IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 17, 2019, accepted September 27, 2019, date of publication October 8, 2019,
date of current version November 11, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2946267*

# Grasping Point Detection of Randomly Placed Fruit Cluster Using Adaptive Morphology Segmentation and Principal Component Classification of Multiple Features

**QIAN ZHANG AND GUOQIN GAO**
School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China

Corresponding author: Guoqin Gao (gqgao@ujs.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51375210, in part by the Zhenjiang Municipal Key Research and Development Program under Grant GZ2018004, in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions, and in part by the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant KYCX17_1780.

**ABSTRACT** Precise and rapid grasping point detection based on machine vision is one of the challenging problems in automatic sorting of randomly placed fruit clusters by robot. Grasping stalk of fruit cluster can improve grasping success probability and reduce fruit damage. For the problem that the segmentation of stalk candidates for randomly placed fruit cluster based on existing morphology algorithm tends to be low precision, an improved image segmentation algorithm based on adaptive morphology is proposed. According to edge distances defined based on minimum distance between edge point and unconnected components in minimum domain, the adaptive convolution kernel is constructed. In addition, a run analysis method with different and unordered labels is designed to reduce calculation time of edge distances. For the problem that it is difficult to describe and classify unconstraint stalk by existing features, an improved region classification algorithm based on principal components of multiple features is proposed. The descriptors based on features of object region are designed and principal components of multiple features are extracted based on variance contribution to improve precision and speed of stalk extraction. The proposed grasping point detection method of randomly placed fruit cluster based on improved morphology image segmentation and region classification algorithms is verified by experiments with grape clusters based on parallel robot sorting system. The results show that, compared with existing methods, the average precisions of segmentation and extraction for stalk increase by 9.89% and 2.17% respectively, the average precision and time of grasping point detection reach 94.50% and 2.01s respectively.

**INDEX TERMS** Image segmentation, grasping, morphology, machine vision, robots, image classification.

## I. INTRODUCTION

The automatic sorting of fruits based on robot technology is of great significance to the automated, large-scale, and accuracy development of agricultural production and agriculture product processing [1]–[3]. During the automatic sorting, the accurate and reliable detection of grasping point is the precondition to realize the accurate, fast, and nondestructive grasping control of robot [4]–[6]. Currently, the main methods detecting grasping point include infrared image analysis, spectrum analysis, and machine vision [7]–[10]. Compared with other methods, the machine vision has the advantages of noncontact, good adaptability, cost-effectiveness, etc. That is more suitable to solve the problem of grasping point detection during automatic sorting of fruits [11]–[13]. Compared with independent fruit [14], such as apple, pear, and pineapple, the grasping point detection of fruit cluster, such as grape, longan, and lychee, is difficult, because of the variety of cluster morphology, a lot of soft berry, complex edge of fruit image [15], [16]. In addition, the automatic sorting of fruit clusters based on parallel robot, which features high accuracy, rigidity, speed, and large load carrying capability,

The associate editor coordinating the review of this manuscript and approving it for publication was Aysegul Ucar.

places greater demands on precision and speed of grasping point detection. In the process of automatic sorting, the fruit cluster is randomly placed on a plane before the robot grasps the fruit cluster. Therefore, this paper focuses on the problem of grasping point detection for randomly placed fruit cluster. It can lay the foundation for accurate, fast, and nondestructive automatic sorting of fruit clusters based on parallel robot and machine vision.

There are few studies on grasping point detection of randomly placed fruit cluster. The existing researches about grasping point detection mainly aim at picking study for the fruit cluster hanging from the vine. The stalk of hanging fruit cluster has the characteristics of locating above the cluster and approximating straight line. Accordingly, the existing methods adopt edge and straight-line detection algorithms to extract stalk. The detected edges or lines are selected to detect grasping point or picking point based on the hanging constraint [17]–[20]. For instance, some scholars detected straight-lines in the fixed region above the cluster and calculated picking point according to the minimum distance between center of gravity and detected lines. The method can detect the picking point of hanging fruit cluster with high success probability, but the detecting precision is affected by the position accuracy of center of gravity and the edges of pedicel, rachis, and others [21]. Compared with the hanging fruit cluster, due to the placement randomness of fruit cluster, the edge distances of stalk, berry, pedicel and others are random, and the shape and position of stalk are unconstraint. Therefore, the existing grasping point detection methods of hanging fruit cluster are not suitable for the randomly placed fruit cluster. This paper presents a grasping point detection method of randomly placed fruit cluster using adaptive morphology segmentation and principal component classification of multiple features to solve the problem that it is difficult to detect grasping point of randomly placed fruit cluster. The precision and speed of grasping point detection are improved to meet the grasping requirement of automatic sorting based on parallel robot.

According to the growth characteristics of fruit cluster, the stalk has good toughness and grasping stalk can reduce damage of fruit cluster and realize stabilized and efficient automatic sorting. So segmentation of stalk candidates and stalk extraction for randomly placed fruit cluster are the key problems of grasping point detection. The existing region segmentation methods that can be used for segmentation of stalk candidates mainly include threshold segmentation, contour segmentation, fuzzy c-mean clustering, and morphology [22]–[24]. Compared with other methods, the morphology is applied widely due to its efficiency [25]. However, the existing morphology algorithms need to preset the size of convolution kernel according to experiments or experience and eliminate non-object regions based on image convolution. Due to the placement randomness of fruit cluster, the edge distances of stalk, berry, pedicel and others are random in image. It is difficult to find an optimal convolution kernel for segmentation of stalk candidates for different randomly placed fruit cluster image. Therefore, for the problem that the segmentation of stalk candidates for randomly placed fruit cluster based on existing morphology algorithm tends to be low precision, an improved adaptive morphology image segmentation algorithm based on edge distances is proposed. Based on separation threshold of edge distances that are defined based on the minimum distance between edge point and unconnected components in minimum domain, the adaptive convolution kernel is constructed. In addition, a run analysis method with different and unordered labels is designed to reduce calculation time of edge distances.

The main methods that can be used for stalk extraction include two categories: region selection and region classification. The region selection is realized by threshold judgement based on region feature. This method is efficient, but the precision and robustness are not satisfactory [26], [27]. The existing region classification methods include two categories: feature-based and pixel-based. There is a problem of information redundancy, though lots of information can be retained by using image pixel as input of classification [28]–[31]. The existing region classification method based on feature has the advantages of high precision and speed, good adaptability [32]. However, due to unconstraint shape and position of stalk of randomly placed fruit cluster, the feature of stalk region and the difference in feature between stalk region and non-stalk region are not obvious. It is difficult to describe and classify stalk by existing features. In addition, the precision and speed of stalk extraction are greatly affected by the size of feature vector and the coupling between features in feature vector. Therefore, an improved neural network region classification algorithm based on principal components of multiple features is proposed. The descriptors based on features of object region are designed and the principal components of multiple features are extracted based on variance contribution.

The improved morphology image segmentation and region classification algorithms are used for the research on grasping point detection of randomly placed fruit cluster to improve the precision and speed. It can lay the foundation for accurate, fast, and nondestructive automatic sorting of randomly placed fruit clusters based on parallel robot and machine vision.

## II. IMPROVED ADAPTIVE MORPHOLOGY IMAGE SEGMENTATION ALGORITHM BASED ON EDGE DISTANCES

### A. RUN ANALYSIS AND EDGE DISTANCE CALCULATION IN MINIMUM DOMAIN

#### 1) RUN ANALYSIS WITH DIFFERENT AND UNORDERED LABELS

The main methods detecting and marking unconnected component include run-based, contour-based, quadtree-based [33]. Due to the factors of complex edge, various holes, and branches in the image of fruit cluster, it is complicated to use the contour-based and quadtree-based methods. The run-based method has the advantages of fast, efficient, and simple,

and it is suitable for detecting and marking unconnected component in the image of fruit cluster. However, the existing run analysis method spends much time to scan image twice and uses equivalent sequence to record the labels of connected components in neighbouring rows. Considering the purposes of detecting the components unconnected with the component including the edge point in domain and marking them with different labels, this paper improves the existing run analysis method.

Compared with the existing run analysis method that needs to scan image twice, the proposed run analysis method with different and unordered labels does not need equivalent sequence and can obtain the unconnected components with different labels by only scanning image once. When the run connects with more than one run in the previous row on the first scanning, the detecting run and all the connected run are marked with the label which is the minimum label of the connected runs in the previous row. It can improve the detecting speed and efficiency. The specific steps are as follows:

(a) The image is scanned row by row. The consecutive nonzero points are combined to compose a run in every row. The start point, end point, and row number of the run are recorded.

(b) Every run in the first row are marked. The labels are the natural number beginning from 0 and increasing in sequence.

(c) Judging whether every run in every row except the first row connects with the run in the previous row with 8 neighborhood and marking the run: If no connection, the detecting run is marked with a new label and the run label in the previous row remain unchanged. If only one run in the previous row is connected, the detecting run is marked with the label same as the connected run in the previous row. If more than one run in the previous row is connected, the detecting run and all the connected run are marked with the label that is minimum label of the connected run in the previous row.

(d) After marking every run in every row, the unconnected components with different and unordered labels are obtained.

### 2) EDGE DISTANCE CALCULATION IN MINIMUM DOMAIN

According to the distribution feature of edges in image, the minimum domain centered on each edge point is used for edge distance calculation. The minimum domain includes the edge point, the component including the edge point, and the components unconnected with the component including the edge point. The minimum distance between the edge point and unconnected components is calculated by cyclical iteration and defined as edge distance of image in minimum domain. The flow diagram is shown in Fig.1.

Every point in image with $m * n$ is selected successively from left to right and from top to bottom firstly. If the pixel of point is nonzero value, the point is marked. Then $k$ marked points are obtained. Then the unconnected components are detected in the domain $M_{ij}$ with $j * j$ centered on every point $i.j = 3, 4, 5 \ldots max_{mn}$, and $max_{mn}$ is the larger one between $m$ and $n$. The $j$ is assigned from 3 to $max_{mn}$ and the unconnected

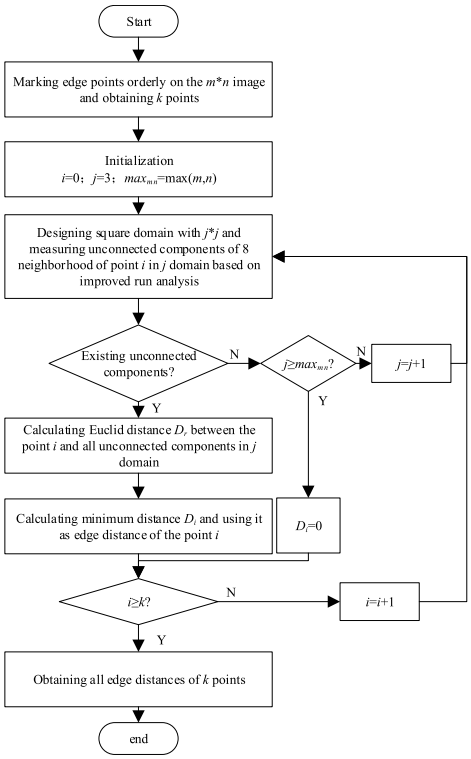



**FIGURE 1.** Edge distance calculation.

components are detected, until the component unconnected with the component including the point $i$ in the domain $M_{ij}$ is existing or $j$ is the $max_{mn}$. The proposed run analysis method with different and unordered labels is adopted to detect the unconnected components. If there is no unconnected components in the domain $M_{ij}$ when $j$ is assigned to $max_{mn}$, the edge distance of the point $i$ is 0. If there are unconnected components when $j = 3, 4, 5 \ldots max_{mn}$, $j$ is assigned to minimum value with which the unconnected components are existing. That $j$ is assigned to minimum value means the $M_{ij}$ is minimum domain. The Euclid distances between the point $i$ and all unconnected components in the minimum domain $M_{ij}$ are calculated. Then the minimum value of the Euclid distances is used as edge distance of the point $i$. The specific steps is as follows.

(a) The maximal label of obtained unconnected components in $M_{ij}$ is $b$. The label of the component including the point $i$ is $C_i$.

(b) According to the order of labels, the distances between the point $i$ and all unconnected components are calculated in sequence by the edge point-by-point method. For every unconnected component, the distance $D_{re}$ between the point $i$ and the edge points of unconnected component are calculated according to (1).

$$D_{re} = \sqrt{(x_i - x_e)^2 + (y_i - y_e)^2} \tag{1}$$

The $r$ is the label of unconnected component and is assigned from 0 to $b$,$e$ is the label of edge point and $e = 0, 1, 2, \ldots, E - 1$,$E$ is the number of edge points, $(x_i, y_i)$ is

the image coordinate of point $i$, $(x_e, y_e)$ is image coordinate of edge point $e$.

(c) The minimum distance between the point $i$ and the edge points of unconnected component is used as the distance $D_r$ between the point $i$ and the unconnected component. And the minimum distance $D_i$ between the point $i$ and unconnected components is used as the edge distance of the point $i$.

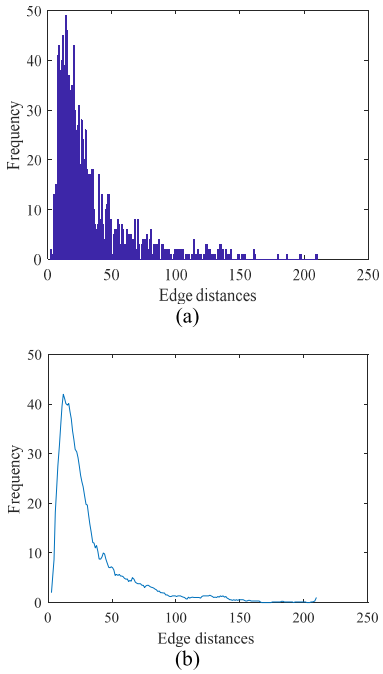



**FIGURE 2.** (a) The distribution histogram of edge distances. (b) The smoothed distribution histogram of edge distances.

### B. IMPROVED ADAPTIVE MORPHOLOGY IMAGE SEGMENTATION ALGORITHM

This paper designs adaptive convolution kernel for segmenting object regions according to the all edge distances between edge point and unconnected components in minimum domain. The edge distances between backgrounds (berries and others) are relatively far and the edge distribution is sparse. The edge distances between stalk candidates (stalks, stalk nodes, rachises and pedicels) are relatively near and the edge distribution is dense. So the separation threshold between edges of stalk candidates and backgrounds can be calculated according to distribution of edge distances. The distribution histogram of edge distances can be obtained by counting the all distances $D_i$ as shown in Fig.2. The $x$-axis represents the $D_i$, and the $y$-axis represents the frequency of $D_i$. After smoothing the distribution histogram, the feature of edge distribution is more apparent. The parts with high frequency and near distance in the distribution histogram mainly correspond to edges of stalk candidates. The parts with low frequency and far distance mainly correspond to edges of backgrounds. So the distance locating on the falling edge and the 78% of peak on the distribution histogram is used as the separation threshold $X_p$ of edge distances.

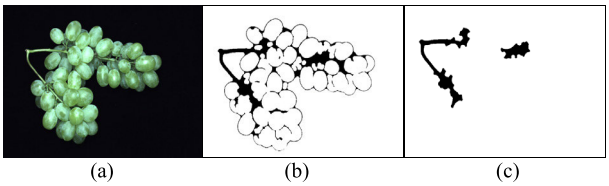



**FIGURE 3.** Results of image segmentation. (a) Original image. (b) The result of Close operation with adaptive convolution kernel. (c) The result of Open operation with adaptive convolution kernel.

The separation threshold $X_p$ of edge distances is used to design adaptive convolution kernel of morphology algorithm. Firstly the Close operation is performed in the edge image of fruit cluster and the convolution kernel size of the Close operation is $X_p * X_p$. The regions of stalk candidates with dense edge distribution are closed and the holes of other regions are retained as shown in Fig.3(a). Secondly the Open operation is performed in the processed image and the convolution kernel size of the Open operation is 1/2 of $X_p * X_p$. The thin edges of big holes are removed and the closed regions of stalk candidates are retained. Finally, the remaining background regions can be selected and removed according to region area. The result of image segmentation is shown in Fig.3(b). This proposed method constructs adaptive convolution kernel according to separation threshold of edge distances, which improves the precision and speed of image segmentation.

## III. IMPROVED NEURAL NETWORK REGION CLASSIFICATION ALGORITHM BASED ON PRINCIPAL COMPONENTS OF MULTIPLE FEATURES

This paper proposes an improved neural network region classification algorithm based on principal components of multiple features. The descriptors that can represent the difference between stalk region and non-stalk region are designed and the multidimensional feature vector is constructed based on the descriptors. The principal components of multiple features are extracted based on variance contribution to reduce dimension and decouple.

### A. CONSTRUCTING, DIMENSIONALITY REDUCTION AND DECOUPLING OF MULTIDIMENSIONAL FEATURE VECTOR

The 20 descriptors based on region features are designed as follows: area of region $a$(2), length of contour $L$ (3), width $W$ (4) and height $H$ (5) of region, area secondary moment $N_{pq}$ (7), center secondary moment $U_{pq}$ (9), anisometry $Ani$(13), bulkiness $Bul$(14) and struct factor $SF$ (15) of equivalent ellipse, rectangularity $Rec$ (16), convexity $Con$(17), circumcircle radius $r_o$, incircle radius $r_i$, length $h_i$ and width $w_i$ of inscribed rectangle, circularity $Cir$(18), mean distance $Mdis$ (19) between contour and center of area, distance deviation $Dd$ (20), roundness $Rou$ (21), and compactness $Com$ (22).

$$a = \sum_{(x,y)\in R} 1 \tag{2}$$

$$L = \sum_{i=2,j=2}^{n_c} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \quad (3)$$

$$W = x_{\max} - x_{\min} \quad (4)$$

$$H = y_{\max} - y_{\min} \quad (5)$$

The $(x, y)$ is the image coordinate of point, $R$ is the region, $n_c$ is the number of point in contour. The geometric feature of region is described by translation-invariant, rotation-invariant, and scale-invariant geometric moment. The $(p + q)$ order geometric moment $M_{pq}$ is expressed as follows:

$$M_{pq} = \sum_{(x,y)\in R} x^p y n^q. \quad (6)$$

The sizes of stalk regions are different. Because of the randomness of region position, the position is independent of the region category. In order to avoid the influence of region size and position on geometric moment, the area and center of region are used to normalize the geometric moment according to (7)-(9).

$$N_{pq} = (1/a) \sum_{(x,y)\in R} x^p y^q \quad (7)$$

$$O_r = (N_{10}, N_{01}) = ((1/a) \sum_{(x,y)\in R} x^1 y^0, (1/a) \sum_{(x,y)\in R} x^0 y^1) \quad (8)$$

$$U_{pq} = (1/a) \sum_{(x,y)\in R} (x - N_{10})^p (y - N_{01})^q \quad (9)$$

The $O_r$ is center of region. The normalized geometric moment is adopted to define the ellipse equivalent to the region, which can describe the orientation and aspect ratio of the region. The center of equivalent ellipse is the same as the center of region. The semi-minor axis $r_2$, the semi-major axis $r_1$ and the orientation $\alpha$ of the semi-major axis with regard to the $x$-axis can be calculated according to geometric moment as follows:

$$r_1 = \sqrt{2(U_{20} + U_{02} + \sqrt{(U_{20} - U_{02})^2 + 4U_{11}^2})} \quad (10)$$

$$r_2 = \sqrt{2(U_{20} + U_{02} - \sqrt{(U_{20} - U_{02})^2 + 4U_{11}^2})} \quad (11)$$

$$\alpha = -(1/2) \arctan(2U_{11}/(U_{02} - U_{20})). \quad (12)$$

Then the anisometry *Ani*, bulkiness *Bul*, and struct factor *SF* can be calculated according to equivalent ellipse as follows:

$$Ani = r_1/r_2 \quad (13)$$

$$Bul = \pi r_1 r_2 / a \quad (14)$$

$$SF = Ani * Bul - 1. \quad (15)$$

The ratio of the area of region to that of rectangle with the same first and second order moments of the region is taken as rectangularity based on geometric moment as follows:

$$Rec = a/a_{Rec}. \quad (16)$$

Regarding a region as a point set, if all points on a straight line connected by any two points in the point set are in the point set, the point set is a convex set. Convex set transformation is performed for region, and the minimum convex set containing all points in the region is regarded as the convex hull of the region. The convexity *Con*, circumcircle, and incircle of the region are calculated by convex hull to describe the shape features of region.

$$Con = a/a_{Con} \quad (17)$$

For the circular feature of region, two descriptors are designed: circularity *Cir* and roundness *Rou*. The circularity *Cir* is calculated based on circumcircle area $a_{Cir}$, which can represent the degree to which the shape of the region is close to that of a circle, as shown in (18). The roundness *Rou* is calculated based on the distance between contour and center of area. The roundness can be represented by mean distance *Mdis* and distance deviation *Dd* according to (19)-(21).

$$Cir = a/a_{Cir} \quad (18)$$

$$Mdis = (1/a) \sum_{(x,y)\in R} \sqrt{(x - N_{10})^2 + (y - N_{01})^2} \quad (19)$$

$$Dd = \sqrt{(1/a) \sum_{(x,y)\in R} (\sqrt{(x - N_{10})^2 + (y - N_{01})^2} - Mdis)^2} \quad (20)$$

$$Rou = 1 - Dd/Mdis \quad (21)$$

The compactness is defined based on the relationship between length and area of region as follows:

$$Com = L^2/4a\pi. \quad (22)$$

The 20 descriptors are adopted to construct multidimensional feature vector $\boldsymbol{M}_F = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_{20})$. The $\boldsymbol{x}_i$ is column vector and represents the descriptors. The elements of $\boldsymbol{x}_i$ represent feature values of sample regions. The number of rows in $\boldsymbol{M}_F$ is the number of sample regions.

Using multidimensional feature vector can improve classification precision, but it is also a time-consuming process. In addition, the correlation between features also has an effect on precision. So this paper extracts the principal components of multiple features based on variance contribution to reduce dimension and decouple for $\boldsymbol{M}_F$. Firstly according to (23) [34], the covariance matrix $\boldsymbol{Cov} = (s_{ij})_{p\times p}$ of $\boldsymbol{M}_F$ is calculated.

$$s_{ij} = (1/(n-1)) \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) i, j = 1, 2, \ldots, p \quad (23)$$

The $\bar{x}_i$ and $\bar{x}_j$ represent the average of $x$ elements in the $i$ and $j$ column of $\boldsymbol{M}_F$ respectively, $n$ and $p$ are the row and column number of $\boldsymbol{M}_F$ respectively.

Then, the eigenvalues $\lambda_i$ and unit eigenvectors $\boldsymbol{b}_i$ of covariance matrix are calculated. The eigenvalues $\lambda_i$ are ranked by size, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_p$. According to (24), the variance contribution $\alpha_i$ is calculated. The variance contribution

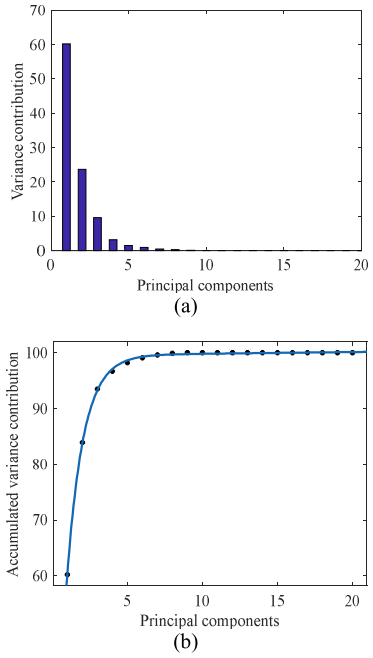



**FIGURE 4.** (a) The variance contribution. (b) The accumulated variance contribution.

can represent the amount of information included in feature vector.

$$\alpha_i = \lambda_i / \sum_{i=1}^{p} \lambda_i \quad i = 1, 2, \ldots, p \tag{24}$$

Taking the images of White Rosa grape clusters as an example, the variance contribution of every component in constructed multidimensional feature vector $\boldsymbol{M}_F$ of stalk region is shown in Fig.4. The first four are 60.2%, 23.7%, 9.6%, and 3.2% respectively. The accumulated variance contribution of first four principal components is 96.7%. It means that the first four principal components include 96.7% of the amount of original information. According to experiments, the principal components including more than 95% of the amount of original information are chosen and used in this paper. According to (25), the new feature vector $\boldsymbol{N}_{Fi}$ can be calculated by dimensionality reduction and decoupling. The $t$ is the number of used principal components. The $\boldsymbol{b}_i$ is the unit eigenvectors of covariance matrix of $\boldsymbol{M}_F$.

$$\boldsymbol{N}_{Fi} = \boldsymbol{b}_i * \boldsymbol{M}_F \quad i = 1, 2, \ldots, t \tag{25}$$

### B. NEURAL NETWORK REGION CLASSIFICATION

In image classification, compared with traditional algorithms such as support vector machines (SVM), k-nearest neighbor (kNN) and k-means, the neural network has more advantages in nonlinear mapping, self-learning and generalization [35]. The neural network can map complex non-linear relationships without building mathematical equations in advance, and distinguish features and details that are difficult to recognize in images. The existing neural networks for image classification mainly include back propagation (BP) neural network, long short-term memory (LSTM) network, recursive neural network (RNN), deep belief network (DBN), deep convolutional network (DCN) and so on. Compared with the shallow neural networks such as LSTM and RNN, the structure of BP neural network with good nonlinear mapping approximation, fault-tolerant and generalization performances is simpler when solving the same problem. In addition, the input of neural network is the multidimensional feature vector constructed with dimensionality reduction and decoupling which can describe the region more comprehensively by less features in this paper. Therefore, compared with the complex deep neural networks that need more training time and samples, such as DBN and DCN, the BP neural network with learning and adaptive abilities can be adopted to classify stalk of randomly placed fruit clusters rapidly and precisely.

After dimensionality reduction and decoupling, the BP neural network of $t$-10-2 is built to classify the new multidimensional feature vector $\boldsymbol{N}_{Fi}$. And the classification of target region based on principal components of multiple features is realized. The neuron number of input layer decreases to $t$ from $p$ by dimensionality reduction and decoupling. The regions are classified into two categories: target and background, so the neuron number of output layer is 2. The neuron number of hidden layer influences classification precision. The overfitting and underfitting may happen, if there are an inapposite neuron number of hidden layer. According to experiments, the neuron number of hidden layer is set to 10. According to (26)-(27) [36], the activation functions of hidden layer and output layer are set based on hyperbolic tangent function and softmax activation function respectively. The $x_i$ is the $i^{\text{th}}$ value in output feature vector of output layer. The $j$ is the number of value in output feature vector used for summation. The output values are mapped to (0,1) and the sum of output values is 1. The initial weights are set to random number of [-1,1].

$$f(x) = \tanh(x) = (e^x - e^{-x})/(e^x + e^{-x}) \tag{26}$$

$$f(x_i) = e^{x_i} / \sum_{j=1}^{n} e^{x_j} \tag{27}$$

## IV. GRASPING POINT DETECTION OF RANDOMLY PLACED FRUIT CLUSTER USING ADAPTIVE MORPHOLOGY SEGMENTATION AND PRINCIPAL COMPONENT CLASSIFICATION OF MULTIPLE FEATURES

For the problem that it is difficult to detect the grasping point of randomly placed fruit cluster rapidly and precisely, a grasping point detection method of randomly placed fruit cluster using adaptive morphology segmentation and principal component classification of multiple features is proposed in this paper as shown in Fig.5. The improved adaptive morphology image segmentation algorithm based on edge distances and improved neural network region classification algorithm based on principal components of multiple features are used

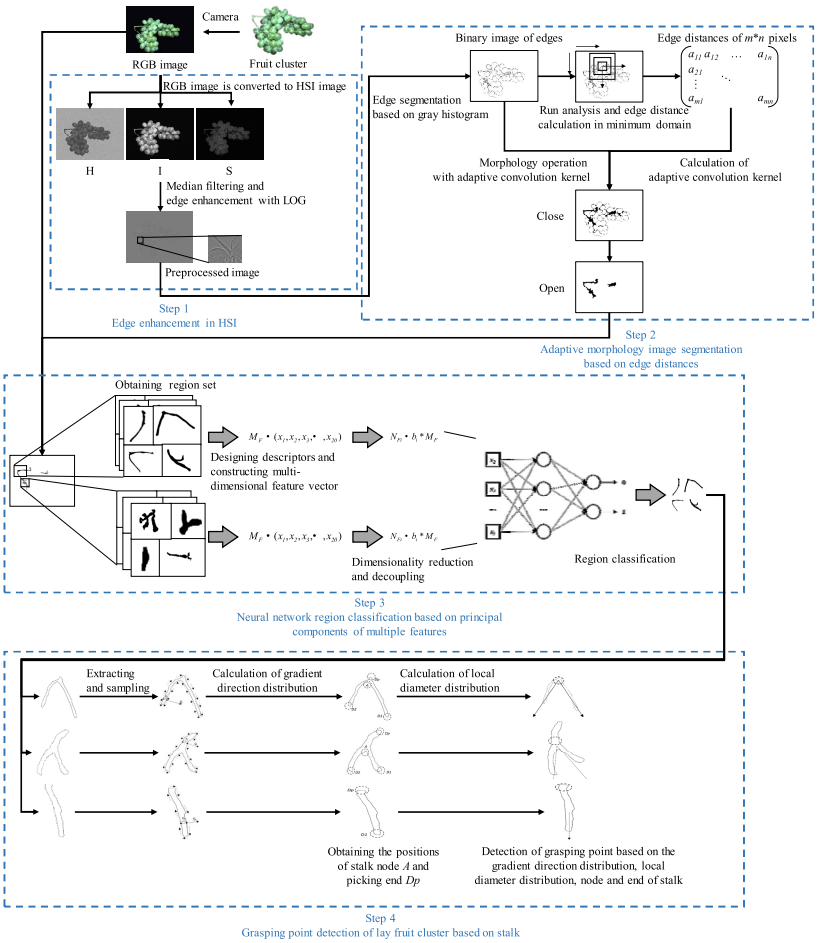



**FIGURE 5.** Grasping point detection method of randomly placed fruit cluster using adaptive morphology segmentation and principal component classification of multiple features.

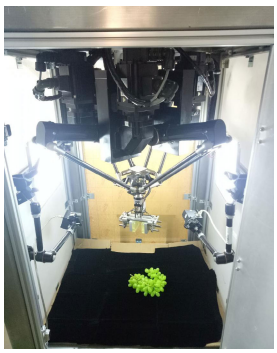


**FIGURE 6.** Visual system of grasping point detection for randomly placed fruit cluster.

for the research on grasping point detection of randomly placed fruit cluster to improve the precision and speed.

### A. IMAGE ACQUISITION

The visual system of grasping point detection for randomly placed fruit cluster is built based on parallel robot sorting system as shown in Fig.6. The research mainly focuses on the physical characteristics of randomly placed fruit cluster that lead to the problem that it is difficult to detect the grasping point based on machine vision. Therefore, the black background is adopted to simplify experimental model and reduce the interference of complex background.

According to the size of workspace for grasping on parallel robot sorting system, the field of view for grasping point detection is set as 400mm×300mm, the feature resolution is set as 0.2mm and the range of object distance is set as 800-900mm. According to (28)-(30), the visual hardware parameters can be calculated. Then the complementary metal oxide semiconductor (CMOS) camera with the pixel size of 2592×1944 and the CMOS size of 1/2.5'' is used. The lens with the focal length of 12mm is adopted to acquire randomly placed fruit cluster image. In addition, in order to reduce the influence of specular reflection on image quality, two flat light sources with color temperature of 6600-7000k are adopted to illuminate the fruit clusters indirectly. For achieving diffuse reflection of light in workspace, two light sources are set to illuminate each other directly and the white

wall of system is set to reflect received light.

$$FOV_i/P_i = \Delta w_i \tag{28}$$

$$1/WD + 1/MD = 1/f \tag{29}$$

$$m = CS_i/FOV_i = MD/WD \tag{30}$$

The $FOV_i$ is the length or width of field of view, $\Delta w_i$ is the feature resolution on the direction of length or width, $WD$ is the object distance, $MD$ is the image distance, $f$ is the focal length, $CS_i$ and $P_i$ is the length or width of CMOS in mm and pixel respectively. Taking the White Rosa grape clusters as an example, the acquired image is shown in Fig.7. The fruit cluster includes berry, stalk, stalk node, rachis, and pedicel.

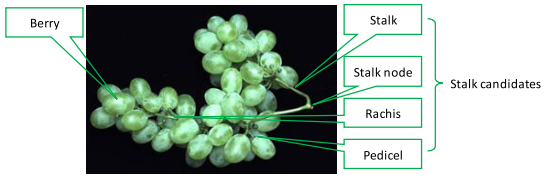



**FIGURE 7.** Fruit cluster.

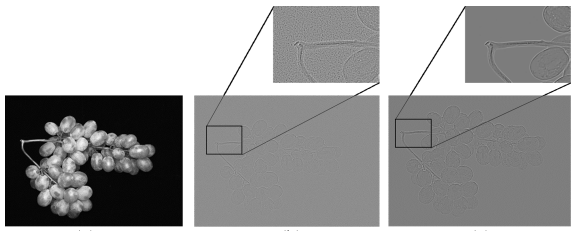



**FIGURE 8.** Preprocessing results. (a) The intensity image I. (b) The result of edge enhancement with conventional method. (c) The result of edge enhancement with improved method.

### B. EDGE ENHANCEMENT IN HSI

The existing image color models mainly include RGB (red green blue), HSV (hue saturation value), HSI (hue saturation intensity), HSL (hue saturation lightness), HSB (hue saturation brightness) and so on. Compared with other color models, the HSI color model can better facilitate the color processing and recognition and the workload for color analysis in HSI color model can reduced greatly. Therefore, considering the color differences among berry, stalk, rachis, and pedicel of fruit cluster, the HSI color model is adopted to increase the contrast between foreground and background in image. In addition, compared with the other components in HSI color model, the intensity image $I$ calculated by averaging RGB components can reduce the influence of uneven illumination and noise. Taking the White Rosa grape clusters as an example, as shown in Fig.8(a), the difference between foreground of grape cluster and background is obvious in $I$ because the calculation of $I$ is similar to mean filtering and can reduce noise in RGB from image acquisition. As shown in Fig.8(b), the noise may be enhanced in the process of edge enhancement. Therefore, before the edge enhancement, the circular convolution kernel with radius of 3 pixels is designed and the median filtering is performed on image with the way of pixel-by-pixel to reduce the salt and pepper noise. Then the Gaussian filter and Laplace operator are combined by (31)-(32) and the Laplacian of Gaussian(LOG) function is adopted to process image. It can reduce the Gaussian noise on the image and enhance edge as shown in Fig.8(c). Compared with Fig.8(b), the noise in Fig.8(c) is reduced obviously. Compared with Fig.8(a), the difference between foreground and background is increased obviously.

$$G_\sigma(x, y) = (1/\sqrt{2\pi\sigma^2}) \exp(-((x^2 + y^2)/2\sigma^2)) \tag{31}$$

$$\begin{aligned} LoG\Delta G_\sigma(x, y) &= \frac{\partial^2}{\partial x^2} G_\sigma(x, y) + \frac{\partial^2}{\partial y^2} Gn_\sigma(x, y) \\ &= ((x^2 + y^2 - 2\sigma^2)/\sigma^4) e^{-(x^2+y^2)/2\sigma^2} \end{aligned} \tag{32}$$

### C. SEGMENTATION OF STALK CANDIDATES AND STALK EXTRACTION FOR RANDOMLY PLACED FRUIT CLUSTER

The threshold segmentation is performed on enhanced image of randomly placed fruit cluster. The gray histogram of image is calculated. The gray value locating on the right of wave peak and the 70% of peak on the gray histogram is used as threshold of global segmentation. After threshold segmentation, as shown in Fig.9, the binary image including the edges of berry, stalk, stalk node, rachis, pedicel and background is obtained. Then the improved adaptive morphology image segmentation algorithm based on edge distances is used for segmentation of stalk candidates.

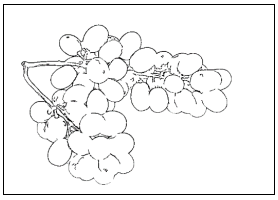


**FIGURE 9.** The binary image of edges.

In order to extract stalk by improved neural network region classification algorithm, it is necessary to label the stalk regions and non-stalk regions in the stalk candidates for training and testing the neural network. The obtained binary image including stalk candidates is used as mask image. Logic AND is operated by obtained mask image and original image. The result is shown in Fig.10(a). Then the threshold processing is performed in the resulting image of Logic AND. The region set of stalk candidates including stalk, rachis, pedicel, and stalk node is obtained as shown in Fig.10(b). The region set of stalk candidates for training and testing neural network is constructed and labeled manually, as shown in Fig.11.

The steepest descent method is adopted to train neural network with the region set of stalk candidates. The parameters of weights and biases are adjusted continually based on back propagation to make error reach minimum. Extracting stalk

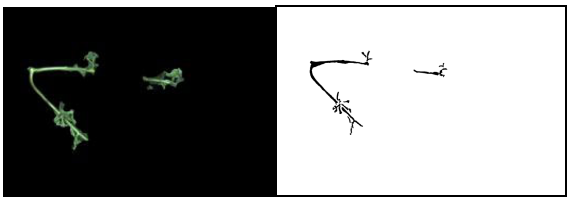



**FIGURE 10.** Stalk candidates. (a) Image of stalk candidates. (b) Regions of stalk candidates.

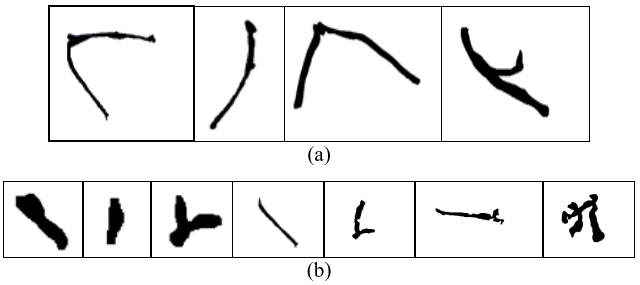



**FIGURE 11.** Region set of stalk candidates. (a) Stalk regions. (b) Non-stalk regions.

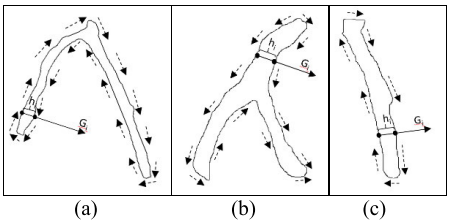



**FIGURE 12.** Gradient direction calculation. (a) Gradient direction calculation for first kind of stalk contour. (b) Gradient direction calculation for second kind of stalk contour. (c) Gradient direction calculation for third kind of stalk contour.

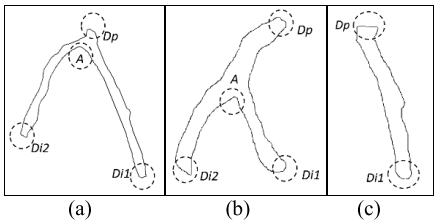



**FIGURE 13.** Stalk contours. (a) The first kind of stalk contour. (b) The second kind of stalk contour. (c) The third kind of stalk contour.

from region set of stalk candidates can be performed based on trained BP neural network.

### *D. GRASPING POINT DETECTION OF RANDOMLY PLACED FRUIT CLUSTER BASED ON STALK CONTOUR*

#### 1) GRADIENT DIRECTION DISTRIBUTION OF STALK CONTOUR

According to extracted stalk region of randomly placed fruit cluster by image segmentation and region classification, the contour of stalk is extracted and the points in the contour are marked orderly by $L_i, i = 0, 1, 2 \ldots, g$. In order to improve the calculation speed, the points $L_i$ are sampled with step length of $\Delta Pi$ based on the length $l$ of finger of robot clamping mechanism according to (33). Then the sampled point set of stalk contour is obtained $L'_i, i = 0, 1, 2 \ldots, g'$. The $g'$ is the quotient of $g/10$.

$$1/\Delta w = \Delta Pi/l \tag{33}$$

The units of $\Delta w$ and $l$ are mm, the unit of $\Delta Pi$ is pixel. The points in the sampled point set are calculated orderly as shown in Fig.12. The gradient vector $\boldsymbol{G}_i$ and gradient direction $a(x, y)$ of contour in the every point $L'_i$ are calculated according to (34)-(35). The gradient direction is expressed based on $x$-axis.

$$\boldsymbol{G}_i = [g_x, g_y] = [\partial f(x, y)/x, \partial f(x, y)/y] \tag{34}$$

$$a(x, y) = \arctan(g_y/g_x) \tag{35}$$

As shown in Fig.12 and Fig.13, the stalk contour can be divided into 3 basic categories. (a) The first kind: picking end $Dp$ is near stalk node $A$. (b) The second kind: picking end $Dp$ is far from stalk node $A$. (c) The third kind: there is no stalk node $A$. The $Di1$ and $Di2$ are the ends in the stalk touching the cluster. The distribution diagram of gradient directions is built considering the point $L'_i$ as $x$-coordinate and the gradient direction $a(x, y)$ as $y$-coordinate. As shown in Fig.14, the points with big change of gradient direction in distribution diagram correspond to the stalk node and ends in stalk. Therefore, according to the rising edges or falling edges in distribution diagram, the positions of stalk node $A$ and picking end $Dp$ can be obtained.

#### 2) LOCAL DIAMETER DISTRIBUTION OF STALK IN GRADIENT DIRECTION

According to the sampled point set $L'_i$ and gradient direction of stalk contour in the every point $L'_i$, the local diameter of stalk in the every point is calculated. It is difficult to fit the various stalk contours by geometry such as rectangle, ellipse, and straight line. In addition, the local diameter of stalk in every contour point is different. It is difficult to accurately calculate the diameter of stalk by using the mean distance between straight lines, mean length or width of rectangles or mean radius of ellipse. Therefore, the distance between pixel peaks in gradient direction of stalk contour is calculated as local diameter of stalk in this paper. Starting from $L'_i$, the pixel change of contour is calculated along the gradient direction $a(x, y)$. The point $K'_i$ that is the first rising edge of pixel change along the gradient direction $a(x, y)$ of detecting point $L'_i$ is regarded as the corresponding point of detecting point $L'_i$. The all corresponding points $K'_i$ related to $L'_i$ are calculated. Then the point pair set $J$ is composed of $(L'_i, K'_i)$. The Euclid distance of every point pair $(L'_i, K'_i)$ is calculated as the local diameter $h_i$ of stalk in $L'_i$. The distribution diagram of local diameters is built considering the point $L'_i$ as $x$-coordinate

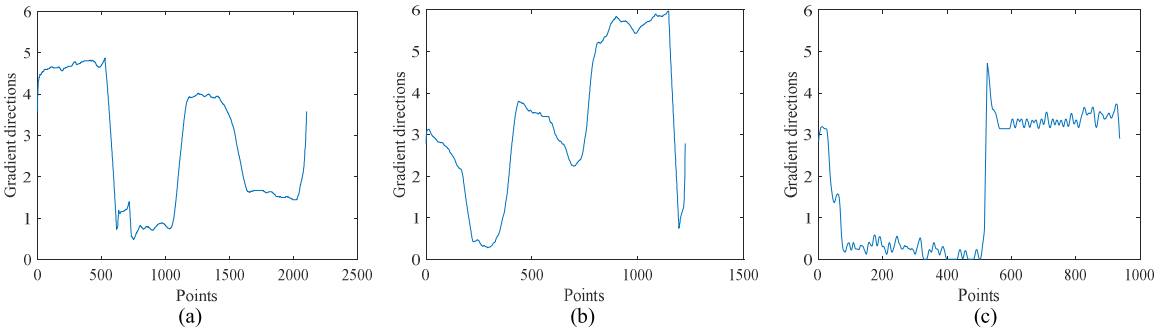



**FIGURE 14.** Distribution diagram of gradient directions. (a) Distribution diagram of gradient directions for first kind of stalk contour. (b) Distribution diagram of gradient directions for second kind of stalk contour. (c) Distribution diagram of gradient directions for third kind of stalk contour.

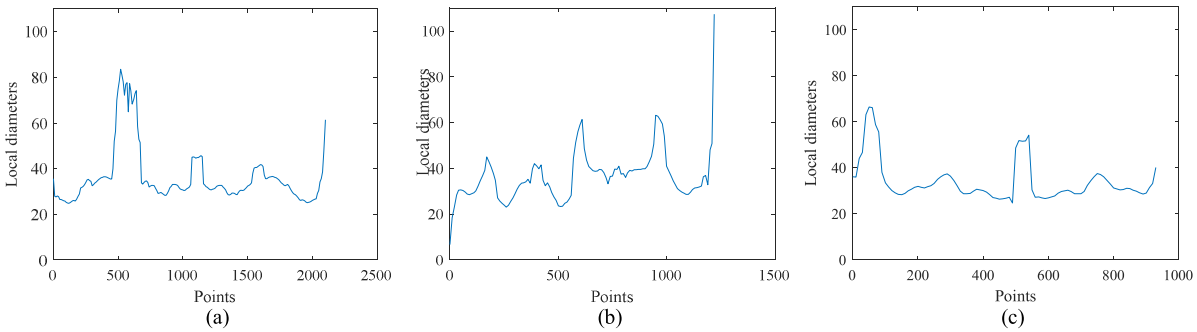



**FIGURE 15.** Distribution diagram of local diameters. (a) Distribution diagram of local diameters for first kind of stalk contour. (b) Distribution diagram of local diameters for second kind of stalk contour. (c) Distribution diagram of local diameters for third kind of stalk contour.

and the local diameter of stalk in gradient direction as $y$-coordinate, as shown in Fig.15.

### 3) GRASPING POINT DETECTION OF RANDOMLY PLACED FRUIT CLUSTER

According to the gradient direction distribution, local diameter distribution, node and end of stalk, the most robust part in stalk is calculated as grasping point. According to the distribution diagram of gradient directions, the stalk parameters including the category of stalk, the positions of stalk node and picking end are analyzed and determined firstly. If there are less than 3 rising edges or falling edges, there is no stalk node. Otherwise, there is stalk node and the positions of stalk node $A$ and picking end $Dp$ need be calculated. Because the direction change of end is obviously greater than that of stalk node, the rising edges or falling edges can be chosen by threshold. The smallest direction change of rising edge or falling edge should correspond to stalk node $A$. The before and after rising edges or falling edges of the stalk node $A$ should correspond to the ends $Di1$ and $Di2$. The rest of the rising edge or falling edge should correspond to $Dp$.

Then the stalk parameters and the local diameter distribution are adopted to detect grasping point. If there is no stalk node, the point with maximum diameter in stalk based on the local diameter distribution is regarded as grasping point. If there is stalk node, the point with maximum diameter in the stalk part between stalk node $A$ and the picking end $Dp$ based on the local diameter distribution is regarded as grasping point.

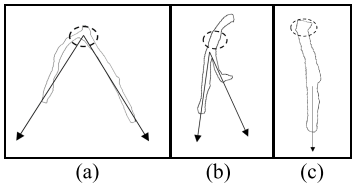



**FIGURE 16.** Force during grasping. (a) Force during grasping for first kind of stalk contour. (b) Force during grasping for second kind of stalk contour. (c) Force during grasping for third kind of stalk contour.

The proposed method considers the gravity and distribution of berries in fruit cluster during grasping. As shown in Fig.16, it makes the force-bearing point of every rachis of fruit cluster mainly focus on the stalk node to avoid the detachment of rachis from stalk caused by the different force in every rachis during grasping. In addition, the robust and burly part of stalk is chosen as grasping point. The stalk node and robust part are combined to realize precise grasping point detection.

The grasping points $\boldsymbol{p}_l$ and $\boldsymbol{p}_r$ in images coming from left and right cameras respectively can be acquired by the presented method. Then the actual grasping point $\boldsymbol{P}_W$ in world

coordinate system can be calculated by (36)-(38) according to $\boldsymbol{p}_l$ and $\boldsymbol{p}_r$.

$$\boldsymbol{P}_W = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{B} \tag{36}$$

$$\boldsymbol{A} = \begin{bmatrix} (u_l m^l_{31} - m^l_{11}) & (u_l m^l_{32} - m^l_{12}) & (u_l m^l_{33} - m^l_{13}) \\ (v_l m^l_{31} - m^l_{21}) & (v_l m^l_{31} - m^l_{22}) & (v_l m^l_{33} - m^l_{23}) \\ (u_r m^r_{31} - m^r_{11}) & (u_r m^r_{32} - m^r_{12}) & (u_r m^r_{33} - m^r_{13}) \\ (v_r m^r_{31} - m^r_{21}) & (v_r m^r_{32} - m^r_{22}) & (v_r m^r_{33} - m^r_{23}) \end{bmatrix} \tag{37}$$

$$\boldsymbol{P}_W = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} m^l_{14} - u_l m^l_{34} \\ m^l_{24} - v_l m^l_{34} \\ m^r_{14} - u_r m^r_{34} \\ m^r_{24} - v_r m^r_{34} \end{bmatrix} \tag{38}$$

The $(u_l, v_l, 1)$ and $(u_r, v_r, 1)$ represent the homogeneous coordinates of $\boldsymbol{p}_l$ and $\boldsymbol{p}_r$ in pixel coordinate system respectively, $(X, Y, Z, 1)$ is the homogeneous coordinate of $\boldsymbol{P}_W$ in world coordinate system, $m^k_{ij}(k = l, r; i = 1, 2, 3; j = 1, 2, 3, 4)$ represents the value in row $i$ and column $j$ of projection matrix, $k$ represents the camera. The above parameters can be required by camera calibration.

## V. RESULTS AND DISCUSSION

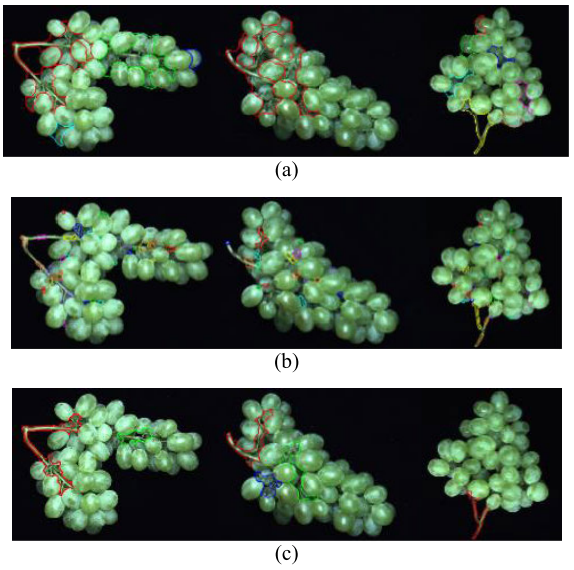
Taking the White Rosa grape clusters as an example, the 200 color images of grape clusters in different morphology are acquired by the visual system of grasping point detection for randomly placed fruit cluster based on parallel robot sorting system. According to distribution of edge distances, the 80 images with bigger interspace between berries are classified into $S$. The 120 images with smaller interspace between berries are classified into $C$ as shown in Fig.17. The hardware of experiment is Intel(R)Core(TM)i7-3770 CPU with 8GB memory and x64 CPU.

(a) (b)

**FIGURE 17.** Representative of fruit cluster categories. (a) $S$. (b) $C$.

### A. SEGMENTATION OF STALK CANDIDATES FOR RANDOMLY PLACED FRUIT CLUSTER

The segmentation experiments of stalk candidates are performed for 200 different images of randomly placed grape clusters based on the existing morphology algorithm with fixed convolution kernel and improved adaptive morphology image segmentation algorithm respectively. The fixed convolution kernel includes small fixed convolution kernel and large fixed convolution kernel. The sizes of small fixed convolution kernel and large fixed convolution kernel are obtained based on the edge distance means of 120 $C$ images and 80 $S$ images respectively.



(a) (b) (c)

**FIGURE 18.** The segmentation results of stalk candidates. (a) The segmentation results of stalk candidates with too large convolution kernel. (b) The segmentation results of stalk candidates with too small convolution kernel. (c) The segmentation results of stalk candidates with proper convolution kernel.

The experimental results are shown in Fig.18, the convolution kernel with improper size affects the segmentation precision of stalk candidates. When the convolution kernel is too large, the background regions are easily segmented as stalk candidates. When the convolution kernel is too small, the obtained regions of stalk candidates are incomplete. Segmenting stalk candidates precisely is difficult in the two situations.

The judgement if the segmentation is correct is performed according to (39). If $ac \geq 70\%$, the segmentation is correct. Otherwise, it is incorrect.

$$ac = (A1 \cap A2)/A2 \tag{39}$$

The $A1$ is the region of stalk candidate obtained by segmentation, $A2$ is the actual region of stalk candidate. The judgement of correctness is performed in every segmentation experiment for every image. From the statistics of results in Table 1, we can see that the average segmentation precision based on improved adaptive morphology image segmentation algorithm reaches 98.54%. Comparing with the existing morphology algorithms with small fixed convolution kernel and large fixed convolution kernel, the average segmentation precision increases by 9.58% and 10.2% respectively. The distributions of edge distances of randomly placed fruit clusters with different interspace between berries are different. The segmentation precisions for randomly placed fruit clusters $S$ and $C$ by using the existing morphology algorithms with fixed convolution kernel are difference. The improved adaptive morphology image segmentation algorithm has a high precision for the images of randomly placed fruit clusters $S$ and $C$. It can improve the precision and applicability of

**TABLE 1.** The experimental results of segmentation of stalk candidates.

| Category | Convolution kernel | Number of images | Number of images with correct segmentation | Average precision | Mean average precision |
|---|---|---|---|---|---|
| *C* | Small fixed | 120 | 113 | 94.17% | 88.96% |
| *S* | convolution kernel | 80 | 67 | 83.75% | |
| *C* | Large fixed | 120 | 98 | 81.67% | 88.34% |
| *S* | convolution kernel | 80 | 76 | 95% | |
| *C* | Adaptive | 120 | 118 | 98.33% | 98.54% |
| *S* | convolution kernel | 80 | 79 | 98.75% | |

**TABLE 2.** The feature vector values of stalk regions in partial samples.

| Samples | *a* | *L* | *W* | *H* | $N_{pq}$ | $U_{pq}$ | *Ani* |
|---|---|---|---|---|---|---|---|
| 1 | 19374 | 2144.61 | 539 | 388 | -1.9E+10 | 1.4E+42 | 1.43 |
| 2 | 4428 | 641.51 | 276 | 57 | -4.1E+07 | -6.2E+28 | 12.95 |
| 3 | 4146 | 600.85 | 188 | 174 | -1E+07 | -4.2E+30 | 8.19 |
| 4 | 1683 | 263.20 | 106 | 25 | -1929540 | -2.1E+24 | 5.39 |
| 5 | 7117 | 910.29 | 289 | 241 | 3.2E+08 | 4.1E+34 | 6.36 |
| Samples | *Bul* | *SF* | *Con* | *Rec* | $r_o$ | $r_i$ | $w_i$ |
| 1 | 10.65 | 14.27 | 0.153 | 0.0622 | 279 | 16 | 240 |
| 2 | 1.31 | 15.92 | 0.594 | 0.6209 | 140.44 | 13.5 | 99 |
| 3 | 1.82 | 13.93 | 0.495 | 0.4883 | 127.14 | 10.5 | 25 |
| 4 | 1.28 | 5.88 | 0.754 | 0.6746 | 53.15 | 9 | 93 |
| 5 | 3.06 | 18.48 | 0.391 | 0.3016 | 182.07 | 12.5 | 13 |
| Samples | $h_i$ | *Cir* | *Mdis* | *Dd* | *Rou* | *Com* | |
| 1 | 13 | 0.0504 | 185.33 | 68.62 | 0.6297 | 18.8916 | |
| 2 | 11 | 0.0658 | 72.68 | 40.88 | 0.4375 | 7.3958 | |
| 3 | 14 | 0.077 | 65.45 | 35.35 | 0.4598 | 6.9294 | |
| 4 | 7 | 0.1871 | 28.66 | 15.31 | 0.4658 | 3.2754 | |
| 5 | 41 | 0.0595 | 93.69 | 53.52 | 0.4288 | 9.2651 | |

segmentation of stalk candidates for randomly placed fruit cluster.

### B. STALK EXTRACTION FOR RANDOMLY PLACED FRUIT CLUSTER

Experimental methods: The 150 images are chosen randomly as sample images. According to the stalk candidates obtained by segmentation from 200 different images, the stalk regions and non-stalk regions in stalk candidates from sample images are used as sample set. There are 150 stalk regions and 870 non-stalk regions in total. The stalk candidates from other 50 images are used as test set to verify precision of stalk extraction. In order to verify the influence of feature vector on the extraction precision, the low-dimensional feature vector, high-dimensional feature vector and 20-dimensional feature vector constructed in this paper are used as inputs of BP neural network. The descriptors of low-dimensional feature vector are chosen randomly from the 20 descriptors proposed in this paper. The descriptors of high-dimensional feature vector include other shape descriptors, such as angle, direction, and size, in addition to the 20 descriptors proposed in this paper.

Experimental results: the data of feature vectors in partial samples are shown in Tables 2 and 3. From the tables we can see that it is difficult to classify the feature vectors directly because of the complexity of data and correlation between features. So the principal components of multiple features are extracted based on variance contribution to reduce dimension and decouple. The BP neural network is built to classify the stalk regions and non-stalk regions.

The experimental results with different feature vectors are shown in Table 4. The trained BP neural network is tested by the feature vectors extracted from test set to calculate the precision of stalk extraction. From Table 4 we can see that the extraction precision based on the low-dimensional feature vector is low because representing region by the low-dimensional feature vector is not comprehensive and easily leads to underfitting in classification. Though representing

**TABLE 3.** The feature vector values of non-stalk regions in partial samples.

| Samples | $a$ | $L$ | $W$ | $H$ | $N_{pq}$ | $U_{pq}$ | $Ani$ |
|---|---|---|---|---|---|---|---|
| 1 | 2789 | 498.86 | 108 | 133 | -4E+07 | 1.4E+31 | 2.47 |
| 2 | 430 | 117.01 | 33 | 28 | 66178 | -3E+19 | 2.14 |
| 3 | 809 | 165.3 | 40 | 53 | 304561 | -8.2E+21 | 2.75 |
| 4 | 215 | 56.04 | 17 | 17 | -2586 | -9.3E+14 | 1.15 |
| 5 | 335 | 77.46 | 31 | 15 | -21400 | -8.9E+16 | 2.33 |
| Samples | $Bul$ | $SF$ | $Con$ | $Rec$ | $r_o$ | $r_i$ | $w_i$ |
| 1 | 4.3 | 9.63 | 0.341 | 0.2202 | 76.78 | 9.5 | 8 |
| 2 | 1.32 | 1.82 | 0.750 | 0.5944 | 18.86 | 6 | 14 |
| 3 | 1.54 | 3.23 | 0.659 | 0.5527 | 30.03 | 8 | 18 |
| 4 | 1.05 | 0.21 | 0.939 | 0.7917 | 9.67 | 6.5 | 10 |
| 5 | 1.05 | 1.44 | 0.946 | 0.7343 | 16.48 | 6.5 | 18 |
| Samples | $h_i$ | $Cir$ | $Mdis$ | $Dd$ | $Rou$ | $Com$ | |
| 1 | 56 | 0.1257 | 50.4 | 18.12 | 0.6405 | 7.1007 | |
| 2 | 9 | 0.3304 | 10.62 | 4.83 | 0.5452 | 2.5339 | |
| 3 | 13 | 0.2442 | 17.35 | 7.84 | 0.5483 | 2.6876 | |
| 4 | 12 | 0.7193 | 7.18 | 1.18 | 0.8357 | 1.1625 | |
| 5 | 11 | 0.3573 | 9.92 | 3.41 | 0.6564 | 1.4251 | |

**TABLE 4.** The experimental results of stalk extraction.

| Feature vectors | Vector processing | Classifier | Precision | Average training time(s) |
|---|---|---|---|---|
| Low-dimensional feature vector | -- | BP neural network | 82.61% | 35.1 |
| | Dimensionality reduction and decoupling | | 86.44% | 18.2 |
| High-dimensional feature vector | -- | | 90.52% | 93.5 |
| | Dimensionality reduction and decoupling | | 93.16% | 46.7 |
| 20-dimensional feature vector | -- | | 96.46% | 63.1 |
| | Dimensionality reduction and decoupling | | 98.63% | 37.2 |

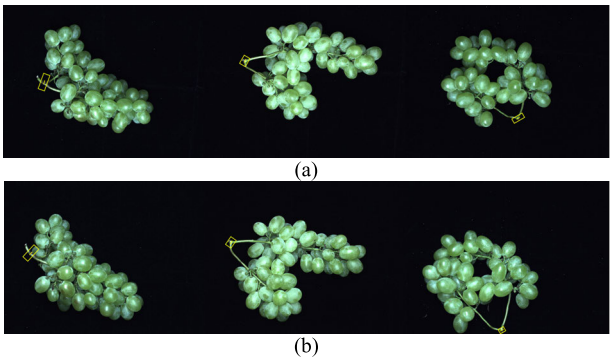


**FIGURE 19.** The results of grasping point detection for randomly placed grape clusters. (a) Images coming from left camera. (b) Images coming from right camera.

region by the high-dimensional feature vector is comprehensive, the overfitting may happen to affect the precision and speed of extraction. Compared with low-dimensional and high-dimensional feature vectors, the 20-dimensional feature vector proposed in this paper can reduce the effect of overfitting and underfitting. The extraction precision increases by at least 5.47%. Compared with the method without dimensionality reduction and decoupling, the extraction precision increases by at least 2.17% and the average training time reduces by at least 25.9s. The improved adaptive morphology image segmentation algorithm based on edge distances can improve the precision and speed of stalk extraction for randomly placed fruit cluster.

### C. GRASPING POINT DETECTION OF RANDOMLY PLACED FRUIT CLUSTER

The experiments of grasping point detection based proposed method are performed for 200 different images of White Rosa grape clusters. As shown in Fig.19, the obtained grasping points of randomly placed grape clusters all are the robust parts with maximal diameter in stalk.

The grasping points from oft-repeated human detection are used as actual grasping points. The accuracy of grasping point detection is calculated by comparing the result of grasping point detection with the actual grasping point. The method of human detection is as follows: (1) Judging if there is stalk node. (2) Detecting the point with maximum diameter in the

**TABLE 5.** The experiment data of grasping point detection.

| Image | $(x_a, y_a, z_a)$ (mm) | $(x_c, y_c, z_c)$ (mm) | $H_{\max}$ (pixel) | Stalk node | $D_e$ (mm) | Time(s) |
|---|---|---|---|---|---|---|
| 1 | (102.888,-64.035,-993.059) | (104.208,-65.182,-990.751) | 79 | √ | 2.896 | 1.95 |
| 2 | (-141.310,-7.243, -1079.297) | (-143.746,-8.884,-1078.528) | 68 | √ | 3.036 | 1.89 |
| 3 | (-137.016,-189.043,-1017.038) | (-137.987,-190.632,-1019.529) | 61 | × | 3.111 | 2.01 |
| 4 | (4.571,-59.475,-1082.091) | (5.390,-58.233,-1083.197) | 75 | √ | 1.853 | 1.99 |
| 5 | (-89.896,-120.026,-937.611) | (-91.174,-120.922,-938.640) | 59 | × | 1.870 | 2.03 |
| 6 | (107.284,18.294,-1024.643) | (108.420,19.067,-1023.706) | 74 | √ | 1.663 | 2.11 |
| 7 | (23.368,26.900,-948.398) | (24.530,25.385,-950.746) | 76 | √ | 3.026 | 1.98 |
| 8 | (-124.555,-130.279,-1024.888) | (-126.349,-128.028,-1026.668) | 84 | √ | 3.384 | 2.05 |
| 9 | (-87.521,166.811,-1028.391) | (-86.201,165.980,-1030.423) | 81 | √ | 2.562 | 2.10 |
| 10 | (-138.025,146.898,-902.969) | (-140.217,145.861,-901.909) | 64 | × | 2.646 | 2.09 |

**TABLE 6.** The experimental results of grasping point detection.

| Category | Number of images | Number of images with correct detection | Average accuracy | Average time(s) |
|---|---|---|---|---|
| $C$ | 120 | 113 | 94.17% | 1.99 |
| $S$ | 80 | 76 | 95.00% | 2.03 |
| Sum | 200 | 189 | 94.50% | 2.01 |

corresponding part by a vernier caliper and regarding it as grasping point. The partial experiment data are shown as Table 5. The $(x_a, y_a, z_a)$ is the coordinate of actual grasping point in world coordinate system. The $(x_c, y_c, z_c)$ is the coordinate of detecting grasping point in world coordinate system. The $H_{\max}$ is the maximum diameter of stalk. The $D_e$ is the detecting error, which can be calculated based on Euclid distance between $(x_a, y_a, z_a)$ and $(x_c, y_c, z_c)$ according to (40).

$$D_e = \sqrt{(x_a - x_c)^2 + (y_a - y_c)^2 + (z_a - z_c)^2} \tag{40}$$

If the $D_e$ less than 5mm, the detection is correct. Otherwise, it is incorrect. As shown in Table 6, the average precision of grasping point detection for 200 different images by proposed method reaches 94.50% and the average time of detection reaches 2.01s.


**FIGURE 20.** Path planning.

The sorting robot of fruit clusters is responsible for implementing the fast grasp-and-place operation, which is arranged in the way as shown in Fig.20. To reduce the vibration and berry dropping that may occur in high-speed robotic sorting [37], the peak acceleration of the clamping mechanism is set as 6 m/s$n^2$. According to the length of fruit cluster and the size of workspace, the distances $S_{AB}$,$S_{BC}$, and $S_{CD}$ are set as 0.4m, 1m, and 0.4m respectively. According to (41), based on the 3-4-5 polynomial, the total time for the clamping mechanism to bring a fruit cluster from $A$ to $D$ is $T_{total} = 2.22$s.

$$T = \sqrt{5.7735S/a_{max}} \tag{41}$$

The $a_{max}$ is the peak acceleration of clamping mechanism, $S$ is the distance.

The system can start grasping point detection when the clamping mechanism grasps the last fruit cluster and arrives at $B$. The grasping point detection should end when the clamping mechanism places the last fruit cluster in $D$ and goes back to $C$. The time of grasping point detection should meet the condition of (42).

$$T_{gd} \le T_{BC} + T_{CD} + T_d + T_{DC} = T_{total} + T_d \tag{42}$$

The $T_{gd}$ is the time of grasping point detection, the $T_{BC}$, $T_{CD}$ and $T_{DC}$ are the time for the clamping mechanism to pass $BC$, $CD$, and $DC$, $T_d$ is the time for the clamping mechanism to place the fruit cluster. According to (41)-(42), $T_{total} + T_d > 2.22$s.

The experimental results show that the proposed method can solve the problem that it is difficult to detect grasping point of randomly placed fruit cluster precisely and rapidly. It not only improves the precision, but also satisfies the time requirement of parallel robot to detect the grasping points of fruit cluster. It can lay the foundation for accurate, fast, and nondestructive automatic sorting of randomly placed fruit clusters based on parallel robot and machine vision.

## VI. CONCLUSION

For the problem that it is difficult to precisely and rapidly extract stalk and detect grasping point of randomly placed fruit cluster based on machine vision, this paper proposes a grasping point detection method of randomly placed fruit cluster using adaptive morphology segmentation and principal component classification of multiple features. The following conclusions can be drawn:

(1) An improved adaptive morphology image segmentation algorithm based on edge distances is proposed. It solves the problem that the segmentation of stalk candidates for randomly placed fruit cluster based on existing morphology

algorithm with fixed convolution kernel tends to be low precision. In addition, the existing run analysis method based on equivalent sequence is improved. It reduces the time detecting and marking unconnected components in the edge distance calculation and improves the efficiency of image segmentation.

(2) An improved neural network region classification algorithm based on principal components of multiple features is proposed. It solves the problem that it is difficult to describe and classify unconstraint stalk by existing features. The principal components of multiple features are extracted based on variance contribution to improve the precision and efficiency of region classification.

(3) The improved morphology image segmentation and region classification algorithms are successfully applied to detect grasping point of randomly placed fruit cluster based on stalk. The proposed method is verified by experiments with 200 images of White Rosa grape clusters based on parallel robot sorting system. Compared with the existing morphology algorithm with fixed convolution kernel, the average segmentation precision of stalk candidates increases by 9.89%. Compared with the existing neural network region classification algorithm, the average precision of stalk extraction increases by 2.17%. Based on proposed method, the average precision of grasping point detection reaches 94.50% and the average time is 2.01s. The experimental results show that the proposed method solves the problem that it is difficult to detect grasping point of randomly placed fruit cluster precisely and rapidly. It can lay the foundation for accurate, fast, and nondestructive automatic sorting of randomly placed fruit clusters based on parallel robot and machine vision.

## REFERENCES


[1] J. Blasco, N. Aleixos, and E. Moltó, "Machine vision system for automatic quality grading of fruit," *Biosyst. Eng.*, vol. 85, pp. 415–423, Aug. 2003.

[2] N. I. Glossas and N. A. Aspragathos, "Fuzzy logic grasp control using tactile sensors," *Mechatronics*, vol. 11, no. 7, pp. 899–920, 2001.

[3] Q. H. Wang, Y. H. Tang, and Z. Xiao, "Grape size detection and online gradation based on machine vision," *Int. J. Agricult. Biol. Eng.*, vol. 10, no. 1, pp. 226–233, 2017.

[4] J. Mo, Z.-F. Shao, L. Guan, F. Xie, and X. Tang, "Dynamic performance analysis of the X4 high-speed pick-and-place parallel robot," *Robot. Comput.-Integr. Manuf.*, vol. 46, pp. 48–57, Aug. 2017.

[5] F. C. Sun, C. F. Liu, W. B. Huang, and J. W. Zhang, "Object classification and grasp planning using visual and tactile sensing," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 46, no. 7, pp. 969–979, Apr. 2016.

[6] C. Blanes, M. Mellado, C. Ortiz, and A. Valera, "Technologies for robot grippers in pick and place operations for fresh fruits and vegetables," *Spanish J. Agricult. Res.*, vol. 9, no. 4, p. 1130, 2011.

[7] K. Przerwa, W. Kasprzak, and M. Stefańczyk, "Infrared image-based 3D surface reconstruction of free-form texture-less objects," in *Proc. CORES*, Wroclaw, Poland, May 2015, pp. 809–820.

[8] T. Yuan, C.-G. Xu, Y.-X. Ren, Q.-C. Feng, Y.-Z. Tan, and W. Li, "Detecting the information of cucumber in greenhouse for picking based on NIR image," *Spectrosc. Spectral Anal.*, vol. 29, pp. 2054–2058, Aug. 2009.

[9] J. Xing, P. Jancsók, and J. De Baerdemaeker, "Stem-end/calyx identification on apples using contour analysis in multispectral images," *Biosyst. Eng.*, vol. 96, pp. 231–237, Feb. 2007.

[10] P. Eizentals and K. Oka, "3D pose estimation of green pepper fruit for automated harvesting," *Comput. Electron. Agricult.*, vol. 128, pp. 127–140, Oct. 2016.

[11] J. Qiao, A. Sasao, S. Shibusawa, and N. Kondo, "Extracting external features of sweet peppers using machine vision system on mobile fruits grading robot," *Int. J. Food Eng.*, vol. 8, no. 3, pp. 1039–1058, 2012.

[12] Seema, A. Kumar, and G. S. Gill, "Automatic fruit grading and classification system using computer vision: A review," in *Proc. IEEE ICACCE*, Dehradun, India, May 2015, pp. 598–603.

[13] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *Int. J. Robot. Res.*, vol. 34, nos. 4–5, pp. 705–724, 2015.

[14] C. Wang, T. Luo, L. Zhao, Y. Tang, and X. Zou, "Window zooming–based localization algorithm of fruit and vegetable for harvesting robot," *IEEE Access*, vol. 7, pp. 103639–103649, 2019.

[15] M. M. Sofu, O. Er, M. C. Kayacan, and B. Cetisli, "Design of an automatic apple sorting system using machine vision," *Comput. Electron. Agricult.*, vol. 127, pp. 395–405, Sep. 2016.

[16] J. Tello, S. Cubero, J. Blasco, J. Tardaguila, N. Aleixos, and J. Ibáñez, "Application of 2D and 3D image technologies to characterise morphological attributes of grapevine clusters," *J. Sci. Food Agricult.*, vol. 96, no. 13, pp. 4575–4583, 2016.

[17] C. Wang, X. Zou, Y. Tang, L. Luo, and W. Feng, "Localisation of litchi in an unstructured environment using binocular stereo vision," *Biosyst. Eng.*, vol. 145, pp. 39–51, May 2016.

[18] S. Cubero, M. P. Diago, J. Blasco, J. Tardaguila, B. Millán, and N. Aleixos, "A new method for pedicel/peduncle detection and size assessment of grapevine berries and other fruits by image analysis," *Biosyst. Eng.*, vol. 117, pp. 62–72, Jan. 2014.

[19] G. Feng, C. Qixin, and N. Masateru, "Fruit detachment and classification method for strawberry harvesting robot," *Int. J. Adv. Robot. Syst.*, vol. 5, no. 1, p. 4, 2008.

[20] L. Luo, Y. Tang, Q. Lu, X. Chen, P. Zhang, and X. Zou, "A vision methodology for harvesting robot to detect cutting points on peduncles of double overlapping grape clusters in a vineyard," *Comput. Ind.*, vol. 99, pp. 130–139, Aug. 2018.

[21] L. Luo, Y. Tang, X. Zou, M. Ye, W. Feng, and G. Li, "Vision-based extraction of spatial information in grape clusters for harvesting robots," *Biosyst. Eng.*, vol. 151, pp. 90–104, Nov. 2016.

[22] S. Cubero, M. P. Diago, J. Blasco, J. Tardáguila, J. M. Prats-Montalbán, J. Ibáñez, J. Tello, and N. Aleixos, "A new method for assessment of bunch compactness using automated image analysis," *Austral. J. Grape Wine Res.*, vol. 21, no. 1, pp. 101–109, 2015.

[23] L. M. Yuan, J. R. Cai, L. Sun, and C. Ye, "A preliminary discrimination of cluster disqualified shape for table grape by mono-camera multi-perspective simultaneously imaging approach," *Food Anal. Methods*, vol. 9, no. 3, pp. 758–767, 2016.

[24] J. Hirvonen and P. Kallio, "Automatic image-based detection and inspection of paper fibres for grasping," *IET Comput. Vis.*, vol. 9, no. 4, pp. 588–594, 2015.

[25] A. Aquino, M. P. Diago, B. Millán, and J. Tardáguila, "A new methodology for estimating the grapevine-berry number per cluster using image analysis," *Biosyst. Eng.*, vol. 156, pp. 80–95, Apr. 2017.

[26] M. J. C. S. Reis, R. Morais, C. Pereira, O. Contente, M. Bacelar, S. Soares, A. Valente, J. Baptista, P. J. S. G. Ferreira, and J. Bulas-Cruz, "A low-cost system to detect bunches of grapes in natural environment from color images," in *Proc. ACIVS*, Ghent, Belgium, Aug. 2011, pp. 92–102.

[27] M. J. C. S. Reis, R. Morais, E. Peres, C. Pereira, O. Contente, S. Soares, A. Valente, J. Baptista, P. J. S. G. Ferreira, and J. Bulas-Cruz, "Automatic detection of bunches of grapes in natural environment from color images," *J. Appl. Logic*, vol. 10, no. 4, pp. 285–290, 2012.

[28] X. Yu, K. Liu, D. Wu, and Y. He, "Raisin quality classification using least squares support vector machine (LSSVM) based on combined color and texture features," *Food Bioprocess Technol.*, vol. 5, no. 5, pp. 1552–1563, 2012.

[29] D. S. Pérez, F. Bromberg, and C. A. Diaz, "Image classification for detection of winter grapevine buds in natural conditions using scale-invariant features transform, bag of features and support vector machines," *Comput. Electron. Agricult.*, vol. 135, pp. 81–95, Apr. 2017.

[30] C. W. Bac, J. Hemming, and E. J. van Henten, "Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper," *Comput. Electron. Agricult.*, vol. 96, pp. 148–162, Aug. 2013.
[31] E. Ivorra, A. J. Sánchez, J. G. Camarasa, M. P. Diago, and J. Tardaguila, "Assessment of grape cluster yield components based on 3D descriptors using stereo vision," *Food Control*, vol. 50, pp. 273–282, Apr. 2015.
[32] S. P. Tang, D. Zhao, W. K. Jia, Y. Chen, W. Ji, and C. Z. Ruan, "Feature extraction and recognition based on machine vision application in lotus picking robot," in *Proc. CCTA*, Beijing, China, Sep. 2015, pp. 485–501.
[33] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations," *Comput. Vis. Image Understand.*, vol. 89, no. 1, pp. 1–23, Jan. 2003.
[34] L. Wu, X. F. Zhu, and T. Tong, "Global and local clustering with KNN and local PCA," *Multimedia Tools Appl.*, vol. 77, pp. 29727–29738, Nov. 2018.
[35] W. Ji, G. Y. Chen, B. Xu, X. L. Meng, and D. Zhao, "Recognition method of green pepper in greenhouse based on least-squares support vector machine optimized by the improved particle swarm optimization," *IEEE Access*, vol. 7, pp. 119742–119754, 2019. doi: 10.1109/ACCESS.2019.2937326.
[36] C. H. Zhao, Y. S. Gao, J. He, and J. Lian, "Recognition of driving postures by multiwavelet transform and multilayer perceptron classifier," *Eng. Appl. Artif. Intel.*, vol. 25, no. 8, pp. 1677–1686, 2012.
[37] J. Liu, S. Tang, S. Shan, and J. Ju, "Simulation and test of grape fruit cluster vibration for robotic harvesting," *Trans. Chin. Soc. Agricult. Mach.*, vol. 47, no. 5, pp. 1–8, 2016.

**QIAN ZHANG** was born in 1991. She received the B.Eng. and M.S. degrees from Jiangsu University, in 2013 and 2016, respectively, where she is currently pursuing the Ph.D. degree with the School of Electrical and Information Engineering. Her research interests include image processing, and machine vision and robots.

**GUOQIN GAO** received the M.S. degree from the Nanjing University of Science and Technology, Nanjing, China, in 1989, and the Ph.D. degree from Jiangsu University, Zhenjiang, China, in 2006, where she is currently a Professor, a Ph.D. Tutor, and the Associate Dean with the School of Electrical and Information Engineering. Her research interests include parallel robot control and sliding mode control.