



Technology and application of intelligent driving based on visual perception

ISSN 2468-2322

Received on 6th August 2017

Revised on 18th September 2017

Accepted on 23rd September 2017

doi: 10.1049/trit.2017.0015

www.ietdl.org

Xinyu Zhang¹, Hongbo Gao² ✉, Guotao Xie², Buyun Gao^{1,3}, Deyi Li⁴

¹Information Technology Center, Tsinghua University, Beijing 100083, People's Republic of China

²State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100083, People's Republic of China

³School of Software, Beijing Institute of Technology, Beijing 100081, People's Republic of China

⁴Institute of Electronic Engineering of China, Beijing 100039, People's Republic of China

✉ E-mail: ghb48@mail.tsinghua.edu.cn

Abstract: The camera is one of the important sensors to realise the intelligent driving environment. It can realise lane detection and tracking, obstacle detection, traffic sign detection, identification and discrimination and visual simultaneous localisation and mapping. The visual sensor model, quantity and installation location are different on different intelligent driving hardware experimental platform as well as the visual sensor information processing module, thus a number of intelligent driving system software modules and interfaces are different. In this study, the software architecture of the autonomous vehicle based on the driving brain is used to adapt to different types of visual sensors. The target segment is extracted by the image segmentation algorithm, and then the segmentation of the region of interest is carried out. According to the input feature calculation results, the obstacle search is done in the second segmentation region, the output of the accessible road area. As driving information is complete, the authors will increase or reduce one or more visual sensors, change the visual sensor model or installation location, which will no longer directly affect the intelligent driving decision, they make the multi-vision sensors adapted to the requirements of different intelligent driving hardware test platforms.

1 Introduction

As early as the 1950s, the USA carried out autonomous vehicle research, in 1950, the United States Barrett Electronics developed the world's first independent navigation vehicle [1]. US Autonomous Vehicle Research originated in the US Department of DARPA (Defense Advanced Research Projects Agency), its research level was the world leader. European countries since the 1980s began to develop autonomous driving technology, made autonomous vehicles an independent individual, and mixed vehicles in the normal traffic flow [2]. In 1987, the Munich Federal Defense Force University, Daimler Benz, BMW, Peugeot, Jaguar and other well-known R & D institutions and automotive companies participated in the implementation of the Prometheus plan (PROMETHEUS, Program for a European Traffic of Highest Efficiency and Unprecedented Safety), which had a significant impact around the world [3]. Since the 1990s, the Advanced Transportation Engine System Association (AHSRA) set up an Advanced Safety Vehicle project every five years to carry out autonomous driving technology research [4]. Chinese autonomous technology research began in the late 80s of last century, supported by the national 'eight six three' plan and the National Defense Science and Technology Commission related research program [5]. Since 2008, with the support of the National Natural Science Foundation of China, China started the future challenge of smart cars, the team number increased year by year, the difficulty of the game increased year by year, the number of completion racing team increased year by year, car companies' enthusiasm in participation was gradually strengthened, which laid a solid foundation of the autonomous technology introduction in China and the rapid progress in autonomous driving technology [6].

Intelligent driving perception technology is mainly through the sensor for autonomous vehicle's internal and external environment information perception. Sensors include visual sensors, radar sensors, navigation and positioning sensors. Among them, the visual sensor is mainly used for lane detection, lane tracking,

obstacle detection, traffic signs identification, including traffic signs detection, traffic signs identification, traffic light recognition and visual simultaneous localisation and mapping (SLAM) technology.

Zhang *et al.* [7] proposed a method of locally improving global satellite navigation system positioning accuracy by using binocular vision using landmark information. Ziegler *et al.* [8] proposed a method of detecting lane lines using visual techniques. Wolcott and Eustice [9] proposed a map-based visual positioning method. Duan [10] did some research on front-end composition, closed-loop detection and back-end optimisation of SLAM based on the graph. Fang [11] proposed the use of invariant moment theory, the occurrence of tilt, rotation and scaling of the invariant nature of the image of the study. Fernando *et al.* proposed a method based on adaptive color histograms [12]. Chen proposed two kinds of the traffic signs detection method. The first method is based on colour and shape information. The second method is based on colour information and Adaboost traffic sign detection technology [13]. Hu [14] proposed a method based on parallel strategy and contour geometry analysis traffic sign detection. Smorawa and Kubanek [15] used the Canny edge detection operator to mark the position of the landmark. Kaempchen *et al.* [16] used cameras and laser scanners to detect and track the position, velocity, orientation and size of other moving objects, and to assess the motion of each object. Li [17] proposed a V parallax algorithm. Zhu *et al.* [18] used the re-projection transformation technology to achieve real-time monitoring of road obstacles.

Compared with other sensors, the visual sensor has the advantages of a large amount of detection information and relatively low price. However, in the complex environment, getting the detection target and the background extracted has the disadvantages of large computational complexity and difficulty to realise the algorithm. At the same time, the visual hardware device and the visual processing algorithm determine the performance of the intelligent driving vehicle's visual perception system. Intelligent driving

vehicle sensor configuration is the foundation to perceive the surrounding environment and obtain their own state, and it is also the most cost of the test platform and the largest difference with other parts. As the different types of intelligent vehicle platform's mechanical structure are different, autonomous vehicle sensor selection and installation location will be different as well, there is no unified program. Some research teams tend to rely on visual sensors to complete environmental awareness, typically represented by the smart car team at VisLab Labs, University of Parma, Italy. Others tend to rely on radar sensors to obtain environmental information, typically represented by Google's autonomous vehicles. They mainly focus on the improvement of the algorithm and technology itself, but how to reduce the impact of driving the decision on the number of sensors, type and installation location changes, so that the technical architecture can adapt to different intelligent vehicle platforms with different sensor configuration and to decouple intelligent decision and sensor information. As driving information is completed, we will need to increase or reduce one or more visual sensors, change the visual sensor model or installation location, which will no longer directly affect the intelligent driving decision. This paper puts forward the technical structure of an autonomous vehicle with driving brain as the core and solves the convenient portability and robustness of visual sensor sensing technology and application.

In this paper, visual perception technology and application are described by the technical framework based on driving brain. In Section 2, the classification of visual sensors is introduced. Section 3 introduces the role of visual sensors in the perception of intelligent driving environment. Section 4 introduces the visualisation based on driving brain. Section 2 introduces environment perception technology by the visual sensor, and Section 5 is the conclusion.

2 Intelligent driving visual sensor classification

2.1 Monocular vision sensor

Domestic brands of industrial cameras are mainly Microvision, Aftvision, Daheng image, foreign brands of industrial cameras are German AVT camera, Canada Point Gray and so on. Visual research in the field of intelligent driving mainly uses AVT Guppy Firewire and Point Grey. The parameters of Point Grey camera and AVT Guppy Firewire camera are shown in Table 1.

2.2 Binocular vision sensor

Binocular vision technology is currently the most mature and the most widely used stereoscopic vision technology, which is mainly used in four areas: robot navigation, micro-operating system parameter detection, three-dimensional measurement and virtual reality. At present, the binocular vision detection distance and precision limitations make its advantages in intelligent vehicle obstacle detection that are still relatively weak, mainly stay in the theoretical research and experimental verification stage. The binocular vision products are shown in Fig. 1. They are MEGA-DCS Megapixel Digital Stereo Head products, nDepthTM Stereo Vision Cores products, TYZX Stereo Vision Cameras products, Bumblebee[®]2 Stereo Vision Camera products, Hydra Stereo Webcam products, Minoru 3D Webcam products, respectively.

2.3 Panoramic vision technology

Panoramic vision technology is divided into single-camera 360° rotary imaging, multi-camera splicing imaging, fisheye camera imaging and folding reflection panoramic imaging, as shown in Fig. 2. Among them, single-camera 360° rotary imaging needs to constantly rotate the camera, requiring a more stable mechanical structure. The system reliability requirements are high; multi-camera splicing imaging requires multiple cameras stitching, the system calibration is more complex, and needs to achieve multi-synchronisation and seamless splicing technology, and has a higher cost. For autonomous platform applications, fisheye camera uses imaging and folding reflective panoramic imaging technology.

3 Role of visual sensor in the intelligent driving environment perception

To achieve the autonomous driving in an unknown environment, a variety of environments are needed to obtain real-time reliable external and internal information. In developing the domestic and international intelligent driving technology, one of the most useful environmental perception sensors is the visual sensor. From the American DAPAR Challenge to the China Intelligent Vehicle Future Challenge, in the practical research and application development, visual perception is mainly used in the lane departure warning (LDW) system, the traffic sign security assistance system, the pedestrian detection safety system, the monocular vision advanced driving assistance system (ADAS) and binocular vision sensing system. Among them, the monocular vision ADAS Mobileye's key technology is based on an EyeQTM and EyeQ2TM processor using only one camera and bundles of multiple applications together. The functions include LDW, vehicle detection based on radar-vision fusion, forward collision warning, headway monitoring and warning, pedestrian detection, intelligent headlight control, traffic sign recognition (TSR), visual adaptive cruise control and so on, as shown in Fig. 3. Its technology has been used in BMW, GM, Volvo, Hyundai, Renault and other vehicles.

Google's autonomous vehicle uses the Point Grey Grasshopper 5 megapixel colour camera to measure the traffic light status, by selecting the 2040 × 1080 region of the pixel area of interest to operate, the horizontal viewing angle of 30°. The farthest detection distance is 150 m, as shown in Fig. 4. The precision positioning technology is applied to assist the detection. The position and height information of each traffic light is marked on the pre-acquired map. The autonomous driving relies on its own precise positioning information to calculate the relative position relationship between the current traffic light and the camera. According to the projection relationship of the image, it will limit the location of each traffic light in the image, thus greatly reduces the interference and improves the efficiency of the algorithm, the preparation rate meets 0.99, as shown in Fig. 5.

Braive Smart Car from The University of Parma, Italy, is equipped with 10 cameras for lane lines, traffic signs and obstacle detection. They installed four cameras at the rearview mirror, formed a long and short binocular vision system, while a single camera also had the function of detecting the road and traffic signs. Exterior mirrors and cameras on both sides of the front are used for lane or crossroads waiting for vehicle detection. The rear of the vehicle is fitted with a set of stereoscopic vision systems consisting of two cameras to detect rear proximity obstacles as shown in Fig. 6.

Table 1 Camera parameters

Camera model	Resolution	Focus, mm	Field angle	Interface	Manufacturers
DragonFly2 camera	1036 × 776	6.4	73.76 × 58.86	IEEE 1394	Point Grey
FireFly MV camera	752 × 480	4	100.43 × 84.15	IEEE 1394	Point Grey
Panoramic Vision camera	752 × 480	3.5	65	IEEE 1394	AVT Guppy Firewire
AVT Guppy Firewire colour camera	752 × 480	4	60	IEEE 1394	AVT Guppy Firewire



Fig. 1 Binocular vision products

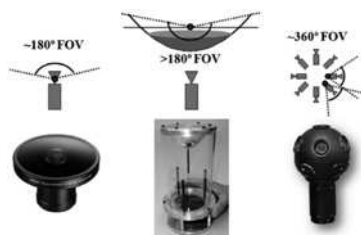


Fig. 2 From left to right, the fisheye lens imaging, folding reflective panoramic imaging and multi-camera imaging, respectively



Fig. 3 ADAS system and pedestrian detection diagram



Fig. 4 Google autonomous car camera layout

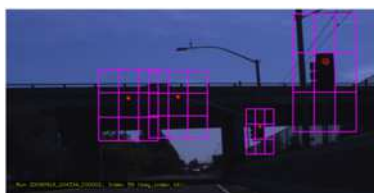


Fig. 5 Google shows the traffic light limit processing area

Porter smart car from The University of Parma, Italy, has a set of two cameras composed of binocular stereoscopic vision system at the front and the rear of the car, the car behind the mirror with three 65 horizontal field angle camera, composed of the 180° panoramic image in front of the vehicle, which makes it possible to do the vehicle detection and tracking in any direction, as shown in Fig. 7.

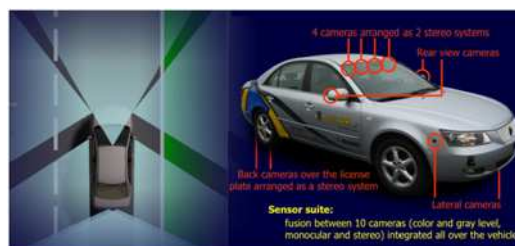


Fig. 6 'Braive' camera layout and field of view



Fig. 7 Porter intelligent car camera layout



Fig. 8 Bowler Wildcat intelligent car camera layout



Fig. 9 CMU BOSS intelligent car camera layout

Oxford University Bowler Wildcat intelligent car used Point Grey Bumblebee2 stereo camera Stereo vision system, pre-collected road images to generate maps, and compared with the historical image, positioning the current vehicle location and navigation, while detecting the obstacles. Horizontal field of view is 65°, the image resolution is 512 × 384, as shown in Fig. 8.

Carnegie Mellon University BOSS intelligent Car uses two Point Grey Firefly colour cameras for road detection, assesses road trends, and detects static obstacles on the road, as shown in Fig. 9.

Binocular vision technology is rarely used in intelligent vehicle obstacle detection, mainly stays in the theoretical research and experimental verification stage. In 2004 and 2005, the GOLD (generic obstacle and lane detection) system of the Parma University in Italy, using the V parallax algorithm, relying on three cameras to form two binocular vision systems with variable baseline lengths to achieve obstacle detection in unstructured roads. University of Oxford Bowler Wildcat intelligent vehicle Point Gray Bumblebee2 stereo camera Stereo vision system used the reflection of panoramic imaging algorithm to achieve visual SLAM.

3.1 Lane information extraction technology

The lane information extraction is mainly done to determine the lateral position of the vehicle in the lane, the direction of the vehicle relative to the lane centreline, the lane geometry, such as the lane width, curvature and so on. The extraction of lane information generally includes two steps: lane detection and tracking. In the absence of prior knowledge limits, lane detection determines the lane boundary in a single image. The lane tracking determines the position of the lane in the sequence image, and the current frame lane position can be limited by the information about the lane boundary position in the previous frame image. The main algorithms are Kalman filter, extended Kalman filter and particle filter.

3.2 Obstacle detection technology

Obstacle detection is an important guarantee for the safe driving of intelligent vehicles. Obstacles appear unpredictable, cannot be based on pre-set electronic map to avoid obstacles, only in accordance with the vehicle in advance of the process of discovery, immediately deal with. Obstacle detection based on monocular vision technology is mainly used to detect vehicles and pedestrians. The main algorithms are feature-based obstacle detection technology and obstacle detection technology based on the optical flow field.

3.3 Identification of traffic signs

The identification technology of traffic signs mainly includes two basic technical aspects. Firstly, the traffic signs are detected, including the traffic signs location and pretreatment. Secondly, the traffic signs are identified, including the feature extraction and classification of traffic signs. The main algorithm of TSR is the traffic mark detection and template matching method based on the colour image. The traffic sign discriminant algorithm is mainly the traffic sign difference method.

3.4 Visual SLAM technology

SLAM has high accuracy in direction measurement, low computational load, flexible and inexpensive, but deep recovery of information. The feature map is the main environmental representation based on the visual SLAM method. Its characteristic is mainly the corner feature. Harris, Kanade-Lucas-Tomasi (KLT) and scale-invariant feature transform (SIFT) features are the point feature detection methods commonly used in visual SLAM. Monocular vision SLAM is mostly based on the extended Kalman filter method. Oxford University has achieved the positioning and navigation by using visual SLAM method and applied it to real-time road test.

4 Vision sensor architecture and environmental awareness and processing

Intelligent vehicle environment awareness technology includes road information, traffic signs information, pedestrian behaviour information and peripheral vehicle information. Information through a variety of vehicle heterogeneous sensor equipment to obtain, there are a variety of cameras, radar (such as laser radar, millimetre-wave radar, ultrasonic radar, infrared radar etc.), GPS receiver, inertial navigation and so on. When a human driver performs a complex driving operation, such as lane changing, it will focus on different areas to ensure lane safety. Radar sensor has many different types; multi-heterogeneous radar sensor can get real-time information through effective extraction, and learn from the characteristics of human cognition. The sensor model, quantity and installation location of different intelligent driving test platform are different, and the sensor information processing module is different, as shown in Fig. 10. The information provided

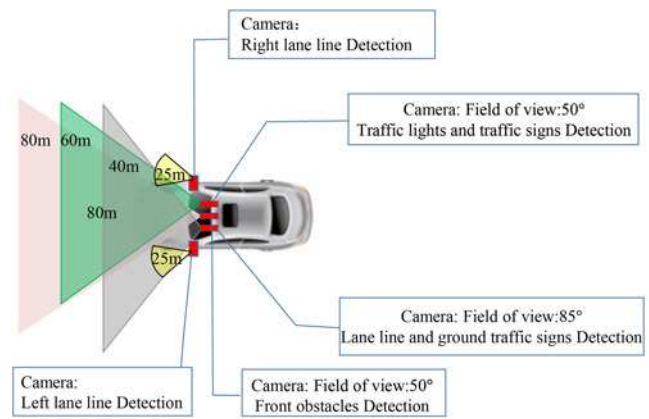


Fig. 10 Intelligent car hardware architecture platform

by different driving maps has no fixed standard of its granularity. The number and interface of the autonomous system software modules are different. To drive the brain as the core, form the driving perception and use the driving cognitive formal language, you can design a common autonomous software architecture. In this architecture, the intelligent decision module is not directly coupled with the sensor information. Through the sensor information and map of a priori information, we can form a comprehensive driving situation to complete intelligent decision making, and use the image segmentation algorithm to extract the target area, and then focus on regional segmentation. According to the input characteristic calculation result, the driver searches for the obstacle in the secondary division area, outputs the travelable road area and performs the information sensing method.

Human drivers integrated the coordination of sensory memory, working memory, long-term memory, computing centre and learning different areas of the brain, complete driving activities. The driving function of the intelligent driving system simulates human-related functions and also requires a series of the above-mentioned key functions and allows these functions to work together, as shown in Fig. 11. These key features include sensor information acquisition and processing, driving map calibration and mapping, driving cognitive formal language and intelligent decision making [19].

Among them, the sensor information acquisition, including camera, laser radar, and millimetre-wave radar, combined positioning system and other information, in which the visual sensor to complete the lane marking identification, traffic signal recognition, traffic signs identification. In the field of image processing, it will involve the object classification and recognition, using SIFT, histogram of oriented gradient (HOG) and other image feature methods to identify, and using support vector machine (SVM) and other machine learning algorithm to complete the classification and identification of objects [20–22]. For intelligent driving test, the application of driving map prior knowledge can improve object recognition performance. Google smart car records the traffic lights, traffic signs and other static traffic elements of the precise location in advance, during the identification of objects, uses the precise location of vehicles, the collection of traffic lights, traffic signs location information mapped to the image, to quickly determine the region of interest, which greatly improves the speed and accuracy of recognition [23]. In addition, simultaneous positioning and mapping (SLAM) can be accomplished using image information, which helps the vehicle acquire its own precise location [24]. As the visual sensor imaging by the perceived object's light or reflection is affected by external lighting conditions greatly, limiting the scale of application of visual sensors.

According to the road segmentation algorithm based on inverse perspective transformation, the working area of the camera is shown in Fig. 12, which can be used to analyse the effective working area of the visual sensor and the radar sensor. Fig. 12

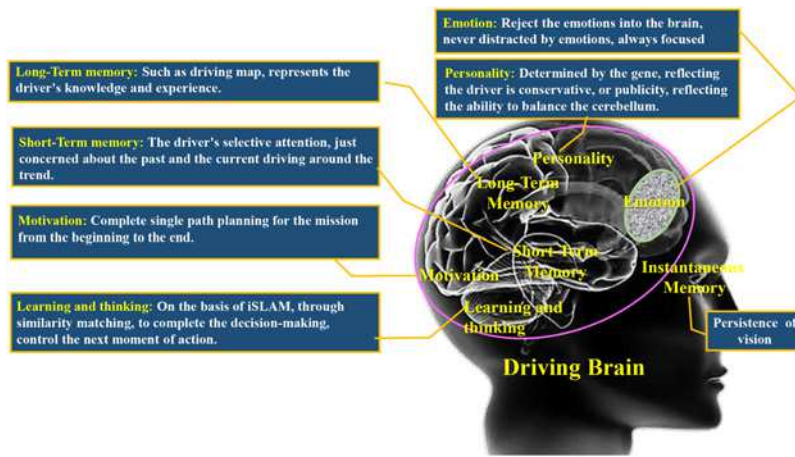


Fig. 11 Driving brain and human brain function technical structure

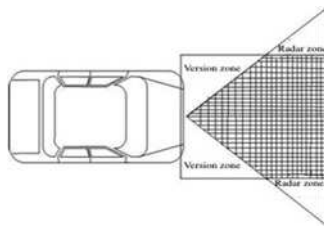


Fig. 12 Schematic diagram of the sensor operating area

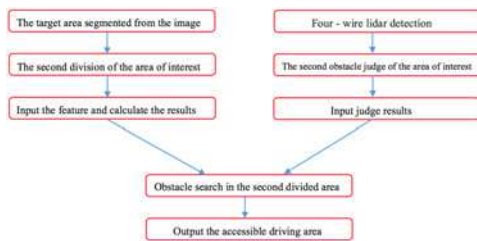


Fig. 13 Graphic sensor information fusion algorithm flow chart

shows the working area of the two sensors. The intersection area is the common area of the camera and the radar sensor.

Unearthed road surface unevenness will cause the body bumps, so the maximum speed of the intelligent vehicle is set at 20 km/h (5.6 m/s), the maximum deceleration of the intelligent driving vehicle is 6 m/s^2 . From the detection of the road, it cannot travel to stop, we need at least 5 m distance. Braking is a gradual increase, taking into account the time consumption of the algorithm and time-consuming response of the braking system, the safety distance is set to 10 m, the algorithm only concerned about the vehicle in front of 10 m within the road area. From the perspective of human cognition, the target area within 10 m of the intelligent driving vehicle is first visually judged. Then, whether there is an obstacle in the target area or not, if not it is considered to be a feasible area, on the contrary, it is considered that the target area in front of the obstacle is a travelable area. Based on the cognitive model algorithm, with reference to the principle of human cognition, we only need to consider the areas that may bring potential threats to the self-driving of autonomous vehicles, and do not need to consider the farther areas, which save the computing resources for the algorithm to meet the real-time principle. Finally, the algorithm should take into account the robustness criterion, so that the decision result of the input decision layer is expected to have better

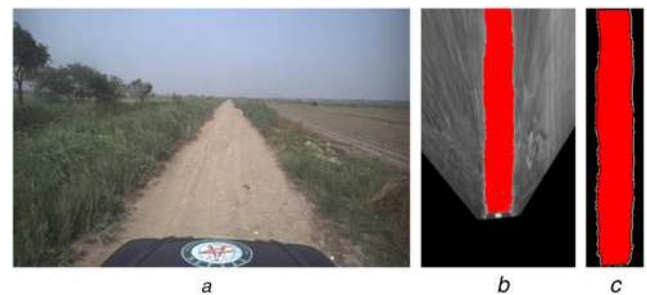


Fig. 14 Initial road area of the map is further extracted from the process diagram

a Original road scene
b Visual output result
c Road area extraction

anti-interference as much as possible. In summary, the process of visual sensor information sensing and processing algorithm is shown in Fig. 13.

4.1 Target area and obstacle information extraction

Smart cars, when driving (especially autonomous exchange lane) in the process of real-time detection of peripheral vehicles, pedestrians and other obstacles, need to determine the travel area to ensure the safety. Obstacle detection is the key and difficult point of environmental perception technology. As the active sensor, it has the characteristics of good direction, high precision and reliable data. The aim of this paper is to use the radar for road identification and the second decision of the obstacle in the area of interest, to obtain a safe path of the road which can be confirmed. Firstly, a target area is determined by the visual sensor. The target area is calculated by the image data obtained by the visual sensor. Then, the target area is searched and the safe and feasible road area is obtained.

4.2 Target area of secondary processing

The original image obtained by the visual sensor is transformed into the road plane, and the initial target road area is obtained by the image segmentation method. In order to meet the real-time criteria, the algorithm focuses on the road area within 10 m of the intelligent vehicle. To ensure the robustness, the judgement result is simple and practical. Therefore, the initial target road area needs to be further extracted and processed, and Fig. 14 shows the extraction process.

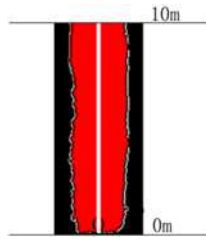


Fig. 15 Area of attention

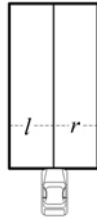


Fig. 16 Target area after secondary processing

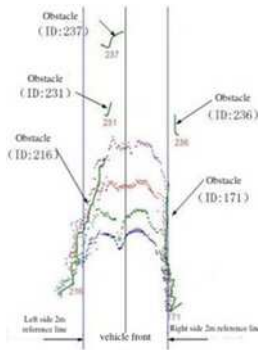


Fig. 17 Initial barrier detection results

The red areas in Figs. 14b and c show the road area obtained by the image segmentation algorithm. According to the camera to calibrate the internal and external parameters, you can calculate the corresponding relationship between the pixel and the actual distance, so you can get the extracted target area length, area, average width and other information. This paper intercepts 10 m within the area of concern for processing. At the same time as the smart car does not necessarily travel in the middle of the road, so the road area is divided into two parts according to the middle of the car's head, as shown in Fig. 15, respectively, calculate its area, according to the relationship between the size of the pixel and the actual distance, and get the average road width of the left and right sides.

Let the last calculated road width average of the left and right sides be l and r , then the final target area is a rectangular area as shown in Fig. 16.

4.3 Secondary judgement of the obstacle

The obstacle detection algorithm based on LIDAR can output the obstacle ID and obstacle number of all clustering results in the radar field of view, and can also obtain the coordinates information of the obstacle centre point. Similarly, only concerned about the radar output within 10 m range of obstacle clustering results, for the further range of the obstacle detection results, it only acts as a basis for tracking comparison. Fig. 17 is the result of the direct output of the obstacle detection algorithm.

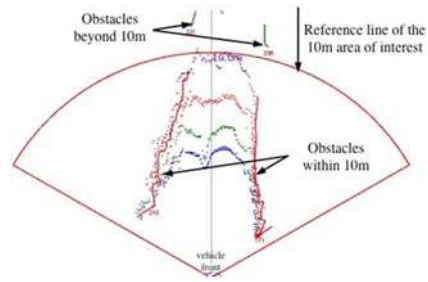


Fig. 18 Obstacle detection after secondary judgement

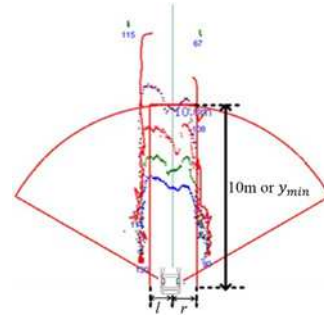


Fig. 19 Map travel road area diagram

Fig. 18 shows the output of the obstacle within 10 m after the second decision. The red sector shows the 10 m field of view of the radar. The obstacle in the 10 m field of view is signed by a red line.

4.4 Accessible road area output

Finally, the sensors need search obstacle for the target area. If there is no obstacle in the target area, it is considered that the target area is a road area that can be travelled; otherwise, the target area before the obstacle is considered to be a feasible road area. The specific output process of the road area is as follows:

First of all, set the camera output of the road width on both sides are l , r . In the Oxy coordinate system of the vehicle plane, four straight lines of $x = -l$, $x = r$, $y = 0$ and $y = 10$ are drawn so as to intersect with the previous two straight lines to obtain a closed rectangular space. The rectangular space is the target area.

Secondly, search for obstacle object, if there is an obstacle object in the target area, extract the value of y all the obstacles in this area. By comparing, get the minimum value y_{min} , draw line $y = y_{min}$, discard the line $y = 10$, get a new rectangular area. The new rectangular area is the accessible road area. The feasible road area obtained by the algorithm is shown in Fig. 19. Since the left and right sides of the road area obtained from the visual information do not contain an obstacle, it is considered that the length of the accessible road area is 10 m or y_{min} metres, and the width of the left and right sides is $(l + r)$ metres, which meets the intelligent vehicle accessible area output.

5 Conclusion

As the machine vision system can quickly obtain a lot of information, easy to integrate design information and processing control information, therefore, in the modern automated production process, people widely use it in machine vision system monitoring, product testing and quality control and other fields. In the intelligent driving vehicle, since the sensor is cheaper than LIDAR and can obtain rich information, the sensor is the intelligent driving vehicle technology research focus, which has

broad application prospects. It is expected that with the maturity and development of machine vision technology, the intelligence of the algorithm will be improved, it will be more and more widely used in intelligent driving vehicles. The future intelligent driving vehicles will mainly rely on machine vision to conquer the environment perception task.

6 Acknowledgements

This work was supported by National Natural Science Foundation of China under grant no. 61035004, no. 61273213, no. 61300006, no. 61305055, no. 90920305, no. 61203366, no. 91420202, no. 61571045, no. 61372148, the National High Technology Research and Development Program ('863' Program) of China under grant no. 2015AA015401, and the National High Technology Research and Development Program ('973' Program) of China under grant no. 2016YFB0100903, and the Junior Fellowships for Advanced Innovation Think-tank Program of China Association for Science and Technology under grant no. DXB-ZKQN-2017-035, and the Beijing Municipal Science and Technology Commission special major under grant no. D171100005017002.

7 References

- [1] Gage, D.W.: 'UGV history 101: A brief history of unmanned ground vehicle (UGV) development efforts', *Unmanned Syst.*, 1970, **13**, pp. 9–32
- [2] Kanade, T., Thorpe, C.: 'CMU strategic computing vision project report: 1984 to 1985' (Carnegie-Mellon University, Pittsburgh, 1986), pp. 10–90
- [3] Williams, M.: 'PROMETHEUS-The European research programme for optimising the road transport system in Europe'. Proc. of IEEE Colloquium on Driver Information, London, UK, 1988, pp. 1–9
- [4] Tsugawa, S., Aoki, M., Hosaka, A., *et al.*: 'A survey of present IVHS activities in Japan', *Control Eng. Pract.*, 1997, **5**, (11), pp. 1591–1597
- [5] Yang, M.: 'Overview and prospects of the study on driverless vehicles', *J. Harbin Inst. Technol.*, 2006, **38**, (8), pp. 1259–1262
- [6] Gao, H.B., Zhang, X.Y., Zhang, T.L., *et al.*: 'Research of intelligent vehicle variable granularity evaluation based on cloud model', *Acta Electron. Sin.*, 2016, **44**, (2), pp. 365–374
- [7] Zhang, Y.R., Guo, C.J., Niu, R.Z.: 'Research on stereo-vision aided GNSS localization for intelligent vehicles', *Comput. Eng. Applic.*, 2016, **52**, (17), pp. 192–197
- [8] Ziegler, J., Lategahn, H., Schreiber, M., *et al.*: 'Video based localization for bertha'. Proc. of 2014 IEEE Int. Conf. on Intelligent Vehicles Symp., Leipzig, Germany, 2014, pp. 1231–1238
- [9] Wolcott, R.W., Eustice, R.M.: 'Visual localization within LiDAR maps for automated urban driving'. Proc. of 2014 IEEE Int. Conf. on Intelligent Robots and Systems, Chicago, USA, 2014, pp. 176–183
- [10] Duan, H.X.: 'Outdoor vision localization and map building for the unmanned vehicle based on binocular vision' (Dalian University of Technology, Dalian, 2015)
- [11] Fang, Q.L.: 'Research on traffic mark recognition method based on unmanned vehicle assisted navigation' (Anhui University, Hefei, 2012)
- [12] Bernuy, F., Solar, D., Parra, I., *et al.*: 'Adaptive and real-time unpaved road segmentation using color histograms and RANSAC'. Proc. of the 9th IEEE Int. Conf. on Control and Automation, Santiago, Chile, 2011, pp. 136–141
- [13] Chen, Z.X.: 'Detection and identification of road traffic signs in urban areas' (University of Science & Technology China, Beijing, 2012)
- [14] Hu, J.C.: 'Research on traffic sign detection and recognition based on stable feature' (Hunan University, Changsha, 2012)
- [15] Smorawa, D., Kubanek, M.: 'Analysis of advanced techniques of image processing based on automatic detection system and road signs recognition', *J. Appl. Math. Comput. Mech.*, 2014, **13**, (1), pp. 103–113
- [16] Kaempchen, N., Schiele, B., Dietmayer, K.: 'Situation assessment of an autonomous emergency brake for arbitrary vehicle-to-vehicle collision scenarios'. *IEEE Transactions on Intelligent Transportation Systems*, 2009, **10**, pp. 678–687
- [17] Li, Y.: 'Research of obstacle recognition based on binocular vision' (Wuhan University of Technology, Wuhan, 2007)
- [18] Zhu, Z.G., Lin, X.Y., Shi, D.J.: 'A real-time visual obstacle detection system based on reprojection transformation', *J. Comput. Res. Dev.*, 1999, **1**, (1), pp. 77–84
- [19] Lowe, D.G.: 'Object recognition from local scale-invariant features'. Proc. of the Seventh IEEE int. Conf. Computer Vision, Kerkyra, Greece, 1999, no. 2, pp. 1150–1157
- [20] Lowe, D.G.: 'Distinctive image features from scale-invariant keypoints', *Int. J. Comput. Vis.*, 2004, **60**, (2), pp. 91–110
- [21] Dalal, N., Triggs, B.: 'Histograms of oriented gradients for human detection'. Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, San Diego, USA, 2005, no. 1, pp. 886–893
- [22] Zhang, X.Y., Gao, H.B., Guo, M., *et al.*: 'A study on key technologies of unmanned driving', *CAAI Trans. Intell. Technol.*, 2016, **1**, (1), pp. 4–43
- [23] Levinson, J., Askeland, J., Dolson, J., *et al.*: 'Traffic light mapping, localization, and state detection for autonomous vehicles'. Proc. of 2011 IEEE Int. Conf. on Robotics and Automation (ICRA), Shanghai, China, 2011, pp. 5784–5791
- [24] Engelhard, N., Endres, F., Hess, J., *et al.*: 'Real-time 3D visual SLAM with a hand-held RGB-D camera'. Proc. of the RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum, Robotdalen, Sweden, 2011, p. 180