Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/6/20230885

Research on environment parameter perception of autonomous vehicle based on multi-vision machine vision

Youwen Zhang

Beijing University of Chemical Technology, Chaoyang District, Beijing, 100029, China

zywty@foxmail.com

Abstract. The current environmental parameter perception method of some autonomous vehicles has the problem that the MOTA (Multiple Object Tracking Accuracy, This measure combines three error sources false positives, missed targets and identify switches) value is too low. A multi-vision machine vision based environmental parameter perception method of autonomous vehicles is designed. The expected distance between the vehicle and the vehicle in front was calculated. The driving data was collected based on multi-eye machine vision, and the target motion process was divided into several infinitesimal sampling cycles. The data preprocessing process of autonomous vehicle was optimized, the relationship between reaction distance and reaction time was described, and the environmental parameter perception method was designed. Experimental results show that the MOTA mean value of the environmental parameter perception method of the autonomous vehicle in this paper is 84.63%, indicating that the environmental parameter perception method of the autonomous vehicle designed with the combination of multi-vision machine vision technology has high accuracy.

Keywords: multivision machine vision, self-driving cars, environmental parameter perception, intelligent driving, driving data, the range of visual field.

1. Introduction

Traditional road traffic systems are difficult to achieve intelligent control and cannot make automatic decisions due to corresponding road traffic conditions. Multi-eye machine vision applies sensing technology to obtain vehicle location and vehicle information for adaptive control of traffic lights at intersections, thus being able to deal with serious asymmetries in queuing traffic flow in different directions. Intelligent transportation technology can not only control the network signal of a single signal light, but also control the signal and induce traffic on the regional road network globally to achieve a balanced traffic flow in the region. The self-driving vehicles in urban roads, intercity expressways and highway environment, which are reflected in the road operation law has a great difference, mainly in the self-care speed, acceleration, the relative distance and relative speed between the self-car and the vehicle in front and the horizontal and vertical speed and other quantitative indicators, different driving behavior on the safety of other vehicles on the road and driving trajectory will produce significant differences. For example, when changing lanes, both the vehicle in front of the current lane and the vehicle in the target lane will be affected. Automated vehicles have started to develop from L1 and L2 to L3, and the number and types of sensors on the vehicle are increasing, and the internal control procedures are becoming more and more complex. Self-driving vehicles have to be able to adapt to harsh weather, changing driving environments and complex driving tasks, which also put higher demands on the environmental parameter sensing methods. The vehicle is in a mode where its relative speed to the vehicle in front of it will remain floating in a certain stable range, which is passed to the vehicles behind the fleet to form a non-free flow of traffic. From the perspective of pattern recognition, establishing a driving behavior pattern recognition model can provide a theoretical basis for autonomous driving system learning, which enables the autonomous driving system to accurately identify the current driving behavior pattern of the vehicle and thus achieve timely risk warnings.

2. Design of environmental parameter sensing method for self-driving cars

2.1. Multi-eye machine vision based driving data acquisition

Lane markings are the most direct basis for characterizing lane information and potential hazard targets, but also to identify the vehicle trajectory reference line. In the lane change and lane keep driving state, the vehicle in the lane of the transverse displacement has significant variability, left and right lane change when the vehicle transverse displacement changes also have obvious differences, the left and right lane line distance are expressed as 11d, rld. follow the gallop and free driving process lane line distance Always keep in the left lane line on both sides (left line distance to the left lane line as the baseline, the wheel is positive when located on its left side, the right side is negative. Similarly, the right line distance to the right lane line as the baseline, the left side of the positive value, and the right side of the negative value), without a large lateral offset. Multi-eye machine vision mainly uses computers to simulate human visual functions, but is not just a simple extension of the human eye, but more importantly has part of the functions of the human brain one by one to extract information from the images of objective things, process and understand them, and finally use them for actual detection, measurement and control. Then the formula for calculating the relative distance of self-driving vehicles is:

$$\Delta L = \frac{(v-v')}{(\delta)} \tag{1}$$

Equation (1), v indicates the speed of the car, v' indicates the speed difference between the car and the car, δ indicates the two collision time. The mutual influence relationship between the self-vehicle and other vehicles on the road is mainly reflected in the relative distance and relative speed, considering that the process of changing lanes produces a strong correlation between the self-vehicle and the surrounding traffic environment, and when identifying its own driving mode, it mainly considers the instantaneous changes in the operating state of the self-vehicle, and has less correlation with factors other than the vehicle. However, by judging the distance between the self-car and the vehicle in front of the current lane, it can indirectly reflect the current maneuvering action taken by the autopilot, which may be stimulated by the slower driving speed of the vehicle in front of it to prompt the autopilot to generate the demand for lane change and thus take the lane change operation [1-2]. Then the equation for the expected following distance between the autopilot and the vehicle in front is:

$$\Delta L' = \frac{\Delta v' \times \Delta \varepsilon}{\varepsilon + R} \tag{2}$$

Formula (2), ε means the relative distance between the car and the car in front, which means the reserved safety distance. With the rapid development of vision sensing technology, computer technology and image processing technology, multi-vision machine vision technology has become increasingly mature and has become an indispensable core technology for modern processing manufacturing. Multi-vision machine vision technology is widely used in various aspects. From medical images to remote sensing images, from industrial monitoring to document processing, and from milli-micron technology to multimedia databases. Assuming that the image has a length of Ppixel point and a width of Q pixel points, the equation for expressing the coordinate system when the machine reads the image automatically is:

$$P' = P - \frac{1}{2}Q \tag{3}$$

$$Q' = Q - \frac{1}{2}P \tag{4}$$

In equations (3) and (4), P, Qrepresents the coordinate system established with the midpoint of the image as the coordinate origin, and represents the coordinate system of the image acquired by the multi-vision machine vision, respectively. It can be said that the occasions that human vision can be used almost with multilocular machine vision, and many occasions where human vision cannot be perceived, such as precise quantitative perception, hazardous scene perception, invisible light object perception, etc. The advantages of multi camera machine vision are more obvious, because there is no systematic classification scheme for scene division in the field of view in the current research situation of vehicle based monocular vision ranging [3-5]. In this paper, the author tries to define the partitioning of the images in the camera field of view, and segmentation will make the modeling research based on the region of interest in this paper faster for image processing. The automatic driving device confirms that it is safe to change lanes by observing the rearview mirror and the surrounding traffic environment, etc. During the driving process, the automatic driving system can effectively determine whether the scenery, vehicles and obstacles reflected in the field of view are dangerous to the vehicle or predict the occurrence of danger, the information that the automatic driving system obtains from the field of view. During normal driving, the auto drive system is within the field of vision, some of which are dangerous to driving safety to a certain extent. It can also play a role in the judgment and decision-making of the car driving system. When the speed of the front car is greater than the speed of the self-car, the relative speed is kept within a certain range, and its relative distance gradually increases as the relative speed increases. At this time, the automatic driving device has a weak demand for lane change, and may not perform lane change operation and keep the status quo.

2.2. Optimizing the data pre-processing process for self-driving cars

After completing the data acquisition step, the data pre-processing process needs to be further developed. When driving under high-speed conditions, the autopilot device has a large visual share of the acquisition of external environmental information, which affects the driving judgment and decision of the autopilot device and thus poses a significant risk to driving safety. Usually, the changing pattern of the visual characteristics of the autopilot device is closely related to the current driving behavior, so it is especially important to investigate the visual characteristics of the autopilot device under different driving behaviors [6-7]. In addition to LIDAR, sensors such as camera and millimeter wave radar are usually equipped with the method of sensing environmental parameters of self-driving cars, and the data output from these sensors are basically generated based on their own coordinate system. To facilitate data processing and algorithm research, the data information acquired by the sensors needs to be transformed into the coordinate system of the vehicle body. The metrics characterizing the visual properties are parameters such as the number of gazes, gaze duration, blink frequency, search breadth, sweep speed, and pupil area. Since the target model is a model function that characterizes the relationship between the state of the target detected by the sensor and time. Generally, the target in the intelligent vehicle environment sensing system can be detected by different sensors to obtain data information. In an environmental sensing system, the detected target is defined as$W$, which is a quantity that varies with time t. The mathematical expression of the data measured by the sensor and the control system quantity is given by:

$$H\big(W(t)\big)\Big|\frac{1}{W(t)-1}, \{S_t\} \tag{5}$$

In Equation (5), H denotes the data measured by the sensor and S denotes the control system quantity. Since the vehicle motion and driving environment parameters are collected at a frequency of 10Hz, the sensor device outputs corresponding to one data every 0.1s, while the visual parameters of the autonomous driving device are collected by the eye-tracking device at a frequency of 60Hz, corresponding to six data outputs every 0.1s. Selecting four of these consecutive frames as reference samples, the autonomous vehicle operation and traffic environment parameters are shown in Table 1:

Table 1. Operating and traffic environment parameters of self-driving vehicles (credit: original).

Parameters/frames	763544	763545	763546	763547
Left line distance /cm	-42	-42	-42	-42
Right line distance /cm	124	124	124	124
Vehicle speed /km. h-1	82.76	82.64	82.55	82.49
Relative Distance /m	148.4	148.7	148.8	149.1
Relative Speed /m2.s-1	2.87	2.93	2.97	3.01
Angular velocity of transverse pendulum /deg.s-1	0.24	-0.36	0.31	0.33
Lateral acceleration /m.s-2	-0.8776	-0.1334	0.1227	0.1443
Longitudinal acceleration / m.s-2	0.3367	0.4221	0.4568	0.4791

In addition, the multi-eye machine vision sensors compose raw statements of data from the detected surrounding obstacles from the sensors and transmit them to the fusion center for association, correlation, classification, tracking, and prediction processing. The advantage of sensor-level data pre-processing structure is that if each sensor uses independent different physical information sources to generate information, that is, each sensor recognizes objects with different mechanisms, and for the same object, the probability of generating false alarms is low. For the 3D point cloud data output from LiDAR, the reference coordinate system is a spherical coordinate system with the radar as the origin, which cannot be directly used for the algorithmic processing of obstacle detection. The point cloud data with the LIDAR as the coordinate origin is transformed into a spatial right-angle coordinate system with the vehicle body as the coordinate origin by coordinate transformation. Since the raw data obtained by each sensor is not pre-processed and transmitted directly to the fusion processor, the sensor-level data pre-processing structure requires high hardware requirements for the fusion processor, such as large storage capacity as well as fast processing speed and other performance, and the obtained raw data is concentrated in the fusion processor for processing, which will lead to an excessive load rate on the communication CAN bus, thus causing some The sensors lose obstacle information at some point, which corresponds to no obstacle being detected. The transformation relationship can be calculated in two steps [8]. The conventional point cloud filtering methods are well encapsulated in the PCL library, and users can choose to call various filters to complete the filtering processing of the point cloud. The process of target motion is divided into several infinitesimal sampling periods, during which the acceleration change is considered to be constant, and the current probability density is expressed by the modified Rayleigh distribution:

$$\mathrm{d}(\mu) = \sum \frac{1}{\varpi} + \bar{\varpi}(\mu) \tag{6}$$

In Formula (6), μ represents the position of the maneuvering target, and $\bar{\varpi}$ represents the mean value of the target acceleration. The central data pre-processing structure and the data pre-processing structure of the liquid level sensor are the biggest differences before entering the fusion processor. The central data pre-processing structure requires each sensor to perform the minimum data processing in advance. The pre-processing structure of the liquid level sensor data is not. The original data obtained by the sensor is directly transmitted to the fusion processor for data processing. Bilateral filter is filtered according to the intensity value of point cloud, which can retain the boundary feature information well, but it takes more time. Statistical filter specifies the threshold for the query point to determine whether the point is an outlier and trim the nonconforming points. The radius filter is similar to the statistical filter, which takes the query point as the radius to draw a circle, at the same time, the number of nearest neighbor points within the specified radius is given, and the qualified points are retained; otherwise, they are regarded as noise points and removed. In the actual process, different filters may have great differences in the degree of point cloud filtering, one can choose to use it according to the specific situation. In order to facilitate model identification and uniformly correspond to the time sequence of parameter data, visual related parameters were preprocessed, and the average value of 6 data corresponding to each 0.ls visual parameter was taken as the sample data of corresponding vehicle operating parameter 0.ls. The visual parameter data corresponding to vehicle operation and traffic environment parameters are shown in Table 4.1 and 4.2. For the processing of missing values, the median of corresponding continuous driving behavior segment is used to supplement.

2.3. Design environmental parameter awareness method

The automatic driving test scenario can be analyzed as a set of several parameters, among which there are certain constraint relations. Scene parameters describe a scene and store information about the weather, illumination, roads, traffic participants, and vehicles in the scene. After determining the parameters that constitute the scene, in order to avoid the subsequent construction of the scene that does not conform to common sense, the constraints between these parameters should be resolved. Four ultrasonic radars are installed on each side around the vehicle, which can detect the 2.5-meter distance in front of the ultrasonic radar to detect obstacles in front of the sensor and are responsible for the detection of obstacles on the side of the vehicle. It is also a supplement to the coverage area not detected by short-distance millimeter wave radar in front and rear. The deceleration and braking characteristics of vehicles are the key characteristics to ensure the safe operation of vehicles, and also have an impact on the braking distance of vehicles and the minimum safety distance of adjacent vehicles under continuous traffic flow. The relationship between vehicle braking distance and vehicle speed, lateral friction between vehicle and road, and road slope is shown in Formula (7):

$$g = \frac{E^2}{254(l \pm K)} \tag{7}$$

In Formula (7), E represents the running speed of the vehicle when braking is implemented, l represents the friction coefficient between the vehicle tire and the road, and K represents the slope of the road. Because the road condition parameters involved in scenes in different regions are quite different, a very large set of scene parameters will be obtained if analyzed together, which will bring great difficulties to the subsequent scene generation. At the same time, due to the different regions of the scene and the tasks to be performed, the constraint relations of different types of scenes will also change. Some constraints exist in one part of the scene, but not in another part of the scene. The multi-vision machine vision camera sensor can detect the obstacles in front of the autonomous vehicle and identify the front lane line [9-10]. Since the millimeter-wave radar sensor cannot classify the types of obstacles in front of us, for example, the obstacles in front of us are pedestrians, cars, bicycles and other objects, the camera sensor can identify and classify the obstacles by extracting their features, which also makes up for the deficiency of the millimeter-wave radar sensor. For example, in the front vehicle emergency braking scenario, the vehicle required to brake must be in front of the vehicle. However, in the tracking scenario, the object does not necessarily need to be in front of the car. The scene parameters can be divided into three categories: environment parameters, road parameters and object parameters. The interaction between the autonomous vehicle and the environment is shown in Figure 1:

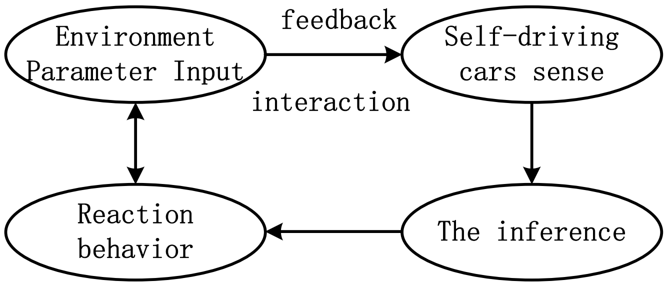

Figure 1. Schematic diagram of the interaction between an autonomous vehicle and the environment.

The self-driving car also needs to recognize the lane ahead to ensure the normal running of the smart car and avoid the self-driving car deviating from the road. The original data of obstacles detected by the camera includes the physical information such as obstacle type, longitudinal distance, longitudinal speed, longitudinal acceleration, transverse distance, transverse speed, width and height. Reaction distance and reaction time have the following relationship:

$$f = 0.278\frac{1}{T} \tag{8}$$

In Formula (8), T represents the speed of the vehicle before braking. The environment parameter describes the environment information in the scene, the object parameter records the information of all traffic participants in the scene, and the road condition parameter contains the information about the road in the scene. Lighting parameter describes the lighting condition in the scene. Since the difference of light source and light intensity will affect the optical sensor of autonomous vehicle, the value range of light parameter must include the distinction between light source and light intensity. The front lane line recognition is based on the front camera sensor, and CAN provide the current lane position and the current vehicle steering offline state. According to the alarm configured by the sensitivity, the driver may unconsciously deviate from the corresponding alarm prompt, and can output various information through the vehicle CAN bus. Considering the lighting condition in the real scene comprehensively, four values of lighting parameters are selected. Not only the distinction of sunlight and street lamps in the light source, but also the different light intensities of day and night. The continuous road is divided into several simple sections, and the restoration of the real scene road in the automatic driving simulation software is realized through the splicing of a variety of simple sections. The road in the scene is divided into several sections. When the road condition parameters of a section of road change, the road is divided into two sections. Iterate over the process until one can no longer partition.

3. The simulation results

3.1. The experiment to prepare

In addition to the LiDAR and vision sensor used to capture the environment information, the whole environment perception system also needs the support of other auxiliary equipment to complete the whole perception task. Keep the camera height fixed and the optical axis parallel to the ground plane, and the acquired image format is: RGB24 color image with a resolution of 1280 ×1024. The onboard battery provides a power supply for all sensors and equipment, and the inverter converts the voltage to the working voltage of the sensor to ensure the normal operation of all parts of the system. The CPU is equipped with Core I7 8700K CPU and TESLA V100 GPU, and 32GB running memory, which can

DOI: 10.54254/2755-2721/6/20230885

quickly process the information obtained by different sensors. The HDL-32E lidar, mounted on an autonomous driving platform, is capable of providing a 360° scan of the vehicle's surroundings while generating 700,000 data points per second. The experimental platform is Intel(R) Core(TM) i7-9700 CPU, 8-core 3.00GHz main frequency, 16G memory, NVIDIA GeForce GTX 1080 Ti GPU, 11 G capacity, CUD A is 11.1.96, Python is 3.7.6, Pytorch is 1.6.1, and JupyterNotebook is 6.0.3. In addition, since the ROS system is based on the Linux kernel operating system, it is necessary to install the Linux system and the corresponding ROS system before the development. There are many versions of Linux, among which the Ubuntu system is widely used for its simple user interface, perfect software source support and good compatibility with hardware devices. The software environment of the experiment in this paper is based on Ubuntu 16.04 system with the corresponding ROS Kinetic version installed and ROS operating environment preliminarily configured. The program development software uses RoboWare Studio.

3.2. The experimental results

In this experiment, multi-object tracking accuracy (MOTA) was used as an index to test the performance of the environmental parameter sensing method for autonomous vehicles. The calculation method of MOTA is:

$$\mathrm{MOTA} = \frac{\sum \varphi(m+n+r)}{\lambda} \tag{9}$$

In Formula (9), m represents the missing number of target tracking at time φ, n represents the number of misjudgments, r represents the number of mismatches, and λ represents the number of targets at time φ. The environmental parameter perception method of autonomous vehicle based on artificial intelligence and the environmental parameter perception method of autonomous vehicle based on data fusion are selected to compare with the environmental parameter perception method of autonomous vehicle in this paper. Test the MOTA values of the three environmental parameter perception methods of autonomous vehicles respectively under the condition of different obstacle distances. The experimental results are shown in Figure 2-4:

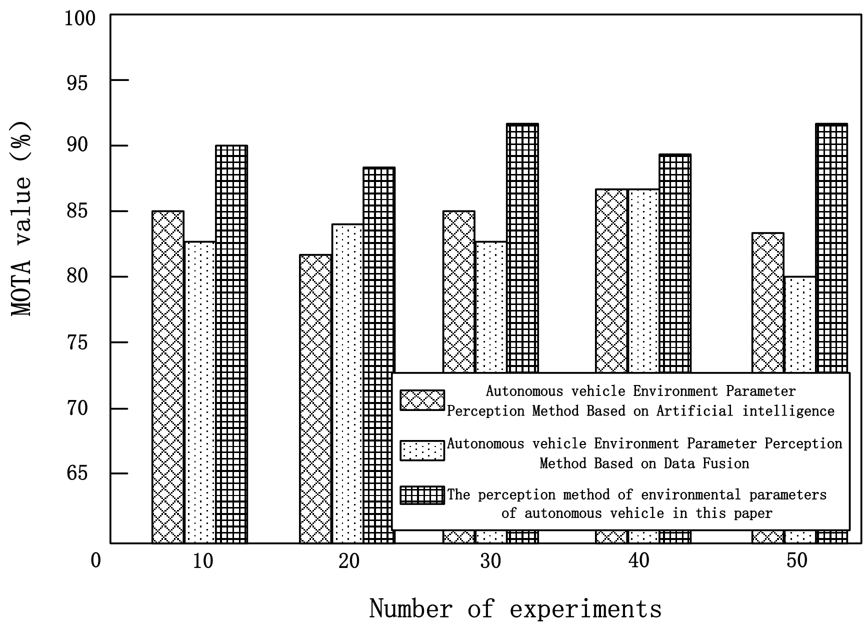

Figure 2. Obstacle distance 100mMOTA value (%).

DOI: 10.54254/2755-2721/6/20230885

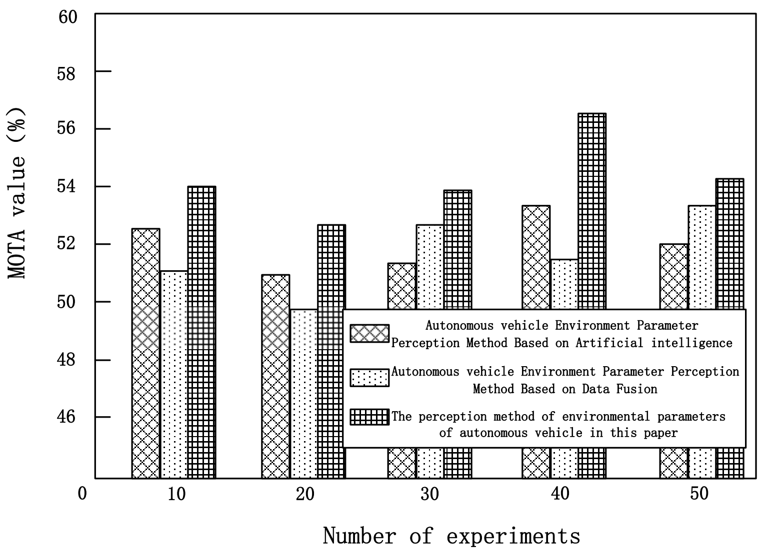

Figure 3. Obstacle distance 300mMOTA value (%).

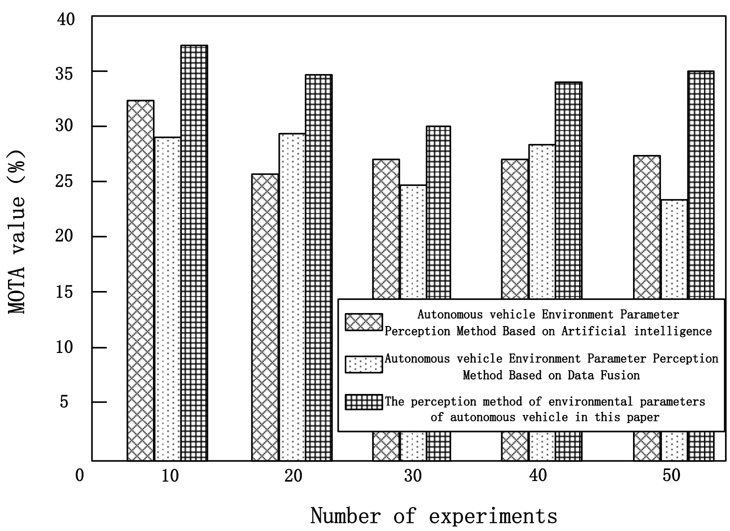

Figure 4. Obstacle distance 500mMOTA value (%).

According to Figure 2-4, there is an inverse relationship between the distance between the vehicle and the obstacle and the MOTA value, that is, the farther away the obstacle is, the lower the MOTA value will be. The mean MOTA of the environmental parameter perception method of the autonomous vehicle in this paper is: 84.63%; the mean MOTA of the artificial intelligence-based autonomous vehicle environmental parameter perception method is: 79.37%; the mean MOTA of the automatic vehicle environmental parameter perception method based on data fusion is: 77.62%.

4. Conclusion

In this paper, after collecting the relevant data of the autonomous vehicle, the corresponding data preprocessing steps are designed, and the obstacle with the closest longitudinal distance in the output original data is taken as the target object, and the selected target object is statically tracked through the time-domain filtering algorithm. The field of view of the vehicle camera is divided into several areas, and only the directly measurable areas useful for driving safety are extracted for analysis and processing after zoning, which greatly reduces the image processing time. Meanwhile, with the help of multi-vision technology, multi-sensor data fusion algorithm is used to improve the detection probability and data accuracy of the front target. In the future, real-time speed measurement should be carried out in the follow-up research, that is, to obtain the inter-frame time difference and inter-frame distance difference of the video, so as to improve the performance of the environmental parameter perception method of the autonomous vehicle in this paper.

References

[1] Custódio T, Alves C, Silva P, et al. A Change of Paradigm for the Design and Reliability Testing of Touch-Based Cabin Controls on the Seats of Self-Driving Cars[J]. Electronics,2021,11(1): 21.
[2] Surbhi, H.D. A Anjali N, et al.. Reliability Assessment of the Planning and Perception Software Competencies of Self-Driving Cars[J]. International Journal of Performability Engineering,2021,17(9): 779-786.
[3] Jiang Tao, ZHANG Guilin, GAO Junpeng. Illumination of a cylinder block transverse hole for machine vision inspection [J]. Chinese Journal of Optics,2020,13(6):1285-1292.
[4] Peng Cong, Liu Bin, ZHOU Gan. Mechanical Fault Detection Based on Machine Vision and Blind Source Separation [J]. Journal of Shanghai Jiaotong University,2020,54(9):953-960.
[5] CAI Wenlong, Zhao Zhen, Li Wenzhong. Positioning of Welding Cup of Aviation Plug Based on Machine Vision [J]. Computer Simulation,2022,39(6):53-56.
[6] Zhang Yihuan, Wang Liang, Jiang Xuhui, et al. An efficient LiDAR-based localization method for self-driving cars in dynamic environments[J]. Robotica,2021,40(1): 38-55.
[7] Justyna M, Pontus L, Johan F. Intuitive and subtle motion-anticipatory auditory cues reduce motion sickness in self-driving cars[J]. International Journal of Human Factors and Ergonomics,2021,8(4): 370-392.
[8] Gillmore C., Tenhundfeld L.. The Good, The Bad, and The Ugly: Evaluating Tesla's Human Factors in the Wild West of Self-Driving Cars[J]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting,2020,64(1): 67-71.
[9] Abhilash P. M., Chakradhar D.. Machine-vision-based electrode wear analysis for closed loop wire EDM process control[J]. Advances in Manufacturing, 2022, 10(1):12.
[10] Zhu Yun, Ling Zhigang, Zhang Yuqiang. Research progress and prospect of machine vision technology [J]. Journal of Graphics,2020,41(6):871-890.