Design of Flexible Spherical Fruit and Vegetable Picking End-effector Based on Vision Recognition

Meng Chen¹, Feng Chen¹*, Wen Zhou¹ and Ruoyu Zuo¹

¹College of Mechanical Engineering, Anhui Science and Technology University, ChuZhou, AnHui, 233100, China

*Corresponding author’s e-mail: chenf@ahstu.edu.cn

Abstract: In order to solve the problems of poor versatility and high clamping damage rate faced by the end-effector of picking robots in today's picking operation, a universal spherical fruit and vegetable picking die end-effector based on vision recognition with adaptive flexible force clamping is designed. This end-effector is a pneumatic three structure, while integrating a small air source device, in the control of a variety of sensors, through the fusion of multiple sensing and algorithm improvement and other ways, so that the mold end-effector in the picking operation, through visual recognition to determine the type of fruit and vegetables, and then provide the optimal gripping force. At the same time, the fingertip is equipped with a pressure sensor to adjust the pressure output in real time to achieve the overall picking process, the pressure is kept constant, and in the judgment of the completion of fruit and vegetable picking, a torque sensor is equipped to presume the completion of the picking operation by twisting the fruit strength. Finally, the actual product was made and several fruit and vegetable simulation experiments were carried out. The experimental structure is good, which verifies that the end-effector has good versatility and flexibility.

1. Introduction

With the deepening of the intelligent process and the vigorous development of facility agriculture, robots have become inevitable instead of manual labor[1,2]. Nowadays, facility agriculture is developing rapidly, and the vigorous development of facility agriculture can not only improve the efficiency of agricultural production, but also promote the process of agricultural modernization, but in the development of facility agriculture, it faces many problems, the first is the massive reduction of labor[3,4]. In response to this problem, countries all over the world are solving this problem from the perspective of intelligence, and most of the various aspects of the operation of facility agriculture have been automated. However, in the harvesting of fruits and vegetables, due to the complexity of the operating environment and the variety of operating objects, it is still not possible to industrialize the research on picking robots, and picking robots are still a difficult research point and research hotspot in the world[5,6,7].

As the most important part of the picking robot, there are many kinds of end-effectors of picking robots, and according to their driving methods, there are pneumatic, linear, SMA and EAP driving methods, etc[8,9]. Although the end-effectors have been researched for a long time and developed in a wide range of types, they are still faced with a large rate of picking damage, poor versatility, low success rate of picking, etc. In general, the end-effectors of picking are still facing many problems in the process of industrialization[10,11,12].

The author designed a pneumatic three-finger end-effector, the die end-effector comes with a small
air source generation device, while integrating a variety of sensors, through the use of a variety of sensors to cooperate with the way to make the overall device more flexible. At the same time in the design of the use of visual recognition, through the judgment of the type of fruit and vegetables to provide clamping force, so that the end-effector in the picking operation, greatly reducing the clamping damage, but also applicable to a variety of oval-shaped fruits and vegetables, to solve the picking end-effector poor versatility, clamping damage serious problems.

2. Overall structure and working principle

The overall structure of this picking jaw is shown in Figure 1. The overall consists of airbag jaws, bracket, compression pump, solenoid valve, motor bracket and coding motor, and a variety of sensors. The overall working process is simple, the coding motor rotation, through the cam driven piston movement, making the compression pump work, the compression cylinder has a pair of reverse one-way valve, to achieve the inflatable port inflatable, the inlet port suction, filling and deflating port through the two-position four-way solenoid valve and the jaws air bag connection, suction when the air bag deformation is open state, when the clamping solenoid valve shift, the air bag inflatable, is bent wrapped state, to achieve flexible bionic clamping.

![Figure 1. Overall structure diagram](image)

3. Control part design

3.1. Control principle

After the structure analysis is completed, a multi-sensor fusion, as well as algorithmic control, fast and non-destructive picking control principle is designed, as shown in the control structure block diagram 2. The overall consists of four sensors: vision sensor, air pressure sensor, pressure sensor and torque sensor composed of the input, and the motor compression cylinder, solenoid valve, and the operation of the robot arm end rotation vice composed of the output, and the processor. The overall control principle is that after the vision module collects the operating image, the processor detects what kind of fruits and vegetables, and after a pressure value is given, the motor compression cylinder is controlled to work and work to quickly reach this given value. At the same time, the pressure sensor to real-time picking force collection, and feedback to the motor rotation, and solenoid valve charging and deflating action, to keep the clamping force to maintain stability, so that the output force is not too large to cause clamping damage, but also to provide a stable force to prevent the picking process of fruit and vegetables off. Here the role of air pressure sensor is to protect the jaws, to prevent overpressure damage, torque sensor to determine whether the picking process between the fruit stalk screwing action is complete.
3.2. Visual recognition module design

First is the front-end visual sensing of fruit and vegetable recognition judgment, this experiment through deep learning to build a fruit and vegetable recognition library, with the YOLOV3 framework built for its training, to achieve the purpose of fruit and vegetable recognition, due to the laboratory environment shooting, fruit and vegetable features are obvious, and fruit and vegetable training library training volume is large, making the overall recognition success rate is high[13,14].

The YOLOv3 target recognition algorithm used in this study is an improved algorithm based on the YOLOv2 algorithm, which makes some adaptive improvements to YOLOv2 in terms of multi-scale recognition, and multi-label classification. It also uses the DarkNet-53 network based on the ResNet network as the feature extractor, making YOLO further improved in the field of visual recognition, and it has become one of the more widely used target recognition algorithms at present[15].

The DarkNet-53 network structure of the YOLOv3 algorithm is shown in Figure 3. The image is first scaled or cropped to a size of 416×416 pixels size during training, using the scale pyramid structure of the FPN network. The later DarkNet-53 feature extraction grid will divide the original image into S × S cells of equal size according to the size of the feature map, where for the feature size size there are three of 13 × 13, 26 × 26, and 52 × 52 sizes. Later, the fusion of shallow sub-features and deep features will be used to obtain more discriminative deep features, where in the regression prediction part, each cell predicts 3 borders with the help of 3 anchor boxes. Meanwhile, YOLOv3 uses logistic regression to predict the probability of containing objects in the anchor frames by judging the magnitude of the overlap between the anchor frames and the real target borders, and YOLOv3

![Figure 2. Overall control block diagram](image)

![Figure 3. The DarkNet-53 network structure of the YOLOv3 algorithm](image)
assigns only one anchor frame to an object during training. Also YOLOv3 uses binary cross-first loss and logistic regression for category prediction during training, and the use of such parties makes it possible for YOLOv3 to classify a target with multiple labels. For target recognition the method is shown in flowchart 4, and the target recognition experiments will be performed below.

Figure 4. Flow chart of the CDSP -YOLO target identification method

The overall training process is to first build a dataset for the detected fruits and vegetables. There are more datasets for fruit and vegetable recognition detection, but there are fewer publicly available datasets and there are not many kinds of fruits and vegetables, so this time, we mainly rely on ourselves to build up the dataset for fruits and vegetables. For this design of five kinds of fruit and vegetable recognition and optimal force clamping on, currently only for the five kinds of fruit and vegetables in the text for image acquisition build, later can be increased by increasing the image set and training, can be analogous to any ellipsoidal fruit and vegetable for applicable recognition picking. In terms of image acquisition for these five fruits and vegetables, since they are more common fruits and vegetables, the image sources for this time mainly rely on two ways: cell phone shooting and online downloading. At the same time, a large amount of data is needed to train the deep neural network, and the increase of training data can significantly improve the recognition efficiency. This time, we also processed the images through image processing techniques, such as flipping, scaling and adding noise, to increase the number of data sets. Finally, 2000 images were obtained, and the images were compressed and cropped into a uniform size of 416×416, and the image gallery is partially displayed as shown in Figure 5.

Figure 5. Pictures of the picture gallery
After the uniform size modification of the established image library, the following will use the data annotation tool labeling to label all the images according to the annotation format of the PASCAL VOC dataset to generate XML type annotation files. For the labeling part of this fruit and vegetable labeling library, all of them are manually labeled. The labeling process is shown in Figure 6, where five types of fruits and vegetables are defined as: apple, pear, orange, tomato and kiwi. Finally, the XML files generated for all the annotated images are put into a folder, and the captured and processed images will be trained below.

The Pytorch framework was used to build the network, and the training was conducted on a lab workstation considering the speed and stability of the training. The workstation is equipped with an Intel i9-9900K @3.6GHz CPU and RTX2080 SUPER graphics card, and the training is run on the GPU. The YOLO algorithm underwent several iterations during training, which took about 2 h in total. After the training was completed, the overall learning results were tested, and the results of the test are shown in Figure 7. The recognition rate for fruits and vegetables with obvious features and no obstructions was kept within 0.98–1. For those with obstructions or fruit mutilations, the recognition was also effective, but the recognition success rate was greatly reduced. In general, this visual recognition part can meet certain usage requirements, but it is still necessary to expand the training library and improve the accuracy of visual recognition continuously by learning from a large number of data sets.
3.3. Multi-sensing fusion control

After the visual recognition of fruit and vegetable categories is completed, the image information is passed to the lower computer TMS320F28069 through the vision controller, and then the lower computer controls the executive side to complete the overall clamping and picking operation, which is completed under the control of multi-sensor fusion. The overall structure block diagram is shown in Figure 8, because in consideration of the later experiments and the use of visual recognition and robotic arm mounted, this control architecture consists of the upper PC, and the TMS320F28069 lower end, and the execution of the acquisition feedback and other modules[16].
One of the core links in this whole control system is the motor and solenoid valve control system under multi-sensing fusion. One of the control logic about the motor, as shown in the motor control system structure diagram in Figure 9. The overall by two closed-loop feedback, the construction of the picking dexterous hand executive action system, the first is under the protection of the air pressure sensor, pressure feedback adjustment system, through a given pressure value and finger picking fruit real-time measurement to compare, to control the work of the inflatable pump motor, and control the opening and closing of the solenoid valve to achieve charging and releasing air pressure, to the purpose of pressure adjustment when picking. In the first link feedback adjustment is completed, the output can be executed screwing or other picking action signal, while the end of the mechanical arm screwing action, is completed under the feedback adjustment control of the torque sensor, through the torque sensor data collection value, to determine the completion of the screwing action, so as to feedback adjustment end execution action. When the overall execution is completed, the final output picking completion signal, for transmission feedback to perform other subsequent work, such as mechanical arm return collection, other action judgment, etc., as this design is mainly for the end of the actuator design, the control loop only value design of the end of the action of this part of the content.

The control system design and the program development and writing of the burn-in were carried out later for the control of the end-effector body, combined with the control hardware diagram and the schematic diagram. This program development and writing mainly relies on MATLAB/Simulink, the module diagram environment. This design also applied Code Composer Studio (CCS) code debugger software, code design suite. Among them, for the overall development process of universal flexible
picking jaws, after the hardware selection is completed and the development environment is built, its main link is simulink control model building, generating C code and importing it into Code Composer Studio for consulting and debugging, and finally downloading it to the control board for observation and debugging.

The main control board is STMS320F28069, and the motor driver module is mainly composed of TB6612FNG driver chip. The VM is connected to the motor drive voltage, VCC is connected to the chip's operating voltage, PWM and IN1, IN2 pins are connected to the GPIO port of the main control board, and the main control board controls the speed of the motor by changing the duty cycle to the PWM pins, and the IN1, IN2 pins control the forward and reverse rotation of the motor. Later, the Simulink control system design will be carried out for the control of the end-effector body, combined with the control hardware diagram and schematic diagram. The modular graphical programming is shown in Figure 10, where the final graphical programming diagram of the motor control under multi-sensing fusion is shown on the left, and the burning and C code generation diagram is shown on the right[17].

![Figure 10. Graphical programming](image1)

After the completion of the graphical programming can be downloaded directly to the controller, it can be directly generated C code file which can be opened and modified with other software and compiled and burned, here we used simulink direct burning, and in CCS to edit and download to the main control board in two ways, where the hardware physical and operational part of the process is shown in Figure 11. During the experiment, we can observe that the coded motor is able to rotate, which indicates that the pin parameters are assigned and the parameters are correctly selected in the program. It is also possible to change the rotation state of the motor when the pressure sensor intervenes, and to realize the phenomenon of institutional rotation when the programmed value is reached.

![Figure 11. Programming and hardware debugging](image2)
4. Physical fabrication and related experiments

After the overall structure design and control part design is completed, the overall structure is physically fabricated, and the overall control principle, the selection of the use of relevant sensors, and for the multi-sensor signal acquisition and processing, and the feedback to the overall control process of the relevant execution unit, the relevant program writing. The actual production process is shown in Figure 12, the prototype support part using 3D printed plastic parts, the overall small and flexible, the overall weight of 1Kg.

4.1. Related experiments
At the same time, several sets of experiments were conducted to verify whether the device could provide the optimal picking force required for picking different fruits and vegetables. The experiments were conducted to observe whether the device could provide the optimal force for different fruits and vegetables.

During the experiment, the visual recognition is followed by signal communication, and the end-effector responds quickly by ramping up quickly to achieve the preset force for a given fruit, the value of which is the value of the optimal clamping force interval obtained above. This time, for the above fruits and vegetables in turn, a number of clamping experiments were conducted, for different fruits and vegetables, the clamping force can be maintained at a stable level near the given value, but due to the hysteresis of the pressure signal acquisition, so that the air pressure continues to fill for a period of time, so that the actual clamping maximum pressure value is slightly greater than the set value.

4.2. Control optimization
In the experiment, we found that the clamping force will be slightly greater than the given value, while for the picking operation, the pressure value will change in the state of dragging or screwing, and the
device will not operate after the initial reach the given pressure value of the contradiction in the algorithm optimization improvement. First of all, real-time acquisition is carried out, and because the collected pressure is more floating, a fuzzy algorithm is added to the control to fuzzify the collected force, so that it can quickly reach a given interval and remain stable in this pressure interval. This makes the clamping force of the end-effector stable within the preset safety interval during the whole harvesting process, which not only reduces the clamping damage on the hardware material but also realizes the purpose of reducing the clamping damage on the control[18].

5 Experimental conclusion
When clamping fruits and vegetables, it is able to recognize them and provide corresponding forces, while the clamping jaws produce better adaptive wrapping, and because of the wrapping feature, the force of the clamping jaws on the fruits and vegetables is greatly reduced. Overall, the dexterous hand is not only able to clamp the fruit without damage, but also produces a smaller contact force with a smaller air pressure input, while providing a better dragging force, which not only completes the function of non-destructive clamping of multiple fruits and vegetables, but also provides support for the non-destructive picking operation of multiple fruits and vegetables. At the same time the device is in good operating condition, can achieve positive and negative pressure without external air source, to achieve control of silicone finger bending and opening action, the corresponding speed is faster, within 2s to achieve the opening action completed, while in the clamping, within 3s to make the cavity to 30KPa air pressure.

Reference:
[1]Clapp J, Newell P, Brent ZW (2018) The global political economy of climate change, agriculture and food systems. J Peasant Stud 45(1):80–88
[2]Jiang M, Hu X, Changa J, Lin Z, Fei R (2020) Does the popularization of agricultural mechanization improve energy-environment performance in China’s agricultural sector? J Clin Prod 276:124210
[3]Yu Z, Xiao L, Chen X et al., (2018). Spatial restructuring and land consolidation of urban-rural settlement in mountainous areas based on ecological niche perspective. Journal of Geographical Sciences, 28(2): 131–151.
[4]Liu Y, Liu Y, Chen Y et al., (2010). The process and driving forces of rural hollowing in China under rapid urbanization. Journal of Geographical Sciences, 20(6): 876–888.
[5]Bac, C. W., van Henten, E. J., Hemming, J., & Edan, Y. (2014). Harvesting robots for high-value crops: State-of-the-art review and challenges ahead. Journal of Field Robotics,31(6), 888–911.
[6]Bergerman, M., Billingsley, J., Reid, J., & van Henten, E. (2016). Robotics in agriculture and forestry. In B. Siciliano, & O. Khatib (Eds.), Springer handbook of robotics (pp. 1463–1492). Cham, Switzerland: Springer.
[7]Clarke, A. (2017). Robotics set to increase farm productivity. Farmers Weekly,1160(May 12th), 72–73.
[8]Tang Zuoliang, Xu Lijia, Wang Yuchao & Xie Hong(2021). Collision-Free Motion Planning of a Six-Link Manipulator Used in a Citrus Picking Robot. Applied Sciences, 11(23), 11336-11336.
[9]Kafiev I & Romanova I(2021). Application of a Mobile Robot for Picking Berries Under Qualitative Uncertainty of Conditions. Journal of Physics: Conference Series, 2096(1).
[10]Zahid Azlan, Mahmud Md Sultan, He Long, Heinemann Paul & Schupp James(2021). Technological advancements towards developing a robotic pruner for apple trees: A review. Computers and Electronics in Agriculture, 189.
[11]Miguel Angel Funes Lora, Edgar Alfredo Portilla-Flores, Raul Rivera Blas & Manuel Faraón Carbajal Romero(2018). Metaheuristic techniques comparison to optimize robotic end-effector behavior and its workspace. International Journal of Advanced Robotic Systems,
[12] Fei Feng, LiNa Tang, JianFeng Xu & YiWei Liu (2016). A review of the end-effector of large space manipulator with capabilities of misalignment tolerance and soft capture. Science China Technological Sciences, 59(11), 1621-1638.

[13] Zhi Feng Xu, Rui Sheng Jia, Hong Mei Sun & Zhe Cui (2020). Light-YOLOv3: fast method for detecting green mangoes in complex scenes using picking robots. Applied Intelligence, 50(prepublish), 1-18.

[14] Chen Jiqing, Wang Zhikui, Wu Jiahua, Hu Qiang, Zhao Chaoyang, Tan Chengzhi & Luo Tian (2021). An improved Yolov3 based on dual path network for cherry tomatoes detection. Journal of Food Process Engineering, 44(10).

[15] Xu Wenkai, Zhao Longgang, Li Juan, Shang Shuqi & Wang Tiewei (2022). Detection and classification of tea buds based on deep learning. Computers and Electronics in Agriculture, 192

[16] Sengupta Ipsita, Gupta Sagar & Ozana Stepan (2021). Dynamic Stability of an Electric Monowheel System Using LQG-Based Adaptive Control. Applied Sciences, 11(20), 9766-9766.

[17] & Chiu ChunKai (2021). Adaptive Fuzzy Sliding Mode Control of Omnidirectional Mobile Robots with Prescribed Performance. Processes, 9(12), 2211-2211.

[18] Wei Xiaozhao, Wang Lei, Ni Xiangdong, Han Shuangman & Li Shen (2021). Speed control strategy for pump-motor hydraulic transmission subsystem in hydro-mechanical continuously variable transmission. Journal of Mechanical Science and Technology, 35(12), 5665-5679.