Mechatronics and Automation Technology
J.-Y. Xu (Ed.)

doi:10.3233/ATDE231092

Key Issues and Countermeasures of Machine Vision for Fruit and Vegetable Picking Robot

Qi XIN[1], Qiang LUO and Hongda ZHU
Department of mechanical engineering, Chongqing three gorges University, Chongqing, 404020, China

Abstract. Fruit and vegetable picking is an important link in China's agricultural production. The existing manual picking method has high labor intensity and low efficiency. The application of picking robots can greatly improve the efficiency of picking operations and reduce the dependence on manpower. Its research has great practical value. The machine vision system of the picking robot is currently a hot and key research area in this field. This article analyzes the key issues faced by the machine vision system research in terms of precise identification of picking objects, spatial positioning, and path planning in complex environments, and classifies and summarizes the latest research results of machine vision at home and abroad.

Keywords. Machine vision; Target recognition; Spatial positioning; Path planning.

1. Introduction

In recent years, the domestic population dividend has gradually disappeared, and the problem of labor shortage has become a bottleneck restricting the development of agriculture, especially the development of labor-intensive industries. Picking robot technology has changed from forward-looking research to practical demand. With the high and new technology represented by computer image processing technology, industrial robot technology and artificial intelligence technology gradually infiltrating into the field of agriculture, the research and development of picking robot has entered a period of rapid development. At present, many enterprises at home and abroad are developing fruit and vegetable picking robots, such as Panasonic in Japan, Harvest CROO Robotics in the United States, FFRobotics in Israel and so on.

The working environment of the agricultural picking robot is very complex, and the picking robot needs to find randomly distributed fruits and vegetables from the chaotic background including branches and leaves, sky and other disturbances [1]. The key to solve this problem is to introduce the machine vision system into the picking robot to make the picking robot have high recognition rate and positioning accuracy, and achieve autonomous navigation in the unstructured environment. realize from search, scanning, identification, positioning to end-effector control and operation, and finally realize the automatic harvest of crops. For example, the intelligent agriculture picking platform

[1] Corresponding Author, Qi XIN, Department of mechanical engineering, Chongqing three gorges University, 666 Tianxin Road, Wanzhou District, Chongqing, China; E-mail: xinqi1998@foxmail.com.

RS-AGR recognizes the target fruit and records the fruit orientation through the vision system based on YOLO algorithm, and obtains a series of fruit spatial coordinate points through coordinate transformation. With the help of autonomous navigation system and SLAM algorithm, the navigation map is constructed to realize the autonomous movement in the orchard natural environment. Based on the research and analysis of the latest research results of agricultural picking robot at home and abroad in recent years, this paper analyzes the difficulties faced by the vision system of agricultural picking robot, and puts forward the corresponding countermeasures.

2. Key Problems in Machine Vision of Picking Robot

2.1. Accurate Recognition of Picking Objects

Under natural light, the accuracy and efficiency of machine vision recognition of fruits and vegetables are affected by many factors, including the uncontrollable scene of the actual picking operation, and the clarity of the shape and color of fruits and vegetables under different light intensity and light angles. the performance of the detection model is limited in different scenes [2]; there are great differences in fruit distribution between the same species and different kinds of fruits, irregular growth shape, and the color of fruits and vegetables is similar to the background color [3]. The fruits cover each other and overlap, and the branches and leaves of the plant cover the fruit. In addition, most of the existing algorithms classify different kinds of fruits in complex environment into the same kind of targets for recognition. If they are picked directly without accurate identification, it is very likely to cause fruit damage, or damage to picking hands and manipulators [4]. Finally, the increase of equipment complexity caused by the improvement of accuracy requires high computing requirements for the hardware equipment of the picking robot, and the system cost is high. These problems have become the main obstacles to the research of visual inspection of fruit and vegetable picking robot

2.2. Positioning of Picking Objects in Spatial Coordinates

After the identification of the picking object, it is necessary to solve the problem of how to locate the picking object. The operating environment of agricultural picking robot is complex, and there are many uncertain environmental factors. Accurate identification and precise location of agricultural picking targets is an important prerequisite for robot picking. The traditional digital image processing methods cannot meet the requirements of high accuracy and real-time performance, and are not robust to the variety of fruit shapes, so it is difficult to meet the accurate and real-time picking requirements of the robot [5]. In addition, when using deep learning technology for target detection, it needs to consume a lot of computer resources, and the accuracy and speed of detection and location still need to be improved. Although scholars at home and abroad have carried out a series of studies in the fields of picking apples, tomatoes, strawberries and cherries, and developed corresponding agricultural picking robots, there are no mature commercial cases at present, and the structure of the algorithm needs to be further optimized to improve the accuracy of fruit target recognition and location.

2.3. Path Planning in Complex Environment

The fruit and vegetable picking robot needs to locate and move to the target area independently after identifying the mature fruit, but it is easy to affect the driving of the robot when it encounters obstacles such as fruit trees and pedestrians with large canopy [6]. This requires the picking robot to have a certain ability of obstacle detection and avoidance, and carry out path planning at the same time in order to effectively avoid accidents and improve operation efficiency [7].

The traditional working area planning path is mostly straight line, which is not conducive to avoid obstacles such as fruit tree canopy. At present, the mainstream navigation in orchard is preset digital map and image processing. This method requires the map to be updated in real time, and cannot be used when there is a deviation between the digital map and the actual road conditions. In order to reduce the difficulty of image processing, road signs will be set up on the pre-designed robot running route, and the robot cannot automatically plan the walking route according to the orchard environment by identifying the road sign and determining the next step.

3. Countermeasures to Solve the Key Problems of Machine Vision

3.1. Target Recognition Technology based on Convolution Neural Network

In order to realize the accurate detection of picking objects by picking robot in complex environment [8], Gao Mengyuan team of Wuhan Textile University proposed a method to detect orchard apples based on Mask RCNN case segmentation model. The algorithm adopts lightweight Mobilenetv3 network and SPP module, which effectively reduces the number of parameters of the model and improves the detection speed. The algorithm effectively detects the apple edge position information and categories in different scenes, achieves high accuracy and recall, and provides an important technical support for the vision system of automatic apple picking robot. In order to facilitate the application of convolution neural network model to picking robot with limited computing power [9], Tang Yi team of Hubei University of Technology proposed a lightweight convolution neural network model, which is used to solve the problem of fast and accurate recognition of citrus picking robot. The model uses DIOU loss function instead of the original loss function, uses MobileNetv3-Small convolution neural network model for feature extraction, adds residual structure, SPP network structure and depth separable convolution layer, simplified spatial pyramid pool network structure and depth separable convolution layer set to improve the recognition accuracy of the model, in order to improve the recognition accuracy and speed of the model. In order to test and improve the performance of the YOLOv3-tiny model, the original YOLOv3-tiny model was built as a comparison model, and the results are shown in table 1. Compared with the original model, the improved YOLOv3-tiny model has higher average recognition accuracy and F_1 score for citrus recognition, and has faster recognition speed and smaller model weight size.

Table 1. Performance comparison between the models

Model name	Average recognition accuracy/%	F_1	Time/ms	Weight size/MB
YOLOv3-tiny	93.28	0.89	62	33.1
Improved YOLOv3-tiny	96.52	0.92	47	16.9

By using Python technology to optimize the efficiency and stability of vision picking robot [10], Su Jingxiao team of Hebei Institute of Engineering and Technology proposed a vision control model based on Python technology, as shown in figure 1. Through the definition of visual model boundary features, threshold setting, target color block display optimization, and the optimization of picking target color block identification and judgment in the core of Python programming, an integrated picking operation control system was established. The experimental results show that this method can effectively reduce the picking error, improve the stability of the system, improve the picking rate and the efficiency of the whole machine, and show the application prospect of Python technology in the intelligent device of agricultural machinery.

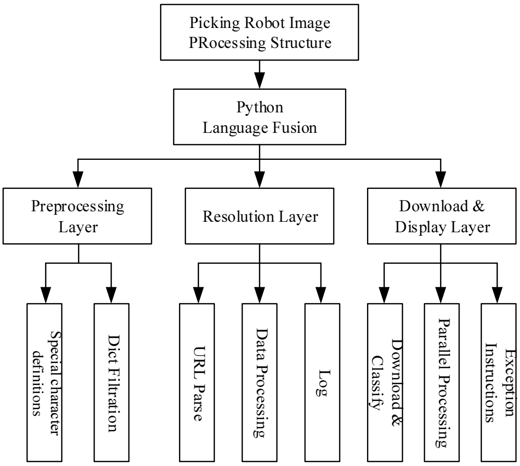

Figure 1. Image processing structure optimization diagram of visual picking robot

In order to solve the problem that it is difficult to accurately identify overlapping fruits [11], Xi Houyin team of Qingdao University of Science and Technology designed a combined recognition and capture method, including median filtering to remove noise, OTSU segmentation algorithm and morphological processing. The real contours of tomatoes were extracted by convex hull algorithm, flooding filling algorithm and Hoff straight line detection, and the ellipse fitting algorithm was used to stereo match the core of the fruit. A prototype robot suitable for tomato fruit picking is designed. The experimental results show that this method is effective, feasible and efficient to identify and capture overlapping fruits. Yan Bin's team from Northwestern University of Agriculture and Forestry Science and Technology proposed a real-time identification method of apple picking method for picking robot based on improved YOLOv5m [12]. Through the improvement of the BottleneckCSP-B feature extraction module and the crossover fusion mode, the algorithm realizes the enhancement of the deep feature extraction ability of the original module and the lightweight improvement of the backbone network, in order to better extract the features of different apple targets, and then improves the cross fusion mode of the feature map input to the medium size target detection layer in the original architecture, and enhances the feature extraction ability of the network to different apple targets in the image, the flow chart is shown in figure 2. This method can accurately identify different apple targets on the fruit tree, distinguish the fruit under different branch occlusion, and avoid the recognition error of the apple in the distance in the image. thus, it provides visual guidance for the manipulator to actively adjust the position to avoid the shading of the branches for fruit picking.

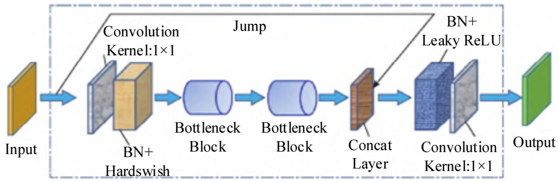

Figure 2. Improved BottleneckCSP (BottleneckCSP-B) block

In order to further analyze the performance of the proposed apple picking method recognition algorithm, the improved YOLOv5m network is compared with the original YOLOv5m, YOLOv3, and EfficientDet-D0 networks on the test set images, and the experiments are shown in table 2.The experimental results show that the algorithm is superior to similar algorithms in terms of recall rate, accuracy, mAP and F_1 values, and achieves higher recognition accuracy and speed in direct picking, circuitous picking and unpickable fruit recognition, which provides important technical support for the picking robot to avoid the blocking of branches and fruits and pick apples in different positions.

Table 2. Performance comparison of various target detection networks

Object detection network	mAP/%	Average recognition time for a single image/s	Number of parameters	Weight size/MB
YOLOv5m	75.3	0.020	0.020	41.3
YOLOv3	58.7	0.053	0.053	235.0
EfficientDet-D0	60.1	0.040	0.040	15.0
Improved YOLOv5m	80.7	0.025	0.025	37.0

Wei Tianyu team of Yangzhou University uses the improved YOLOv5s model to detect and locate hot pepper [13]. Based on the image database established by combining the scenes of light intensity, light angle, branch and leaf occlusion and fruit overlap, the bi-directional feature pyramid network is used to improve YOLOv5s's feature fusion network to realize depth feature extraction, so as to enhance the information expression ability of the network and improve the detection accuracy. Through the experimental results in different scenes, the method shows high recognition accuracy, adapts to different lighting conditions, occlusion and so on, and can meet the positioning accuracy requirements of pepper picking.

3.2. Target Location Technology based on 3D Vision

In order to solve the problem that it is difficult to locate the target of fruit and vegetable picking robot [14], Gao Shuai team of Jiangsu Vocational and Technical College of Agriculture and Forestry put forward a method which combines improved YOLOv3 algorithm with 3D vision technology to realize accurate target recognition and precise positioning, and the conversion between target coordinate system and robot coordinate system is completed by calibration, as shown in figure 3. Through experimental analysis, the improved algorithm has higher recognition accuracy and positioning accuracy, and can better complete the follow-up picking work, which has important reference value for the development of agricultural robots.

Zhou Hao's team of South China Agricultural University uses a binocular camera to capture the left and right images of Camellia oleifera fruit, and uses the target detection network YOLOv4-tiny to detect the left and right image of Camellia oleifera fruit [15]. Different from the traditional stereo matching technology of binocular camera image, this method extracts the region of interest of Camellia oleifera image according to the prediction box generated by YOLOv4-tiny network, and adaptively carries out stereo matching to solve the parallax, and finally obtains the Camellia oleifera fruit picking point. The experimental results show that this method can detect Camellia oleifera fruit accurately and in real time, and provide reference for the key technologies of visual perception of picking robot working in orchard environment. Feng Shuo's team from Henan Institute of Industry and Technology proposed a Kiwi picking robot arm control system based on convolution neural network [16]. The system uses the principle of binocular vision to realize the stereo positioning and accurate picking of the target fruit. The experimental results show that the picking robot manipulator can accurately reach the target position according to the control instructions, and the error between the actual arriving position and the visual positioning coordinates is less than 4.5%, which can meet the picking requirements of the picking robot.

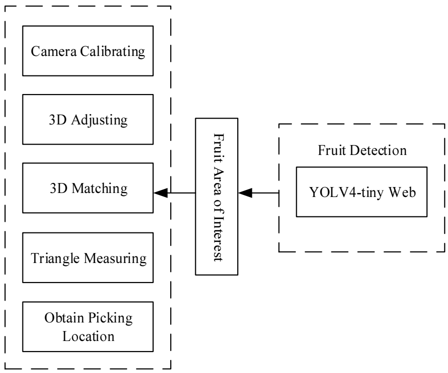

Figure 3. Schematic diagram of the 3D visual perception module

The Takeshi Yoshida team at the Tsukuba Agricultural Robotics Research Center in Japan used RGB-D cameras to automate the picking of pears and apples on V-shaped scaffolding [17]. Because the point cloud obtained by the RGB-D camera in the outdoor environment may be inaccurate, it is proposed that the picking robot not only use the three-dimensional information obtained by the RGB-D camera, but also use two-dimensional images and camera information to identify fruit. The experimental results show that the proposed method meets the accuracy required by the picking robot to pick fruits continuously.

3.3. Autonomous Navigation Technology Represented by Image Fusion

Global path planning mainly includes finding a path that connects the starting point and the target point to avoid collision with obstacles at the same time. The path length, algorithm time and space complexity are the main indicators to evaluate the performance of the global path planning algorithm.

Yang Xiaohui's team from Baise Vocational College proposed a picking robot navigation and walking system based on image processing [18]. Based on image processing, the system takes the dividing line between planting ridges and weeds as the navigation trajectory, realizes the fitting of discrete navigation pixels, and transforms the navigation equation in the image into the actual spatial coordinate system to get the actual navigation equation. The test shows that the system can effectively restrain the influence of the change of light intensity, which shows that the navigation method based on image processing can improve the stability and accuracy of the picking robot. Zhang Xiaoliang's team from Jiaozuo normal College introduced research on autonomous navigation technology of picking robot based on image fusion technology [19]. Through the fusion processing of multiple real-time captured images, this method can improve the quality of captured images, as shown in figure 4. The experimental results show that IHS image fusion algorithm has obvious advantages and can improve the accuracy of autonomous navigation of picking robot.

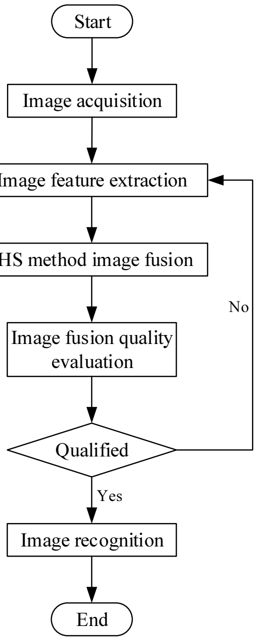

Figure 4. Visual navigation image fusion flowchart

Cui Yongjie team of Northwestern University of Agriculture and Forestry Science and Technology proposed an improved method (Straight-RRT) based on sampling state to guide random tree expansion in real time to improve the navigation efficiency of kiwifruit picking robot [20]. The method divides the sampling state by evaluation index and threshold to guide the expansion of the random tree in real time, introduces the dynamic threshold and optimizes the nearest node selection mechanism to enhance the adaptability of the algorithm to different environments and quickly avoid irregular obstacles. To verify the effectiveness of the improved algorithm, the Straight-RRT method is compared with the RRT algorithm, the target gravity-RRT algorithm and the RRT-Connect algorithm in the environment shown in figure 5, and the results are shown in table 3 and table 4. The experimental results show that the improved algorithm has better adaptability and planning efficiency in kiwifruit orchard environment, and provides a solution for improving the navigation efficiency of kiwifruit picking robot.

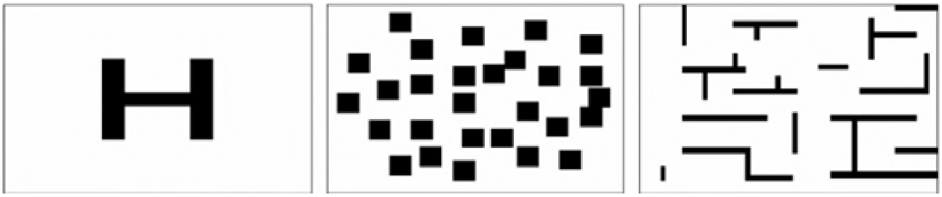

Figure 5. Simulation environment

Table 3. Comparison of different object detection network search times

Map	RRT	Target gravity-RRT	RRT-Connect	Straight-RRT
1	0.17	0.078	0.024	0.020
2	0.2285	0.085	0.025	0.022
3	0.587	0.238	0.162	0.121

Table 4. Comparison of iterations of different object detection networks

Map	RRT	Target gravity-RRT	RRT-Connect	Straight-RRT
1	637	432	356	218
2	871	508	273	199
3	1268	1305	1296	773

Xuan Feng team of Henan Vocational and Technical College proposed an obstacle detection and avoidance algorithm based on monocular vision and artificial potential energy field [21]. This method can effectively collect and detect the information of obstacles, and optimize the path by using the artificial potential energy field method according to the distance between the obstacles and the target area, so as to realize the autonomous movement of the picking robot. The test results show that the robot can avoid obstacles and plan the most efficient path to the target working area by using monocular vision and artificial potential field method. This shows that monocular vision and artificial potential field methods can improve the intelligent level of picking robot. In order to solve the problem that obstacles such as large canopy of fruit trees and pedestrians are easy to affect robot driving, Hu Guangrui team of Northwestern University of Agriculture and Forestry proposed an inter-row navigation path optimization method based on improved artificial potential field method [22]. In this method, the solid-state lidar carried by the mobile picking robot is used to obtain 3D point cloud information between orchard rows, LSM, Hough transform and RANSAC methods are used to extract the ridge line and initial path, and the improved artificial potential field method is used to optimize the path to avoid obstacles. Three frames were selected from the collected orchard point cloud according to the presence or absence of pedestrian obstacles, and the initial path extracted by RANSAC, the optimized path of the traditional artificial potential field method and the optimized path of the improved artificial potential field method were used as the navigation path, and the shortest distance between the canopy and the pedestrian obstacle point cloud from the navigation path was calculated, and the results are shown in table 5.The experimental results show that the method can optimize the path in real time to avoid obstacles, has good anti-noise ability and real-time ability, and provides a technical reference for the autonomous navigation of mobile picking robot in orchard environment.

Table 5. Shortest distance between three navigation paths and obstacle point cloud

Obstacle type	Selected sample frames	Shortest distance between navigation path and obstacle point cloud D/m		
		Random Sample Consensus (RANSAC)	Traditional artificial potential field	Improved artificial potential field method
Canopy obstacles	1	0.788	0.858	0.853
	2	0.501	-	0.856
	3	0.705	0.894	0.907
Canopy and pedestrian obstacles	1	0.324	-	0.778
	2	0.123	-	0.666
	3	0.156	0.838	0.863

4. Conclusion

This paper summarizes the key problems and existing solutions in the research of machine vision of picking robot. In the aspect of target recognition, the convolution neural network algorithm based on deep learning can not only identify the target fruit and background image, but also improve the detection accuracy and prediction frame accuracy of partially occluded targets through case segmentation. it is suitable for target recognition in natural environment. In the aspect of locating the acquisition target, 3D stereo vision matching and color camera are usually used to locate the target. This method can be used in indoor or outdoor scenes, and is less affected by the changes of lighting conditions and target characteristics. It meets the higher positioning accuracy requirements of the picking robot. In order to improve the accuracy of the picking robot navigation system, the image fusion technology is introduced into the image processing of navigation vision. Through the fusion of unclear or exposed images, the image acquisition quality of the picking robot is improved.

Acknowledgments

This research was supported by Chongqing three Gorges University Graduate Research and Innovation Fund, YJSKY23020, Research and Application of Intelligent Strawberry picking Technology.

References

[1] Su J, Zhang B, Zhang Q. Optimization of operation of visual picking robot based on Python technology [J]. Research on Agricultural Mechanization, 2023, 45 (08): 206-210
[2] Wang F, Cui D, Li L. Target recognition and positioning algorithm for picking robot based on deep learning [J]. Electronic Measurement Technology, 2021, 44 (20): 162-167
[3] Hua X, Li H, Zeng J, et al. A Review of Target Recognition Technology for Fruit Picking Robots: From Digital Image Processing to Deep Learning [J]. Appl. Sci., 2023, 13: 4160
[4] Xiao F, Wang H, Li Y, et al. Object Detection and Recognition Techniques Based on Digital Image Processing and Traditional Machine Learning for Fruit and Vegetable Harvesting Robots: An Overview and Review [J]. Agronomy, 2023, 13: 639

[5] Tang Y, Chen M, Wang C, et al. Recognition and Localization Methods for Vision-Based Fruit Picking Robots: A Review [J]. Front Plant Sci., 2020, 11: 510
[6] Jin Y, Gao Y, Liu J, et al. Research on deep vision servo-eye coordination planning of picking robot [J]. Transactions of the Society for Agricultural Machinery, 2021, 52 (06): 18-25+42
[7] Li T, Qiu Q, Zhao C, et al. Task planning of dwarf dense planting orchard multi-arm picking robot [J]. Transactions of the Society of Agricultural Engineering, 2021, 37(02): 1-10
[8] Gao M, Ma S, Dong Y, et al. Visual positioning and detection of apple picking robot based on example [J]. Jiangsu Agricultural Sciences, 2022, 50 (03): 201-208
[9] Tang Y, Yang G, Wang Y. Improved YOLOv3-tiny lightweight citrus identification method for picking robot [J]. Science Technology and Engineering, 2022, 22 (31): 13824-13832.
[10] Su J, Zhang B, Zhang Q. Optimization of operation of visual picking robot based on Python technology [J]. Research on Agricultural Mechanization, 2023, 45 (08): 206-210
[11] Xi H, Zhang D, Zhou T, et al. Research on the method of picking robot to identify and grasp overlapping tomato fruits [J]. Agricultural Mechanization Research, 2021, 43 (12): 17-23+50
[12] Yan B, Fan Pan, Wang M. et al. Real-time recognition of apple picking method of picking robot based on improved YOLOv5m [J]. Transactions of the Society for Agricultural Machinery, 2022, 53 (09): 28-38+59
[13] Wei T, Liu T, Zhang S, et al. Identification and positioning method of pepper picking robot based on improved YOLOv5s [J]. Journal of Yangzhou University (Natural Science Edition), 2023, 26 (01): 61-69
[14] Gao S, Liu Y, Gao J, et al. Research on target recognition and localization of agricultural picking robot based on YOLOv3 algorithm and 3D vision [J]. Chinese Journal of Agricultural Mechanization, 2022, 43 (12): 178-183
[15] Zhou H, Tang Y, Zou X, et al. Research on key technologies of visual perception of agricultural picking robot [J]. Agricultural Mechanization Research, 2023, 45 (06): 68-75
[16] Feng S, Li Q. Research on control system of robotic arm of kiwifruit picking robot——Based on computer vision and convolutional neural network [J]. Journal of Agricultural Mechanization Research, 2022, 44 (10): 25-29
[17] Yoshida T, Kawahara T, Fukao T. Fruit recognition method for a harvesting robot with RGB-D cameras [J]. Robomech J, 2022
[18] Yang X. Application of image processing in path planning of kiwifruit picking robot [J]. Journal of Agricultural Mechanization Research, 2021, 43 (04): 37-41
[19] Zhang X. Research on autonomous navigation technology of picking robot based on image fusion technology [J]. Agricultural Mechanization Research, 2020, 42 (11): 193-196
[20] Cui Y, Wang Y, He Z, et al. Global path planning of kiwifruit picking robot based on improved RRT algorithm [J]. Transactions of the Society for Agricultural Machinery, 2022, 53 (06): 151-158
[21] Xuan F, Zhang X. Obstacle detection and path planning of picking robot based on monocular vision [J]. Research on Agricultural Mechanization, 2021, 43 (11): 29-33
[22] Hu G, Kong W, Qi C, et al. Optimization of navigation path of mobile picking robot in orchard environment [J]. Transactions of the Society of Agricultural Engineering, 2021, 37 (09): 175-184