Design of structured-light vision system for tomato harvesting robot

Feng Qingchun[1], Cheng Wei[2], Zhou Jianjun[1], Wang Xiu[1*]

(1. *Beijing Research Center of Intelligent Equipment for Agriculture, Beijing 100097, China*;
2. *Zhejiang University of Technology, Hangzhou 100085, China*)

Abstract: In order to improve the operating precision of the harvesting robot, a vision system for intelligently identifying and locating the mature tomato was designed. The active detection method based on structured-light stereo vision was expected to deal with the problem of variable illumination and target occlusion in the glasshouse. The maximum between-cluster variances of hue (H) and saturation (S) value were adopted as the threshold for color segmentation, which weakened the impact on the image caused by the light intensity variation. Through the limit on the pixel size and circularity of the candidate areas, the vision system recognized the fruit area and removed the noise areas. The fruit's 3D position was computed on the basis of spatial relationship between the laser plane and the camera, when the linear laser was projected on the centre area of the mature fruit. The blue view-scanning laser stripe pixels on the mature fruit were extracted according to its Cb color characteristic. As the field test results show, the measurement error on the fruit radius is less than 5 mm, the centre distance error between the fruit and camera is less than 7 mm, and the single axis coordinate error is less than 5.6 mm. This structured-light vision system could effectively identify and locate mature fruit.

Keywords: harvesting robot, tomato, linear structure-light, 3D measurement, feature extraction
DOI: 10.3965/j.ijabe.20140702.003

Citation: Feng Q C, Cheng W, Zhou J J, Wang X. Design of structured-light vision system for tomato harvesting robot. Int J Agric & Biol Eng, 2014; 7(2): 19－26.

1 Introduction

Automatic harvesting robots for tomato have been researched since the 1990s[1-3]. The development of the robots aims at decreasing cost and lowing labor-intensity for picking mature fruits. As one key component of the robot, the vision unit is used to identify and locate the mature fruits, playing the most important role for the robotic harvesting under the agricultural non-structure conditions. However, the vision unit has been proved the bottleneck of the commercial application of such robots, because it is difficult to accurately obtain the fruit information in agricultural conditions. The barriers include variable illumination conditions in greenhouses, complex background of each fruit, the targets' random distribution, and their irregular cluster-shapes.

There are several techniques to obtain 3D information from a scene, which can be classified into passive and active methods[4]. While passive methods use only cameras to acquire and analyze the images, active methods require, as well, devices that project some type of light over the scene. The most widely known passive method is binocular vision[2,4-7] which is more handy and easier to use, but the main disadvantage is the problem of feature points matching. Occurred more to the occluded and fascicled fruits, this problem decreases locating accuracy and even causes failed harvesting. The use of structured light is one of the most widely used active methods for industrial application[8] and it consists of the projection of light patterns onto the scene in order to infer the geometry of an object by analyzing the pattern projection over the object's surface. In addition, the high intensity light and outstanding color is easier to identify in the glasshouse conditions with variable illumination, which can remedy defects of the vision

Received date: 2013-10-21 **Accepted date:** 2014-03-29
Biographies: Feng Qingchun, Engineer, mainly engaged in agricultural robot and machine vision. Email: fengqc@nercita.org.cn; **Cheng Wei**, Master student, Email: vic.cheng@foxmail.com; **Zhou Jianjun,** PhD, Email: zhoujj@nercita.org.cn.
*** Corresponding author: Wang Xiu,** Professor, PhD, Beijing Research Center of Intelligent Equipment for Agriculture, Beijing 100097, China; mainly engaged in intelligent equipment for fertilizer and pesticide on agriculture. Tel: 86-10-51503346, Email: wangx@nercita.org.cn.

system only with cameras. Tarrío used a matrix of spot laser and a stereoscopic vision system to locate the occluded fruits[4]. Feng designed a laser vision system without camera for the apple harvesting robot[9].

A new vision system for tomato's robotic harvesting is described in this paper, which contains a color camera, a vertical-moving laser projector and the controller. Also, the method for extracting mature fruits from the background is introduced. Based on the linear structured-light vision principle, a mature fruit's 3D position was located according to the laser's pixel coordinates and scanning displacement. Moreover, the field test for evaluating system performance was conducted.

2 Material and methods

2.1 Working condition for harvesting robot

The working environment for the tomato harvesting robot is shown in Figure 1. The middle line of the rails is 600 mm away from the plants of both sides. The harvesting machine is supposed to move on the rails to pick tomatoes on both sides. In the newly applied tomato cultivation model, tomato stems hang on strings attached to the ceiling of the greenhouse. Through releasing the string along the process, tomatoes can always appear in the height range between 100 mm and 600 mm above the rails. Therefore the design objective for the robotic vision system is to ensure it can accurately identify and locate mature fruits existing in this scope.

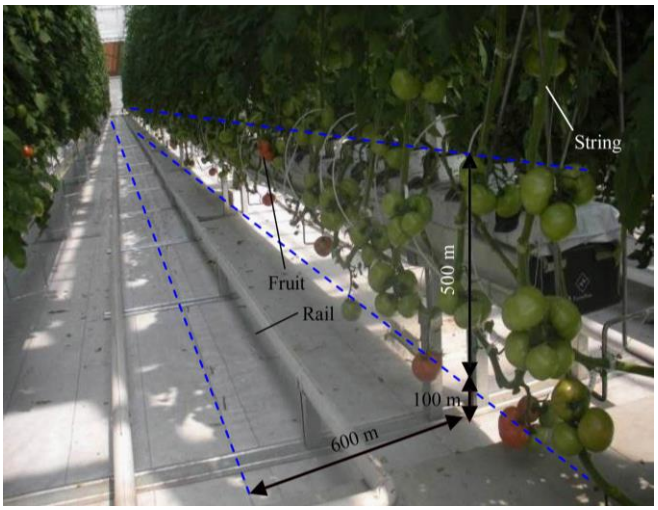
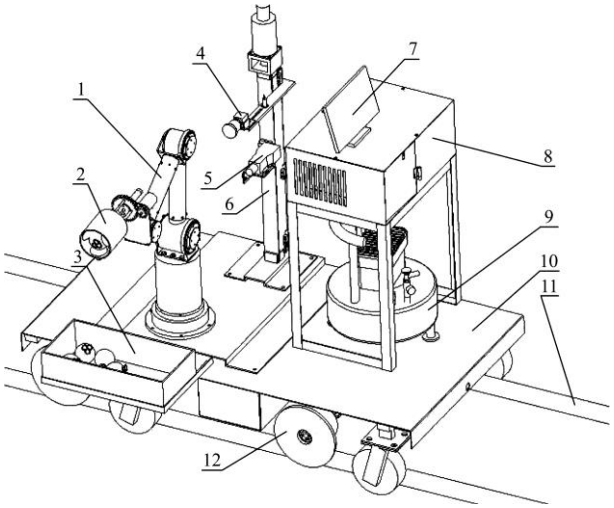


Figure 1 Working condition for harvesting robot

2.2 Harvesting robot model

The new model of the tomato harvesting robot system is shown in Figure 2, which contains a vision unit for identifying mature fruits, a jointed manipulator of four freedom degrees, a railed vehicle and the system controller. The vision unit, which consisted of a color camera, a linear laser projector and an electric slider, was used to identify and locate mature fruits. The controller receives and processes image data from the vision unit, then sends a control signal to the manipulator, the end-effector and the vehicle through CAN bus. The vehicle carries the whole system moving on the rail at a speed of 2 m/s, until it receives the stop signal from the controller when the mature fruits are identified in the camera field. The jointed manipulator positions the end-effector to approach the target fruit, and moves it back to the container after the fruit is separated from the plant. A twist-typed end-effector is adopted to pick the fruit, and the air pump supplies it with a high-pressure gas source.



1. Manipulator 2. End-effector 3. Fruit container 4. Camera 5. Laser projector 6. Slider unit 7. Human interface 8. Controller 9. Air pump 10. Vihecle 11. Rail 12. Rail wheel

Figure 2 Tomato harvesting robot model

2.3 Linear structured-light vision system

As shown in Figure 3, the vision system of the tomato harvesting robot consists of a color camera, a laser projector, a motor-driven slider unit and the controller, which is installed in the robot through the base. A Basler FL3-U3-13S2C-CS camera with 1 328 × 1 048 pixels resolution and a 5 mm focus lens is adopted for acquiring the fruit image, and it is fixed with the supporting pole. The laser projector can cast a linear stripe of 500 mm length and 5 mm width, and it moves

vertically with the slider driven by the motor, to make the linear laser scan the whole view field of the camera. A coder linked to the motor detects the motor's angular displacement, so that the movement of the laser can be controlled precisely.

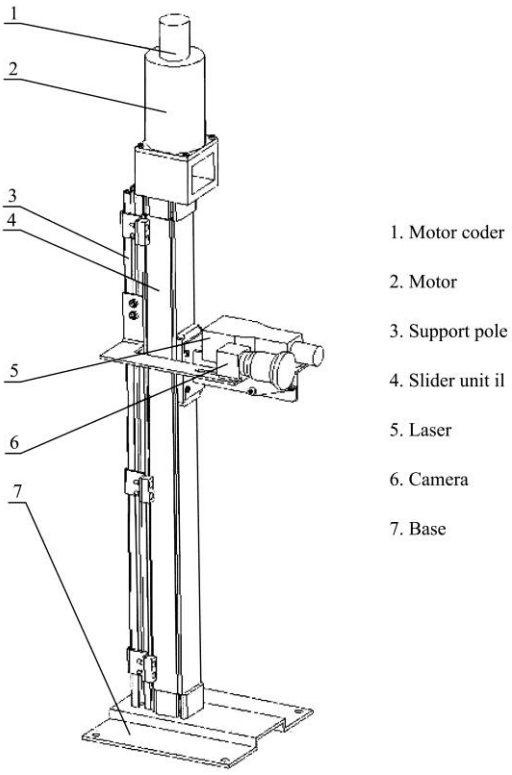


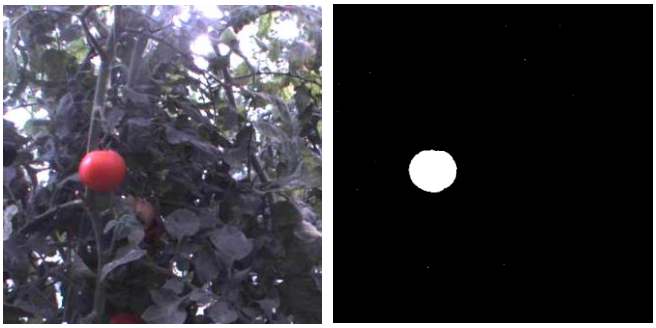
Figure 3 Model of structured-light vision system

The workflow of the vision system is represented as follows. Firstly, when the camera acquires a picture of the plant, the controller detects if mature fruits exist in the current view field. Secondly, once the mature fruit area is recognized, the linear laser will be powered and begin to scan the whole camera view. If the laser stripe is recognized to have arrived at a mature fruit's center area, the image coordinates of the laser and the scanning displacement will be saved. Finally, based on the measurement principle of the linear structured-light vision, the three-dimensional position of the mature fruit can be located.

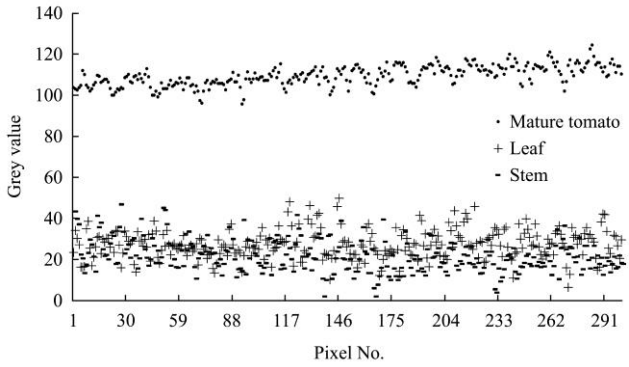
2.4 Mature fruit extraction

2.4.1 Color feature analysis

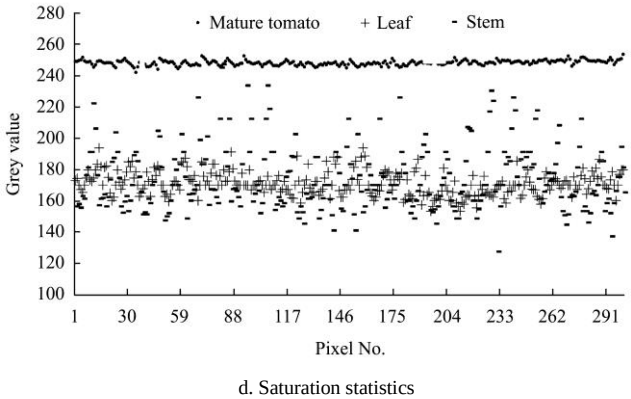
The image shown in Figure 4a is taken by the vision system in the greenhouse under the luminous intensity of 60 klux. In order to weaken the chromatic variation, which is caused by the unstable sunlight and the various degrees of shadow, the color image is transformed from RGB mode into HIS mode and the grey value of HIS is stretched into the range between 0 and 255. H (Hue) and S (Saturation) values of 300 pixels from mature fruit, leaf and stem are respectively counted. As shown in Figure 4c and Figure 4d, the H and S grey values of stem and leaf gather together, and keep an obvious distance from the mature fruit values.



a. Original image b. Identification result



c. Saturation statistics



d. Saturation statistics

Figure 4 Mature fruit extraction

2.4.2 Fruit feature identification

Based on the analysis of the color property of mature tomato, the thresholds, the maximum interclass variances of H and S values, are used for the color image segmentation. As shown in Equation (1), if μ is the average grey of the whole image, and $\omega(k)$ and $\mu(k)$ represent, respectively, the quantity ratio and the average

grey of the target pixels, $\sigma(k)$ is the grey variance between the target and the background for the threshold value κ. The optimal threshold K is the one corresponding to max$\sigma(k)$. The thresholds for Figure 4a are computed as $Thr_S = 54$, $Thr_H = 179$.

$$\sigma^2(k) = \frac{[\mu\omega(k) - \mu(k)]^2}{\omega(k)[1-\omega(k)]} \quad (1)$$

After the color image segmentation, the size S_i and the circularity ratio r_i of the candidate area in the binary image are counted. The fruit area should meet the following requirements as Equation (2); if not, the white area will be deleted as noise pixels. The final result is shown in Figure 4b.

$$\begin{cases} S_i > 7000 \\ r_i > 0.8 \end{cases} \quad (2)$$

2.5 Fruit location

2.5.1 Principle of structured-light vision

According to the principle of linear structured-light stereo vision, if the laser stripe is projected on the objects of different distances from the camera, it will show as the segments in the different lines in the image from the camera. The 3D position between the target and the camera could be calculated according to the pixel coordinates of the laser on its surface if the spatial relationship between the laser plane and camera is calibrated.

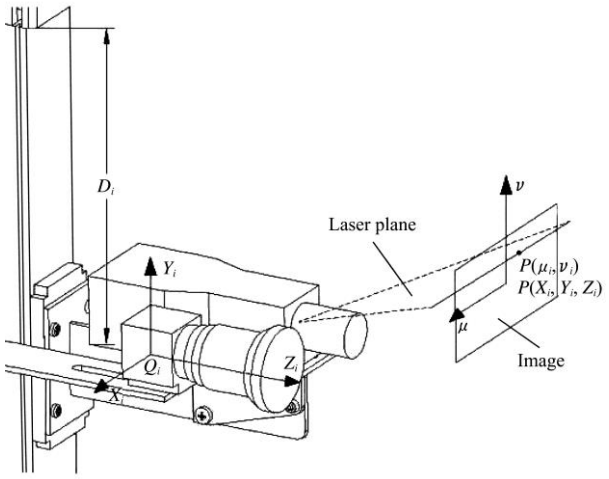


Figure 5 Coordinate definition

As shown in Figure 5, the world coordinate $O_iX_iY_iZ_i$ is established at the optical centre of the camera. The image coordinates of point $P(X_i, Y_i, Z_i)$ is (μ_i, v_i). The equation of the laser plane is represented as $aX_i + b(Y_i - D_i) + cZ_i + d = 0$, when the laser stripe lies on the point $P(X_i, Y_i, Z_i)$ after scanning over the distance D_i from the initial position. Based on the principle of structured-light vision, (X_i, Y_i, Z_i) can be obtained from the Equation (3).

$$\begin{cases} \rho \begin{bmatrix} \mu_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & \mu_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \\ aX_i + b(Y_i - D_i) + cZ_i + d = 0 \end{cases} \quad (3)$$

a, b, c, d are the equation coefficients of the laser plane under the world coordinates, and α_x, α_y, μ_0, v_0 are the internal parameters of the camera, which all could be determined through calibration[10].

While the linear laser is scanning from the bottom up in the camera view field, the automatic recognition for the stripe in the mature fruits area is being executed. The laser's scanning displacement D_i and pixels position (μ_i, v_i) will be saved once the light stripe moves into the central area of the mature fruit.

2.5.2 Parameters calibration

The main task of the vision system calibration was to determine all the above parameters in Section 4.1. In this paper, a 9 × 9 planar checkerboard was used for calibrating the vision unit. Firstly, the vision unit took a set of images of the board in different postures, when the laser projector stayed at the initial position ($D_i = 0$). As shown in Figure 6, two images (with and without laser stripe) of the same posture were acquired, through switching the laser projector manually. Then 20 checkerboard images without the laser stripe (as Figure 5a) were input into Matlab7.0 Calib-toolbox; the camera internal parameters (Equation (4)) can be obtained through processing the images in the Calib-toolbox.

$$\begin{cases} \alpha_x = 1397.34846, \ \alpha_y = 1397.73001 \\ \mu_0 = 666.29098, \ v_0 = 512.52465 \end{cases} \quad (4)$$

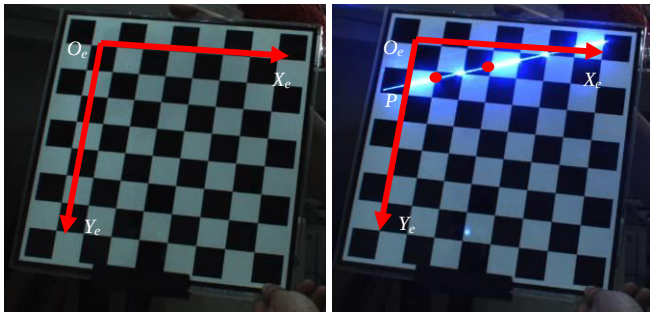
Besides, every transition matrix H_{ei}^c between the checkerboard coordinate $O_{ei}X_{ei}Y_{ei}Z_{ei}$ and the camera coordinate $O_iX_iY_iZ_i$ were also determined from the Matlab Toolbox. The checkerboard coordinate was built on its surface – as Figure 6-a, and the camera coordinate was consistent with the world coordinate. The images with light stripe were used for calibrating the laser plane under the $O_iX_iY_iZ_i$ coordinate. The checkerboard coordinate of

the laser points $P^{ei}(X_{ei}, Y_{ei}, 0)$ from different images could be measured manually, and their world coordinate $P(X_i, Y_i, Z_i)$ under the $O_iX_iY_iZ_i$ coordinate could be calculated as Equation (5).

$$P = H_{ei}^{c} P^{ei} \tag{5}$$

As all the point coordinates $P(X_i, Y_i, Z_i)$ were contained in the laser plane, the plane (Equation (6)) at the initial position could be determined based on the least-squares fitting method.

$$Z_i = -2.662651X_i - 2.269665Y_i + 534.039941 \tag{6}$$



a.Image without laser stripe b.Image with laser stripe

Figure 6 Checkboard image

Finally, having put the calibration results into Equation (3) and solved, the three-dimensional coordinates of the mature fruit (X_i, Y_i, Z_i) were calculated using the following Equation (7).

$$\begin{cases} X_i = \dfrac{1.8\times10^4\mu_i + 41.6\mu_i D_i - 6900D_i + 2.9\times10^5}{22.13\mu_i + 5.084\nu_i - 5.19\times10^4} \\ Y_i = \dfrac{-2.16\times10^3\mu_i + 5.04\mu_i D_i - 2560D_i + 9.20\times10^5}{22.13\mu_i + 5.084\nu_i - 5.19\times10^4} \\ Z_i = \dfrac{57.9D_i + 2.49\times10^7}{22.13\mu_i + 5.084\nu_i - 5.19\times10^4} \end{cases} \tag{7}$$

2.5.3 Laser stripe extraction

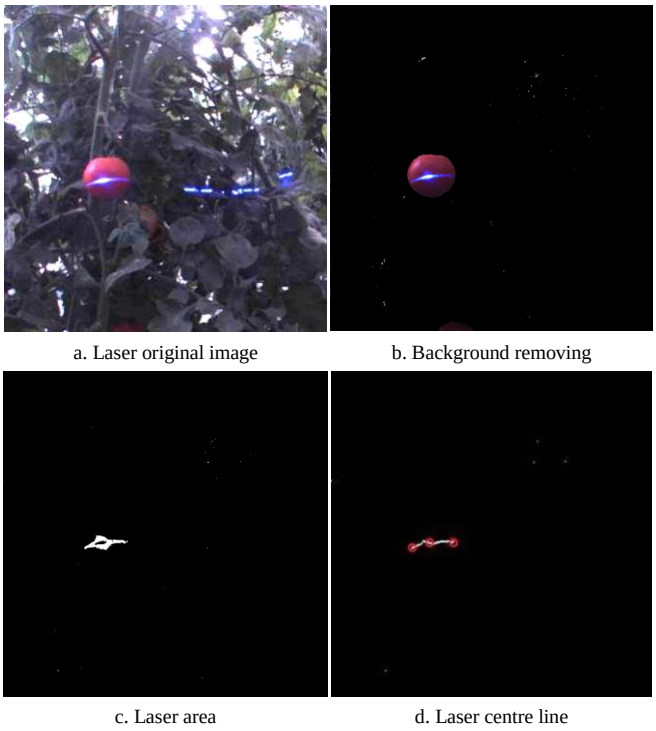
A linear laser generator of 470 nm wavelength (blue) was adopted for the vision system, so as to highlight the light stripe from the green leaves and red fruit. During scanning in the view field, a real-time algorithm was executed to identify if the laser stripe moved at the central area of 10 pixels width. The laser extraction method was as the following steps:

1) Image of the laser projected on a mature tomato is shown in Figure 7a. And the background can be removed according to the result of recognizing the mature fruit area in Section 2 as Figure 7b.

2) In order to make the laser hue more obvious and weaken the effect of unstable luminous intensity in the glasshouse, the laser image was transformed into YCbCr from RGB model. And the pixels with color component Cb>150 were considered as the laser ones. The image segmentation was implemented as Figure 5c.

3) $p(u_i, v_i)$ represents the image coordinate of laser pixel, and $p(\mu_i, v_{m\ldots n})$ are all the laser pixels in the same column. So all the centre points P_{centre} of laser pixels on the fruit could be denoted as Equation (8), and the centre line was estimated as Figure 5d.

$$P_{centre}(u_i, \sum_{v=m}^{n} v_i \Big/ (n-m)) \tag{8}$$



a. Laser original image b. Background removing

c. Laser area d. Laser centre line

Figure 7 Laser centre line extraction

The 3D coordinates ($X_{1\ldots3}$, $Y_{1\ldots3}$, $Z_{1\ldots3}$) of the three points (left, middle and right) of the laser centre line on the mature fruit, were calculated according to Equation (7). If the tomato fruit was approximated as a sphere in Figure 8, the radius R and the centre point $P(X, Y, Z)$ of the fruit were computed as Equation (9).

$$\begin{cases} X = \dfrac{(Z_2 - Z_3)(X_1^2 + Z_1^2) + (Z_3 - Z_1)(X_2^2 + Z_2^2) + (Z_1 - Z_2)(X_3^2 + Z_3^2)}{2X_1(Z_2 - Z_3) + 2X_2(Z_3 - Z_1) + 2X_3(Z_1 - Z_2)} \\ Y = (Y_1 + Y_2 + Y_3)/3 \\ Z = \dfrac{(X_2 - X_3)(X_1^2 + Z_1^2) + (X_3 - X_1)(X_2^2 + Z_2^2) + (X_1 - X_2)(X_3^2 + Z_3^2)}{2X_1(Z_2 - Z_3) + 2X_2(Z_3 - Y_1) + 2X_3(Z_1 - Z_2)} \end{cases} \tag{9}$$

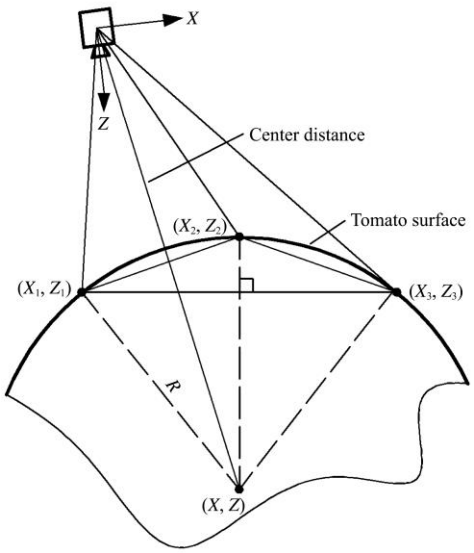


Figure 8 Approximate-shape of tomato

3 Results and discussion

3.1 Field test results

The experiment to test the precision of the vision system was carried out between 15:00 and 18:30 on August 14, 2013 in Beijing Special Vegetable Garden. Twenty mature fruits, which were empirically regarded as the ones fit for harvesting by man, were automatically identified and located under the light illuminance between 30 klus and 60 klus. The leaves in the height range with mature fruits had been cut artificially to maximize the nutrient uptake of the fruits, so the fruits were less likely to be occluded by leaves. However, a large portion of the tomato fruits which crowded together in the same cluster,still existed after the robotic harvest. Twelve fruits among the 20 were occluded by one another in our test, and the others grew on separate stems.

Figure 9 Field test for vision system

The results such as the pixel size, circularity, radius of the fruits and the distance between the fruits' center and camera were noted. In the meantime the radius and distance were also measured in person and noted to compare with the robotic test results. Finally the experimental result is listed in Table 1.

Table 1 Test results

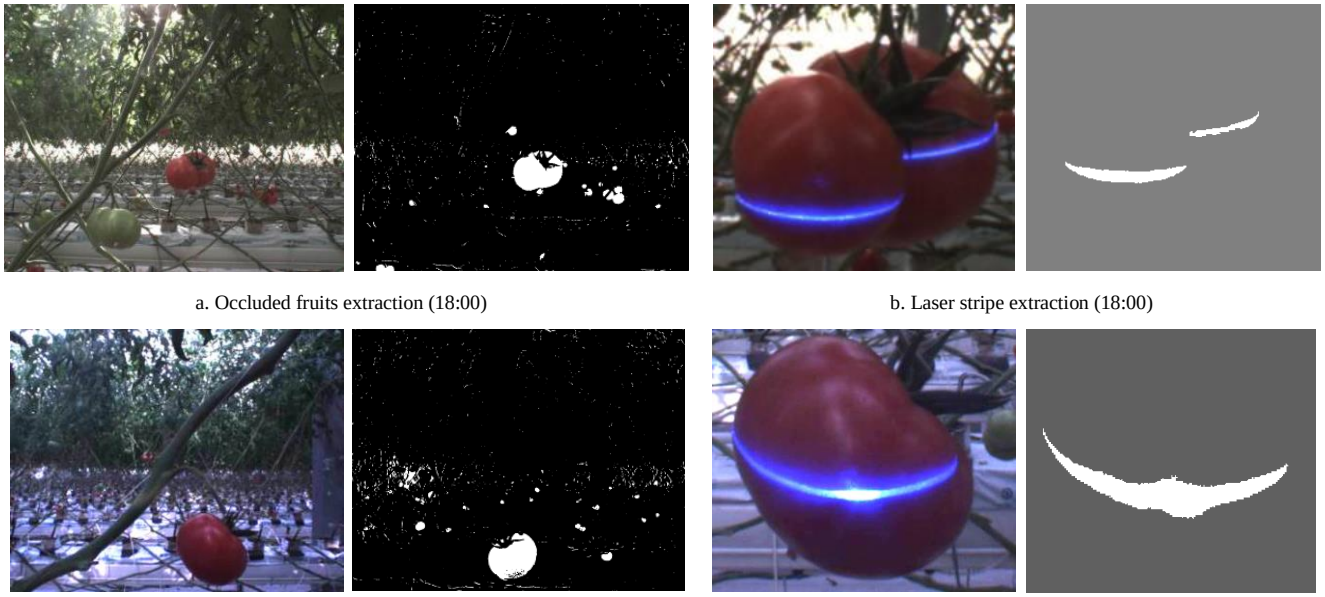
No.	Size/Circularity (pixel)	Radius(mm)		Distance(mm)	
		Auto	Manual	Auto	Manual
1	7532/0.82	28	30	789	785
2	7883/0.86	29	31	812	815
3	7945/0.81	33	31	823	819
4	8213/0.89	30	28	785	789
5	8165/0.84	28	29	769	772
6	8249/0.84	32	27	789	783
7	8069/0.91	35	30	859	855
8	7120/0.91	23	22	783	787
9	7329/0.87	26	25	773	776
10	7278/0.89	26	28	821	825
11	7867/0.82	28	31	786	791
12	7923/0.85	29	30	792	789
13	8230/0.81	33	32	829	825
14	8253/0.82	32	30	845	842
15	8135/0.80	35	30	859	855
16	7249/0.85	30	28	876	883
17	7154/0.88	25	26	853	857
18	7163/0.88	29	31	875	869
19	7598/0.87	28	31	863	867
20	7893/0.82	30	32	843	846

3.2 Discussion

The images with occluded or single fruits, which were acquired at various time, and the processing results are shown in Figure 10. The segmentation based on hue (H) and saturation (S) threshold values could accurately recognize the mature fruit area from the background under the various illumination conditions. Moreover, because of the discontinuous variation of depth information from the fruits crowded with each other, the laser stripe projected on them was interrupted obviously in the image. According to this, the vision unit could detect if there were more than one fruit in a candidate area after image segmentation. During our test, based

on the calculated result of the mature fruits' coordinates, the one closest to the camera was first picked from the occluded fruits by a person, then the rest were identified and located again, till all the fruits of the same cluster were picked.

As the results show, all the 20 mature tomatoes could be recognized according to the criteria on the pixel size (more than 7000) and circularity (more than 0.8). However, error of less than 5 mm exists on fruit radius, which is attributed to the non-spherical shape of fruits, and the error is greater for larger fruits due to the used method which fitted a sphere through three points. The measurement error of the center distance is less than 7 mm, and the 3D coordinate error of the single axis less than 5.6 mm. The 3D location error of the linear light vision system is mainly due to the laser's diffuse reflection on the fruit's surface, which makes the laser stripe in the image appear uneven in width, so that the centre line of the laser could not be figured out precisely. Even so, the radius error is acceptable for the automatic harvesting, and the location accuracy could meet the tomato robot's harvesting needs[11].



a. Occluded fruits extraction (18:00) b. Laser stripe extraction (18:00)

c. Single fruit extraction (15:30) d. Occluded fruits extraction (15:30)

Figure 10 Field test images and processing results

4 Conclusions

In order to meet the needs for robotic harvesting of tomatoes, a vision system was designed to identify and locate the mature fruits based on linear structured-light vision. The maximum between-cluster variances of H and S grey value were adopted as the threshold for color image segmentation, which could weaken the interference from the changing sunlight in the glasshouse and distinguish the mature fruits from the background. Besides, the noise after segmentation was removed through the limitation on the pixel size and circularity. The laser stripe pixels on the mature fruit image area were extracted according to the Cb value. Finally, the radius and centre position of the fruit were measured through spatially locating the three laser points on the fruits' surface by the structure-light vision. As the experimental results show, the system measurement error on fruit radius is less than 5 mm, error on the centre distance from fruit to camera is less than 7 mm and the single axis coordinate error is less than 5.6 mm. Above all, the method based on HIS color segmentation and structured-light stereo vision could effectively identify and locate the clustered fruits under various illumination conditions.

Acknowledgments

We thank the funding for this research supported by the National High Technology Research and Development Program of China (2013AA100307)

[References]

[1] Lü X L, Lü X R, Ni S C. Design and research of the computer identification method of picking tomatoes in field. Chinese Journal of Sensor and Actuators, 2010; 23(1): 82-86.

[2] Jiang H Y, Peng Y S, Ying Y B. Measurement of 3-D locations of ripe tomato by binocular stereo vision for tomato harvesting. ASABE Annual International Meeting, 2008. Paper Number: 084880.

[3] Ji P, Wang J, Chen H B. The tomato identification and geometric size measurement based on image. Journal of Anhui Agricultural Sciences, 2012; 40(33): 16426-16428.(in Chinese with English Abstract)

[4] Tarró P, Bernardos A M, Casar J R, Besada J A. A harvesting robot for small fruit in bunches based on 3-D stereoscopic vision. Computers in Agriculture and Natural Resources, 4th World Congress Conference, Florida, 2006: 270-275.

[5] Yuan T, Li W, Tan Y Z, Yang Q H, Gao F, Ren Y X. Information acquisition for cucumber harveating robot in greenhouse. Transactions of the Chinese Society for Agricultural Machinery, 200; 40(10): 151-155. (in Chinese with English Abstract)

[6] Feng Q C, Yuan T, Ji C, Li W. Feedback control based close scene for robotic cucumber harvesting. Transactions of the Chinese Society for Agricultural Machinery, 2011; 42(2): 154-157. (in Chinese with English Abstract)

[7] Kawollek M, Rath T. Machine vision for three-Dimensional modelling of gerbera jamesonii for automated robotic harvesting. International Conference on Sustainable Greenhouse Systems Greensys 2004, Leuven, 2005: 757—764.

[8] Xu K, Yang C L, Zhou P, Liang J. 3D detection technique of surface defects for steel rails based on linear lasers. Chinese Journal of Mechanical Engineering, 2010; 46(8): 1-5.

[9] Feng J, Zeng L H, Liu G, Ma X D, Pang S J, Zhou W. Construction of laser vision system for apple harvesting robot. Transactions of the Chinese Society of Agricultural Engineering, 2013; 29: 32-37. (in Chinese with English Abstract)

[10] Feng Q C, Liu X N, Jiang K, Fan P F, Wang X. Development and experiment on system for tray-seedling on-line measurement based on line structured-light vision. Transactions of the Chinese Society of Agricultural Engineering, 2013; 29(21): 143-149.

[11] Kendo N, Nishitsuji Y, Ling P P, Ting K C. Visual feedback guided robotic cherry tomato harvesting. Transactions of the American Society of Agricultural Engineers, 1996; 39(6): 2331-2338.