

# Location and model reconstruction algorithm for overlapped and sheltered spherical fruits based on geometry

*International Journal of Advanced  
Robotic Systems*

January-February 2022: 1–11

© The Author(s) 2022

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/17298814211056788

journals.sagepub.com/home/arx

**Li Li , Zhikang Ouyang, Wenqian Tian and Wei Sun**

## Abstract

For fruit picking robot, it is an essential prerequisite for achieving fruit picking using machine vision technology to accurately identify the fruits growing in the natural environment. This article presents a vision system of fruit picking robot to perform fruit location and three-dimensional model reconstruction. Firstly, combining the features of color and shape of fruit to reconstruct the actual contour of overlapped and sheltered fruits. Secondly, the least square method was used to reconstruct the three-dimensional model of each fruit according to the spatial coordinates corresponding to image location. Finally, fruit picking experiments in the laboratory environment are used to verify the feasibility of the proposed vision system. Three parameters including Segmentation Error, Intersection Over Union, and False Negative Rate are used to evaluate the performance of the algorithm. The average Segmentation Error, Intersection Over Union, and False Negative Rate of the fruit location algorithm based on geometry were 6.36%, 87.9%, and 6.72%, respectively. The experimental results showed that the average computation time of the algorithm is 3.2 s and the reconstructed three-dimensional model matched the size and position of fruits in the actual scene. The research results can be applied to the vision system of fruit picking robot.

## Keywords

Fruit picking robot, overlapped and sheltered fruits, vision system, fruit location, 3D model reconstruction

Date received: 13 April 2021; accepted: 12 October 2021

Topic: Vision Systems

Topic Editor: Antonio Fernandez-Caballero

Associate Editor: DJ Lee

## Introduction

With the development of science and technology, agricultural equipment develops toward the direction of automatic and intelligent.<sup>1</sup> Fruit picking robot, as one important kind of agricultural equipment, is of great significance to solve the problems of labor shortage, low productivity, and high-production cost in the agricultural production.<sup>2</sup> In the process of fruit picking, the robot captures the image in real time through the camera, and the field of view includes the fruits, branches, leaves, and so on. What is more, the target of fruit picking robot is easy to be affected by uneven illumination, overlapped and sheltered, which leads to the problem of false detection.<sup>3,4</sup> Therefore, for the fruit

picking robot, whether it can accurately identify the fruit target and quickly locate the picking point and then complete the fruit picking on the basis of ensuring the undamaged fruit are important evaluation standards.<sup>5,6</sup>

Image segmentation is an important processing task for agricultural robots because the subsequent identification

School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China

### Corresponding author:

Li Li, School of Mechatronic Engineering and Automation, No. 99 Shangda Road, BaoShan District, Shanghai City, China.

Email: lili.shu@shu.edu.cn



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

processes by the robots are based on the results of image segmentation, such as spatial location<sup>7,8</sup> and three-dimensional (3D) model reconstruction.<sup>9,10</sup> In the process of fruit picking using machine vision technology, we need to obtain the region of fruit in the image captured by camera and then acquire the specific coordinate of the fruit in the world coordinate system. The acquisition of visual information can apply different equipments, such as RGB or RGB-D camera,<sup>11,12</sup> 3D laser scanner,<sup>13,14</sup> and thermal camera.<sup>15</sup> This article focuses on the algorithm of fruit recognition and location based on RGB camera. In the research of target location algorithm, there are mainly two different methods: traditional image processing algorithms and machine learning algorithms. Zhuang et al. adopted traditional algorithms to extract the local texture information and then made the final decision based on a histogram intersection kernel-based support vector machine according to the local texture information.<sup>16</sup> Tao and Zhou extracted the improved 3D descriptor with the fusion of 3D geometry features and color features according to the preprocessed point cloud data, and then the optimized support vector classifier was used to identify the branches, leaves, and fruits.<sup>17</sup> Williams et al. proposed a new fruit scheduling system, in the part of fruit detection and location, a fully convolutional network was utilized to perform object segmentation, and then the position of each fruit was acquired through stereo matching.<sup>18</sup> Majeed et al. extracted the RGB and depth information from the acquired point cloud data to remove the background trees and then used a convolutional neural network (SegNet) to identify the trunk and branches of the tree.<sup>19</sup> The current state-of-the-art deep learning approaches require a trade-off between detection rate and processing time.<sup>18</sup> Using traditional machine vision method to realize object recognition is mainly based on the features of color, texture, and shape. It is the combinations of processes, such as color segmentation, thresholding, masking, and edge detection. For instance, Wang et al. transformed the RGB color space to Lab color space and then adopted K-means clustering algorithm to recognize the occluded apples.<sup>20,21</sup> Xiong et al. combined the improved fuzzy clustering method (FCM) and random signal histogram to remove the background of the nocturnal image in YIQ color model and then used the Otsu algorithm to identify the fruit from the stem base.<sup>22,23</sup> Chaivivatrakul and Dailey proposed a study of texture-based fruit detection for green fruits (bitter melon and pineapple) on plants in the field and recognized the green fruits in natural environment based on feature classification and region extraction.<sup>24</sup> Rizon et al. combined the morphological operator and texture analysis to isolate the overlapped and sheltered mango fruit and then used randomized Hough transform (RHT) to determine the fruit region and the picking point.<sup>25</sup> Luo et al. extracted the region of the overlapping grape clusters based on K-means clustering algorithm and separated the region pixels of double overlapped grape clusters based on the contour

intersection points and then detected the cutting point of each grape cluster according to the geometric constraint.<sup>26</sup> Fu et al. distinguished the fruits calyx from the skin based on color differences and obtained the contact points between the adjacent fruits by analyzing the edge information and then determined the borders of each fruit according to the contact points.<sup>27</sup> Song et al. adopted the convex hull theory in the segmentation of overlapped fruits and obtained the effective edge and intersection point of overlapped apples and then reconstructed the actual contour of each fruit based on effective information.<sup>28,29</sup> Lu and Sang detected the contour fragments of fruit target and the corners within the edges and then combined the valid contour fragments by analyzing the concavity or convexity, bending degree, and length to reconstruct the actual contour of occluded fruits.<sup>30</sup> Kelman and Linker presented a method for detecting the fruit in the tree using shape analysis and then obtained the edges that conformed to the geometric features and located each fruit according to these merged eligible edges.<sup>31</sup> Miao et al. proposed a combined algorithm based on Otsu algorithm and watershed algorithm to recognize and segment the overlapped objectives under natural environments.<sup>32</sup>

In this article, the grasping of spherical fruit in the natural environment is taken as the research object, and the recognition, location, and model reconstruction algorithm of the fruit picking robot are studied. Through the research of image processing algorithm, a new method is proposed to reconstruct the actual contour by extracting the effective edge of each fruit. According to the results of image location, the 3D coordinates are obtained based on the binocular camera, and then the 3D model of each fruit is reconstructed according to the spatial coordinates.

## Description of the location and 3D model reconstruction algorithm

In this article, it distinguishes the fruit target and background according to color feature and then reconstructs the actual contour of overlapped and sheltered fruits according to shape feature. The location and model reconstruction algorithm for overlapped and sheltered fruits consists of five steps:

- (i) Segmenting the fruits from the complex background after image denoising;
- (ii) Obtaining the simply connected domain of single object from overlapped and sheltered fruits;
- (iii) Acquiring the pixel coordinates set of outer contour of each fruit using eight-connected boundary tracking algorithm and extracting the effective edge from the non-actual contour of overlapped and sheltered fruits based on geometry;

- (iv) Reconstructing the actual contour of each fruit using least square method according to the effective information;
- (v) Obtaining the spatial coordinates corresponding to image location based on binocular stereo vision and then reconstructing the 3D model of each fruit.

### Recognition of overlapped and sheltered fruits

The original image acquired by camera in natural environment includes the fruits, branches, leaves, and so on. In the process of fruit recognition and location, it is necessary, firstly, to segment the fruits from complex background. Clustering algorithms, such as fuzzy c-means and K-means, have been widely used because of its good effect in the field of background segmentation. However, the accuracy of clustering algorithm is highly dependent on the clustering parameters and the improper selection of clustering parameters may result in the failure of segmentation.

In the natural environment, the color of ripened fruit is mostly close to red or orange and the background color is mostly close to green, blue, and other cold colors.<sup>33</sup> The difference of color between fruit and background is obvious, and therefore the image color segmentation is also one of the effective methods for background segmentation. In this article, the normalized color difference is used to segment the fruits from the complex background. The algorithm based on normalized color difference is expressed as

$$S = \frac{255(2R + B)}{2(R + G + B)} \quad (1)$$

where R, G, B are color components in RGB color space.

Perform image segmentation according to equation (2)

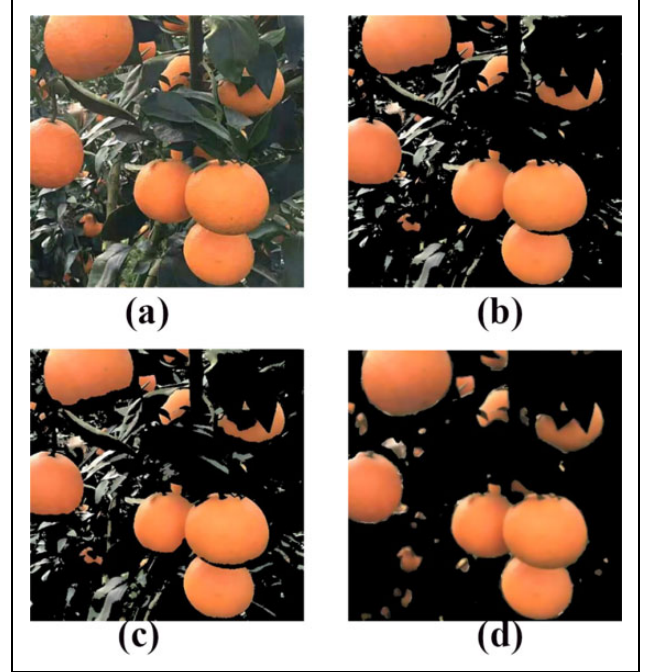
$$I_{(x_i, y_i)} = \begin{cases} I_{(x_i, y_i)} & S \geq T \\ 0 & S < T \end{cases} \quad (2)$$

where  $I_{(x_i, y_i)}$  is the pixel value of image coordinates  $(x_i, y_i)$ ,  $T$  is the segmentation threshold.

As shown in Figure 1, three algorithms can segment fruits from the complex background successfully. However, it is difficult to determine the optimal parameters to obtain the best segmentation effect in the real scenario. Therefore, the algorithm based on normalized color difference is adopted to recognize the overlapped and sheltered fruits growing in the natural environment in this article.

### Segmentation of overlapped and sheltered fruits

In the natural environment, the problem of fruits overlapping exists widely, which severely affects the recognition performance of the fruit picking robot. Therefore, it is



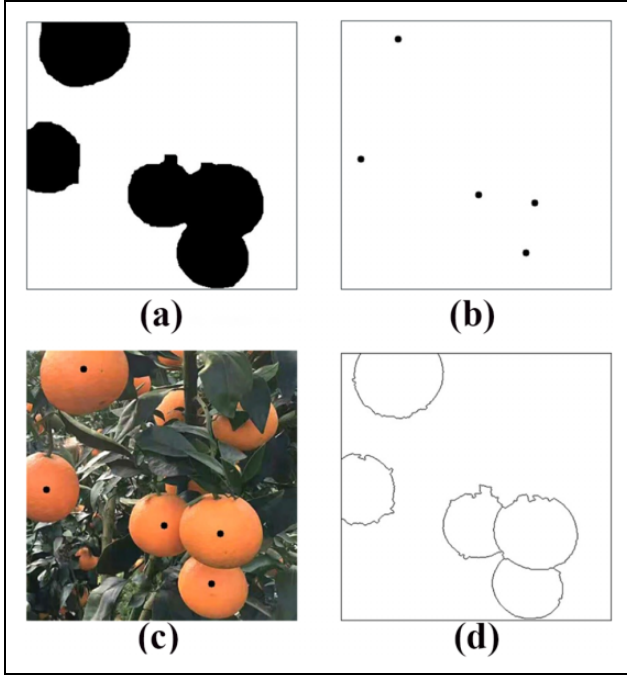
**Figure 1.** Recognition results of different methods: (a) original image; (b) K-means clustering algorithm; (c) FCM clustering algorithm; and (d) normalized color difference. FCM: fuzzy clustering method.

essential to identify the accurate position of each fruit from overlapped fruits and then pick them in turn.

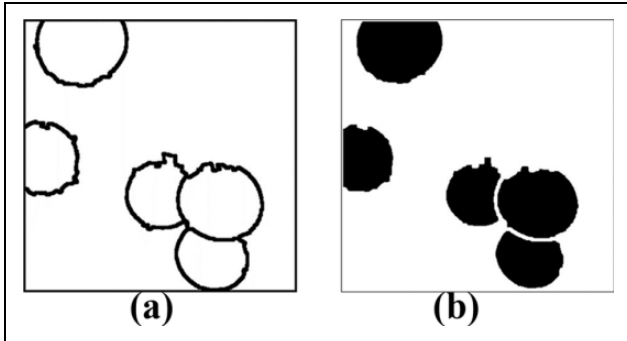
Distance transform is a global operation on binary image which will generate a gray image, the value of pixel represents the distance between the nonzero pixel and the nearest zero pixel in an image. After normalizing the gray image, the brightest point in the image indicates the nonzero pixel farthest from the zero pixel, which is the marker of foreground. The binary image is obtained after image preprocessing, such as mathematical morphology, area threshold, and binarization. The effect is shown in Figure 2(a).

In this article, it combines the local peak value of distance transform and watershed algorithm to achieve segmentation of overlapped fruits. Distance transform is used for the binary image, the effect is shown in Figure 2(b). Then the segmentation boundaries are obtained by watershed algorithm, the effect is shown in Figure 2(d). After morphological dilation and preprocessing operations, we can obtain the simply connected domain of single fruit, the effect is shown in Figure 3.

As shown in Figure 3(b), the algorithm basically realizes the segmentation of overlapped fruits. However, due to disturbance of the uncertain factors, we obtained the non-actual contour of overlapped and sheltered fruits. Therefore, this article presents a new method to obtain the effective edge by eliminating invalid pixels and then reconstructs the actual contour of each fruit according to the effective edge. It will be discussed in more detail in later sections.



**Figure 2.** Process of the segmentation of overlapped and sheltered fruits: (a) binary image; (b) distance transform; (c) markers of the fruit; and (d) watershed algorithm.

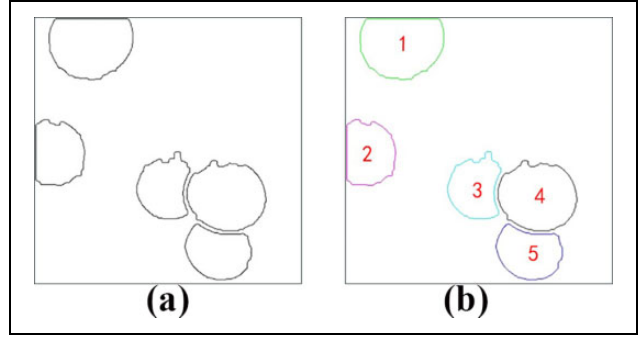


**Figure 3.** Process of obtaining the simply connected domain of single fruit: (a) morphological dilation and (b) simply connected domain.

### Extraction of effective edge

Because the shape of spherical fruit is close to ellipse in the machine vision image, therefore, the actual contour of the fruit can be reconstructed after the edges of similar circular arc (effective edge) are extracted. After obtaining the simply connected domain of single fruit, this article adopts Canny edge detector to extract the outer contour of simply connected domain, and the pixel coordinates set of outer contour of each fruit is obtained using eight-connected boundary tracking algorithm, the effect is shown in Figure 4. The pixel coordinates set belonging to the same object is represented by the same color.

The method proposed in this article can eliminate the invalid pixels formed by factors like overlapping,



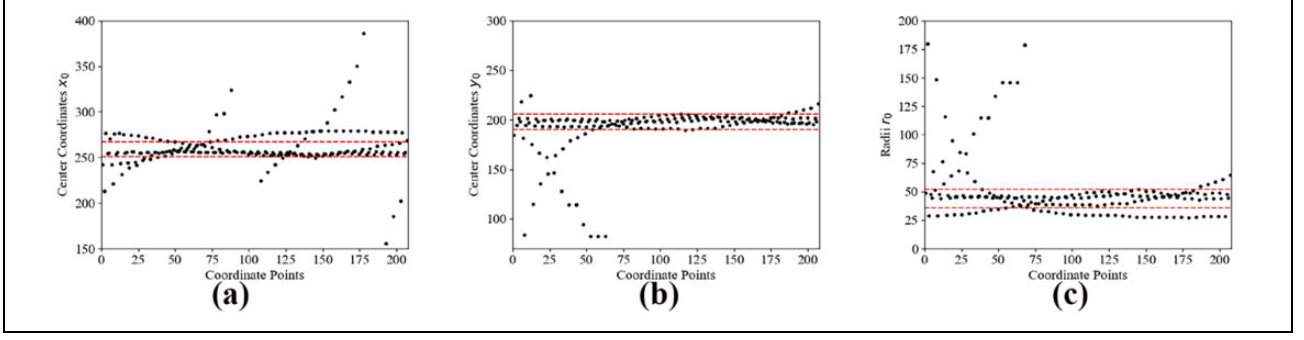
**Figure 4.** Process of obtaining the pixel coordinates set of outer contour: (a) Canny edge detector and (b) eight-connected boundary tracking.

occlusion, and uneven illumination and then reconstruct the actual contour of fruit with ellipse to locate each fruit according to the effective information. The algorithm includes the following steps:

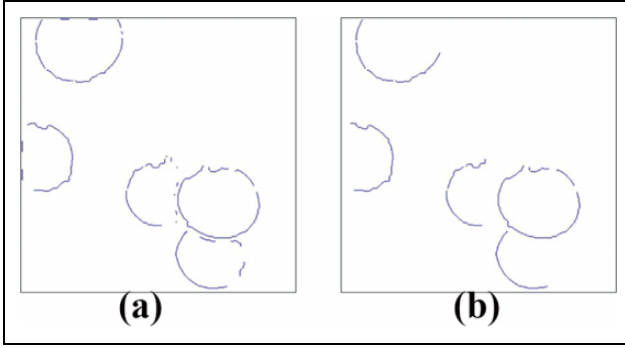
- (i) Acquiring the set  $\{x_1, x_2, \dots, x_n\}$  of outer contour pixel coordinates of the simply connected domain by eight-connected boundary tracking algorithm;
- (ii) Dividing the set into several groups  $\{x_1, x_{i+1}, \dots, x_{ki+1}\}$ ,  $\{x_2, x_{i+2}, \dots, x_{ki+2}\}$ ,  $\{x_m, x_{i+m}, \dots, x_{ki+m}\}$  at an interval of  $i$ ;
- (iii) Traversing the pixel coordinates set after grouping, according to the serial number 0, 1, 2; 1, 2, 3; ... to obtain three pixels in turn in each group, and then obtaining the center coordinates  $(x_0, y_0)$  and radii  $r_0$  of the circles determined by these three points according to equations (3) to (5);
- (iv) Obtaining the distribution interval of all center coordinates  $(x_0, y_0)$  and radii  $r_0$  and then acquiring the most concentrated interval that contains largest number of center coordinates  $(x_0, y_0)$  and radii  $r_0$  respectively, and the pixels that are not within the most concentrated interval will be regarded as invalid pixels and eliminated;
- (v) Adopting eight-connected boundary tracking algorithm to detect the discrete contour edges after eliminating invalid pixels, and obtaining the pixel coordinates set of all discrete contours belonging to the same object;
- (vi) Recording the number of pixels contained in each discrete contour and then obtaining the edges of similar circular arc (effective edge) according to selection principle

$$x_0 = \frac{(y_1 - y_2)[(x_1^2 - x_3^2) + (y_1^2 - y_3^2)] - (y_1 - y_3)(x_1^2 - x_2^2)}{2[(y_1 - y_2)(x_1 - x_3) - (x_1 - x_2)(y_1 - y_3)]} - \frac{(y_1 - y_3)(y_1^2 - y_2^2)}{2[(y_1 - y_2)(x_1 - x_3) - (x_1 - x_2)(y_1 - y_3)]}$$

(3)



**Figure 5.** Distribution interval of all center coordinates and radii: (a) distribution interval of  $x_0$ ; (b) distribution interval of  $y_0$ ; and (c) distribution interval of  $r_0$ .



**Figure 6.** Process of obtaining the effective edge: (a) eliminating invalid pixels and (b) eliminating invalid edges.

$$y_0 = \frac{(x_1 - x_3)[(x_1^2 - x_2^2) + (y_1^2 - y_2^2)] - (x_1 - x_2)(x_1^2 - x_3^2)}{2[(y_1 - y_2)(x_1 - x_3) - (x_1 - x_2)(y_1 - y_3)]} - \frac{(x_1 - x_2)(y_1^2 - y_3^2)}{2[(y_1 - y_2)(x_1 - x_3) - (x_1 - x_2)(y_1 - y_3)]} \quad (4)$$

$$r_0 = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \quad (5)$$

where  $(x_1, y_1)$ ,  $(x_2, y_2)$ ,  $(x_3, y_3)$  are the coordinates of three pixels selected by the serial number in turn.

As shown in Figure 5, it represents the distribution interval of the center coordinates  $(x_0, y_0)$  and radii  $r_0$ , and the red dotted line in the graph represents the most concentrated interval. Recording the serial number of pixels that are not in the range of most concentrated interval and eliminating these invalid pixels, the effect is shown in Figure 6(a). It can be seen that most of the invalid pixels have been eliminated. The eight-connected boundary tracking algorithm is used again, and then the edges of similar circular arc are extracted according to the selection principle of equation (6). The edges with pixels less than 20 are eliminated, the effect is shown in Figure 6(b).

$$\text{Edge} = \begin{cases} 255 & n \cdot \text{smax} < \text{num} \\ 0 & \text{num} \leq n \cdot \text{smax} \end{cases} \quad (6)$$

where  $\text{smax}$  is the second maximum value of the number of pixels contained in each discrete contour;  $\text{Edge}$  is the pixel value of each discrete contour;  $\text{num}$  is the number of pixels contained in each discrete contour;  $n$  is the ratio coefficient, and  $0 < n \leq 1$ .

#### Location of overlapped and sheltered fruits

The least square method is a mathematical optimization method. It can solve the appropriate fitting function of input data by minimizing the sum of error square, which can be used to obtain the unknowns from a known set of data. Therefore, according to the extracted effective edge information, we can reconstruct the actual contour with ellipse based on the least square method.

The general expression of ellipse can be described by the vector form

$$F_a(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} = 0 \quad (7)$$

where  $\mathbf{x} = [x^2, xy, y^2, x, y, 1]^T$ ,  $\mathbf{a} = [a, b, c, d, e, f]^T$ , and  $b^2 - 4ac < 0$ .

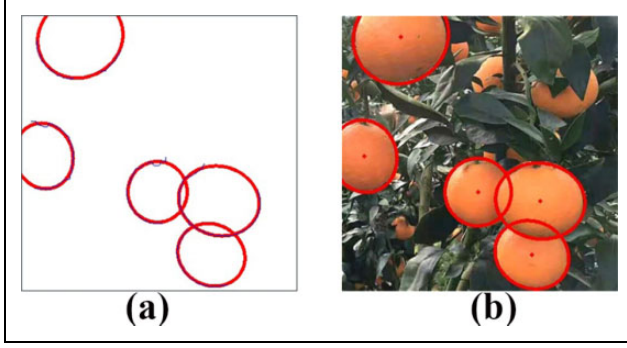
To ensure the effectiveness of the solution, the ellipse-specific constraint  $b^2 - 4ac < 0$  should be considered. It is noteworthy that the coefficients  $\mathbf{a}$  can be scaled, because the  $\alpha \cdot \mathbf{a}$  expresses the same conics as  $\mathbf{a}$  in the condition of  $\alpha \neq 0$ . Therefore, the ellipse-specific constraint can be converted to equality constraint under a proper scaling.<sup>34</sup> The solution of ellipse fitting expression can be expressed as

$$\min_a \|\mathbf{D}\mathbf{a}\|^2 \quad \text{subject to} \quad \mathbf{a}^T \mathbf{C} \mathbf{a} = 1 \quad (8)$$

where the design matrix  $\mathbf{D}$  of the size  $N \times 6$ , and the constraint matrix  $\mathbf{C}$  of the size  $6 \times 6$ .

$$\mathbf{D} = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_i^2 & x_i y_i & y_i^2 & x_i & y_i & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_N^2 & x_N y_N & y_N^2 & x_N & y_N & 1 \end{bmatrix} \quad (9)$$





**Figure 7.** Location result of overlapped and sheltered fruits: (a) pixels fitting and (b) reconstruction of fruit contour.

$$C = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (10)$$

Constructing Lagrange function and calculating the partial derivative, and then the simplified function can be obtained as

$$S^{-1}Ca = \frac{1}{\lambda}a \quad (11)$$

where  $\lambda$  is the Lagrange multiplier and  $S = D^T D$ .

$a$  is the eigenvector of  $S^{-1}C$ , it can be obtained by singular value decomposition. After extracting the effective edge from the non-actual contour, we can obtain the center, orientation, and the two semi-axis of the corresponding ellipse according to equation (11) deduced from equation (7). Then the actual contour is reconstructed with ellipse to locate each fruit, as shown in Figure 7. It can be seen that the result of contour reconstruction based on geometry is close to the actual shape of the fruit.

It should be noted that corrosion is carried out to smooth contour in the process of image preprocessing. Therefore, the axes of ellipse are set to 1.1 times of the calculated value. The subsequent images are processed in the same way.

### Model reconstruction of overlapped and sheltered fruits

It is a crucial task to build the 3D model of fruits, which can provide effective size and spatial parameters for the mechanical arm and end effector. In this article, according to the result of image location, obtaining the corresponding spatial coordinates based on binocular camera, and then reconstructing the 3D model of each fruit based on the spatial coordinates.

In the space rectangular coordinate system, the spherical equation can be expressed as

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r_0^2 \quad (12)$$

Constructing the equation

$$E(x_0, y_0, z_0, r_0) = \sum_{i=1}^N ((x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2 - r_0^2) \quad (13)$$

Calculating the partial derivative, and then the simplified function can be obtained as

$$x^3 \bar{x} - 2x_0 \bar{x} \bar{x} + x_0^2 + \bar{x} y^2 \bar{x} - 2y_0 \bar{x} \bar{y} \bar{x} + y_0^2 + \bar{x} z^2 \bar{x} - 2z_0 \bar{x} \bar{z} \bar{x} + z_0^2 = r_0^2 \quad (14)$$

$$\bar{x}^2 y \bar{y} - 2x_0 \bar{x} \bar{y} \bar{y} + x_0^2 + \bar{y}^3 \bar{y} - 2y_0 \bar{y}^2 \bar{y} + y_0^2 + \bar{y} z^2 \bar{y} - 2z_0 \bar{y} \bar{z} \bar{y} + z_0^2 = r_0^2 \quad (15)$$

$$\bar{x}^2 z \bar{z} - 2x_0 \bar{x} \bar{z} \bar{z} + x_0^2 + \bar{y}^2 z \bar{z} - 2y_0 \bar{y} \bar{z} \bar{z} + y_0^2 + \bar{z}^3 \bar{z} - 2z_0 \bar{z}^2 \bar{z} + z_0^2 = r_0^2 \quad (16)$$

$$\bar{x}^2 - 2x_0 \bar{x} + x_0^2 + \bar{y}^2 - 2y_0 \bar{y} + y_0^2 + \bar{z}^2 - 2z_0 \bar{z} + z_0^2 = r_0^2 \quad (17)$$

where  $\bar{x}^3 = \frac{1}{n} \sum_{i=1}^n x_i^3$ ,  $\bar{x} y^2 = \frac{1}{n} \sum_{i=1}^n x_i y_i^2$ , and so on.

According to equations (14) to (17), we can obtain equation (18) as

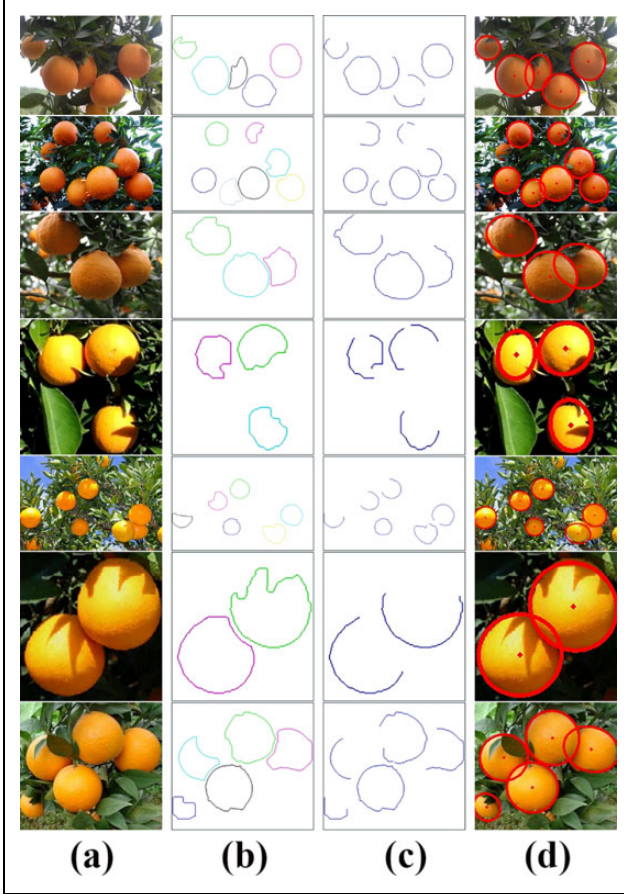
$$\begin{bmatrix} \bar{x}^2 - \bar{x}^2 & \bar{x} \bar{y} - \bar{x} \times \bar{y} & \bar{x} \bar{z} - \bar{x} \times \bar{z} \\ \bar{x} \bar{y} - \bar{x} \times \bar{y} & \bar{y}^2 - \bar{y}^2 & \bar{y} \bar{z} - \bar{y} \times \bar{z} \\ \bar{x} \bar{z} - \bar{x} \times \bar{z} & \bar{y} \bar{z} - \bar{y} \times \bar{z} & \bar{z}^2 - \bar{z}^2 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \bar{x}^3 - \bar{x} \times \bar{x}^2 + \bar{x} y^2 - \bar{x} \times \bar{y}^2 + \bar{x} z^2 - \bar{x} \times \bar{z}^2 \\ \bar{x}^2 y - \bar{x}^2 \times \bar{y} + \bar{y}^3 - \bar{y} \times \bar{y}^2 + \bar{y} z^2 - \bar{y} \times \bar{z}^2 \\ \bar{x}^2 z - \bar{x}^2 \times \bar{z} + \bar{y}^2 z - \bar{z} \times \bar{y}^2 + \bar{z}^3 - \bar{z} \times \bar{z}^2 \end{bmatrix} \quad (18)$$

The center coordinates  $(x_0, y_0, z_0)$  can be obtained by solving equation (18), and then the radii  $r_0$  can be obtained by substituting the center coordinates into equation (17). These parameters can be used to reconstruct the 3D model of each fruit. It will be verified in the fruit picking experiment.

### Results and analysis

We select 60 images of oranges captured by mobile phone in the orchard scene to test the performance of the

algorithm. Sixty images of overlapped and sheltered fruits are divided into three groups, including 20 in direct sunlight condition, 20 in backlighting condition, and the others in uniform illumination condition. Three parameters including Segmentation Error (SE), Intersection Over Union (IOU), and False Negative Rate (FNR) are used to evaluate the performance of the algorithm in the condition of sunlight, backlighting, and uniform illumination.<sup>20</sup>



**Figure 8.** Some representative experimental results: (a) original image; (b) eight-connected boundary tracking; (c) effective edges extraction; and (d) reconstruction of fruit contour.

1. SE represents the error rate of segmentation and is calculated by equation (19)

$$SE = \frac{|S_i - S|}{S} \times 100\% \quad (19)$$

where  $i = 1, 2, 3$ ,  $S_1$ ,  $S_2$ ,  $S_3$  mean the area fitted by the algorithm in direct sunlight, backlighting, and uniform illumination.  $S$  means the real fruit area, which is obtained by artificial marking.

2. IOU represents the rate of pixels segmented correctly of fruits and background and is calculated by equation (20)

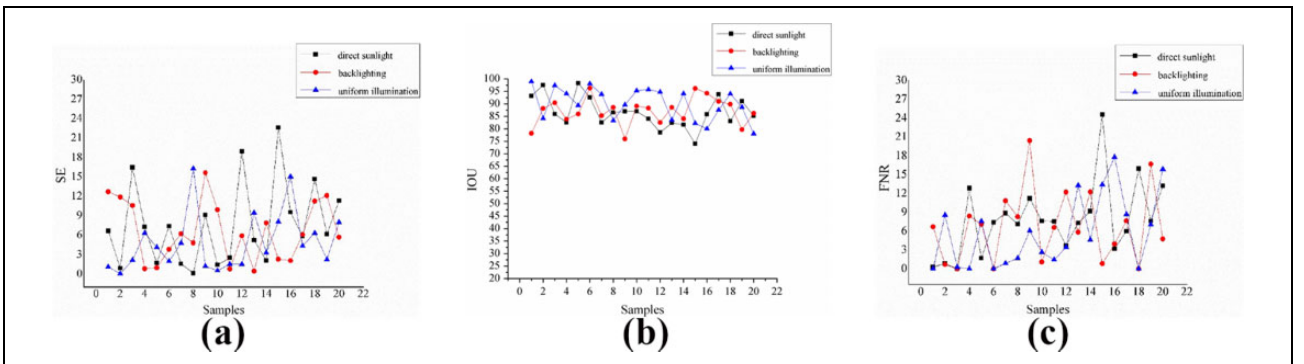
$$IOU = \frac{S_i \cap S}{S_i \cup S} \times 100\% \quad (20)$$

3. FNR represents the rate of pixels classified mistakenly of fruits and is calculated by equation (21)

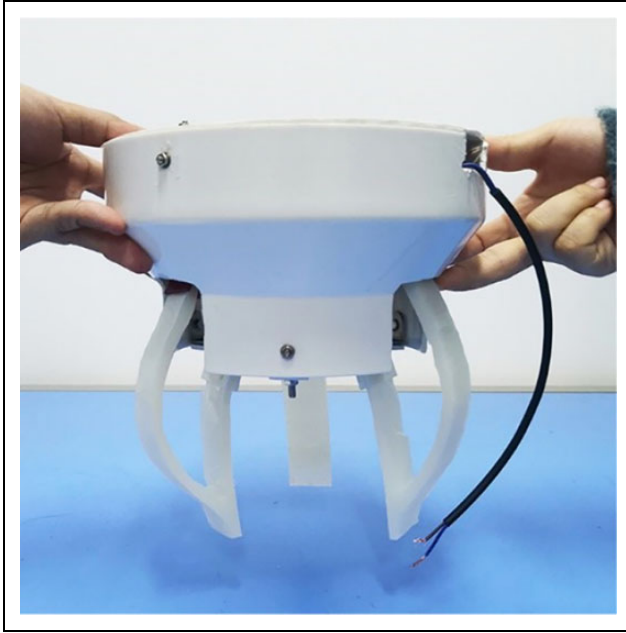
$$FNR = \frac{|S - S_i \cap S|}{S} \times 100\% \quad (21)$$

To demonstrate the effect of the algorithm clearly, Figure 8 shows seven representative results of algorithm, and the calculation results of SE, IOU, and FNR parameters of 60 images are entirely shown in Figure 9 with curve values. Based on the test of 60 images, we come to the conclusion that the computation time of algorithm is related to the area of fruits in the image, and the average computation time of 60 images is 3.2 s. As shown in the Figure 8, due to uncertain factors such as sunlight, branches, and leaves, partial effective edge of the fruit is transformed into invalid edge, the location algorithm can extract the effective edge from the non-actual contour and then reconstruct the actual contour of overlapped and sheltered fruits with ellipse. It can be seen that the results are basically fitting the actual shape of the fruit.

According to the test result, we can obtain that the average SE are 7.58%, 6.60%, 4.92%, the average IOU are 86.62%, 87.12%, 90.13%, and the average FNR are 7.81%, 6.71%, 5.66% in direct sunlight condition, backlighting condition, and uniform illumination condition. It



**Figure 9.** Diagram of (a) SE, (b) IOU, and (c) FNR in direct sunlight, backlighting, and uniform illumination.



**Figure 10.** Flexible grasping manipulator.

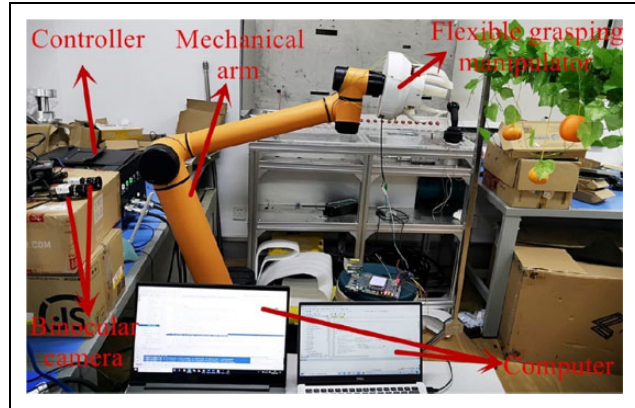
can be seen that the contour reconstruction algorithm based on geometry is effective.

## Fruit picking experiment

### Experimental platform

The fruit picking experimental platform is shown in Figure 11. In the figure, the image acquisition equipment is Flea3 FL3-U3-20E4C camera (1600 × 1200) produced by Point Grey company (FLIR Systems, Wilsonville, Oregon, USA), and the camera lens is HS0814J produced by Myutron incorporation (Nishikoiwa, Edogawa-Ku, Tokyo, Japan). Image processing equipment is a portable computer with Intel(R) Core(TM) i7-8550U @1.80 GHz, 64 bit with 8 GB RAM. The algorithms are written in Python version 3.7. The radii of three fruits are between 20 mm and 30 mm, and the size of fruits are medium, small, and large from left to right. The motion mechanism is six-degree-of-freedom mechanical arm AUBO-i10 produced by AUBO company (Lianshihu West Road, Mentougou District, Beijing) and the flexible grasping manipulator developed by our group. The external structure of flexible grasping manipulator is shown in Figure 10.

The flexible grasping manipulator consists of three flexible fingers made of silica gels. It has the characteristics of continuous motion, large range deformation, and high flexibility. Therefore, it is generally believed that the uncertainty in the process of picking fruit can be compensated by the compliance of flexible grasping manipulator. What is more, the designed three-finger flexible grasping manipulator can provide enough grasping force while ensuring safe interaction to achieve grasping of the fruit target.



**Figure 11.** Experimental platform diagram.

Stereo rectification is used to the original images according to the MATLAB calibration toolbox before spatial positioning, and then combining principle of binocular stereovision and stereo matching algorithm to obtain the 3D information of the center point of fruit in the camera coordinate system. It should be noted that the background area may be mistaken as the target area at the edge of the fruit and the failure of pixel matching in the process of obtaining spatial coordinates, which would obtain obvious noise data. We eliminate the noise data after obtaining the most concentrated interval according to the distribution interval of spatial coordinates.

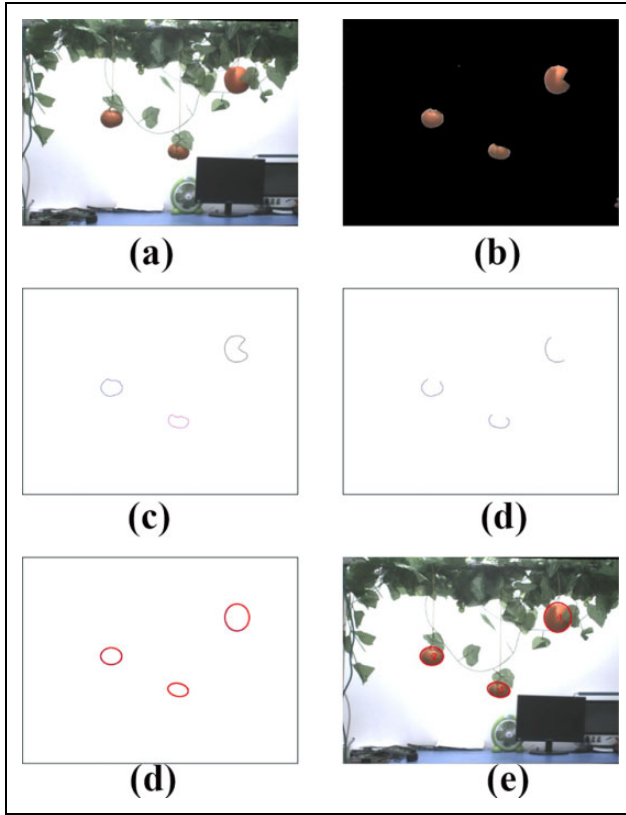
In the fruit picking experiment, to convert 3D coordinates of the center point of fruit in the camera coordinate system into the coordinates in the manipulator coordinate system. We obtain the 3D coordinates in the camera coordinate system and manipulator coordinate system and then adopt the built-in function of OpenCV to solve the optimal 3D affine transformation matrix  $\mathbf{H}$ .

### Algorithm verification

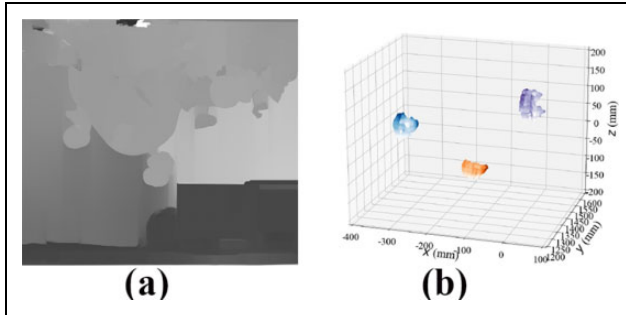
After stereo rectification of original images, taking the image captured by left camera as an example, the process of image location is shown in Figure 12. It should be noted that the position of fruits is not overlapped because the picking planning algorithm of overlapped and sheltered fruits is not involved in this article, but the effectiveness of image location algorithm has been proved fully in previous chapters.

We adopt binocular camera to obtain the disparity map according to the stereo matching algorithm and then obtain the spatial coordinates corresponding to image location. The disparity map is shown in Figure 13(a). Then the spatial coordinates after removing outliers are substituted into equation (18) to obtain the corresponding spherical center coordinates and radius. The effect is shown in Figure 13(b). Then we can reconstruct the 3D model of fruits using spheres in spatial location. The effect is shown in Figure 14.



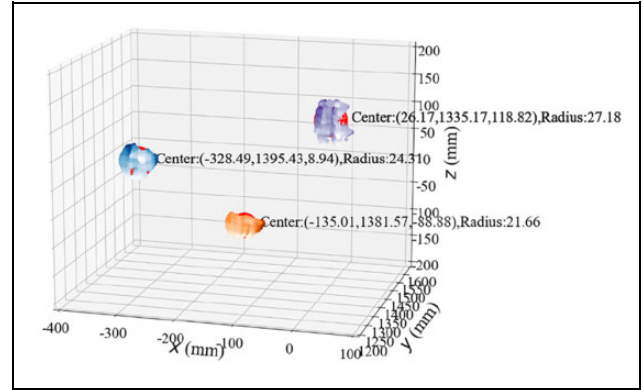


**Figure 12.** Process of image location: (a) left camera image; (b) normalized color difference; (c) eight-connected boundary tracking; (d) effective edges extraction; (e) pixels fitting; and (f) contour reconstruction of fruit.

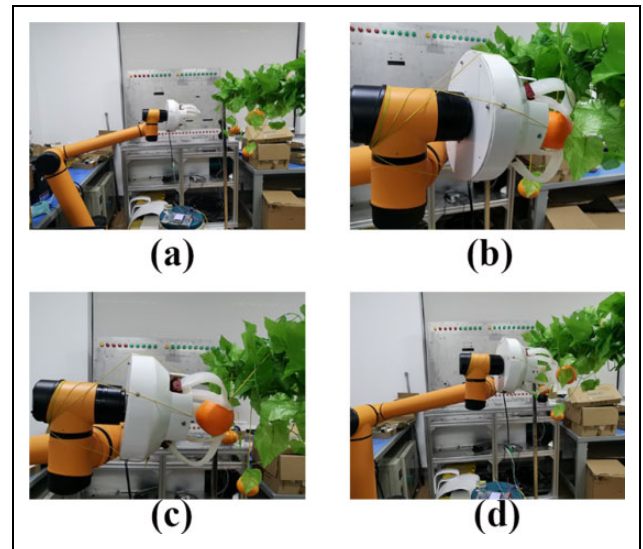


**Figure 13.** Spatial location of fruits target: (a) disparity map and (b) the spatial coordinates after removing outliers.

As shown in Figure 14, the coordinates of spherical center of reconstructed 3D model are  $(-328.49, 1395.43, 8.94)$ ,  $(-135.01, 1381.57, -88.88)$ ,  $(26.17, 1335.17, 118.82)$ , and the radii are 24.31, 21.66, 27.18; the result of 3D model reconstruction basically matched the size and position of fruit in the actual scene. According to the result of spatial position and 3D model reconstruction, it can provide effective size and spatial information for the mechanical arm and flexible grasping manipulator. When the base of mechanical arm is fixed, it takes about 90 s for the mechanical arm and the flexible grasping manipulator



**Figure 14.** Reconstruction of 3D model.



**Figure 15.** Process of fruit picking: (a) initial state; (b) contacting with the fruit; (c) picking the fruit; (d) task finished.

to complete the picking task of three fruits in turn from the initial position. This article takes the large one fruit as an example and not involve the process of cutoff stalk. The diagram of fruit picking experiment is shown in Figure 15.

The fruit picking experiment is carried out in an ideal environment at present, however, in the real scenario, there are many factors will affect the task of the fruit picking such as the instability of mobile platform and natural self-movements of the fruit. These factors should be considered in the follow-up outdoor experiment. In terms of image processing, precision and fast detection of the 3D coordinates will be studied in the follow-up research, and in terms of actuator, the grasping accuracy and speed should be increased.

## Conclusions

Location and model reconstruction of overlapped and sheltered fruits in unstructured natural environment is an essential prerequisite for fruit picking robot to achieve successful

picking. In this article, on the basis of traditional image processing algorithm, an effective edge extraction method based on geometry is proposed to reconstruct the actual contour of fruit, and then reconstructing the 3D model of each fruit based on least square method. Sixty fruit images are used to verify the feasibility of the proposed algorithm, and the experimental results show that the average SE, IOU, and FNR are 6.36%, 87.9%, and 6.72%, respectively, in natural environment. The vision system of fruit picking robot is verified in laboratory environment, the spatial distribution matches the fruit position in the actual scene, and the end effector could implement picking task successfully. It indicated that the fruit location and model reconstruction algorithm have preferable performance, and it can be applied to the design of vision system of fruit picking robot, which have certain guiding significance and engineering application prospect.

Considering the research direction of the research group, and the limitations of image samples and hardware devices, the traditional image processing algorithms were adopted in our project at present. We believe that the effective edge extraction method based on geometry in this article provides a new idea for the contour reconstruction of overlapped and sheltered fruits in the natural environment. Even using machine learning algorithms to recognize the fruits, the method can also be used to reconstruct the actual contour of the fruits. We tried to use Fully convolutional network (FCN) model provided through Github to segment fruits and branches. The 216 fruits in 60 test images were tested, and we recognized 185 fruits successfully. Although the evaluation criteria of the two methods are not consistent, it also shows that the method based on machine learning has great potential.


### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Li Li  <https://orcid.org/0000-0001-5320-2536>

### References

1. Zhuang JJ, Hou CJ, Tang Y, et al. Computer vision-based localisation of picking points for automatic litchi harvesting applications towards natural scenarios. *Biosyst Eng* 2019; 187: 1–20.
2. Xiong Y, Peng C, Grimstad L, et al. Development and field evaluation of a strawberry harvesting robot with a cable-driven gripper. *Comput Electron Agric* 2019; 157: 392–402.
3. Qiu YQ, Huang ZN, Rath T, et al. Review on application of machine vision in intelligent agricultural production. *Mech Res Appl* 2019; 32(2): 202–206.
4. Yu JJ, Rao HH, Luo ST, et al. Status and development trend of machinery picking of Nanfeng Citrus. *J Chin Agric Mech* 2017; 38(2): 135–138.
5. Xu WY, Chen H, Su Q, et al. Shadow detection and removal in apple image segmentation under natural light conditions using an ultrametric contour map. *Biosyst Eng* 2019; 184: 142–154.
6. Wang Y, Yang Y, Yang CH, et al. End-effector with a bite mode for harvesting citrus fruit in random stalk orientation environment. *Comput Electron Agric* 2019; 157: 454–470.
7. Lindner L, Sergiyenko O, Rivas-Lopez M, et al. Exact laser beam positioning for measurement of vegetation vitality. *Ind Rob* 2017; 44(4): 532–541.
8. Tang YC, Li LJ, Feng WX, et al. Binocular vision measurement and its application in full-field convex deformation of concrete-filled steel tubular columns. *Measurement* 2018; 130: 372–383.
9. Ivanov M, Sergiyenko O, Tyrsa V, et al. Influence of data clouds fusion from 3D real time vision system on robotic group dead reckoning in unknown terrain. *IEEE/CAA J Autom Sin* 2020; 7(2): 368–385.
10. Keerthy K, Tomas K, Simon P, et al. 3D-vision based detection, localisation and sizing of broccoli heads in the field. *J Field Robot* 2017; 34: 1505–1518.
11. Arad B, Balendonck J, Barth R, et al. Development of a sweet pepper harvesting robot. *J Field Robot* 2020; 37(6): 1027–1039.
12. Kirk R, Cielniak G, and Mangan M. L\*a\*b\*Fruits: a rapid and robust outdoor fruit detection system combining bio-inspired features with one-stage deep learning networks. *Sensors* 2020; 20: 275.
13. Sergiyenko O and Tyrsa V. 3D optical machine vision sensors with intelligent data management for robotic swarm navigation improvement. *IEEE Sens J* 2021; 21(10): 11262–11274.
14. Sergiyenko O, Ivanov M, Tyrsa V, et al. Data transferring model determination in robotic group. *Rob Auton Syst* 2016; 83: 251–260.
15. Bulanon DM, Burks TF, and Alchanatis V. Study on temporal variation in citrus canopy using thermal imaging for citrus fruit detection. *Biosyst Eng* 2008; 101(2): 161–171.
16. Zhuang JJ, Luo SM, Hou CJ, et al. Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications. *Comput Electron Agric* 2018; 152: 64–73.
17. Tao YT and Zhou J. Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking. *Comput Electron Agric* 2017; 142: 388–396.
18. Williams H, Jones MH, Nejati M, et al. Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms. *Biosyst Eng* 2019; 181: 140–156.

19. Majeed Y, Zhang J, Zhang X, et al. Apple tree trunk and branch segmentation for automatic trellis training using convolutional neural network based semantic segmentation. *IFAC PapersOnLine* 2018; 51(17): 75–80.
20. Wang DD, Song HB, Tie ZH, et al. Recognition and localization of occluded apples using K-means clustering algorithm and convex hull theory: a comparison. *Multimed Tools Appl* 2016; 75(6): 3177–3198.
21. Wang DD, Song HB, Yu XL, et al. An improved contour symmetry axes extraction algorithm and its application in the location of picking points of apples. *Span J Agric Res* 2015; 13(1): e02–005.
22. Xiong JT, Lin R, Liu Z, et al. The recognition of litchi clusters and the calculation of picking point in a nocturnal natural environment. *Biosyst Eng* 2018; 166: 44–57.
23. Xiong JT, He ZL, Lin R, et al. Visual positioning technology of picking robots for dynamic litchi clusters with disturbance. *Comput Electron Agric* 2018; 151: 226–237.
24. Chaivivatrakul S and Dailey MN. Texture-based fruit detection. *Precis Agric* 2014; 15(6): 662–683.
25. Rizon M, Yusri N, Kadir M, et al. Determination of mango fruit from binary image using randomized Hough transform. In: *Proceedings of SPIE. Eighth international conference on machine vision (ICMV 2015)*, Vol. , No. 987503, 8 December 2015. DOI: 10.1117/12.2228511.
26. Luo LF, Tang YC, Lu QH, et al. A vision methodology for harvesting robot to detect cutting points on peduncles of double overlapping grape clusters in a vineyard. *Comput Ind* 2018; 99: 130–139.
27. Fu LS, Tola E, Al-Mallahi A, et al. A novel image processing algorithm to separate linearly clustered kiwifruits. *Biosyst Eng* 2019; 183: 184–195.
28. Song HB, Zhang CD, Pan JP, et al. Segmentation and reconstruction of overlapped apple images based on convex hull. *Trans Chin Soc Agric Eng* 2013; 29(3): 163–168.
29. Song HB, He DJ, and Pan JP. Recognition and localization methods of occluded apples based on convex hull theory. *Trans Chin Soc Agric Eng* 2012; 28(22): 174–180.
30. Lu J and Sang N. Detection of citrus fruits within tree canopy and recovery of occlusion contour in variable illumination. *Trans Chin Soc Agric Mach* 2014; 45(4): 76–81+60.
31. Kelman E and Linker R. Vision-based localisation of mature apples in tree images using convexity. *Biosyst Eng* 2014; 118(1): 174–185.
32. Miao ZH, Shen YC, Wang XH, et al. Image recognition algorithm and experiment of overlapped fruits in natural environment. *Trans Chin Soc Agric Mach* 2016; 47(6): 21–26.
33. Chu GL, Zhang W, Wang YJ, et al. A method of fruit picking robot target identification based on machine vision. *J Chin Agric Mech* 2018; 39(2): 83–88.
34. Fitzgibbon A, Pilu M, and Fisher RB. Direct least square fitting of ellipses. *IEEE Trans Pattern Anal Mach Intell* 1999; 21(5): 476–480.