumption, and a relatively lower price. Consequently, visual sensors are widely utilized in autonomous vehicles. The research presented in this paper primarily focuses on the application and development prospects of visual sensors in autonomous driving.

2. Visual perception technology

2.1. Types of cameras in visual perception systems

2.1.1. CCD image sensor

The first CCD was invented in Bell Labs in the United States in 1970 [5]. Common CCDs are divided into linear type and planar type according to the arrangement shape of pixels. The planar type is mainly used in advanced digital [5]. When the optical signal shines on the CCD sensor, the CCD sensor uses the photoelectric effect to convert the optical signal into a charge signal by the drive circuit [5]. After these signals are collected, they are converted into analog signals for output [5].

2.2.2. CMOS image sensor

CMOS was created at the Jet Propulsion Laboratory of the National Aeronautics and Space Administration CMOS image sensors have higher integration than CCD image sensors and the power consumption is only about 1/10 of that of ordinary CCD image sensors [6]. CMOS has the advantages of low manufacturing costs, a simple structure, a high yield rate, and a favorable price. At the present stage, CMOS image sensors mainly include active pixel sensors and digital pixel sensors [6].

2.2. Comparison Between the LiDAR System and Vision Sensor

The LiDAR system is furnished with a narrow-band laser and a receiving system. Its working principle is to emit light pulses and once an object is detected, the light is reflected and then received and processed by the receiver to obtain information data [7].

The LIDAR can provide accurate distance information with few inaccuracies [7]. Although the accuracy of the information such as the distance and speed of an object obtained by the LiDAR is particularly high, it can only receive the surface characteristics of the object. In comparison, with the assistance of algorithms, visual sensors can more effectively recognize features such as the color, texture, and shape of obstacles, thereby acquiring more extensive and comprehensive information. However, a single visual sensor is incapable of accomplishing depth detection, and it is requisite for two or more visual sensors to collaborate to conduct distance measurements.

The cost of the camera is lower than that of the lidar.

Under harsh lighting and weather conditions, LiDAR is more reliable than visual sensors.

3. Applications of visual perception in autonomous driving

3.1. Environmental Perception

3.1.1. Road recognition and lane line detection

Visual lane recognition is of great importance in autonomous driving. The road recognition method for intelligent vehicles mainly includes the detection of the vanishing point in road images, the road segmentation by integrating multiple visual features, and the 3D reconstruction of the road scene and obstacle detection using binocular vision [8-9].

For the detection of the vanishing point in road images, using a four-direction Gabor filter to extract the direction of road texture and reduce the complexity of texture calculation. By using confidence to filter effective pixel points to eliminate the influence of weak texture pixels on the results [8]. The basic vanishing point detection of road images is shown in Fig. 1. By extracting the main direction of road texture and combining voting accumulation, the initial position of the vanishing point is obtained [8].

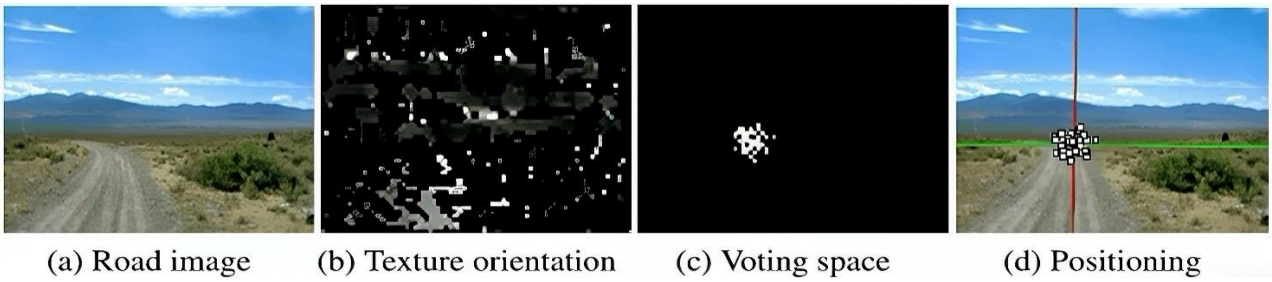

Fig. 1 Basic steps of vanishing point detection [8]

In the aspect of road segmentation, based on the linear road model constrained by the vanishing point, by integrating the three visual features of road texture, road surface color, and non-road surface color, the road segmentation is converted into the estimation of Bayesian posterior probability density [8]. Specifically, the RGB image is converted to the HSI color space, and an algorithm is used to compare the pixel data with the basic road surface pixels to obtain the basic model[8]. The computational efficiency is enhanced through integral image acceleration and GPU parallel acceleration techniques [8].

By introducing the adaptive image processing algorithm, different processing methods are adopted according to the images of different brightness [3]. For high-brightness images, discrete wavelet transform is used for processing, and for low-brightness images, the retina algorithm is employed[3]. Through this approach, the brightness of the dark areas is increased, the brightness of the bright areas is reduced, and more image information is obtained [3].

3.1.2. Detection, segmentation, and recognition of traffic signs

Using deep learning algorithms for traffic sign recognition has become the main method for traffic sign recognition [9]. The visual sensor continuously takes pictures of the road conditions. The image is processed through operations such as denoising, enhancement, and grayscale conversion to improve the features of the traffic signs [9]. For example, the image of the visual sensor is enhanced, the RGB color space is transformed into the HSI color space, and the traffic signs are segmented, processed, compared and analyzed according to the algorithm [9]. The detection result is shown in Fig. 2. Deep learning algorithms are used to extract the features of the image, and a model trained with a large amount of data is used to judge whether there is a traffic sign and determine its type. Finally, the recognition result is output.

Fig. 2 Detection and segmentation of traffic signs [9].

3.1.3. Target Recognition

By introducing the 3D object detection algorithm, the visual sensor can also obtain three-dimensional target information [4]. By adding sequential feature association and depth cues to the 3D object detection algorithm and through neural network training, the monocular 3D object detection algorithm based on sequential association and depth cues can complete the recognition task very well, and its operating efficiency is better than that of the traditional object detection algorithm [4]. The effect is shown in Fig. 3.

Fig. 3 Examples of detections [4].

3.2. Vehicle positioning

In an enclosed environment, the feature map of the environment is saved inside the driverless vehicle and the key feature information is recorded [10]. During the driving process, the environmental image information is continuously obtained through the visual sensor for processing, and the image features are extracted and matched with the features in the offline map [10]. Positioning is achieved through matching the feature information [10].

4. Challenges faced by visual perception technology

4.1. Effects of Light Changes

Under different times and weather conditions, the changing illumination will affect the clarity of the image. If the sunlight shines directly on the visual sensor, it will cause a strong glare and affect the result of the visual perception.

In the nighttime or in environments with low light, the performance will decline significantly. Even with lighting equipment, the perception performance is limited.

4.2. Interference of Harsh Weather on Visual Sensors

Adverse weather conditions such as rain, snow, and fog can have a strong impact on visual perception. The raindrop on the camera can blur the image; the dense fog will reduce visibility and make it difficult to identify objects.

5. Solutions and Future Development Directions

5.1. Advantages of Fusion of Vision and Other Sensors

The collaborative work of multiple vision sensors can obtain more information. Compared with a single vision sensor, multiple vision sensors can greatly improve reliability. Sensor fusion plays a significant role in enhancing overall stability. For example, under extreme weather conditions, Global Positioning System (GPS), LiDAR, and millimeter-wave radar can effectively compensate for the performance shortcomings of visual sensors in such environments. Different types of sensors provide various types of information. By fusing the data of these sensors, vehicles can obtain more comprehensive environmental information, providing strong support for achieving accurate navigation and decision-making. This multi-sensor fusion method fully utilizes the advantages of different sensors and overcomes the limitations of a single sensor in specific environments, thus greatly improving the reliability and adaptability of the vehicle system.

5.2. Progress in Hardware and Software Technology

In the field of autonomous driving, visual sensors play a crucial role. To further enhance the performance of autonomous driving systems, it is necessary to optimize and improve visual sensors in multiple aspects.

Enhancing the precision of visual sensors is of utmost importance. By continuously exploring advanced technical means and design concepts, this research aims to improve the sensor's ability to capture environmental information around it, so that it can more accurately obtain the position, shape, color, and other characteristics of target objects. At the same time, vigorously researching and developing more advanced image processing algorithms and making full use of the latest research results in the fields of computer vision and pattern recognition to conduct in-depth analysis and processing of the image data collected by visual sensors, thereby significantly improving the ability of visual sensors to process information and providing more accurate and reliable environmental perception data for autonomous driving systems.

Optimizing the cost and power consumption of visual sensors is one of the key factors to realize the wide application of autonomous driving technology. On the premise of ensuring sensor performance, more efficient manufacturing processes and materials are adopted to reduce production costs. At the same time, by optimizing the sensor's circuit design and signal processing algorithms, power consumption is reduced and energy utilization efficiency is improved to meet the strict requirements of autonomous driving vehicles for energy management.

Using high-precision manufacturing equipment and a strict quality control system to ensure that the sensor can maintain stable performance in various complex working environments. In addition, through reliability testing and life assessment of the sensor, continuously improving the manufacturing process and materials to extend the service life of the sensor and reduce maintenance costs.

Improving the data perception and processing ability of the sensor is one of the core tasks to improve the performance of autonomous driving systems. Using advanced sensor technology and data fusion algorithms to achieve efficient fusion and processing of multi-source data and improve the sensor's perception range and precision of environmental information. At the same time, aiming at the autonomous driving requirements in extreme environments optimizes the sensor's recognition ability under harsh weather, low light, and other conditions. Adopting special optical materials and enhancing signal processing algorithms and other means, ensure that autonomous driving vehicles can drive safely and reliably under various extreme environments.

6. Conclusion

In summary, visual perception constitutes a crucial element in the realm of autonomous driving. Various types of cameras, namely CCD and CMOS image sensors, are employed in visual perception systems. Although the LiDAR system offers precise distance information, visual sensors, aided by algorithms, can more effectively recognize the features of obstacles.

Visual perception finds extensive application in multiple aspects of autonomous driving, encompassing environmental perception and vehicle positioning. Nevertheless, it encounters challenges, such as the influence of light changes and the interference from adverse weather conditions.

To surmount these challenges, the fusion of vision with other sensors proves advantageous, as it enables the acquisition of more comprehensive environmental information and enhances overall stability. Furthermore, advancements in hardware and software technology are imperative, including the enhancement of the precision, cost-effectiveness, power consumption, stability, and data perception and processing capabilities of visual sensors.

In general, the development of visual perception technology in autonomous driving necessitates continuous research and innovation to elevate its performance and reliability, thereby ensuring the safe and efficient operation of autonomous vehicles. This article shows the application and important position of visual sensors in autonomous driving and help researchers feel the application scenarios of visual sensors more intuitively.

References

[1] Zhang Xu. Research on the method of traffic sign recognition based on machine learning. Shaanxi: Xi'an University of Technology, 2019.

[2] Huang Chunxi, Xie Weiyin, Huang Qihao, etc. Effect of advanced driver assistance systems on fatigue levels of heavy truck drivers in prolonged driving tasks. Journal of Tongji University (Natural Science),2024,52(6):846-855.

[3] Zhang Wenhai. Research on obstacle detection and road recognition of intelligent vehicles based on vision. Chongqing: Chongqing Jiaotong University, 2022.

[4] Gao Tianze. Multi-dimensional object detection and tracking in autonomous driving based on monocular vision. Heilongjiang: Harbin Institute of Technology, 2022.

[5] Sun Li, Deng Min. CCD The current situation and future development of CCD image sensor, Modern Manufacturing Technology and Equipment 2017, (08): 160-161.

[6] Ni Jinghua, Huang Qiyu. CMOS image sensor and its development trend. OME Information, 2008, 25(5): 33-38.

[7] Luo Qian, Jin Qijun. Current situation and development analysis of intelligent vehicle lidar sensing technology. Science and Technology & Innovation,2020, (17): 4-6+9.

[8] Shi Jinjin. Research on vision based road recognition and obstacle detection for intelligence vehicles. Heilongjiang: Harbin Institute of Technology, 2016.

[9] Yang Xueqi. Research on lane line detection and traffic sign recognition of outdoor intelligent vehicle based on machine vision. Jilin: Changchun University of Technology, 2022.

[10] Yan Xiaoying. Application of robot vision navigation and path planning in automated workshop. Papermaking Equipment & Materials,2024,53(7):83-85.