Robotics and Autonomous Systems 114 (2019) 134–143

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Dual-arm cooperation and implementing for robotic harvesting tomato using binocular vision

Xiao Ling [a], Yuanshen Zhao [b,*], Liang Gong [a], Chengliang Liu [a], Tao Wang [a]

[a] School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China
[b] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

HIGHLIGHTS

- A dual-arm cooperation approach using stereo vision applied for robotic harvesting tomato.
- Tomato detection algorithm achieved the speed of 10 fps .
- The target tomato localized in the 3D scene simulation environment.
- The positioning error of robot hand–eye coordination was less than 10 mm.
- The success rate of robotic harvesting tomato in certain conditions was 87.5%.

ARTICLE INFO

Article history:
Received 5 October 2018
Received in revised form 23 January 2019
Accepted 30 January 2019
Available online 6 February 2019

Keywords:
Binocular vision
Dual-arm cooperation
Tomato detection
Three-dimensional scene reconstruction
Harvesting cycle time

ABSTRACT

Dual-arm cooperation is considered as an available approach to improve the poor efficiency by autonomous robotic harvesting. While, cooperating arm movements using visual information is a key challenge for harvesting robots working in non-structured environments. In this paper, we develop a dual-arm cooperative approach for a tomato harvesting robot using a binocular vision sensor. Firstly, a tomato detection algorithm combining AdaBoost classifier and color analysis is proposed and employed by the harvesting robot. Then, a fast three-dimensional scene reconstruction method is obtained in the simulation environment by using point clouds acquired from a stereo camera. Integration of tomato detection, target localization, motion planning and real-time control for dual-arm movements, the dual-arm cooperation for robotic harvesting can be implemented. To validate the proposed approach, field experiments were conducted with the potted tomatoes in greenhouse. Over 96% of target tomatoes were correctly detected with the speed of about 10 fps. The positioning error of robot end-point of less than 10 mm was achieved for large scale direct positioning of the harvesting robot. With the vacuum cup grasping and wide-range cutting, the success rate of robotic harvesting achieved 87.5%. Meanwhile, the harvesting cycle time excluding cruise time was less than 30 s. These results indicate that the dual-arm cooperative approach is feasible and practical for robotic harvesting in non-structured environments.

1. Introduction

In order to overcome labor shortage in the agricultural industry, robots are considered as alternatives that can undertake manual and tedious tasks. Especially, harvesting work is not only tiresome but also skill demanding. This problem escalates particularly in the harvesting season. Hence, research in harvesting robot systems has become a focus compared with other agricultural robots such as transplanting, cultivating, spraying and trimming [1]. Harvesting robots are required to sense the complex environment and use that information, together with a goal, to perform harvesting actions [2]. After the concept of harvesting robot proposed by Schertz and Brown [3], harvesting robots had ushered the era of rapid development [4–7]. Although they hold ample promise for the future, the present overall performance of harvesting robots is often incomparable to manual operation [7,8].

The traditional harvesting robots' poor efficiency is due to the single manipulator. As we known, single manipulator is not able to perform some complex tasks or to perform simple tasks in the required time [9]. Thus, an approach towards the goal of efficient robotic harvesting was developed for a multi-arm robotic harvester [10–12]. Some papers conducted the multiple manipulators cooperative methods [13–16]. However, The influences of visual information on multiple cooperative manipulators attracted no attention.

The current work on the application of visual information for harvesting robot is focused on the hand–eye coordination for single

* Corresponding author.
E-mail address: ys.zhao@siat.ac.cn (Y. Zhao).

https://doi.org/10.1016/j.robot.2019.01.019
0921-8890/

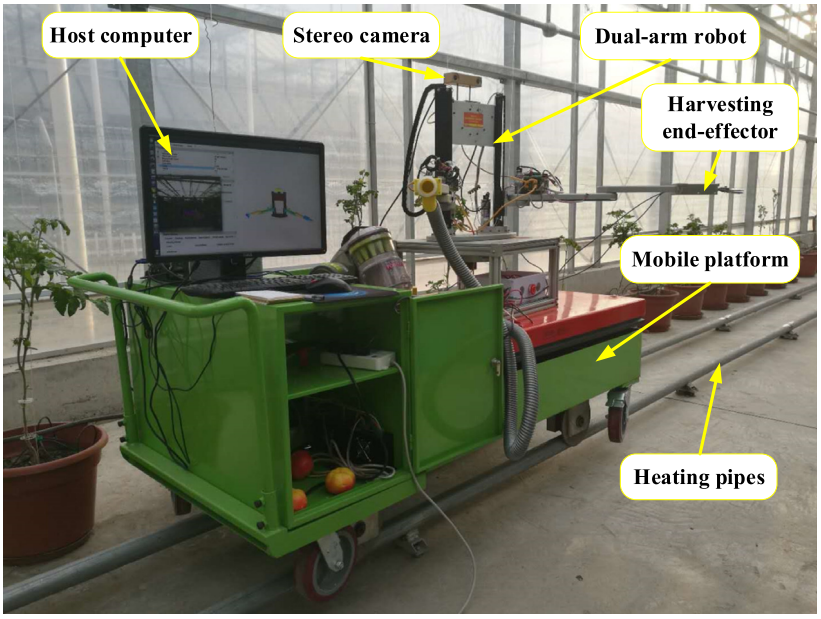

Fig. 1. Overview of the developed dual-arm harvesting robot.

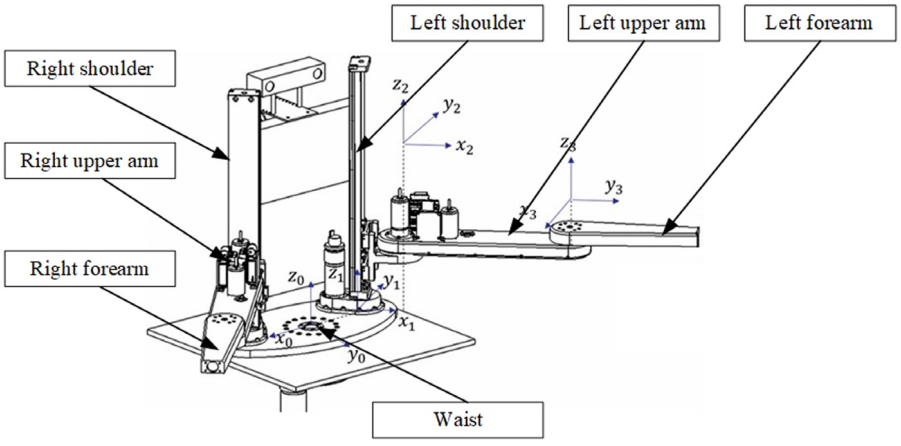

Fig. 2. Schematic of the dual-arm harvesting robot.

manipulator. The researches of hand–eye coordination for single-arm harvesting robot which ranges from fruit detection kinematics to visual servoing approach have received attention in literature ([17–19]). However, to date, analyzed of hand–eye coordination for dual-arm harvesting robot have not been reported.

To provide a realistic case study for the analysis of hand–eye coordination for dual-arm harvesting, the analyses of dual-arm cooperative in relation to visual information and other parameters involved (design of end-effectors, dual-arm structure and crop environment) have been conducted in this paper.

The system structure, including hardware and software, of the developed dual-arm harvesting robot system is introduced in Section 2. Then, the algorithms developed and used to implement the dual-arm cooperation for harvesting robot are described in Section 3. Experiments, along with testing results are described and presented in Section 3.1. Finally, the conclusion is given in Section 4.

2. Materials and methods

2.1. Hardware system

A dual-arm robot was developed for harvesting tomato in a greenhouse. The robot hardware adopts the modular design, which consists of a mobile platform, a dual-arm robot, end-effectors, a stereo camera and a host computer. As shown in Fig. 1, the robot was deployed in the Venlo type greenhouse in which the mobile platform can move on the heating pipes along crop rows. The mobile platform first moves automatically according to a certain step length. If there is no target tomato within the camera field of view, the mobile platform moves to the next working position under the control of the host computer. A stereo camera (Bumblebee2, Point Grey, Vancouver, Canada) with a resolution of 480 (H) × 640 (W) at 10 fps is mounted on the top of the robot to acquire the proper view of the field. An industrial personal computer (Advantech, Suzhou, China, Intel(R) Core(TM) i5-3610ME CPU, 4.0 GB RAM, Operating

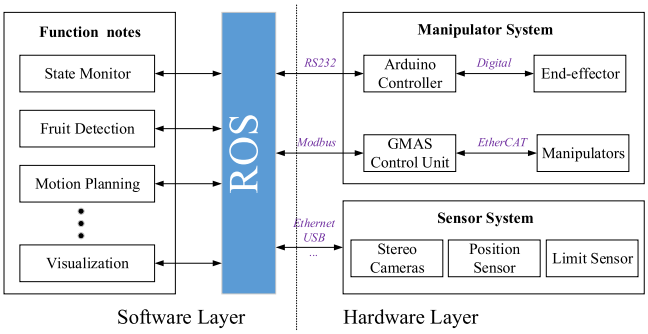

Fig. 3. The ROS based software architecture of tomato harvesting robot.

on Ubuntu 12.04) is used as the host computer dedicated to implement dual-arm cooperation for robotic harvesting. End-effectors were designed to be modular and exchangeable. Two types of end-effectors, cutting gripper and vacuum cup, are mounted on the left and right arm of the robot respectively. The two types of end-effectors ensure that the dual-arm manipulator can work cooperatively, by which the vacuum cup grasps the target fruit and the cutting gripper detaches the fruit from the plant.

A schematic of the dual-arm robot system is shown in Fig. 2. It is composed of two mirrored 3-DOF Cartesian type manipulators. Each manipulator is similar in design to the SCARA manipulator which has one prismatic joint and two revolute joints. The aim of adapting SCARA-like manipulator is to simplify the motion planning comparing with other harvesting robots of higher DOF [20–22] or parallel manipulator schemes [23,24]. Each joint is controlled by DC servo motors which are supported by digital drivers (G-Solo Whistle). Communication between the control unit and digital driver adopts the EtherCAT bus, chosen for its high transmission speed and reliability.

2.2. Software framework

The software architecture of the developed harvesting robot is displayed in Fig. 3. With the use of functional implementation in independent modules, the software framework is flexible and adaptive. To achieve dual-arm cooperative harvesting operation, three required modules are employed in the software [25]. Fruit detection, motion planning and motion control are arranged in three primitives of sensing, planning and acting [26]. All the primitives are implemented in the state machine [27]. Furthermore, the ROS-based software framework is open that allows adding addition functionalities such as visualization.

2.3. Overview of the proposed dual-arm cooperative approach

The dual-arm cooperative approach for robotic harvesting was develop based on open-loop control framework. As shown in Fig. 4, the control framework consists of five steps: scanning, tomato detection, 3D scene reconstruction, right arm grasping, and left arm detaching. In the first step, with the movement of mobile platform on the heating pipes, the stereo camera scans the crop row. The RGB image is transmitted to the host computer in real-time. In the second step, ripe tomatoes are detected by an algorithm combining a AdaBoost classifier and color analysis using the RGB image. In the third step, a 3D scene of field is reconstructed in the ROS visualization environment using point cloud data acquired from the stereo camera. The 3D position of a tomato object is obtained according to the relationship between 2D image pixel coordination and 3D point cloud coordinate. After computing the inverse kinematics, dual-arm motion parameters are acquired, which include target joint angles of each manipulator. The fourth step is that the vacuum cup-type end-effector grasps the target fruit. In the last step, the left arm detaches the fruit from the plant. The harvested tomato is collected by the right arm.

2.4. Tomato detection

Tomato detection is the foundation of visual control based dual-arm cooperation for robotic harvesting, which provides the correct object for robot operation. In dual-arm cooperation control framework, the function of tomato detection is conducted by the binoculars vision sensor Bumblebee2. Bumblebee2 stereo camera constructs from two CCD sensors (Sony ICX424 CCD, $1/3''$, $7.4\,\mu m$) can capture two RGB images using its left and right camera. In order to improve the speed of tomato detection, we process only the image capture from the right camera. There are challenges when the robot is using color camera as the visual perception device. Therefore, a robust tomato detection algorithm combining AdaBoost classifier and color analysis is employed in the developed harvesting robot system. The flow diagram of the proposed tomato recognition algorithm is shown in Fig. 5, which can be encapsulated in four general steps:

1. Sliding a sub-window on the entire image.
2. Extracting Haar-like features within each sub-window.
3. Detecting tomatoes using AdaBoost classifier.
4. Color analysis using average pixel value (APV) classifier on the I-component image.
5. Merging the detection results.

The proposed detection approach allocates a sliding sub-window (a square patch) for every pixel in an image, and then determined whether the pixels within the sub-window are classified as ripe tomato. In this work, 40×40, 80×80, and 120×120 pixels of sub-window sizes, according to the expected tomato diameters, are selected to scan the entire image. Twenty, forty, and sixty pixel increments are used as the sliding step in the scanning process with respect to the sub-window sizes.

The AdaBoost classifier is used to decide whether the sub-window contains tomato or background [10–12]. The AdaBoost is an algorithm for constructing a strong classifier from a linear combination of weak classifiers. The weak classifier could be thought of as a simple threshold operation on a special Haar-like feature [28]. The iterative training procedure for obtaining AdaBoost classifier is called weak learning. Weak learning is designed to select the weak classifier which could best separate the positive and negative samples.

For each feature, the weak learning process determines the optimal threshold value, such that a minimum number of samples are misclassified. A weak classifier is described by Eq. (1).

$$h_t\left(x_j\right) = \begin{cases} 1, & x_j < \theta_t \\ -1, & \text{otherwise} \end{cases} \tag{1}$$

where $h_t(x_j)$ is the weak classifier, x_j is the absolute value of the jth feature, t is a running index which denotes the iteration count, and θ_t is the threshold which can be calculated from Eq. (2).

$$\theta_t = \arg\ \min \varepsilon_t \tag{2}$$

where ε_t is the error rate of the tth iterative process. The error rate ε_t is the sum of weights of features which are misclassified, and

$$\varepsilon_t = \sum_{j=1}^{M} D_t(j),\ \text{for}\ y_j \neq h_t\left(x_j\right) \tag{3}$$

where $D_t(j)$ is the weight of the jth feature that is misclassified, y_j is the reference value of jth Haar-like feature which is set as 1 or

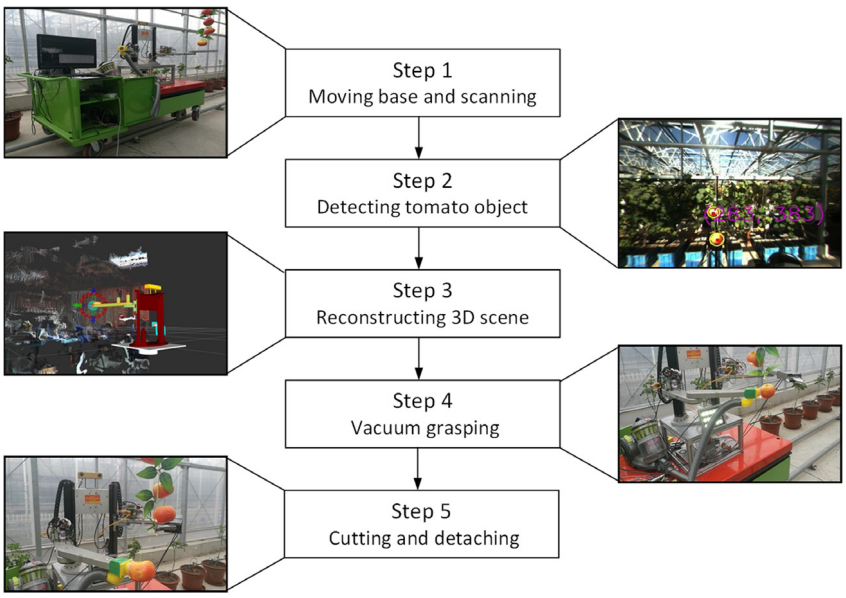

Fig. 4. The workflow of proposed algorithm in this study.

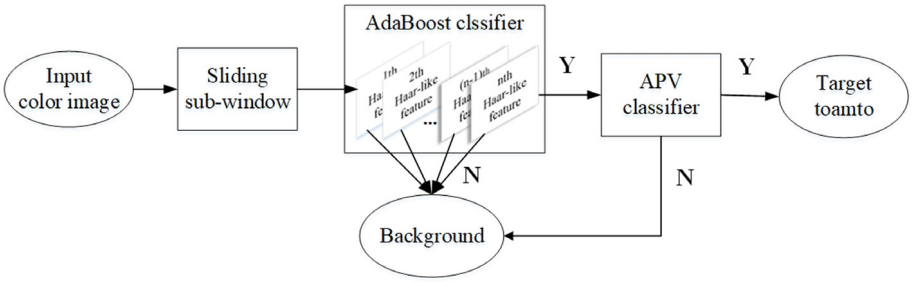

Fig. 5. Illustration of the tomato detection by combining AdaBoost classifier and APV classifier.

-1, M is the number of samples. The weight of feature needs to be updated with each iterative process, that is

$$D_{t+1}(j+1) = D_t(j) \exp\left(-a_t y_j \theta_t\right) / Z_t \tag{4}$$

where Z_t is the normalizing factor and is obtained by Eq. (5).

$$Z_t = \sum_{j=1}^{M} D_t(j) \tag{5}$$

Haar-like features that were selected in early rounds of the training process had error rates between 0.1 and 0.4. In practice, no single weak classifier can perform the classification task with low error. These weak classifiers are multiplied by their weights and combine linearly to construct a strong classifier. The final strong classifier is described by Eq. (6).

$$H(x) = \begin{cases} 1, & \sum_{t=1}^{T} a_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} a_t \\ -1, & \text{otherwise} \end{cases} \tag{6}$$

where, $H(x)$ and $h_t(x)$ denote the function of strong classifier and weak classifier, T is the number of weak classifiers or features, a is the weight of weak classifier and it is updated in each round of iteration. The update rule of a is

$$a_t = \frac{1}{2} \ln \frac{1-\varepsilon_t}{\varepsilon_t} \tag{7}$$

The employment of AdaBoost classifier might result in high false negative rate. For solving this problem, we combining a novel classifier to the AdaBoost classifier for reducing the false negative rate. The novel classifier is obtained by Color analysis, which is called average pixel value (APV) classifier.

For obtaining the APV classifier, a total of 1500 samples which included 344 ripe tomato images and 1156 background images were converted to the YIQ color spaces. Then I-component images of each samples were exacted.. For finding the difference between ripe tomatoes and background, the APV of I-component images is calculated from Eq. (8).

$$\mathrm{v} = \frac{1}{N} \sum_{i=1}^{N} p_i \tag{8}$$

where N is the number of pixels within a sample image, and p_i is the absolute value of differences in the ith pixel, v is the average pixel value of each component.

The average pixel value distribution of I-component samples is shown in Fig. 6. Although the APVs of the tomato and background samples overlap, a majority of distinctive pixels between background samples and tomato samples in the I-component image can be observed. Fig. 6 shows that when the threshold of average pixels value is set to 15, about 95% background are removed and 98% ripe tomato templates are successfully extracted. Thus, the threshold of average pixels setting 15 is designed as the APV classifier.

2.5. 3D scene reconstruction

After the harvesting robot detected the ripe tomato, the position of target tomato required to measure. The 3D scene reconstruction

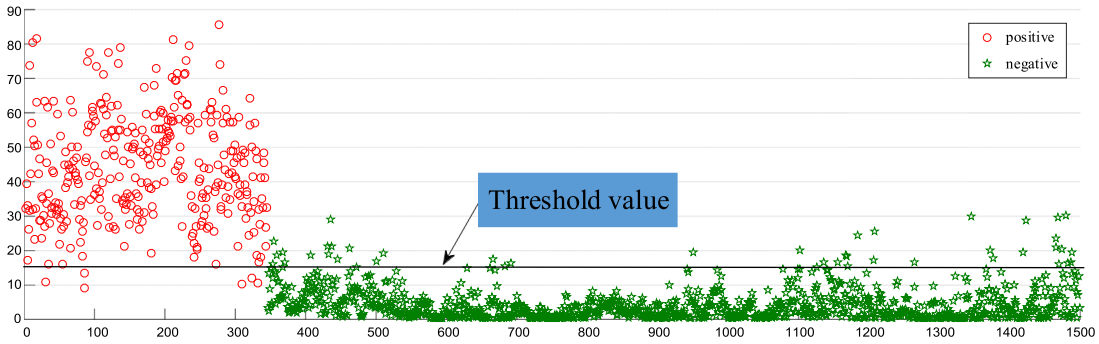

Fig. 6. The APV classifier using the threshold value of I-component samples.

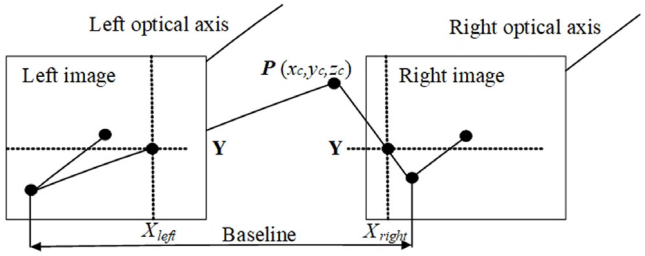

Fig. 7. Principle of point cloud acquisition using stereo camera.

is a fast method to obtain spatial position for dual-arm harvesting robot. 3D scene reconstructed is conducted based on point cloud re-factoring from the point cloud data acquired from the stereo camera. The principle of point cloud acquisition using Bumblebee2 camera is shown in Fig. 7. A spatial point $P(x_c, y_c, z_c)$ is imaged in the two parallel mounted CCD sensors. The image coordinates of P in the left image and right image can be described as P_{left} (X_{left}, Y_{left}) and P_{right} (X_{right}, Y_{right}).

As two CCD sensors are parallelly mounted, Y_{left} is equal to Y_{right}. According to the triangulation principle, the relationship between $P(x_c, y_c, z_c)$, P_{left} (X_{left}, Y_{left}) and P_{right} (X_{right}, Y_{right}) can be described by

$$\begin{cases} X_{left} = f \times \dfrac{x_c}{z_c} \\ X_{right} = f \times \dfrac{(x_c - B)}{Z_c} \\ Y = f \times \dfrac{y_c}{z_c} \end{cases} \tag{9}$$

in which B is the baseline distance between the two parallel CCD sensors, f is the CCD sensor focal length.

Through transformation, the 3D coordinates of $P(x_c, y_c, z_c)$ can be obtained from

$$\begin{cases} x_c = \dfrac{B \times X_{left}}{D} \\ y_c = \dfrac{B \times Y}{D} \\ z_c = \dfrac{B \times f}{D} \end{cases} \tag{10}$$

where D is the disparity between left image and right image. It can be calculated as

$$D = X_{left} - X_{right} \tag{11}$$

The point cloud data is a set of points that represent the disparities. Through importing the point cloud data to RViz, which is a 3D visualization tool in ROS, the true field scene can be reconstructed in a simulation environment. To import the 3D model of robot to the simulation environment, a Unified Robot Description Format (URDF) based robot model is employed. It can be transformed from the 3D model designed in SolidWorks with sw2urdf plug-in. The target tomato coordinate can then be published in this simulation environment, as shown in Fig. 8. Therefore, the dual-arm robot can have the relative position of target tomato in the robot coordinate.

It can be anticipated that the location of cutting point is different from the location of the tomato center. However, it can be inferred from the 3D location of target tomato and its radius. According to statistics of tomatoes growth in the greenhouse, the ripe tomato radiuses varied from 30 mm to 60 mm. For ensuring all the tomato stems can be cut off, the maximum radius value was used to determine the position of cutting point. Assuming that the 3D position of target tomato is (X, Y, Z), the position of cutting point (X_c, Y_c, Z_c) can be obtained as

$$\begin{cases} X_c = X \\ Y_c = Y + 0.06 \\ Z_c = Z + 0.06 \end{cases} \tag{12}$$

where X, Y, Z, X_c, Y_c and Z_c are measured in meter.

2.6. Dual-arm robot kinematics

The dual-arm harvesting robot has 2 mirrored 3-DOF arms. Each robotic arm is designed as SCARA-like manipulator that has one prismatic joint and two rotational joints. Fig. 9 shows the two rotational joints of SCARA-like manipulator.

The SCARA-like manipulator's inverse kinematics are used to calculate the joint angles $[\theta_1\ \theta_2]$ required to position the end-effector at the coordinates (x, y). That is

$$\theta_1 = \tan^{-1}\left(\frac{y}{x}\right) + \beta, \quad \beta = \cos^{-1}\left(\frac{{l_1}^2 + {l_2}^2 - r^2}{2l_1 \times r}\right) \tag{13}$$

$$\theta_2 = \pi - \alpha, \quad \alpha = \cos^{-1}\left(\frac{{l_1}^2 + {l_2}^2 - r^2}{2l_1 \times l_2}\right) \tag{14}$$

$$\mathrm{r} = \sqrt{(x^2 + y^2)} \tag{15}$$

where l_1 is the length of upper arm (330 mm), l_2 is the total length of forearm (250 mm) and end-effector.

Collision avoidance is another key difficulty in motion planning, especially for the dual-arm harvesting robot. In this framework, collision avoidance is performed Moveit (toolkit for ROS). By loading the robot collision matrix which consists of a collection of simple rules to form separate body parts to the integrated toolkit, the dual-arm motion planning may avoid the special trajectory. The collision matrix of dual-arm harvesting robot is shown in Table 1, in which the tick symbol stands for the possibility of collision between two parts of robot.

Table 1
Dual-arm harvesting robot collision matrix.

Waist	Left shoulder	Left upper arm	Left forearm	Right shoulder	Right upper arm	Right forearm	
×	×	×	√	×	×	√	Waist
	×	×	×	×	√	√	Left shoulder
		×	×	√	√	√	Left upper arm
			×	√	√	√	Left forearm
				×	×	×	Right shoulder
					×	×	Right upper arm
						×	Right forearm

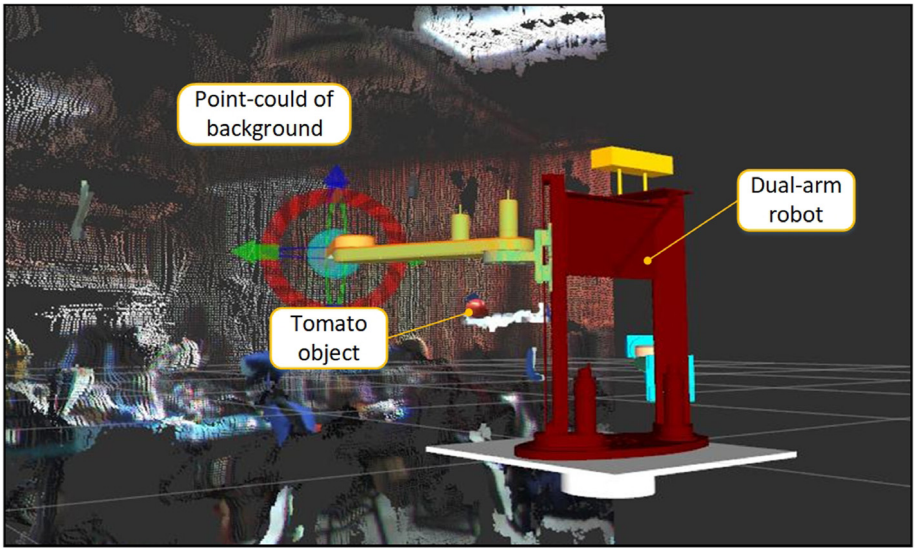

Fig. 8. The 3D scene reconstructed in the simulation environment for localizing the target tomato.

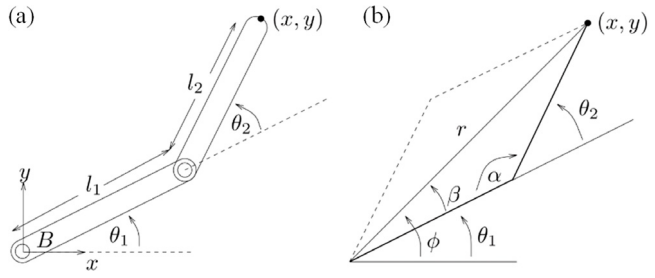

Fig. 9. The kinematics of manipulator for dual-arm harvesting robot.

2.7. Motion control

Motion control on the dual-arm harvesting robot is conducted by a multi-axis controller GMAS that supports EtherCAT and CANopen serial communication bus standard. The EthCAT bus is adopted closed loop connection mode which the controller GMAS can connect to each motor driver. The GMAS controller operates in Linux and communicates with the host computer through TCP/Modbus protocol. Fig. 10 shows the flowchart of motion control in the GAMS controller. There are five main function modals; *idle*, *home*, *move*, *stop* and *poff*, comprising the controlling program. After initialization, the GMAS controller loops through reading the register value transmitted from the host computer. The GMAS controller may select different functional motion according to the value of its register. If the function is selected is *move*, it may control each joint according to target joint angles. After the manipulator finishes moving, GMAS controller will transmit the current parameters of each joint to the host computer and then wait for the next instruction.

3. Results and discussion

3.1. The accuracy and speed of tomato detection

The tomato detection algorithm in the experiment was programmed in Python with a 32-bit Intel® Pentium® G640 2.8 GHz CPU. Experiments were conducted on the test image set that include 60 sample images acquired by Bumblebee2 with of 640 × 480 pixel resolution. There were 171 target tomatoes exiting in those sample images. In evaluating the experiments, the test images used contain complexity background such as varying illumination, partially occluded or overlapped. For the proposed detection algorithm, about 96.5% of the actual number of the ripe tomatoes were successfully detected against complexity backgrounds. About 3.5% of the ripe tomatoes were not detected. As shown in Fig. 11, even though in the condition of varying illumination or overlap, the ripe tomatoes were correctly detected. If the leaves occlusion area than 50% of the tomato area, the target tomato might not be detected.

Since detection time is an important performance parameter for the application, the processing time of the detection algorithm was also evaluated. Table 2 shows the detection time when applied to the testing images from test sets. During these experiments, the overall detection time varied approximately from 39 ms to 122 ms for one image. The average detection time was about 85 ms, which are short enough for the harvesting robot system to operate in real-time.

3.2. The control precision of hand–eye coordination

In the proposed dual-arm cooperation framework, the error in 3D localization of target tomato leads to errors in the reconstructed simulation scene and may influence the final positioning of the

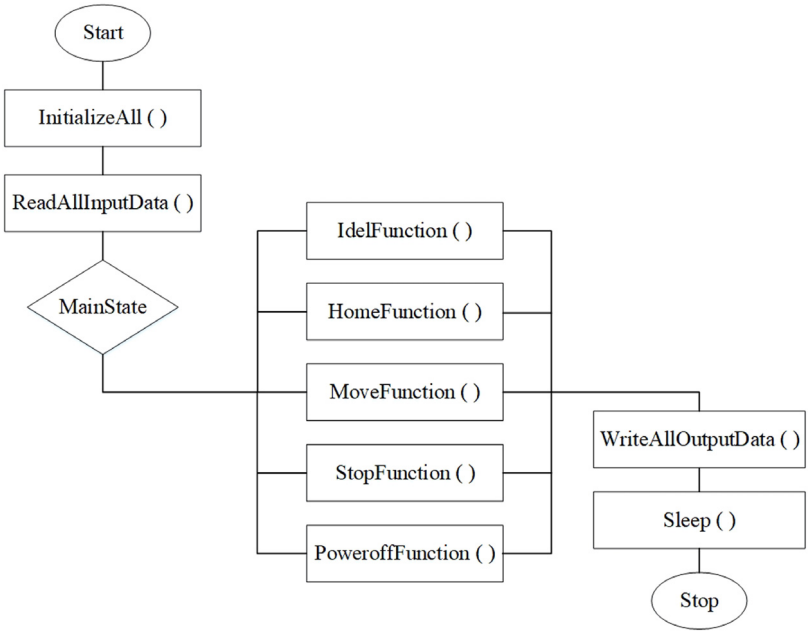

Fig. 10. The flowchart of motion control program in GMAS controller.

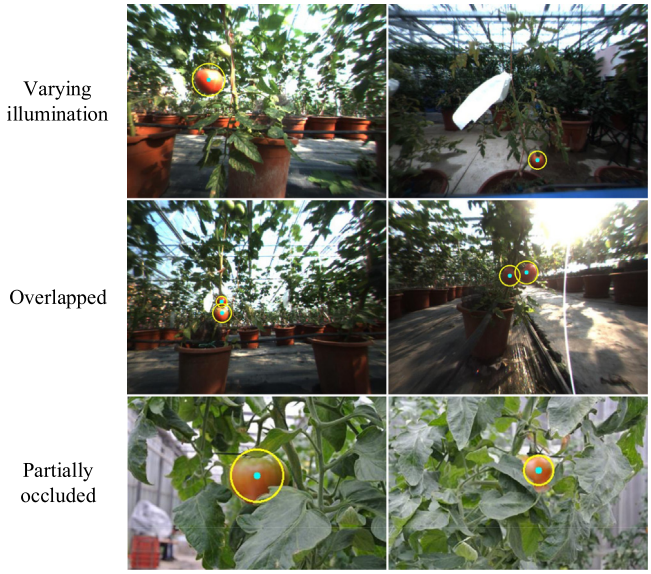

Fig. 11. Tomato detection results under varying illumination, partially occluded or overlapped.

robot arms. To analyze the impact of the 3D localization error when positioning the robot arm using stereo vision, the precision of 3D localization was examined. In the developed dual-arm harvesting robot workspace, the useful distance of depth direction for localizing the target tomato using stereo camera is between 200 mm to 800 mm. Four sets of distances were selected to test the precision of localization: 200 mm, 400 mm, 600 mm, and 800 mm. In each distance, ten simulated tomatoes in different position were randomly selected as test samples. This measurement was performed in each Cartesian axis and the total error was calculated as the Euclidean norm of these measurements. That is

$$\|(X_e, Y_e, Z_e)\| = \sqrt{X_e^2 + Y_e^2 + Z_e^2} \tag{16}$$

where X_e, Y_e, Z_e are the average deviations of measurement in each axis respectively and $\|(X_e, Y_e, Z_e)\|$ is the Euclidean norm.

Table 2
The detection time of each test samples.

Sample number	Time (ms)	Sample number	Time (ms)	Sample number	Time (ms)	Sample number	Time (ms)
1	104.570	16	115.804	31	80.947	46	67.343
2	97.754	17	111.175	32	98.840	47	61.362
3	94.352	18	93.938	33	75.748	48	58.144
4	99.009	19	107.981	34	93.864	49	50.969
5	94.177	20	100.070	35	95.823	50	52.081
6	116.364	21	102.716	36	89.555	51	47.178
7	99.606	22	103.668	37	48.891	52	86.593
8	96.358	23	95.814	38	41.109	53	82.128
9	108.072	24	89.070	39	48.383	54	77.868
10	101.560	25	87.629	40	60.072	55	97.510
11	106.020	26	111.173	41	39.704	56	91.471
12	98.555	27	98.034	42	41.908	57	91.600
13	104.246	28	114.118	43	42.968	58	83.610
14	90.814	29	101.585	44	55.253	59	54.812
15	121.819	30	107.456	45	51.265	60	54.824

The average deviations of measurement in x-direct, y-direct and z-direct can be obtained as

$$\begin{cases} X_e = \dfrac{\sum_{i=1}^{10} |e_{xi}|}{10} \\ Y_e = \dfrac{\sum_{i=1}^{10} |e_{yi}|}{10} \\ Z_e = \dfrac{\sum_{i}^{10} |e_{zi}|}{10} \end{cases} \tag{17}$$

where i is the index of samples, e_{xi}, e_{yi}, e_{zi} are the absolute error of each axis in the ith sample. The absolute error for each axis is the difference between the measurement values and those obtained from the stereo camera.

Table 3 shows the results of 3D localization using stereo vision with the developed detection algorithm for different distances. The results indicate that the localization error is sensitive to the error in the z-direction. The error in robot workspace is less than 2.5 mm. It is known that image processing in tomato object detection may also influence results of 3D localization. The sensitivity of tomato size to sliding distance may lead the error into the final measurement result.

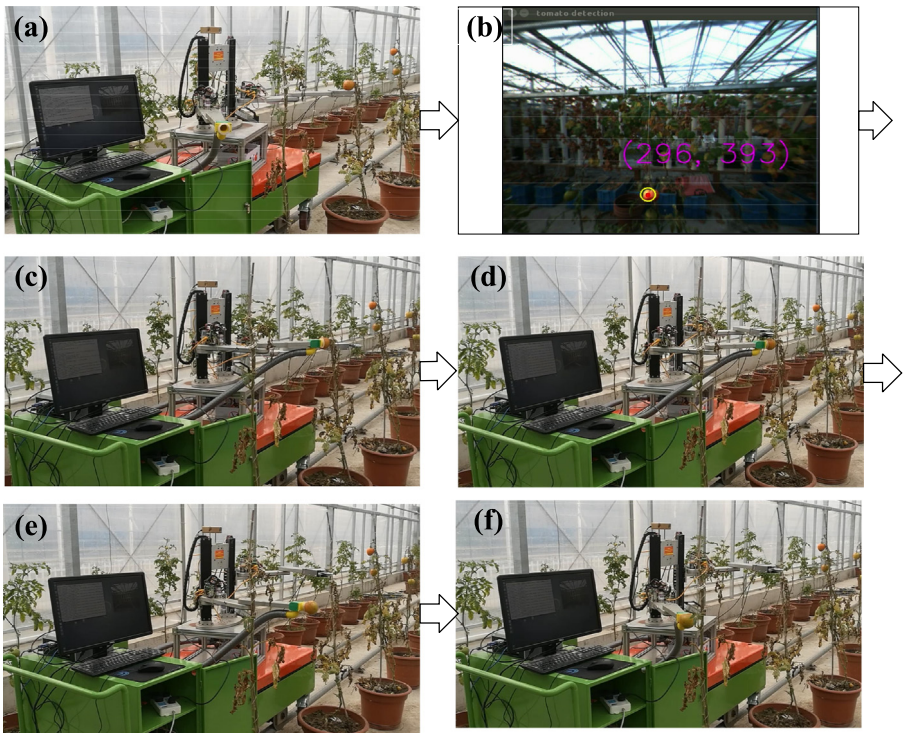

Fig. 12. The main processes and results of experiment for harvesting a tomato.

Table 3
The errors distribution of 3D localization in different distances.

Distances	Errors of each axis			Localization errors
(mm)	X_e (mm)	Y_e (mm)	Z_e (mm)	(mm)
200	0.2	0.3	1.8	1.84
400	0.3	0.3	1.9	1.95
600	0.3	0.4	2.0	2.06
800	0.5	0.5	2.3	2.41

Tests of hand–eye coordination were conducted by autonomously positioning the end-point of harvesting robot on a target point that simulated the target tomato position detected by the stereo camera. The positioning error, defined as the distance between the target point and the real location of the robot end effector, was used to quantify the performance of eye–hand coordination. Ten positions in the workspace of harvesting robot were selected as test samples.

According to the results of experiments, the positioning errors of the end-point of robot arm using the hand–eye coordination approach are randomly distributed in the range from 5 mm to 10 mm, which are direct consequences of random errors in the positioning measurement. Compared with the 3D localization test, the positioning error is caused by the error of simulation and real robot movements. However, this error is acceptable with regard to various conditions in the non-structured environment.

3.3. The performance of proposed algorithm

The robotic harvesting experiments were conducted in the greenhouse scenario with the potted tomatoes. In the experiments, eighty ripe tomatoes were randomly selected as the harvesting objects for the robot. These target tomatoes should be satisfied certain conditions contained ripe, single growing, and locating in the workspace of robot. Totally, seventy target tomatoes were harvested by the robot. One of the target tomatoes fell down caused by plant shaking when the robot harvested the other tomato from the same plant. The experiment results showed that the success rate of robotic harvesting is 87.5%. The mean harvesting times which exclude the cruise time for robot moving on the heating pipes is about 29 s/fruit. The cruise time is influenced by the distance of each potted tomatoes.

Fig. 12 illustrates the main processes and results of the experiment for harvesting a tomato. In Fig. 12(a), the harvesting robot moves on the heating pipes and scans the crop row. The harvesting robot detects the ripe tomato in the current view with the speed of 10 fps. Fig. 12(b) shows the target tomato detection processing. The harvesting robot scans the visual field when moving on the heating pipes. The robot stops moving and implemented harvesting operations when the target tomato enters into the working space of robot. Fig. 12(c) shows that the right arm is approaching the target tomato and the left arm is moving to the cutting position. In Fig. 12(d), the cutting gripper conducts the detachment operation. Meanwhile, the right arm has been grasped the target tomato used vacuum cup. The advantage of the dual-arm cooperative approach is avoiding the tomato shaking when cutting the stem. Thus, the harvesting success rate is enhanced compared with the single arm harvesting robot. In Fig. 12(e), the harvested tomato is collected by the right arm. As shown in Fig. 12(f), the robot moves to the next potted tomato when the harvesting operations have been completed.

4. Conclusion and future work

Ripe tomato detection success rate was about 95% for the proposed tomato detection algorithm which combined the AdaBoost classifier and APV classifier. And 5% of the ripe tomato were miss detection because of the leave occluded. When the leaves occlusion area than 50% of the tomato area, the target tomato might not be detected. The detection algorithm also owned good robustness, which it can meet the challenges of environment interference factors such as varying illumination, partially occluded or overlapped. The speed of detection algorithm was about 10 fps, which was enough for the harvesting robot system to operate in real-time.

The hand–eye coordination for dual-arm harvesting robot was also feasible. The 3D location of target tomato was measured through using the point cloud data acquired from stereo camera. The measuring error in robot workspace is less than 2.5 mm. The sensitivity of tomato size to sliding distance may lead the error into the final measurement result. The error distributed of hand–eye coordination was random, which ranged from 5 mm to 10 mm. which are direct consequences of random errors in the positioning measurement. The error of hand–eye coordination was caused by two main factors which were 3D location error and robot motion error.

The dual-arm cooperation framework using the visual information for harvesting tomato achieved the successful harvesting rate of 87.5%. Meanwhile, the harvesting cycle time excluding cruise time was less than 30 s. However, the performance of dual-arm cooperative control framework was relation to the certain conditions such as the pose of fruit, the distance of crops and the control of end-effectors. The future work may focus on the improvement of successful harvesting rate under uncertain condition.

Acknowledgments

This work was mainly founded by the National Natural Science Foundation of China (No. 51775333), the National Key Technology R&D Program of China (No. 2015BAF13B00) and partially supported by the National High-Tech R&D Program of China (863 Program No. 2013AA102307). The authors would also want to give thanks to Sunqiao modern agricultural park of Shanghai for the experiment samples. The authors also thanked the key laboratory of Urban Agriculture Ministry of Agriculture which provided the experiment site for robot harvesting.

References

[1] Y. Edan, S.F. Han, N. Kondo, Automation in agriculture, in: Handbook of Automation. Part G, Springer, 2009, pp. 1095–1128.
[2] Y. Edan, E. Gaines, Systems engineering of agricultural robot design, IEEE Trans. Syst. Man Cybern. 24 (8) (1994) 1259–1265.
[3] C.E. Schertz, G.K. Brown, Basic considerations in mechanizing citrus harvest, Trans. ASAE 34 (1968) 3–346.
[4] F. Taqi, F. Al-Langawi, H. Abdulraheem, M. El-Abd, A cherry-tomato harvesting robot, in: IEEE 18th International Conference on Advanced Robotics (ICAR), 2017, pp. 463–468.
[5] L. Wang, B. Zhao, J. Fan, X. Hu, S. Wei, Y. Li, Q. Zhou, C. Wei, Development of a tomato harvesting robot used in greenhouse, Int. J. Agric. Biol. Eng. 10 (4) (2017) 140–149.
[6] Q. Qiu, L. Tian, X. Qiao, K. Jiang, Q. Feng, Selecting candidate regions of clustered tomato fruits under complex greenhouse scenes using RGB-D data, in: IEEE 2017 3rd International Conference on Control, Automation and Robotics, 2017, pp. 389–393.
[7] T. Grift, Q. Zhang, N. Kondo, K.C. Ting, A review of automation and robotics for the bio-industry, J. Biomech. Eng. 1 (1) (2008) 37–54.
[8] C.W. Bac, E.J. Van Henten, J. Hemming, Y. Edan, Harvesting robots for high-value crops: state-of-the-art review and challenges ahead, J. Field Robot. 31 (6) (2014) 888–911.
[9] J.J. Roldan, J. Cerro, D.R. Garzon, P.A. Garcia, M. Garzon, J. Leon, A. Barrientos, Robots in agriculture: state of art and practical experiences, Serv. Rob. (2018) 67–90.
[10] Y.S. Zhao, L. Gong, Y.X. Huang, C.L. Liu, A review of key techniques of vision-based control for harvesting robot, Comput. Electr. Agric. 127 (2016) 311–323.
[11] Y.S. Zhao, L. Gong, B. Zhou, Y.X. Huang, C.L. Liu, Detecting tomatoes in greenhouse scenes by combining AdaBoost classifier and color analysis, Biosyst. Eng. 148 (2016) 127–137.
[12] Y.S. Zhao, L. Gong, Y.X. Huang, C.L. Liu, Robust tomato recognition for robotic harvesting using feature images fusion, Sensors 16 (2) (2016) 173–185.
[13] K.J. Ahlin, A.P. Hu, N. Sadegh, Apple picking using dual robot arms operating within an unknown tree, in: ASABE 2017 Annual International Meeting, Spokane, Washington, 2017, pp. 1–11.
[14] F. Roure, G. Moreno, M. Soler, D. Faconti, D. Serrano, P. Astolfi, G. Bardaro, A. Gabrielli, L. Bascetta, M. Matteucci, GRAPE: ground robot for vineyard monitoring and protection, in: Third Iberian Robot Conference, 2017, pp. 249–260.
[15] B. Zion, M. Mann, D. Levin, A. Shilo, D. Rubinstein, I. Shmulevich, Harvested-order planning for a multiarm robotic harvester, Comput. Electr. Agric. 103 (2014) 75–81.
[16] J.R. Davidson, C.J. Hohimer, C. Mo, M. Karkee, Dual robot coordination for apple harvesting, in: ASABE 2017 Annual International Meeting, Spokane, Washington, 2017, pp. 1–9.
[17] K. Kapach, E. Barnea, R. Mairon, Computer vision for fruit harvesting robots-state of the art and challenges ahead, Int. J. Comput. Vis. Robot. 2 (1/2) (2012) 4–34.
[18] S.S. Mehta, T.F. Burks, Vision-based control of robotic manipulator for citrus harvesting, Comput. Electr. Agric. 102 (2014) 146–158.
[19] R. Barth, J. Hemming, E.J. Van Henten, Design of an eye-in-hand sensing and servo control framework for harvesting robotics in dense vegetation, Biosyst. Eng. 146 (2016) 71–84.
[20] C.W. Bac, T. Roorda, R. Reshef, S. Berman, J. Hemming, E.J. Van Henten, Analysis of a motion planning problem for sweet-pepper harvesting in a dense obstacle environment, Biosyst. Eng. 146 (2016) 85–97.
[21] R. Arikapudi, A.D. Petiteville, S.G. Vougioukas, Model-based assessment of robotic fruit harvesting cycle times, in: 2014 ASABE-CSBE/SCGAB Annual International Meeting, Montreal, Quebec Canada, 2014, pp. 1–9.
[22] E.J. Van Henten, E.J. Schenk, L.G. van Willigenburg, J. Meuleman, P. Barreiro, Collision-free inverse kinematics of the redundant seven-link manipulator used in a cucumber picking robot, Biosyst. Eng. 106 (2) (2010) 112–124.
[23] E. Coronado, M. Maya, A. Cardenas, O. Guarneros, D. Piovesan, Vision-based control of a delta parallel robot via linear camera-space manipulation, J. Intell. Robot. Syst. 85 (1) (2016) 93–106.
[24] H.T. Lin, M.H. Chiang, The integration of the image sensor with 3-DOF pneumatic parallel manipulator, Sensors 16 (7) (2016) 1026–1043.
[25] R. Barth, J. Baur, T. Buschmann, Y. Edan, T. Hellstrom, T. Nguyen, O. Ringdahl, W. Saeys, C. Salinas, E. Vitzrabin, Using ROS agricultural robotics – Design considerations and experiments, in: Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture, Madrid, Spain, 2014, pp. 509–518.
[26] R.R. Murphy, Introduction to AI Robotics, MIT Press, 2000, pp. 488–492.
[27] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, A.Y. Ng, ROS: an open-source robot operating system, in: Workshop on Open Source Software ICRA, 2009.
[28] P. Viola, M. Jones, Rapid object detection using a boost cascade of simple features, in: Proceedings of Conference on Computer Vision and Pattern Recognition, Vol. 1, 2001, pp. 511–518.

Xiao Ling is a PhD student at the School of Mechanical Engineering at Shanghai Jiao Tong University. She received the B.S. degree on Mechatronic Engineering from the Shandong Agricultural University, Shandong in 2011. She received the M.S. degree in Mechatronic engineering from Shanghai Jiao Tong University in 2013. She is actively working in the fields of robotics, controlling and data processing.

Yuanshen Zhao is an assistant research fellow at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. He received the PhD degree on Mechanical Engineering at Shanghai Jiao Tong University. He received the B.S. degree on Mechatronic Engineering from the North University of China, Shanxi in 2010. He received the M.S. degree on Mechatronic engineering from the University of Shanghai for Science and Technology in 2012. He is actively working in the fields of agricultural robot, computer vision and machine learning.

Liang Gong is an associate professor for Mechatronic Engineering at Shanghai Jiao Tong University. He received his M.S. and Ph.D. degrees on Mechatronics Engineering from Shanghai Jiao Tong University in 2004 and 2010, respectively. His research interests include intelligence robot and machine

Chengliang Liu is a professor for Mechanical Engineering at Shanghai 346 Jiao Tong University. He received his B.S. degree on Mechanical Engineering from Shandong University in 1985. He obtained his M.S. and Ph.D. degrees on Mechanical Engineering from Southeast University in 1991 and 1999, respectively. Currently he is Yangtze River Scholar Distinguished Professor and director of the Institute of Mechatronics & Logistics Equipment. His research interests include intelligence maintenance and remote monitoring for equipment, intelligence agriculture equipment and robotics.

Tao Wang is master student at school of mechanical engineering at Shanghai Jiao Tong University. His research interest is in robot control technology.