Research Review on Flexible Manipulator Picking Based on Machine Vision

Li Shen
School of Information Science and Engineering, Shandong Agricultural Engineering College, Zibo, 255300, China
lishen_work@outlook.com

ABSTRACT

With the continuous advancement of agricultural intelligence, greenhouse agriculture has become a core area of development, and automated harvesting is a key requirement for improving its production efficiency. However, the complexity and variability of the greenhouse environment and the insufficient adaptability of traditional rigid robotic arms have led to increasing attention to research on harvesting using flexible robotic arms based on machine vision. This paper reviews the key technologies and research progress of machine vision and flexible robotic arms. In the machine vision section, most current work focuses on fruit target detection and localization, improving image algorithms or innovating feature extraction methods, and even introducing deep learning methods to improve the target detection rate and robustness of fruits under different lighting conditions and occlusion conditions. For the flexible robotic arm itself, researchers. Improvements are mainly made from two perspectives: motion planning and control methods, to better cope with the complex working environment and constraints inside greenhouses and minimize damage to plants and fruits during operation. Although progress has been made, many problems in this regard still need to be solved. For example, it is necessary to make the vision system more robust, and it is difficult to dynamically model the flexible robot arm, so the control accuracy is not good enough. In the future, technologies such as multi-modal large-scale models and digital twins can be used to mix information from multiple sensors. This can improve the adaptability of the sensing system and the intelligent control level of the robot arm. If we summarize and summarize what we have found, it can provide insights and guide future research.

Keywords: Complex Greenhouse Environment; Machine Vision; Flexible Robotic Arm; Harvesting Research; Compliant Control.

1. INTRODUCTION

In recent years, the focus of global agricultural development has gradually turned to intelligence. This is the only way to solve the shortage of migrant workers and crop failure, and it is also the only way to promote the sustainable development of agriculture. Zhao Chunjiang and others believe that agricultural robot is an important tool of intelligence, and the development of agricultural robot technology directly affects the development of modern agriculture in China [1]. Green house is an important tool for producing agricultural facilities. They are widely used in France and all over the world because they can provide good growth conditions and year-round production. However, the special environment in the garden also brings some difficulties to fruit harvesting. Compared with open fields, uneven lighting usually occurs in green houses. The combination of natural light and artificial lighting can easily lead to overexposure or shadow of the captured image. In addition, the crops in the garden are usually densely planted, and the distance between plants is very small, which will affect the movement of harvesting equipment. In addition, the blockage of branches and leaves will also affect the accuracy of target detection. The traditional rigid harvesting robot is difficult to harvest accurately in a small range because of its limited range of motion and hard contact characteristics. They may also damage fruits or plants in the process of operation, which does not meet the requirements of non-destructive and flexible harvesting of green space. At present, it is the main bottleneck to prevent the industrialization of greenhouse harvesting robot [1].

How to achieve efficient and automatic harvesting of plants in a greenhouse where everything is chaotic? This paper introduces some important technologies developed in recent years to solve this problem, such as machine vision and flexible robot weapons. Flexible robot arms, because they are flexible, can easily change positions in a small space, and do little harm to plants during harvesting. Machine vision technology can accurately detect the position of fruit and provide the position information to the robot arm so that it can work [2]. By combining these two technologies, we can solve the problems of the old hard harvesting machinery, meet the needs of greenhouse harvesting and help the development of greenhouse agriculture.

With the rapid development of intelligent agricultural technology and the wide application of flexible robots in agriculture, the related research results have shifted from the preliminary research of single technology to the exploration of system integration and specific application scenarios. This paper mainly focuses on the relationship between “complex environment in greenhouse-flexible manipulator-fruit picking”, and summarizes the following three aspects: the first part is about fruit identification and detection, and multi-sensor fusion based on machine vision; The second part is the modeling and motion planning of the flexible robot arm, the complex control and the design of the end effector; The third part is the overall structure of the acquisition system and its application. After reviewing the above contents, we can determine the current gap in this field and the direction that needs further research, thus laying the foundation and pointing out the direction for future research.

2. OVERVIEW OF CORE THEORETICAL FOUNDATIONS

In order to fully understand how this field has evolved, we need to understand these basic theories before we understand them. According to the relevant literature reviewed so far, the relevant theoretical research mainly focuses on environmental perception in complex greenhouses, fruit target detection, dynamics and control of flexible robotic arms, which are the basic theoretical basis for intelligent picking.

2.1 Complex Greenhouse Environment Perception Theory

The complexity of the greenhouse environment is the first and most fundamental problem faced by harvesting robots, namely, the uncertainty of light changes, fruit shading, and uneven spatial distribution of crops. Current research on these problems mainly focuses on three aspects, establishing the foundation for greenhouse environment perception methods [3,4].

The first is the light transmission model, which is the key theoretical basis for solving the problem of uneven greenhouse light. Sun Zhiqiang et al. established the influence of fog droplets and shading on the transmission of different wavelengths of light based on a combination of natural light and artificial supplementary lighting, thus providing a spectral reference for visual algorithms[5]; Zhang Xiaodong et al. further proposed a simulation method based on Monte Carlo light transmission using multispectral vision, which can effectively eliminate image brightness changes caused by light changes, improving fruit detection accuracy by 22%[6].

Secondly, there is the occlusion geometric probability model. In cases where dense planting of greenhouse crops causes mutual occlusion between fruits, Zhao Chunjiang et al. proposed a method to estimate the probability of occlusion based on crop morphological characteristics. By studying the relative positions of branches, leaves, and fruits, they obtained an image showing the probability of occlusion. Based on this, the visual recognition system can be more inclined to find areas less likely to be occluded, improving detection accuracy and reducing the number of missed fruits [1].

Finally, the crop spatial distribution field model combines 3D point clouds and crop growth models to obtain a spatial distribution field that represents the position and density of fruits, branches, and leaves in 3D space, which is beneficial for global path planning by robotic arms [7]. These three models combined constitute a stochastic perception method, which can transform the uncertainty in greenhouses into a measurable probability distribution. This provides a clear environmental prior constraint for subsequent visual algorithms and robotic arm control, and is currently the main approach for greenhouse environmental perception.

2.2 Machine Vision Target Recognition Theory

Based on the theory of perception of complex environments, machine vision systems need to develop target recognition methods that can overcome the effects of lighting and occlusion. Accurate detection of fruit targets is the foundation for automated harvesting. Research in this area has been working to improve the detection rate and robustness of fruit targets under various conditions. Early target detection methods mainly relied on combining manually extracted features, using color, shape, and texture features to distinguish between fruit and background. This method can achieve good results in environments with uniform lighting and no obstacles. However, in greenhouses, due to the large changes in lighting and the occlusion caused by branches and leaves, the features become unstable and the robustness is poor [8].

With the development of deep learning technology, the methods for fruit target recognition have been revolutionized, ushering in a new era of end-to-end learning in recognition research [9]. From the initial Convolutional Neural Network

(CNN) to the later Faster R-CNN two-stage detection method, the YOLO series of single-stage detection methods, and methods based on the Transformer attention mechanism [10], people have improved the network to learn features and achieve better results [11]. The YOLO series, due to its good accuracy and speed, is widely used in actual harvesting [12]. The Transformer, using a self-attention mechanism, can better model the global contextual relationships between dense fruits and has better robustness in situations with significant occlusion [13].

In addition to the above reasons, meta-learning and domain transfer are also widely used in greenhouses due to the limited labeled data and the difficulty in adapting different types of fruit [9]. Meta-learning adopts a “task-meta-knowledge” learning approach, which can use a small amount of labeled data to quickly adapt to new fruit types [14]; while domain adaptation can make the data distribution between the training set and the test set more similar, thereby reducing the decrease in model accuracy caused by factors such as lighting and occlusion, and improving the robustness of the classifier [15].

2.3 Dynamics and Control Theory of Flexible Robotic Arms

Based on the perception of complex greenhouse environments and accurate fruit recognition, flexible robotic arms need to establish adaptive unstructured dynamic models and control strategies to achieve efficient and non-destructive harvesting operations. The precise control of flexible robot arm needs to establish an accurate dynamic model and adopt a reasonable control method. This is an important basis for successful non-destructive harvest [16]. For dynamic modeling, the two most commonly used methods are Euler-Bernoulli and Timoshenko. Euler-Bernoulli assumes that the link is always perpendicular to the neutral axis, regardless of shear deformation and rotational inertia. This simplifies the problem theoretically. It provides a good approximation for the low frequency vibration of thin flexible arm. However, in greenhouse harvesting, the flexible arm must start, stop and change direction quickly, which is easy to produce higher frequency vibration. In these cases, Euler-Bernoulli will have a significant deviation [17]. Timoshenko model adds correction terms to shear deformation and rotational inertia. It more accurately reflects the dynamic behavior of the arm under the condition of short arm and high frequency vibration, making it more suitable for the needs of greenhouse harvest. This is a widely used modeling method in this field [18].

Complex control is an important research point to meet the demand of nondestructive recovery. At present, there are three main methods: impedance control, mixed force position control and model predictive control. Impedance control realizes complex contact with fruit by changing the stiffness, damping and other impedance parameters of robot arm. It is suitable for large-scale fruit harvesting with known characteristics. However, when fruit maturity and variety are different, fixed impedance control can easily lead to fruit damage or harvest failure [19]. Hybrid force-position control separates position control from force control, and applies force in different directions respectively. This is helpful to improve the accuracy of harvesting position, and at the same time ensure the consistency of force, making it more suitable for the operating environment of dense crops [20]. Model predictive control uses online optimization to predict the motion path of the arm, avoid collision with obstacles, and ensure safe operation in a closed space, but it requires a lot of calculation and high hardware requirements [21].

3. RESEARCH PROGRESS OF CORE TECHNOLOGIES

Based on the above theory, researchers at home and abroad have done a lot of research on visual perception and control of flexible robot arms. They have made some achievements and put forward solutions suitable for the greenhouse environment.

3.1 Research Progress in Machine Vision Technology

3.1.1 Optimization of Fruit Target Recognition Algorithm

To address the problem of low recognition rates caused by dense fruit shading and uneven lighting in greenhouses, researchers mainly rely on existing mainstream deep learning methods, and then improve them in combination with the greenhouse environment.

For instance segmentation models, Chen Jing et al. improved the RoIAlign module of the Mask R-CNN model to address the problem of dense occlusion in greenhouse tomatoes. They changed the sampling method to improve the segmentation effect of overlapping fruit edges. The improved model achieved a mAP of 95.45% on the tomato dataset, an improvement of 2.29 percentage points compared to the original version, which is beneficial for robotic arms to grasp more accurate fruit contour information [22]. This effectively solves the problem of excessive computation caused by deep neural networks repeatedly learning the same gradients, and further improves the detection accuracy of the model while reducing the computational load. The improved Mask R-CNN structure is shown in Figure 1.

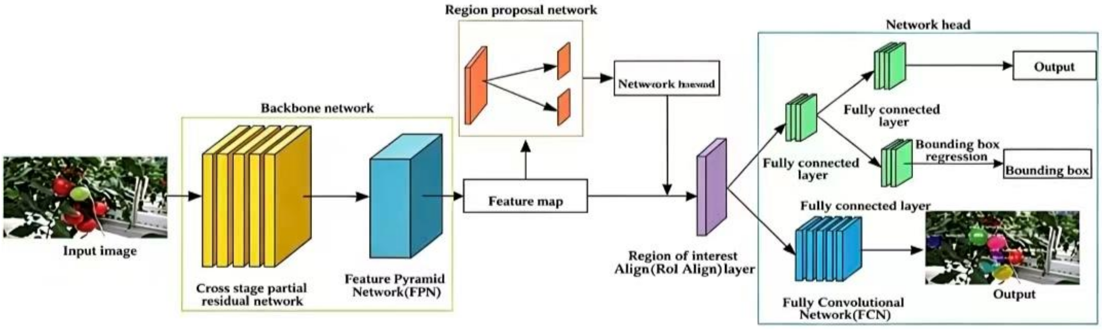

Figure 1. Improved Mask R-CNN Network Structure (This figure is adapted from Ref [22]).

Regarding real-time detection models, Xia Xianfei et al. summarized that, due to the high real-time requirements of harvesting operations, some scholars have improved the YOLO series to achieve effective detection of fruits and vegetables: by adjusting the network structure and loss function, the detection speed is improved while maintaining good detection performance, thereby achieving better real-time performance and accuracy to meet the needs of online harvesting [2].

For severe occlusion situations, the Transformer model is widely used due to its advantages. Fu Lanhui et al. proposed a multi-size robot grasping system MSGS based on a depth camera. It uses depth image acquisition and related processing methods to identify and locate objects of different sizes, which can effectively avoid interference caused by complex environments. The entire system has good timeliness and robustness to meet the requirements of the production line[23]. The experimental platform is shown in Figure 2.

Figure 2. MSGS Experimental Platform (This figure is adapted from Ref [23]).

In addition, attention mechanism is also the key technology to improve the robustness of the model. After calculating the relationship between different features, the important features in the target area are given more weight. This can effectively eliminate the interference of leaves or light changes. Zhang Xiaodong et al. used multispectral vision to identify physiological stress and diseases of plants. By studying and combining various spectral information, they improved the accuracy and robustness of plant identification. It has played an important role in agriculture [6].

3.1.2 Precise Fruit Positioning and Multi-Sensor Fusion Technology

After fruit identification, accurate spatial positioning is very important for harvest. At present, there are three main positioning methods: monocular depth estimation, binocular parallax positioning and RGB-D point cloud recording.

The monocular depth estimation method only needs one RGB image to obtain the corresponding depth. They are easy to build and low cost, which makes them suitable for light harvesting tasks. However, because a single image does not contain much information, the results of depth estimation may have great errors, for example, when the skin of fruit is smooth or the light changes greatly, which makes it difficult to achieve centimeter-level positioning accuracy [24]. Binocular parallax positioning uses the principle of stereoscopic vision; It obtains the depth value by calculating the distance between the corresponding points in the left and right images. It can achieve centimeter-level positioning in areas with a lot of textures, but there are often matching problems in areas with occlusion or less textures. At the same time, stereo matching needs a lot of calculation and powerful hardware [25]. The RGB-D point cloud recording scheme combines the semantic information of RGB images and the 3D point cloud data of depth camera. It can get the exact location of the fruit. The positioning error is usually 2cm, which is quite firm when it is blocked or the light changes. This is the most commonly used method in all kinds of harvesting robots [26].

A single visual sensor is insufficient to meet the all-round needs of greenhouses. In situations such as nighttime supplemental lighting or foggy weather, a single sensor is prone to perception failure. Therefore, multi-sensor fusion has become the main means to improve overall robustness. Current research mainly utilizes the advantages of visible light, multispectral, and lidar sensors to form a heterogeneous fusion perception method: visible light cameras can acquire appearance attributes such as fruit color and texture; multispectral sensors can obtain plant growth status through obstacles; and lidar can provide high-precision three-dimensional coordinate point cloud data under no-light conditions. The combination of the three sensors can cover the perception requirements under various conditions [27-30].

For multi-sensor data fusion, Kalman filtering and factor graph optimization are currently commonly used methods.

Kalman filtering can fuse observation information from multiple sensors in a timely manner, continuously update the fruit position, remove the influence of noise, and has good robustness in mobile harvesting [31]; while factor graph optimization achieves the goal of consistency optimization among multiple sensors by minimizing the overall system error, thereby improving positioning accuracy [32]. Combining the two can balance positioning speed and accuracy to a certain extent. In some harsh environments, such as low light or with a large amount of water droplets, the method of using visible light, multispectral, and lidar data for mixed filtering optimization can improve the fruit detection rate to 12%-18%, which effectively solves the problem of perception failure [27].

Through multi-sensor fusion, the robustness of the vision system has been significantly improved, initially solving the problem of perception failure of a single sensor under changes in light or severe weather, and providing more reliable environmental information for the precise operation of the robotic arm.

3.2 Advances in Flexible Robotic Arm Harvesting Technology

3.2.1 Motion Planning Algorithm Optimization

The limited space caused by the dense planting of greenhouse crops places high demands on the motion planning of flexible robotic arms. They must not only avoid obstacles such as stems and leaves, but also ensure a smooth path to reduce robotic arm vibration, and have a high harvesting speed. To this end, people have designed many different methods to solve the above problems.

Zhang Yali et al. proposed an improved version of the APF-FMT* algorithm, which retains the obstacle avoidance based on the Artificial Potential Field (APF) method while combining the large-scale search characteristics of the Fast Moving Tree (FMT*) algorithm. By varying the distance of the repulsive field, it avoids the situation where the APF method is prone to getting trapped in local optima. It also improves the sampling method to narrow the search range, thereby accelerating the algorithm's convergence speed [33]. For simulation environments with narrow channels, its search time can be reduced by more than 75% compared to methods such as RRT*, and it can better solve the path planning problem of agricultural robots.

To solve the path jitter and excessive vibration of the flexible arm caused by the traditional RRT algorithm, researchers proposed an RRT-B-spline smoothing method. This method adds a B-spline smoothing component to the global path search of RRT to interpolate and optimize the initial path, improving path smoothness. It also adds a dynamic obstacle prediction function to avoid changes in crop position in a timely manner. In a cucumber greenhouse with a plant spacing of 0.8m, this algorithm reduces the time required by approximately 27% compared to the original RRT, by about 0.5s, and the vibration of the end effector is limited to less than or equal to 1mm, thus preventing fruit damage due to path jitter [34].

To address the requirement for collaborative work in multi-arm harvesting systems, Liu Jizhan et al. introduced a non-uniform B-spline method in multi-axis coordinated motion trajectory planning. Under constrained control of the vertices, they obtained collision prevention and high-order continuity of the flexible arm motion trajectory, which was then applied to multi-arm collaborative harvesting in greenhouses [35]. In the corresponding greenhouse experiments, the collision rate between the arms was very low, and the algorithm's running speed was also greatly improved, meeting the requirements of large-scale multi-arm collaborative harvesting in greenhouses [36].

It should be noted that the above three methods come from different studies, and their performance results are obtained under different experimental conditions. Specifically, the results of APF-FMT come from the simulation experiment in the narrow channel environment, the results of RRT-B spline method are obtained in the greenhouse experiment, and the results of multi-arm cooperative planning method verify that multiple manipulators work together in the experiment. Due to the differences in experimental conditions, such as the simulated and real-world environment, the distance between cultural queues, the degree of freedom of manipulators and task scenes, these results cannot be directly compared with each other. This paper only demonstrates the improvement of each method under its own conditions to show the influence of different optimization methods of flexible manipulator motion planning, without judging which one is better or worse.

3.2.2 Compliant Control Strategy and End-Effector Design

Compatibility control is the key technology to harvest fruits without destroying them. In practical application, impedance control with fixed parameters is difficult to adapt to the mechanical characteristics and maturity level of different types of fruits. If the impedance is too high, the fruit will be easily crushed, while if the impedance is too low, the harvest will fail. In order to solve these problems, Hong Kun and others proposed a self-tuning method of impedance parameters based on deep reinforcement learning. In this method, the adjustment of impedance parameters is regarded as an optimization problem in continuous action space. On this basis, a multi-objective reward function is defined, including the maximum contact force, the probability of successful acquisition and the mobility of action. Then the deep Q network is used to obtain the best strategy of online impedance parameters. Experiments show that this method can adjust the impedance value in real time according to the fruit condition, greatly reducing the maximum contact force exerted on the fruit during picking and reducing the peel damage rate. This effectively solves the problem of grasping fruits with unknown hardness flexibly [37].

At the same time, the sampling frequency of force feedback plays a very important role in the robustness of complex control systems. In the application of crop robot with forced feedback control, low sampling frequency will lead to large delay. This causes the phase margin of the system to become smaller, resulting in low-frequency oscillation and force tracking error. On the other hand, if the sampling frequency is too high, it will greatly increase the calculation workload and require more powerful hardware. The research results in frequency field show that when the sampling frequency of feedback force is less than 200Hz, the phase margin of the system will decrease rapidly when the frequency decreases, which makes the system prone to unstable vibration. When the sampling frequency is higher than 500Hz, the phase margin of the system can exceed 65, which is already beneficial to complex industrial control. If the sampling rate is further increased, the phase marginal gain will remain approximately the same, but the computational resources used will increase proportionally. Therefore, 200~500Hz is the best sampling range to balance system stability and calculation load, so as to realize high-precision force control during fruit harvesting [38].

The effective end is the part in contact with the fruit, and its design directly changes the success rate of picking and the rate of fruit damage. The current research mainly focuses on improving flexibility and adaptability. Zhao Yunwei and others created a flexible pneumatic finger based on Beloff structure. They established the relationship between pressure and deformation, and found how air pressure changed the bending angle and contact force of fingers. This can control the grip strength of fingers well. In the experiment of picking fruits and vegetables, this greatly reduced the damage to fruits and improved the success rate of picking [38,39]. Liu Jishan and others put forward a kind of end effector whose stiffness can be adjusted. It can automatically change the contact stiffness with the change of fruit hardness to adapt to the mechanical properties of different types of fruits. In the experiment of cucumber picking, this greatly reduced the damage rate of fruit [40]. Xia Xianfei et al. reviewed that underactuated grippers have good effects on picking spherical fruits such as tomatoes due to their simple structure and strong adaptability. For tomatoes with a diameter of 50-80mm, the recovery rate is over 95%, which is more suitable for large-scale harvesting [2].

Through the complex control strategy and the design of the effector, the adaptability of the robot arm to various mechanical properties of fruits is effectively improved, and the fruit damage rate in the picking process is greatly reduced. This better solves the great challenge of efficient nondestructive testing in greenhouse environment and provides key technical support for the practical application of automatic testing system.

4. RESEARCH CHALLENGES AND FUTURE DEVELOPMENT TRENDS

4.1 Core Challenges Faced by Existing Research

Although some achievements have been made in the current research, there are still three main problems from the perspective of realizing industrialization.

Firstly, the environmental robustness of machine vision system needs to be improved; At present, most recognition models are trained under standard laboratory conditions. Under the condition of extra illumination in greenhouse at night, the difference between artificial light source and natural spectrum will lead to serious domain shift, which will lead to a significant decrease in recognition accuracy.

Secondly, the control accuracy of the flexible manipulator needs to be improved, and the flexible manipulator is prone to strong nonlinear vibration when it starts and stops quickly and contacts with fruits. It is difficult to effectively eliminate these vibrations by current methods. The research of Zhao Yunwei and others shows that the vibration duration of the flexible arm is more than 0.5s after high-speed movement, and the end position deviation reaches 5mm, which is much greater than 1mm of the rigid robot arm. This leads to the inaccurate picking position and increases the possibility of damaging the fruit, which is the biggest problem affecting the picking efficiency.

Finally, the high energy consumption and high cost of the system hinder its industrial application. At present, the harvesting system used to realize the functions of visual recognition, path planning and real-time control consumes about 15 kWh of electricity when it runs continuously for 8 hours, far exceeding the expected ideal value of 8 kWh. This brings very high operating costs and affects the battery life of the mobile harvesting robot running on the battery. The total cost of modular vision and flexible robot arm accounts for more than 70% of the total equipment cost. The cost per unit is usually more than 100 thousand yuan, which is too expensive for most farmers, which has become a big problem in the large-scale use of this technology.

4.2 Future Research Trends and Prospects

In view of the above problems, according to the current research and technology development trend, the future development trend in this field mainly includes the following three aspects.

First of all, the zero-shot fruit recognition method based on multi-mode wide model is an important progress in the field of vision. Traditional methods need a lot of labeled data for training, and they can't deal with different types and conditions in greenhouses. However, different information, such as visible light, multi-spectral data and depth, can be combined by using multi-mode width model. This allows the establishment of knowledge transfer between different fields, so as to realize the zero-shot identification of fruits that have never been seen before, thus improving the robustness and generalization ability of the whole system. In addition, the attention mechanism in the large-scale model can automatically focus on the important part of the fruit, and even if there are a lot of jams, it can achieve high recognition accuracy.

Secondly, using digital twins for online motion planning can change the way of controlling flexible robot weapons. Because the greenhouse environment is always changing, the flexible arm must be able to respond very quickly and adjust its path within a few milliseconds. Online motion planning can establish a connection between the actual harvesting process and the virtual model, thanks to digital twins. People can see the path of the flexible arm in the virtual world in advance, find the possible collision and improve the path parameters. This saves a lot of time and money for testing. In addition, with the help of low-delay communication technology, the digital twin model can also receive information from the actual equipment very quickly. Therefore, it can change the movement of the flexible arm and harvest safely and flexibly in a complex environment.

Thirdly, the vibration reduction technology of flexible arm driven by intelligent materials can solve the problem of insufficient control accuracy of flexible arm. The vibration of about 20Hz generated by the flexible arm during harvesting is the main reason for its poor positioning accuracy. However, by integrating sensors and actuators, self-expanding smart materials can monitor their own hardness and then change the hardness to achieve active vibration reduction. According to related reports, the flexible arm made of this smart material can reduce the vibration at 20Hz by more than 40%, thus significantly improving the harvesting efficiency and positioning accuracy, prolonging the service life of the flexible arm and reducing the maintenance cost.

In a word, the combination of these three technologies can realize efficient and non-destructive greenhouse harvesting methods, fully meet the actual needs, solve the current problems, greatly reduce the equipment investment, and make the greenhouse harvesting robot more accurate, unattended and extensible.

REFERENCES

[1] Zhao Chunjiang, Fan Beibei, Li Jin & Feng Qingchun. (2023). Progress, challenges and trends of agricultural robot technology. Smart Agriculture (Chinese and English), 5(04), 1-15.

[2] Xia Xianfei, Jia Lei, Zhang Shuo, Gong Qingshuo, Lü Juntong, Shen Cheng & Liang Mingxu. (2026). Research status and development trend of fruit and vegetable picking robots. Transactions of the Chinese Society of Agricultural Engineering, 42(02), 14-28.
[3] Wang Cong, Ma Chengwei & Tong Guohong. Construction of a solar direct radiation shadow model in a solar greenhouse. Transactions of the Chinese Society of Agricultural Engineering, 1-8.
[4] Zhang Yufeng, Yang Jing, Deng Hanbing, Zhou Yuncheng & Miao Teng. (2024). Semantic segmentation model of greenhouse tomato image based on RGB and depth dual modality. Transactions of the Chinese Society of Agricultural Engineering, 40(02), 295-306.
[5] Sun Zhiqiang, Tang Senyang, Luo Xingfa, Dong Jiwei & Xu Ning. (2025). Research and application status of path planning for agricultural inspection robots. Agricultural Equipment and Vehicle Engineering, 63(01), 18-24.
[6] Zhang Xiaodong, Mao Hanping, Zuo Zhiyu, Gao Hongyan & Sun Jun. (2011). Diagnosis of water stress in rapeseed based on multispectral vision technology. Transactions of the Chinese Society of Agricultural Engineering, 27(03), 152-157.
[7] Rong Jiacheng, Xiong Ziming, Yuan Jiace, Li Wei, Wang Pengbo & Yuan Ting. (2026). Design and experiment of tomato picking robot with composite robotic arm configuration. Transactions of the Chinese Society of Agricultural Machinery, 57(05), 61-70+94.
[8] Ouyang An, Huang Huixing, Zhang Anqi, Wang Xiaochan, Zhang Yongnian, Wang Qingyu & Xu Jiajun. (2026). Research progress and prospects of fruit and vegetable picking robots. Transactions of the Chinese Society for Agricultural Machinery, 57(05), 19-38+49.
[9] Song Huaibo, Shang Yuying & He Dongjian. (2023). Research progress on deep learning recognition technology for fruit targets. Transactions of the Chinese Society of Agricultural Machinery, 54(01), 1-19.
[10] Ren Shaoqing, He Kaiming, Girshick Ross & Sun Jian. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE transactions on pattern analysis and machine intelligence,39(6),1137-1149.
[11] Li Hui, Zhang Jun, Yu Shuochen & Li Zhixin. (2025). Research progress on picking recognition algorithm of spherical fruit based on deep learning. Journal of Fruit Science, 42(02), 412-426.
[12] Zhang Zhen, Zhou Jun, Jiang Zizhen, & Han Hongqi. (2024). Apple recognition method in natural orchardenvironment based on improved YOLO v7 lightweight model. Transactions of the Chinese Society for Agricultural Machinery, 55(3), 231-242,262.
[13] Zheng, T., Zhu, Y., Liu, S., Li, Y., & Jiang, M. (2025). Detection of citrus in the natural environment using Dense-TRU-YOLO. International Journal of Agricultural and Biological Engineering, 18(1), 260-266.
[14] Finn, C., Abbeel, P., & Levine, S. (2017, July). Model-agnostic meta-learning for fast adaptation of deep networks. In International conference on machine learning (pp. 1126-1135). PMLR.
[15] Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... & Lempitsky, V. (2016). Domain-adversarial training of neural networks. Journal of machine learning research, 17(59), 1-35.
[16] Yang Min & Liu Keping. (2011). Research progress on dynamic modeling and control method of flexible robotic arm. Journal of Changchun University of Technology (Natural Science Edition), 32(01), 7-13.
[17] Zohoor, H., & Kakavand, F. (2013). Timoshenko versus Eulerâ Bernoulli beam theories for high speed two-link manipulator. Scientia Iranica, 20(1), 172-178.
[18] Liu Shihao, Wang Fei, Lin Mao, Liu Junxiao & Yao Keheng. (2024). Analysis and experiment on motion characteristics of mango harvesting robot arm. Transactions of the Chinese Society of Agricultural Engineering, 40(01), 58-67.
[19] Geng Xinlong & Zhang Xu. (2024). Research on end-effector compliant grasping force control method of harvesting robot under impedance model. Industrial Instrumentation & Automation Device, (04), 52-56.
[20] Zhang Yuxin, Guo Yue, Gao Yapeng, Wang Wanting & Li Haifang. (2026). Flexible control of apple pickers based on active disturbance rejection and model prediction. Transactions of the Chinese Society of Agricultural Engineering, 42(01), 37-47.
[21] Liu Zhe, Song Rui & Zou Tao. (2018). End-effector force tracking control algorithm of grinding robot based on model predictive control. Journal of Shandong University (Engineering Science), 48(01), 42-49.
[22] Long Jiehua, Zhao Chunjiang, Lin Sen, Guo Wenzhong, Wen Chaowu & Zhang Yu. (2021). Improved Mask R-CNN method for segmenting tomatoes of different maturity in greenhouse environment. Transactions of the Chinese Society of Agricultural Engineering, 37(18), 100-108.
[23] Fu Lanhui, Qiao Weilai, Deng Fuqin, et al. MSGS: Multi-specification robot grasping system based on depth camera[J]. Frontiers of Data & Computing, 2026, 8(01): 195-206.

[24] Zheng Taixiong, Jiang Mingzhe & Feng Mingchi. (2021). Review of vision-based target recognition and localization method of harvesting robot. Journal of Instrumentation, 42(09), 28-51.
[25] Yang Xuejun, Zhong Jinbiao, Lin Kaiyan, Wu Junhui, Chen Jie & Si Huiping. (2025). Research progress on binocular stereo vision technology and its application in smart agriculture. Transactions of the Chinese Society of Agricultural Engineering, 41(01), 27-39.
[26] Fan Xiangpeng, Zhang Yanqi, Zhou Shuo, Ren Meifei, Wang Yuwei & Chai Xiujuan. (2025). Identification and localization method of tomato picking robot based on improved YOLOv8s and RGB-D information fusion. Transactions of the Chinese Society of Agricultural Engineering, 41(15), 106-116.
[27] Li Daoliang, Zhao Ye & Du Zhuangzhuang. (2025). Research progress on multimodal fusion technology methods and applications in the agricultural field. Transactions of the Chinese Society of Agricultural Machinery, 56(01), 1-15.
[28] Chen Tianci, Zhang Shiang, Fu Genping, Chen Jiazheng & Zhu Lixue. (2025). Review of research progress on tropical fruit picking robots. Transactions of the Chinese Society of Agricultural Machinery, 56(05), 184-201.
[29] Wang Nan, Li Zhen, Li Jiameng, Zhang Yuan, Sun Hong & Li Minzan. (2023). Research on chlorophyll detection system of crop plant fusion with multispectral imaging and deep learning. Transactions of the Chinese Society of Agricultural Machinery, 54(S2), 260-269.
[30] Yang Yang, Han Huayu, An Dong, Wang Yu, Tang Wu, Liu Jinghui & Zhou Yan. (2024). Design and experiment of intelligent pruning system for fruit trees based on lidar. Transactions of the Chinese Society of Agricultural Machinery, 55(07), 47-56+123.
[31] Qiao Yichen, Ning Zeting, Feng Qingchun, Wu Jianwei, Chen Liping, Shen Congju & Li Hongwen. (2026). Error calibration method for fruit tree picking robot based on machine learning. Transactions of the Chinese Society of Agricultural Machinery, 57(05), 39-49.
[32] He Chuangxin, Feng Wei, Li Yunhui, Ou Fang, Li Nan, Miao Zhonghua & Han Zengde. (2025). Orchard robot localization method based on factor graph fusion map. Agricultural Mechanization Research, 47(09), 15-21.
[33] Zhang Yali, Mo Zhenjie, Tian Haoxin, Lan Yubin & Wang Linlin. (2024). Path planning algorithm of agricultural robot based on improved APF-FMT*. Journal of South China Agricultural University, 45(03), 408-415.
[34] Li Yang, Zhang Lei, Li Pengfei, Wang Xiaohua & Wang Wenjie. (2023). A method for robot motion planning based on improved RRT combined with B-spline. Computer Integrated Manufacturing Systems, 29(01), 254-263.
[35] Liu Jizhan, Liu Wei, Mao Hanping & Li Pingping. (2014). Transplanting sequence and path analysis of robot for column cultivation. Transactions of the Chinese Society of Agricultural Engineering, 30(05), 28-35.
[36] Jiang, Y., Liu, J., Wang, J., Li, W., Peng, Y., & Shan, H. (2022). Development of a dual-arm rapid grape-harvesting robot for horizontal trellis cultivation. Frontiers in Plant Science, 13, 881904.
[37] Hong Kun. (2014). Research on compliant control technology of dexterous arm (Master’s thesis, Shenyang University of Technology).
[38] Zhao Yunwei, Geng Dexu, Liu Xiaomin & Sun Guodong. (2019). Kinematic analysis and experiment of pneumatic flexible fruit and vegetable picking manipulator. Transactions of the Chinese Society of Agricultural Machinery, 50(08), 31-42.
[39] Zhao Yunwei, Huang Haoran, Liu Xiaomin, Zhang Weiyi & Geng Dexu. (2023). Design and experiment of untethered pneumatic adaptive soft gripper for picking cones. Transactions of the Chinese Society of Agricultural Machinery, 54(09), 74-84.
[40] Liu Jizhan, Xu Xiuqiong & Li Pingping. (2014). Analysis and experiment of laser cutting of fruit stalks during fruit harvesting. Transactions of the Chinese Society for Agricultural Machinery, 45(01), 59-64.