Rapid Recognition and Automatic Localization of Picking Points for Table Grapes in Natural Environments (Postprint)

Authors: Zhu Yanjun, Du Wensheng, Chunying Wang, Liu Ping, Li Xiang

Date: 2023-08-14T00:00:00+00:00

Abstract

[Objective/Significance] Rapid identification and precise localization of table grapes in natural environments is a prerequisite for realizing automatic grape harvesting by robots. [Method] This study proposes an automatic positioning method for table grape picking points based on an improved K-means clustering algorithm and contour analysis method. First, a weighted grayscale threshold is adopted as the similarity criterion for the clustering algorithm, and based on this, a K-means clustering algorithm with adaptive adjustment of the K-value is proposed to achieve rapid and effective identification and detection of table grapes; then, the proposed contour analysis method is utilized to obtain the fruit stem axis and the region of interest for picking points, and geometric methods are employed to achieve rapid and accurate localization of table grape picking points; finally, the algorithm proposed in this study is experimentally validated using 917 collected table grape images. [Results and Discussion] The success rate of the picking points localized by the proposed algorithm for table grapes being within an error of less than 12 pixels from the optimal picking points is 90.51%, with an average localization time of 0.87 s, achieving rapid and accurate localization of table grape picking points. Fifty simulation experiments were conducted under trellis and pergola planting methods respectively. The results show that the success rate of picking point localization for trellis-planted purple grapes is 86.00%, the success rate of identification and localization for pergola-planted purple grapes reaches 92.00%, the success rate of picking point localization for trellis-planted green grapes is 78.00%, and the success rate of identification and localization for pergola-planted green grapes is 80.00%. The overall experimental results are satisfactory. [Conclusion] This study can provide technical support for table grape harvesting robots to achieve precise grape picking.

Full Text

Rapid Recognition and Picking Points Automatic Positioning Method for Table Grape in Natural Environment

ZHU Yanjun[1], **DU Wensheng**[1,2], **WANG Chunying**[1], **LIU Ping**[1], ***LI Xiang***[3]

[1]Shandong Provincial Key Laboratory of Horticultural Machinery and Equipment/Intelligent Agricultural Machinery and Equipment Laboratory, College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China

[2]College of Construction Machinery, Shandong Jiaotong University, Jinan 250357, China

[3]National Key Laboratory of Wheat Improvement, College of Life Sciences, Shandong Agricultural University, Taian 271018, China

Abstract

[**Objective/Significance**] Rapid recognition and precise positioning of table grapes in natural environments is a prerequisite for robotic automatic harvesting. [**Methods**] This study proposes an automatic positioning method for table grape picking points based on an improved K-means clustering algorithm and contour analysis. First, a weighted gray threshold is adopted as the similarity criterion for the clustering algorithm, forming the basis for an adaptive K-value adjustment mechanism in the K-means clustering algorithm to achieve rapid and effective recognition and detection of table grapes. Then, a proposed contour analysis method is used to obtain the fruit stem axis and region of interest for the picking point, enabling rapid and accurate positioning of table grape picking points through geometric methods. Finally, the proposed algorithm is experimentally validated using 917 collected table grape images. [**Results and Discussion**] The results demonstrate that the proposed algorithm achieves a success rate of 90.51% for table grape picking point localization with an error of less than 12 pixels from the optimal picking point, with an average positioning time of 0.87 seconds. Simulation experiments conducted under hedgerow and trellis cultivation systems show that the success rate for purple grape picking point localization is 86.00% under hedgerow cultivation and 92.00% under trellis cultivation, while the success rates for green grape picking point localization are 78.00% and 80.00% respectively. [**Conclusion**] The overall experimental performance is satisfactory, indicating that this study can provide technical support for table grape harvesting robots to achieve precise grape picking.

Keywords: table grape; K-means clustering algorithm; contour analysis method; fruit stem axis; picking point; picking robot

1 Introduction

In recent years, China' s grape cultivation area and production have increased annually, yet table grape harvesting remains predominantly manual, with slow development of harvesting robots [1-3]. Automatic positioning of picking points is crucial for precise harvesting by grape-picking robots. However, the complex natural environment, irregular shapes of table grapes, similar colors between fruit stems and branches/leaves, and environmental factors such as foliage and lighting make grape picking point localization extremely challenging.

Numerous scholars have conducted research on grape recognition and harvesting. As early as 1995, Kondo et al. [4-6] proposed a multifunctional grape agricultural robot that utilized spectral reflectance and shape characteristics for grape detection. However, spectral reflectance is significantly affected by external humidity and weather conditions, resulting in complex constraints and substantial fluctuations in recognition accuracy in natural environments. Liu et al. [7] combined color and texture features for grape recognition experiments, achieving 88% accuracy. Reis et al. [8] identified and classified red and white grapes based on color images, with success rates of 97% and 91% respectively. Lei and Lu [9] utilized the U component of the YUV color model for grape bunch classification and employed a corner detection algorithm for picking point localization. Perez-zavala et al. [10] used texture and shape features combined with clustering algorithms for grape recognition. However, due to the similar pixel values between grapes and branches/dry leaves, coupled with external lighting and environmental factors, the specific color criteria exhibit large fluctuations, resulting in low generalizability of these methods. Berenstein et al. [11] achieved grape segmentation and localization using statistical measurements and shape matching algorithms, with a recognition success rate of 90%. Liu et al. [12] recognized and segmented overlapping grapes in natural environments, achieving an 89.71% success rate. Miao et al. [13] employed edge detection and contour fitting for detecting and measuring overlapping grape berries, with an average error rate of 1.5%, and further achieved picking point localization for overlapping grapes [14].

Compared to picking point localization, grape recognition and segmentation are relatively simpler. However, the variable morphology of grapes makes picking point localization far more challenging than grape segmentation and recognition. Some researchers have investigated picking point localization for cluster-like fruits such as grapes. Luo et al. [15] utilized the H component of the HIS color space for grape recognition and determined picking points based on the centroid, achieving an 88.03% localization success rate. Xiong et al. [16] rotated the R component of the RGB color channels and combined an improved Chan-Vese (C-V) level set method for grape background separation, using minimum bounding rectangles and Hough line detection for picking point localization, with recognition accuracy ranging from 80% to 92.5%. However, target recognition and segmentation using single color channels is severely affected by external disturbances, reducing accuracy in natural environments. Xiong et al. [17]

further studied disturbed grapes in unstructured environments, obtaining fruit stems through Otsu threshold segmentation and using point-line combination and Hough line fitting for picking point recognition. However, due to similar colors between grape stems and fruits, dry leaves, branches, and leaves, threshold segmentation struggled to achieve desired results, with picking point localization accuracy of only 80%. Luo et al. [18] proposed a binocular stereo vision-based method for obtaining spatial information of grape clusters, achieving rapid grape bunch localization but with only 87% accuracy. Yuan [19] proposed a hand-eye system for trellis grape harvesting robots based on depth vision, using depth cameras for image acquisition and background separation, and separating grape stems from bunches through depth point clouds. However, the target grapes required no occlusion and approximately consistent bunch heights, conditions difficult to satisfy in natural environments.

In recent years, deep learning has brought new opportunities for fruit recognition and picking point localization. Luo et al. [20] improved the YOLOv5s algorithm for detecting hairy and bagged peaches. Shang et al. [21] used improved YOLOX for rapid dragon fruit detection in natural environments. Zhou et al. [22] employed Mask R-CNN to separate backgrounds for red globe grape bunches and performed bunch segmentation through Hough transform. Ning et al. [23] utilized Mask R-CNN for grape stem recognition and segmentation, finally employing a region growing algorithm for precise stem segmentation and picking point localization, achieving 99.43% accuracy. Sun [24] improved Cascade R-CNN and E-Net networks for accurate segmentation of overlapping tomato clusters. However, deep learning methods require substantial time and effort for image annotation and demand high equipment performance during network training and mobile deployment, resulting in high costs.

To address these issues, this study proposes a table grape picking point automatic positioning method based on an improved K-means clustering algorithm and contour analysis to enhance the robustness and accuracy of table grape picking point localization, providing theoretical support for table grape harvesting robots to achieve precise grape picking.

2 Materials and Methods

2.1 Image Acquisition

From 07:30 to 11:30 on October 18, 2019, table grape images were collected at the Jinniu Mountain Base of Shandong Fruit and Vegetable Research Institute (117.18°E, 36.16°N) under sunny weather with temperatures ranging from 12°C to 23°C [Figure 1: see original paper]. The grape varieties included Summer Black, Moldova, and Youyong. Images were captured using the rear camera of a MI8 smartphone at approximately 12 megapixels (3024\$$\times$4032), $withhorizontalandverticalresolutionsof72dpi, aperturevalueoff f/1.8, andexposuretimeof1/10$ $neighborinterpolation, resultinginapproximately78,000pixels$(242$\times$\$322) to improve algorithm speed and practicality.

3 Research Methods

This study proposes a table grape picking point automatic positioning method based on an improved K-means clustering algorithm and contour analysis. First, a weighted gray threshold is used as the similarity criterion for the clustering algorithm, forming the basis for an adaptive K-value adjustment mechanism in the K-means clustering algorithm to achieve rapid and effective segmentation of table grapes. Second, an adaptive threshold algorithm is applied to obtain binary images of table grapes, followed by morphological processing, hole filling, and other denoising operations to ultimately obtain precise target grape images and their contours based on maximum connected domains. Then, an improved contour analysis method is used to obtain the fruit stem axis and picking point region of interest, and geometric methods are employed to achieve rapid and accurate positioning of table grape picking points. The algorithm flowchart is shown in Figure 2 [Figure 2: see original paper].

3.1 Grape Clustering Segmentation

3.1.1 Table Grape Clustering The K-means clustering algorithm [25-28] is a prototype-based partitioning algorithm that evaluates similarity among clusters in an image based on Euclidean distance—the closer the distance, the greater the similarity, and vice versa. The algorithm has two significant defects: first, the K-value is randomly assigned as a fixed value, while each image has unique characteristics, making a fixed K-value unable to meet the clustering requirements for every image. Second, Euclidean distance is a custom or fixed value that cannot adaptively change according to image variations, and calculating Euclidean distances between points consumes considerable time, making rapid table grape detection difficult.

To address these issues, this study improves K-means by proposing an algorithm with adaptively changing K-values and weighted gray threshold Dis as the evaluation criterion, enabling rapid and effective recognition and segmentation of table grapes. The improved clustering flowchart is shown in Figure 3 [Figure 3: see original paper].

Based on the clustering principle of K-means, the RGB image is first converted to grayscale, and an initial K-value is set. The grayscale image is then divided into K clusters according to grayscale levels, with any point having the most pixels at a grayscale level in a single cluster serving as the initial cluster center. After the first clustering, the planar coordinate information of each cluster center is obtained. To improve clustering accuracy, image rasterization is completed according to formula (1) to achieve initial clustering.

$$\text{Gray} = \frac{\max_i - \min_i}{\text{cluster}} \quad \text{where} \quad \text{mod}(\text{Gray}, K) = 0$$

where $\max_i$ is the maximum grayscale value in image i; $\min_i$ is the minimum grayscale value in image i; and cluster is the number of grayscale levels per cluster. Note that when $\mathrm{mod}(\mathrm{Gray}, K) \neq 0$, the Kth cluster contains $\mathrm{mod}(\mathrm{Gray}, K)$ grayscale levels.

After image rasterization, initial clusters and initial cluster centers are obtained (Figure 4 [Figure 4: see original paper]). Formulas (2) and (3) calculate the proportion of pixels in each initial cluster relative to total pixels and the grayscale difference between adjacent initial cluster centers.

$$\mathrm{Per}_i = \frac{\mathrm{Num\_cluster}(i)}{\mathrm{Num\_img}} \times 100\%, \quad i \in [1, K] \tag{2}$$

$$D_i = \mathrm{Gray}_{i+1} - \mathrm{Gray}_i, \quad i \in [1, K-1] \tag{3}$$

where Num_cluster(i) is the pixel count of a single cluster; Num_img is the total pixel count of the image; D_i is the grayscale threshold between adjacent cluster centers; Gray_i is the grayscale value of the ith cluster center; and Per_i is the percentage of pixels in each cluster relative to total image pixels.

To enhance clustering robustness, formulas (4) and (5) calculate the weighted gray threshold Dis.

$$\mathrm{Dis}_1 = \sum_{i=1}^{K-1} D_i \times (\mathrm{Per}_i + \mathrm{Per}_{i+1}) \tag{4}$$

$$\mathrm{Dis}_2 = D_1 \times \mathrm{Per}_1 \tag{5}$$

$$\mathrm{Dis} = \mathrm{Dis}_1 + \mathrm{Dis}_2$$

By comparing the grayscale difference D_i between two adjacent clusters with the weighted gray threshold Dis:

- When $D_i > \mathrm{Dis}$, the two clusters are determined to be different classes;
- When $D_i \leq \mathrm{Dis}$, the two clusters are determined to be the same class, the clusters are merged, the midpoint of the two clusters becomes the new cluster center, and the K-value is updated.

This process continues until the weighted gray threshold Dis no longer changes, at which point clustering ends. Formula (6) calculates Je_j, and the cluster division is output, with clustering results shown in Figure 5 [Figure 5: see original paper].

$$Je_j = \sum_{i=1}^{N(K_j)} |g_i - u_j|^2 \tag{6}$$

where Je_j is the sum of squared errors for each cluster, with smaller values generally indicating higher data compactness within cluster j and better clustering effect; g_i is the grayscale value of data points within cluster j; u_j is the grayscale value of cluster j' s center; and $N(K_j)$ is the number of sample data points in cluster j.

3.1.2 Target Grape Acquisition Analysis of clustering results (Figure 5 [Figure 5: see original paper]) reveals that grape branches, dry leaves, and other elements have similar colors and grayscale values to table grapes, resulting in significant noise within grape clusters. This study employs adaptive Otsu thresholding to segment the clustered grayscale image, obtaining a binary image (Figure 6(a) [Figure 6: see original paper]). Morphological opening and closing operations, hole filling, and extraction of the maximum connected domain are applied to obtain a precise binary image of the target grape (Figure 6(b) [Figure 6: see original paper]).

3.2 Contour Analysis Method

3.2.1 Contour Analysis The Canny operator is used to extract the contour of the table grape region (Figure 7(a) [Figure 7: see original paper]). The left and right extreme points A and B of the contour are identified, and all points on the contour are treated as mass points. The centroid C of the table grape is calculated using formula (7) (Figure 7(b) [Figure 7: see original paper]).

$$X_m = \frac{\sum m_i x_i}{\sum m_i}, \quad Y_m = \frac{\sum m_i y_i}{\sum m_i} \tag{7}$$

where the image' s top-left corner is the origin, with the vertical direction as the positive Y-axis and the horizontal direction as the positive X-axis; X_m is the horizontal coordinate of the centroid; Y_m is the vertical coordinate of the centroid; x_i and y_i are the horizontal and vertical coordinates of contour points; and m_i is the mass of each point (default mass of 1 for contour points).

In the absence of external interference, table grapes hang in the air due to gravity, and the fruit stem axis passes through the centroid. The fruit stem must lie between extreme points A and B. Therefore, the region of interest for the fruit stem axis is obtained based on extreme points A and B, ensuring the fruit stem is located within the contour region of interest (Figure 7(c) [Figure 7: see original paper]).

3.2.2 Fruit Stem Axis Acquisition Since the fruit stem axis passes through centroid C, points on the region of interest contour are connected sequentially with centroid C from left to right to form a bundle of candidate lines (Figure 8(a) [Figure 8: see original paper]), which must contain the actual fruit stem axis. Each candidate line divides the table grape region into left and right sub-regions (Figures 8(b) and 8(c) [Figure 8: see original paper]). The left sub-region is rotated 180° around the candidate line to obtain the image shown in Figure 8(d) [Figure 8: see original paper]. Based on the planar coordinate information from Figures 8(c) and 8(d), the pixel similarity P value [29] is calculated using formula (8). Higher similarity between the flipped left region and right region yields a smaller P value; the line corresponding to the minimum P value is identified as the fruit stem axis.

$$P = \sum_{(x_l, y_l)} \sum_{(x_r, y_r)} |(x_l, y_l) - (x_r, y_r)| \tag{8}$$

where P is the similarity value between the right region and the flipped left region of the candidate line; (x_l, y_l) represents pixel coordinates of the flipped left region; and (x_r, y_r) represents pixel coordinates of the right region.

3.2.3 Target Grape Picking Point Localization The fruit stem axis is determined through similarity P value analysis (Figure 9(a) [Figure 9: see original paper]). To accurately locate picking points and reduce environmental interference, the picking point region of interest is determined based on the intersection point D between the grape contour and fruit stem axis, the left and right extreme points A and B, and the centroid C. The length ROI_L is set to $0.8 \times L_{AB}$, and the height ROI_H is set to $0.25 \times L_{CD}$ (Figure 9(c) [Figure 9: see original paper]), where L_{AB} is the horizontal distance between extreme points A and B, and L_{CD} is the vertical distance between centroid C and intersection point D.

The central point O of the picking point region of interest is calculated using formula (9):

$$X_O = \frac{\sum_{i=1}^{4} x_i}{4}, \quad Y_O = \frac{\sum_{i=1}^{4} y_i}{4} \tag{9}$$

where X_O and Y_O are the horizontal and vertical coordinates of point O, and x_i and y_i are the horizontal and vertical coordinates of the four vertices of the bounding rectangle.

Two scenarios exist for locating picking point E: (1) When center O lies outside the fruit stem axis, the point E on the fruit stem axis with the shortest distance to point O is determined as the final picking point; (2) When center O lies on the fruit stem axis, point O is directly taken as the picking point. The localization results are shown in Figure 10 [Figure 10: see original paper].

4 Results and Analysis

4.1 Experimental Data and Error Analysis

Pixel positioning error [15] is used to analyze picking point localization accuracy. The positioning precision is evaluated through picking point pixel positioning error e, calculated using formula (10):

$$e_x = \min|X-x|, \quad e_y = \min|Y-y|, \quad e = \sqrt{e_x^2 + e_y^2} \tag{10}$$

where X and Y are the horizontal and vertical coordinates of the optimal picking point pixel region; x and y are the horizontal and vertical coordinates of the calculated picking point; e_x and e_y are the row and column coordinate errors between the calculated picking point and the optimal picking point pixel region; and e is the picking point pixel positioning error.

To ensure data validity, the optimal picking point coordinate range is selected from the fruit stem line region within the picking point region of interest. Based on points D and M, the optimal picking point coordinate range is manually set with the midpoint of D and M as the base point within a 6$×$30 pixel range. The error between picking points located by the proposed method and manually determined optimal picking points is calculated.

To validate the proposed method, recognition experiments were conducted on the collected table grape images, with 10 samples randomly selected for detailed analysis. Table 1 shows the clustering information statistics for the 10 images. Analysis reveals that the overall clustering effect is optimal when the initial K-value is in the range [18, 22]. Therefore, the initial K-value is set within this interval. The improved K-means method significantly improves clustering speed while maintaining comparable clustering effectiveness to the original algorithm.

Table 2 presents the picking point localization information for the 10 target grapes. Notably, in the 8th experiment, the upper boundary of the fruit stem region of interest exceeded the positive vertical coordinate region, resulting in a negative vertical coordinate for point M, but this did not affect accurate picking point localization. Figure 11 [Figure 11: see original paper] shows the 10 experimental results.

Given that harvesting robot end-effectors have certain tolerance ranges, a localization is considered successful when the error between the calculated table grape picking point and the optimal picking point is within 12 pixels. Analysis of Table 2 shows that only one picking point error exceeds 12 pixels. The average localization time for the 10 samples is 0.74 seconds. Among the 917 images tested, 87 failed in picking point localization, yielding an overall success rate of 90.51% with an average localization time of 0.87 seconds. Statistical analysis of failed cases indicates that lighting is the primary factor causing localization failure, while foliage occlusion and grape bunch overlap are secondary factors.

4.2 Comparative Analysis

A comparative analysis of the proposed algorithm was conducted, primarily comparing the improved K-means algorithm with the original K-means algorithm (Figure 12 [Figure 12: see original paper]). Figures 12(a) and 12(b) show clustering and localization failure and success cases for the original K-means algorithm, respectively. Using Euclidean distance as the clustering criterion results in numerous iterations and complex calculations, leading to long clustering times. Statistics show the original K-means algorithm has an average clustering time of 2.37 seconds. The fixed K-value introduces significant uncertainty in clustering results, with the same image producing different clustering and localization results due to the fixed and random nature of the K-value, making it difficult to adapt to various environments for table grape detection and localization and affecting subsequent positioning accuracy.

In contrast, the proposed algorithm' s image rasterization saves considerable time while completing K-value assignment and deployment appropriate for each image. Using gray threshold as the evaluation criterion significantly improves clustering speed while maintaining accuracy, providing a foundation for rapid picking point localization. Figure 12(c) shows the clustering and localization results using the improved K-means algorithm.

4.3 Simulation Tests

Simulation tests were conducted on a Denso robotic arm platform with a 32-bit microcontroller as the control core (Figure 13 [Figure 13: see original paper]) to validate the proposed method in a simulated environment. The results are shown in Figure 14 [Figure 14: see original paper] and Table 3 . Fifty simulation tests were performed for each of the four conditions: hedgerow purple grapes, trellis purple grapes, hedgerow green grapes, and trellis green grapes. The success rates were 86.00% for hedgerow purple grapes, 92.00% for trellis purple grapes, 78.00% for hedgerow green grapes, and 80.00% for trellis green grapes, with overall satisfactory results. Further analysis reveals that hedgerow grapes are more affected by foliage interference (occlusion, support, etc.), resulting in slightly lower success rates compared to trellis grapes. Green grapes are more severely affected by lighting and foliage occlusion (Figure 14(d) [Figure 14: see original paper]), leading to significantly lower picking success rates compared to purple grapes, though still meeting localization requirements for most scenarios.

5 Conclusion and Discussion

This study proposes a table grape picking point automatic positioning method based on an improved K-means clustering algorithm and contour analysis, achieving rapid and accurate positioning of table grapes in natural environments:

1. **Improved K-means clustering algorithm**: A weighted gray threshold is adopted as the similarity criterion, enabling adaptive K-value adjustment to meet clustering requirements for table grapes in natural environments and achieve rapid and effective segmentation.

2. **Proposed contour analysis and point-line combination method**: Contour analysis obtains the fruit stem axis region of interest, and similarity P-value is used to locate the fruit stem axis line. Combined with

contour analysis and point-line methods, the picking point region of interest is obtained, enabling rapid and accurate picking point localization.

3. **Algorithm validation**: Using Summer Black, Moldova, and Youyong table grape varieties, the proposed method achieves an average success rate of 90.51% with errors less than 12 pixels from optimal picking points, and an average localization time of 0.87 seconds.
4. **Simulation validation**: The hedgerow purple grape picking point localization success rate is 86.00%, trellis purple grape recognition and localization success rate reaches 92.00%, hedgerow green grape picking point localization success rate is 78.00%, and trellis green grape recognition and localization success rate is 80.00%, demonstrating overall satisfactory performance. The proposed method can rapidly and accurately achieve automatic table grape picking point localization, providing a theoretical basis for table grape harvesting robots to precisely harvest grapes.

This study has not yet addressed the issue of overlapping table grapes, which will be the focus of future research, along with in-depth investigation of lighting and foliage occlusion conditions.

Conflict of Interest Statement: The authors declare no conflicts of interest related to this research.

References

[1] JIE Y J. Research status, problems and countermeasures of fruit and vegetable picking robot[J]. Times agricultural machinery, 2018, 45(9): ID 42.

[2] YUAN Z Y, CHEN J, GUO P, et al. Analysis on the development status, problems and countermeasures of fruit and vegetable picking robot[C]// International Academic Annual Conference of Chinese Society of Agricultural Machinery. Beijing, China: Chinese Society of Agricultural Machinery, 2012.

[3] ZHANG J, LI Y W. Research situation, problems and solutions of fruit-vegetable picking robots[J]. Journal of machine design, 2010, 27(6): 1-5.

[4] KONDO N, SHIBANO Y, MOHRI K, et al. Basic studies on robot to work in vineyard (Part 1) manipulator and harvesting hand[J]. Journal of the Japanese society of agricultural machinery, 1993, 55(6): 85-94.

[5] KONDO N, SHIBANO Y, MOHRI K, et al. Basic studies on robot to work in vineyard (Part 2) discriminating, position detecting and harvesting experiments by using visual sensor[J]. Journal of the Japanese society of agricultural machinery, 1994, 56(1): 45-53.

[6] MONTA M, KONDO N, SHIBANO Y, et al. Basic studies on robot to work in vineyard (Part 3) measurement of physical properties for robotization and manufacture of berry thinning hand[J]. Journal of the Japanese society of agricultural machinery, 1994, 56(2): 93-100.

[7] LIU S, WHITTY M. Automatic grape bunch detection in vineyards with an SVM classifier[J]. Journal of applied logic, 2015, 13(4): 643-653.

[8] REIS M J C S, MORAIS R, PERES E, et al. Automatic detection of bunches of grapes in natural environment from color images[J]. Journal of applied logic, 2012, 10(4): 285-290.

[9] LEI W X, LU J. Visual positioning method for picking point of grape picking robot[J]. Jiangsu Journal of Agricultural Sciences, 2020, 36(4): 1015-1021.

[10] PEREZ-ZAVALA R, MIGUEL T T, AUAT C F, et al. A pattern recognition strategy for visual grape bunch detection in vineyards[J]. Computers and Electronics in Agriculture, 2018, 151: 136-149.

[11] BERENSTEIN R, SHAHAR OBEN, SHAPIRO A, et al. Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer[J]. Intelligent service robotics, 2010, 3(4): 233-243.

[12] LIU P, ZHU Y J, ZHANG T X, et al. Algorithm for recognition and image segmentation of overlapping grape cluster in natural environment[J]. Transactions of the Chinese society of agricultural engineering, 2020, 36(6): 161-169.

[13] MIAO Y B, HUANG L L, ZHANG S. A two-step phenotypic parameter measurement strategy for overlapped grapes under different light conditions[J]. Sensors, 2021, 21(13), ID 4532.

[14] LUO L F, TANG Y C, LU Q H, et al. A vision methodology for harvesting robot to detect cutting points on peduncles of double overlapping grape clusters in a vineyard[J]. Computers in industry, 2018, 99: 130-139.

[15] LUO L F, ZOU X J, XIONG J T, et al. Automatic positioning for picking point of grape picking robot in natural environment[J]. Transactions of the Chinese society of agricultural engineering, 2015, 31(2): 14-21.

[16] XIONG J T, LIU Z, LIN R, et al. Green grape detection and picking-point calculation in a night-time natural environment using a charge-coupled device (CCD) vision sensor with artificial illumination[J]. Sensors, 2018, 18(4): ID 969.

[17] XIONG J T, HE Z L, TANG L Y, et al. Visual localization of disturbed grape picking point in non-structural environment[J]. Transactions of the Chinese society for agricultural machinery, 2017, 48(4): 29-33, 81.

[18] LUO L F, TANG Y C, ZOU X J, et al. Vision-based extraction of spatial information in grape clusters for harvesting robots[J]. Biosystems Engineering, 2016, 151(11): 90-102.

[19] YUAN Y. Hand-eye system design and close-shot stem recognition for robotic harvesting of trellis grape clusters based on depth sensing[D]. Zhenjiang: Jiangsu University, 2019.

[20] LUO Q, RAO Y, JIN X, et al. Multi-class on-tree peach detection using improved YOLOv5s and multi-modal images[J]. Smart Agriculture, 2022, 4(4):

84-104.

[21] SHANG F N, ZHOU X C, LIANG Y K, et al. Detection method for dragon fruit in natural environment based on improved YOLOX[J]. Smart agriculture, 2022, 4(3): 120-131.

[22] ZHOU W J, ZHA Z H, WU J. Maturity discrimination of "Red Globe" grape cluster in grapery by improved circle Hough transform[J]. Transactions of the Chinese society of agricultural engineering, 2020, 36(9): 205-213.

[23] NING Z T, LUO L F, LIAO J X, et al. Recognition and the optimal picking point location of grape stems based on deep learning[J]. Transactions of the Chinese society of agricultural engineering, 2021, 37(9): 222-229.

[24] SUN B Y. Research on tomato fruit target detection and tomato string picking point location technology based on deep learning[D]. Tianjin: Tianjin University of Technology, 2021.

[25] JIANG M, SUN S S, HE D J, et al. Recognition and localization method of occluded apples based on K-means clustering segmentation algorithm and convex hull theory[J]. Smart agriculture, 2019, 1(2): 45-54.

[26] CHEN H, MA J, CHEN S H, et al. Automatic extraction of effective points from fringe projection profilometry based on improved K-means clustering and Otsu algorithms[J]. Laser journal, 2020, 41(9): 40-46.

[27] XIA Y F. Cotton HSV image segmentation based on K-means clustering and two-dimensional otsu[J]. Computer engineering & software, 2020, 41(7): 170-173.

[28] CAO S S, CHEN X X, MIAO P, et al. Image segmentation method based on optimization of PSO algorithm and K-means clustering algorithm[J]. Computer and modernization, 2020(1): 22-27.

[29] YU Y, ZHANG K L, YANG L, et al. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN[J]. Computers and electronics in agriculture, 2019, 163: ID 104846.

Note: Figure translations are in progress. See original paper for figures.