Computers and Electronics in Agriculture 122 (2016) 118–123

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

A method of segmenting apples at night based on color and position information

Xiaoyang Liu *, Dean Zhao, Weikuan Jia, Chengzhi Ruan, Shuping Tang, Tian Shen

School of Electrical & Information Engineering, Jiangsu University, Zhenjiang 212013, China

ARTICLE INFO

Article history:
Received 5 November 2015
Received in revised form 23 January 2016
Accepted 23 January 2016
Available online 5 February 2016

Keywords:
Machine vision
Apple-harvesting robot
Color space
Color and position information
Night images

ABSTRACT

This paper proposes a method to segment apples on trees at night for apple-harvesting robots based on color and position of pixels. Images of apples acquired under artificial light with low illumination at night include less color information than daytime images, so it is necessary to take position of pixels into consideration. The new method has two main steps. Firstly, color components of sampled pixels in RGB and HSI color space are used to train a neural network model to segment the apples. However, the segmentation results are incomplete and not able to guide apple-harvesting robots accurately, because partial edge regions of apples are dark in shadows and difficult to be recognized due to uneven illumination. Secondly, the color and position of pixels around segmented regions and pixels on the boundary of segmented regions are taken into consideration to segment the edge regions of apples. The union of two segmentation results is the final result. The complete recognition can increase the accuracy of location by about 6.5%, which verified the validity and feasibility of the method.

1. Introduction

Fruit-harvesting robots have attracted a wide attention in recent years (Plebe and Grasso, 2001; Tanigaki et al., 2008; Hayashi et al., 2010; Bachche and Oka, 2013). They are used to replace humans to harvest fruits on trees. It's obvious that the application of fruit-harvesting robots can reduce the labor cost and improve work efficiency. The emphasis of the research is the machine vision system, which is the core of the robots (Bulanon et al., 2004).

The features of images are treated as the basis of recognizing and segmenting fruits in complex natural background. Ji et al. adopted region growing based on color features to segment images, then extracted color and shape features of segmented images (Ji et al., 2012). The classification algorithm of support vector machine was applied to classify the extracted features and recognize apples. Lu and Sang adopted RB chromatic aberration map and normalized red map of original RGB images to segment images preliminarily (Lu and Sang, 2015). Then the contour fragments of segmented images are used to separate occluded citrus fruits. Teixidó et al. proposed a method to detect red peaches (Teixidó et al., 2012). Sampled pixels of leaves, branches and peaches are used to construct linear color models in RGB color space. Color distance from each pixel to different constructed models is the standard to classify pixels.

Experiments in references listed above were conducted under natural illumination in the daytime. If fruit-harvesting robots can work at night, the average work efficiency during a day will be improved prominently. Although researches on fruit harvesting at night are rare, some scholars have conducted researches on estimating fruit yield based on images that acquired at night. Font et al. acquired red grape images in vineyards under artificial lighting at night (Font et al., 2014). The specular reflection peaks from the spherical surface of the grapes are detected and the morphological method is applied to define the intensity peaks. Nuske et al. proposed a method to estimate yield by detecting green grapes in vineyards at night (Nuske et al., 2014). Cameras and illumination were installed on a vehicle to acquire images. The keypoints of images were detected firstly and classified on the basis of six-dimensional color vector from three RGB channels and three Lab channels and three texture features from three broad classes of features. Payne et al. used night time mango images at 'stone hardening' stage to estimate crop yield (Payne et al., 2014). Variance filter, gray scale, border limited mean, hessian filter and color components Cb and Cr are applied to recognize the mangos step by step.

The apple is selected as research object of the paper because it is one of popular fruits. This paper is aimed at solving the problem of image recognition at night and guiding robot to harvest apples

* Corresponding author.
E-mail address: liuxiaoyang@ujs.edu.cn (X. Liu).

http://dx.doi.org/10.1016/j.compag.2016.01.023
0168-1699/

rather than estimate yield. So, how to detect and locate apples on the tree at night is the key of the paper. Image recognition is divided into two steps. The first step is to recognize apples preliminarily on the basis of color information of images. The second step is to recognize the edge regions of apples based on the mixture of color information and position information from edge regions and not-edge regions. The experiment in this paper shows that the method is effective and feasible, which can increase the accuracy of location by about 6.5%.

2. Image recognition preliminary

Artificial light is applied to acquire images at night. Compared with natural illumination in the daytime, artificial illumination at night is uneven and in a limited range with low intensity because sunshine is approximately parallel and has higher intensity. Even if fruits are shaded by leaves, the illumination intensity on surface of the fruits is about 200 lx or more in a sunny day. However, the uneven illumination at night may cause regions in shadow under 10 lx. The camera can take photos clearly at 200 lx, but it is difficult to capture original color and details information of objects under 10 lx. Although the disadvantages of artificial light will damage partial color information of images, the color is still important feature of images because it still has enough information. Therefore, color information is used to recognize images preliminarily, which is divided into three small steps: image acquisition, color features extraction and neural network training.

2.1. Image acquisition

Image acquisition is the first and important step. The quality of images will exert effect on subsequent image processing. The orchard in which the images were captured is located in Feng Country, Xuzhou City, Jiangsu Province, China. The variety of apples in this orchard is Fuji apples. Canon IXUS 275HS is used to take images, of which the sensor is color complementary metal oxide semiconductor (CMOS). In order to reduce the volume of data, image resolution is set as $640 * 480$ pixels with jpg format. Every pixel includes 24 bits RGB (Red, Green and Blue) color information.

Incandescent lamps are used as artificial light for its good color rendering property. Incandescent light and sunshine are similar and the spectrums of them are both continuous. The color rendering index of incandescent lamps closes to $100R_a$ that is the maximal, which can make objects present their original color with little deviation. The color temperature is about 2900 K. The type of incandescent lamps is 12 V and 40 W because 12 V portable power source is adopted in field orchard. Two lamps were fixed on both ends of 1 m wide support, which can provide higher illumination intensity than only one lamp and weaken the shadow to a certain extent.

2.2. Color features extraction

The color information is significant and distinct feature of images that includes abundant valuable information. In order to analyze and apply the information adequately, some pixels in images are extracted as samples. These samples are representative for the main objects in images. These objects include apples, leaves, branches and other objects.

The color of different parts of one object is not identical. The pixels of apples are extracted from different regions of the surface of apples. The pixels of leaves are extracted from the obverse and the reverse sides of leaves. Fig. 1 shows an example about color feature extraction. These pixels are displayed in RGB color space as Fig. 2 shows. Fig. 2 also shows that the color of apples is different

Fig. 1. Color feature extraction example.

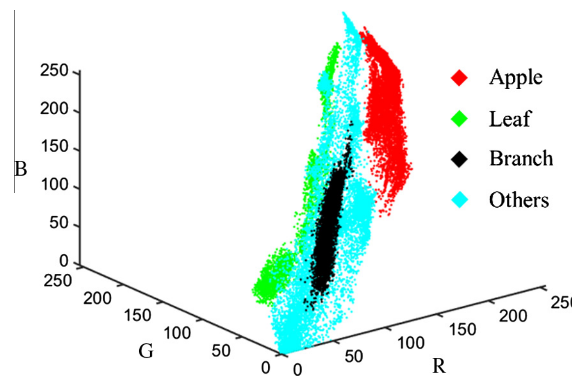

Fig. 2. Extracted pixels in RGB color space.

from the color of other objects obviously, which demonstrates the feasibility of image recognition.

Besides RGB color space, HSI (Hue, Saturation and Intensity) color space is a common choice to describe color. Intensity is separable from the color space and HSI color space is closer to human visual system than RGB color space. So, HSI color space is also chosen to describe colors in the experiment. The conversion formula (Weeks and Hague, 1997) from RGB to HSI is displayed as follows:

$$\left.\begin{array}{l} H = \begin{cases} \theta, (B \leqslant G) \\ 360 - \theta, (B > G) \end{cases} \\ S = 1 - \frac{3}{(R+G+B)}[\min(R, G, B)] \\ I = \frac{1}{3}(R + G + B) \end{array}\right\} \tag{1}$$

In formula (1), $\theta = \arccos\left\{\frac{[(R-G)+(R-B)]}{2[(R-G)^2-(R-G)(G-B)]^{1/2}}\right\}$.

In formula (1), R, G, B color components is normalized and angle θ is measured with respect to the red axis of the HSI color space. S and I are in the range [0, 1] and H can be normalized to [0, 1] by dividing by 360° (Gonzalez, 2009). HSI color space is described by a circular cone model. In order to display samples in HSI color space expediently, the circular cone is converted to rectangular coordinate system, as shown in Fig. 3. Fig. 3 also shows that the color of samples from apples is different from other samples in HSI color space.

2.3. Neural network training and data classification

In order to apply the color information of samples, a feed-forward back propagation neural network (BPNN) is established. BPNN is one of classical artificial neural networks, which has been applied to data classification widely (Hecht-Nielsen, 1989). The

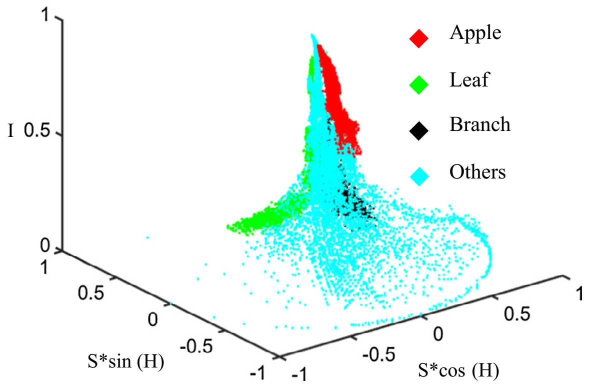

Fig. 3. Extracted pixels in HSI color space.

BPNN is applied to classify all pixels of images to two categories based on color information in this paper, which consists of an input layer, an output layer and a hidden layer.

R, G, B, H, S and I color components of samples are used to be the inputs of the BPNN. So, the input layer has 6 input interfaces. The size of output data is 1, which shows whether a pixel is an apple pixel. The size of hidden layer is defined by an empirical formula (Gao, 1998), as shown in formula (2).

$$s = \sqrt{0.43mn + 0.12n^2 + 2.54m + 0.77n + 0.35} + 0.51 \quad (2)$$

In formula (2), s: size of hidden layer, m: size of input layer, n: size of output layer. The s needs to be rounded to an integer. The formula is determined by surface fitting based on a great deal of instances. The empirical formula has been verified and applied in other researches (Sun et al., 2005; Shen and Zhang, 2010). According to the formula, m is 6 and n is 1, so s is 5. So far, the topological structure is determined as $6 \times 5 \times 1$.

The 6 color components of extracted pixels are used to train the BPNN by 10 times. The mean squared error (*MSE*) of each trained BPNN is displayed in Table 1. The trained BPNN with minimum *MSE* is selected to classify the pixels in images. The output data is rounded to 0 or 1 and reshaped to binary images. The binary images are just preliminary recognition results.

3. Image modification

The results of preliminary image recognition also need to be processed by morphological method, such as filling holes and filtering small connected regions. The processed image plus corresponding original image is shown in Fig. 4, which shows that the edge regions of apples are not recognized only based on the color information of images. The color of the edge regions is close to the color of some other objects in images due to uneven artificial light with low illumination. If the aim is to estimate yield, the recognition of the edge regions is able to be ignored. However, the aim of image recognition in this paper is to guide robot to harvest apples. It's necessary to modify the results and improve accuracy of image recognition.

3.1. Sampling pixels on the edge of apples

The color information of the edge regions of apples is not enough to separate the edge from images according to the experience above, so the position information of pixels on the edge regions is also taken into consideration. The position information is inspired by an assumption that if a pixel belongs to a certain object, the pixels around it are likely to belong to the objects, because an object in an image consists of lots of pixels that are adjacent in two-dimensional plane. In the experiment, if a pixel is around the preliminary recognized region, it has more possibility to belong to apples. According to the assumption, satisfactory pixels are sampled. The specific example is clarified as follows.

In Fig. 5, A is the center of recognized region of an apple; C is a pixel around the edge region of the apple and B is the point of intersection between the line from A to C and boundary of the recognized region. The position information and color information of B and C are applied to recognize the edge regions.

Supposed that the coordinate of C is (x_1, y_1), the coordinate of B is (x_2, y_2), and the color value in RGB color space of C is (R_1, G_1, B_1), the color value in RGB color space of B is (R_2, G_2, B_2). In order to decrease data, it's necessary to simplify the data. The position relationship between B and C is able to be described by Euclidean distance, as shown formula (3) and the color relationship is able to be described by the difference of B and C that is $(R_1–R_2, G_1–G_2, B_1–B_2)$.

Table 1
MSE of each trained BPNN.

Time	1	2	3	4	5	6	7	8	9	10
MSE	0.00159	0.00170	0.00087	0.00154	0.00126	0.00111	0.00150	0.00124	0.00092	0.00131

Fig. 4. Superposed images.

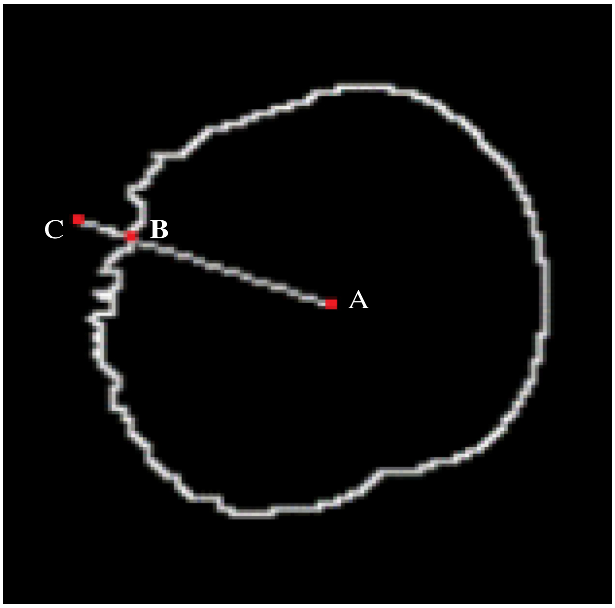

Fig. 5. The position relationship of pixels on edge regions.

$$l = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{3}$$

where l is the Euclidean distance from B to C.

3.2. Neural network training and data classification

In order to classify pixels around recognized regions of apples, (l, R_1–R_2, G_1–G_2, B_1–B_2) of each sample is applied to train a BPNN. The structure of BPNN has an input layer, a hidden layer and an output layer. The size of input layer is 4, which depends on the size of input vector. The size of hidden layer is 4, which is determined by an empirical formula, as shown in formula (2). The output data is rounded to 1 or 0, which represent whether pixels belong to the edge regions.

The 4 input data are used to train the BPNN by 10 times. The mean squared error (*MSE*) of each trained BPNN is displayed in Table 2. The trained BPNN with minimum *MSE* is selected to classify the pixels on the edge regions.

It's necessary to determine the range of the edge regions of apples. The pixels in the range have more possibility to belong to the edge regions. Compared with the long strips branches, the shape of apples is suborbicular in the plane. Therefore, the circle is used to determine the range. In the experiment, the geometrical center of the recognized region is used as the center of a circle. A circle is drawn on the superposed image based on the center and a certain radius, as shown in Fig. 6a. If the radius is the maximum distance from points on the contour of the recognized region to the center, the circle just include 97.12% of the whole apple based on statistical data of 20 images. If the radius is the maximum distance plus the width of 5 pixels, the circle can include 99.72% of the whole apple. Therefore, the radius is determined as the maximum distance plus the width of 5 pixels, which is an empirical value.

The pixels in the circle except for the recognized region are considered as pixels that are more possible to belong to the edge regions. Vector (l, R_1–R_2, G_1–G_2, B_1–B_2) extracted from each pixel in the range and its corresponding intersection is used as the input data of the BPNN. The output data is rounded to 0 and 1 and reshaped to a binary image, as shown in Fig. 6b. The sum of two results of preliminary image recognition and recognized edge regions is the final result.

4. Analysis and summary

4.1. Feasibility analysis

Color feature is easy to be extracted and has remarkable difference, which is applied to recognize fruits under natural environment widely (Tabb et al., 2006; Stajnko et al., 2009; Hočevar et al., 2014). In this paper, the method that is adopted to classify the pixels by BPNN based on their color is an application of color features. One of simple and direct applications is chromatic aberration method, which is fast recognition method (Lü et al., 2014). The chromatic aberration is described by the formula (4).

$$OUT = R - G \tag{4}$$

In formula (4), R: normalized R color component, G: normalized G color component, OUT: output result. The formula is proposed based on the analysis of the color of different objects. The OUT is

Table 2
MSE of each trained BPNN.

Time	1	2	3	4	5	6	7	8	9	10
MSE	0.00205	0.00389	0.00324	0.00447	0.00246	0.00363	0.00182	0.00216	0.00367	0.00410

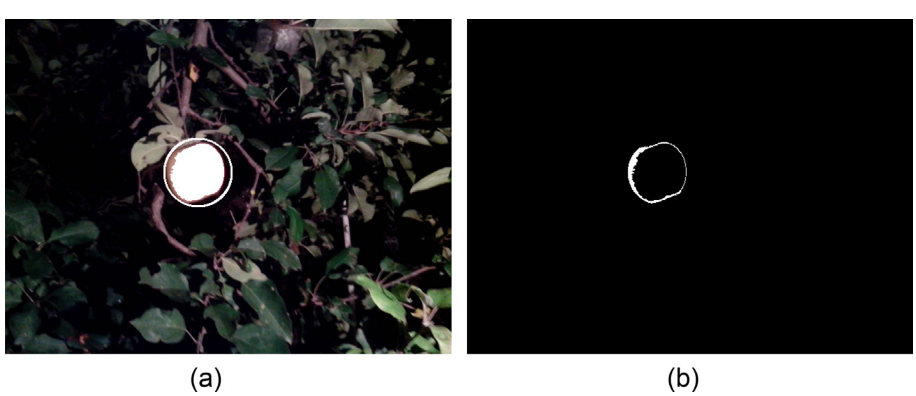

Fig. 6. The classification of pixels on the edge regions.

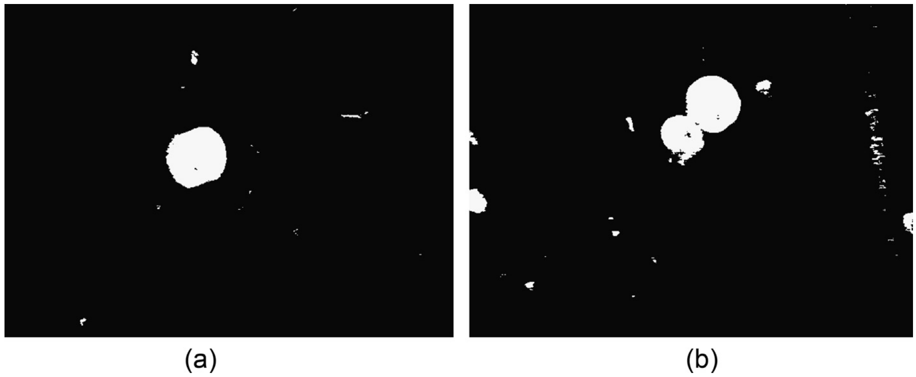

Fig. 7. The result of improved chromatic aberration method.

Table 3
20 ratios of the offset.

No.	Value (%)	No.	Value (%)	No.	Value (%)	No.	Value (%)
1	7.27	6	2.73	11	8.26	16	9.45
2	6.75	7	6.23	12	7.14	17	7.14
3	3.30	8	5.23	13	2.48	18	6.56
4	4.90	9	4.56	14	3.38	19	8.35
5	9.13	10	10.71	15	6.84	20	9.57

a grayscale image and OTSU method (Otsu, 1975) is applied to transform it to binary image. The recognition result is the binary image after necessary morphological calculation. However, the recognition accuracy of the method is not high and the method can't adapt the night illumination.

In order to improve the applicability of the chromatic aberration method, an improved chromatic aberration method is proposed (Dean et al., 2015). It is my previous work and an exponent, γ, is employed for the improved formula (5).

$$OUT = R^{\gamma} - G \tag{5}$$

The parameter γ is used to stretch the R component to increase difference between R and G. Comparing different recognition results caused by different value of γ, the value of it is determined as 1.5. Fig. 7 shows the binary output results. They have obvious and more noise, which shows that the method proposed in this paper is more accurate and has less error rate. The experiment verifies the feasibility of applying the method to fruit recognition at night further.

4.2. Accuracy analysis

The accuracy of using BPNN to classify pixels based on color information is not enough to guide the robot to harvest apples, because partial edge regions of apples is not recognized for uneven and low illumination. The absence of unrecognized regions will lead to the offset of center of fruit. In order to recognize the edge regions, pixels around recognized regions are classified by BPNN based on their color and position information. The result of complete recognition is the region of preliminary recognition plus corresponding edge region.

In order to describe the improvement of accuracy, the center of complete recognition is compared with the center of preliminary recognition. 20 apples unsheltered by branches or leaves are selected because it is convenient to describe the offset of centers without other interferences. The offset of the two centers is described by formula $Value = l/r \times 100\%$. In the formula, l is the distance between the two centers; r is the radius of the circle that has equal area to the result of complete recognition; *Value* is the ratio of the offset.

The data is displayed as Table 3 and the average of the ratio of offset is 6.499%. These data shows that if the edge regions are not recognized, the accuracy will be affected. The effect probably leads to the failure of harvesting apples. In a word, compared with the preliminary recognition, the complete recognition can increase the accuracy by about 6.5%.

4.3. Experiment summary

Fig. 8 shows all main steps to recognize the images at night. The method proposed in this paper is aimed at solving image recognition of apple-harvesting robot under artificial light with low illumination at night. The method recognizes images by two BPNN. The first BPNN is used to recognize fruits preliminarily, which recognizes the main regions of apples by color information of all pixels. The second BPNN is used to modify preliminary recognized results to improve the accuracy of recognition, which recognizes the edge regions of apples by color and position information of pixels around recognized regions.

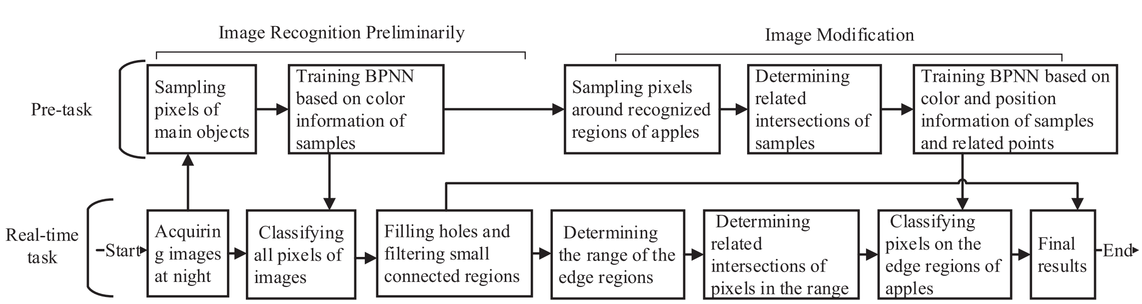

Fig. 8. The flow diagram of image recognition.

Sampling, training and classifying are main steps to apply BPNN in practice. However, the main problems are how to determine the topological structure and how to extract characteristic values of samples as input data. In this paper, the BPNN adopted classical topological structure with 3 layers and the number of nodes in hidden layer is determined by empirical formula. The color components of RGB color space and HSI color space are used as input data of the first BPNN. The color of apple pixels in shadow regions is damaged and it is similar to the background and branches especially. If the color components of these pixels are extracted as training data of the first BPNN, error recognition will be caused. Therefore, the color of each pixel around recognized regions and its related intersection and the distance of them are extracted as vector (l, R_1–R_2, G_1–G_2, B_1–B_2), which is applied as input data of the second BPNN.

All main steps can also be divided to two parts from another perspective: pre-task and real-time task. Pre-task is shown as the top half of Fig. 8, which includes pixels sampling and BPNN training. Real-time task is the other half, which is to use two trained BPNN to recognize images one by one. The pre-task doesn't consume the apple-harvesting robot work time, which are finished in advance. The BPNN not only finishes time-consuming training task before working in field, but also can replace artificial modeling to model automatically.

The experiment shows that the method is effective and feasible and can improve the accuracy of recognition at night. Especially, the recognition of the edge regions can increase the accuracy of location by about 6.5%. However, it also has some disadvantages that need to be improved further. The method is difficult to adapt to the change of illumination. Once the illumination changes, the BPNN has to be trained again.

Acknowledgements

This work is supported by the National Nature Science Foundation of China (No. 31571571); Priority Academic Program Development of Jiangsu Higher Education Institutions; The Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20133227110024); Ordinary University Graduate Student Research Innovation Projects of Jiangsu Province (No. KYLX14_1062).

References

Bachche, S., Oka, K., 2013. Performance testing of thermal cutting systems for sweet pepper harvesting robot in greenhouse horticulture. J. Syst. Des. Dynam. 7 (1), 36–51.

Bulanon, D.M., Kataoka, T., Okamoto, H., et al., 2004. Development of a real-time machine vision system for the apple harvesting robot. SICE 2004 Annual Conference, vol. 1. IEEE, pp. 595–598.

Dean, Z., Xiaoyang, L., Yu, C., et al., 2015. Image recognition at night for apple picking robot. Trans. Chin. Soc. Agric. Mach. 3, 003.

Font, D., Pallejà, T., Tresanchez, M., et al., 2014. Counting red grapes in vineyards by detecting specular spherical reflection peaks in RGB images obtained at night with artificial illumination. Comput. Electron. Agric. 108, 105–111.

Gao, D.Q., 1998. On structures of supervised linear basis function feedforward three-layered neural networks. Chin. J. Comput. 21, 80–86.

Gonzalez, R.C., 2009. Digital Image Processing. Pearson Education India.

Hayashi, S., Shigematsu, K., Yamamoto, S., et al., 2010. Evaluation of a strawberry-harvesting robot in a field test. Biosyst. Eng. 105 (2), 160–171.

Hecht-Nielsen, R., 1989. Theory of the backpropagation neural network. Neural Networks, 1989. IJCNN., International Joint Conference on. IEEE, pp. 593–605.

Hočevar, M., Širok, B., Godeša, T., et al., 2014. Flowering estimation in apple orchards by image analysis. Precision Agric. 15 (4), 466–478.

Ji, W., Zhao, D., Cheng, F., et al., 2012. Automatic recognition vision system guided for apple harvesting robot. Comput. Electr. Eng. 38 (5), 1186–1195.

Lü, Jidong, Zhao, Dean, Ji, Wei, et al., 2014. Fast tracing recognition method of target fruit for apple harvesting robot. Trans. Chin. Soc. Agric. Mach. 45 (1).

Lu, J., Sang, N., 2015. Detecting citrus fruits and occlusion recovery under natural illumination conditions. Comput. Electron. Agric. 110, 121–130.

Nuske, S., Wilshusen, K., Achar, S., et al., 2014. Automated visual yield estimation in vineyards. J. Field Robotics 31 (5), 837–860.

Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica 11 (285–296), 23–27.

Payne, A., Walsh, K., Subedi, P., et al., 2014. Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and night time imaging. Comput. Electron. Agric. 100, 160–167.

Plebe, A., Grasso, G., 2001. Localization of spherical fruits for robotic harvesting. Mach. Vis. Appl. 13 (2), 70–79.

Shen, X., Zhang, J., 2010. Research of intrusion detection based on the BP networks and the improved PSO algorithm. Comput. Eng. Sci. 6, 34–36.

Stajnko, D., Rakun, J., Blanke, M., 2009. Modeling apple fruit yield using image analysis for fruit color, shape and texture. Eur. J. Horticult. Sci., 260–267

Sun, Y.P., Liu, M., Wu, C., 2005. Intelligent prediction algorithm of trim beam number in colored textile manufacturing process. Control Eng. China 12 (6), 523–526.

Tabb, A.L., Peterson, D.L., Park, J., 2006. Segmentation of apple fruit from video via background modeling. ASABE Paper, 063060.

Tanigaki, K., Fujiura, T., Akase, A., et al., 2008. Cherry-harvesting robot. Comput. Electron. Agric. 63 (1), 65–72.

Teixidó, M., Font, D., Pallejà, T., et al., 2012. Definition of linear color models in the RGB vector color space to detect red peaches in orchard images taken under natural illumination. Sensors 12 (6), 7701–7718.

Weeks, A.R., Hague, G.E., 1997. Color segmentation in the HSI color space using the k-means algorithm. Electronic Imaging'97. International Society for Optics and Photonics, pp. 143–154.