Computers and Electronics in Agriculture 112 (2015) 68–74

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Location of apples in trees using stereoscopic vision

Si Yongsheng [a,b], Liu Gang [b,*], Feng Juan [a]

[a] College of Information Science & Technology, Agricultural University of Hebei, Baoding 071001, China
[b] Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China

ARTICLE INFO

Article history:
Received 12 April 2014
Received in revised form 9 January 2015
Accepted 9 January 2015
Available online 2 February 2015

Keywords:
Fruit harvesting
Image segmentation
Stereo matching
Random ring method
Machine vision

ABSTRACT

In order to design a robot which can automatically recognize and locate apples for harvesting, a machine vision system was developed. Three algorithms used in the vision system to recognize and locate apples were described in this study. An apple recognition algorithm with color difference $R-G$ and color difference ratio $(R-G)/(G-B)$ was presented. If a pixel met $R-G>0$ and $(R-G)/(G-B)>1$, either the pixel was identified as an apple, else the pixel was background. The fruit shape features were extracted from contour images based on random ring method (RRM). A matching algorithm based on area and epipolar geometry was discussed to locate the apples. The apples with similar areas were matched according to the principle of ordering constraint by calculating the maximum value of cross-correlation function of vertical projections. The experiment results showed that the proposed recognition method could eliminate the influences of shade and soil. Over 89.5% of fruits were successfully recognized from 160 tested images. The circle centers and radii were extracted precisely based on the random ring method. The errors were less than 20 mm when the measuring distance was between 400 mm and 1500 mm.

1. Introduction

The harvesting of fruit such as apples, oranges or tomatoes is a monotonous task and is currently being performed by hand. At the same time, labor costs are rising inversely proportionate to the declining farmer population. These are the two main reasons why numerous researchers have studied robotic fruit harvesting. Since the first tomato harvesting robot was developed (Kawamura et al., 1984) at Kyoto University about 30 years ago, many projects related to fruit harvesting have been carried out worldwide (Slaughter and Harrell, 1987; Harrell et al., 1990; Kondo et al., 1992; Hayashi et al., 2002; Safren et al., 2007). The first major task of a harvesting robot is to recognize the fruit and determine its location (Van Willigenburg et al., 2004).

Many studies concerning vision systems for robotic fruit harvesting have been reported and have been reviewed by Li et al. (2011). Typically, the key element of these vision systems is a camera which produces images from harvesting scenes. The detection of fruits on those images is performed by means of image processing algorithms. Since the image is two-dimensional, the range or distance from the camera to image scene is unknown. To estimate this distance, another range sensor such as an ultrasonic sensor or a laser range sensor is used (Bulanon et al., 2004; Hannan et al., 2004), and this method can be referred to as a combination of camera and a range sensor. The binocular stereo vision method is another common method used to calculate three dimensional (3D) locations of fruits. The binocular stereo vision method is able to calculate the distance of fruits from the camera using two images captured by two cameras placed at different viewpoints, based on the triangular measurement principle. In most studies, two cameras were placed together in parallel or with a certain angle (Hayashi et al., 2002; Ting et al., 2009), while in some studies, one camera was mounted on the vehicle and the other was mounted on the end-effecter or Manipulator (Van Henten et al., 2002). The binocular stereo vision method has been widely used in different fruit robotic harvesting: apples (Sun et al., 1994), oranges (Plebe and Grasso, 2001) and tomatoes (Kondo et al., 2008). The above mentioned methods using cameras as the main sensing devices show the trend of being affected by illumination variations. To avoid or reduce the influence of natural lighting conditions, the active triangulation method was used by some researchers. The active triangulation system employed a laser projector able to project two laser beams with different wavelengths. The laser beams were projected onto the objects and received by position sensitive devices (PSDs). Then, by scanning the laser, the beams and fruits were distinguished from the background by the difference in spectral-reflection characteristics between the two laser beams. This method was used in several studies for the detection and 3D location of different fruits such as apples (Bulanon

* Corresponding author. Tel.: +86 10 62736741.
E-mail address: pac@cau.edu.cn (G. Liu).

http://dx.doi.org/10.1016/j.compag.2015.01.010
0168-1699/

et al., 2004), cherry tomatoes (Subrata et al., 1997), cucumbers (Fujiura et al., 2000), cherries (Tanigaki et al., 2008) and oranges (Jiménez et al., 2000). More machine vision systems for automatic fruit harvesting have been reviewed by Li et al. (2011).

All the methods mentioned above have several problems which have not been sufficiently solved by previous researchers. The range sensors can only return a distance of a single point. The method of range sensor cannot provide the exact 3D locations for all the points in the harvesting scene, only some specified points. The advantage of the active triangulation method is that the range data may be obtained without much computation. However, this method only applies when objects are sufficiently close, so that the reflected light may be reliably observed (Kondo and Ting, 1998). Moreover, its speed is very slow, and thus not suitable for real-time applications. The stereo vision method does possess remarkable potential, but it also has some difficulties which have yet to be overcome: (a) variation in lighting conditions may result in poor quality images; (b) the correspondence problem or stereo matching is difficult to overcome.

Harvesting speed is a key index among the performance parameters of a robot for practical use. One reason why these harvesting robots have yet to be commercialized is that the operation speed of one of these robots is equal to or less than that of human (Kondo et al., 2008). Among all the techniques proposed for the automatic location of fruit, stereo vision can satisfy the real-time requirements of automatic harvesting systems, as it only involves the acquisition of two images per scene, without any mechanical movements of other devices (Tarrío et al., 2006). In this study, stereo vision was selected as the fruit location method.

The harvesting robots work in an uncontrolled environment. In addition to the great variation in illumination as mentioned above, there are some other difficulties in uncontrolled environments which have yet to be overcome, such as the occlusions and overlapping of fruits. An apple harvesting robot consisting of a manipulator, an end-effecter, a visual sensor and a vehicle has been developed at the Precision Agriculture Research Centre (PAC) at China Agricultural University. The robot has been tested for harvesting in an orchard. The objectives of this study were as follows: (1) to develop an apple recognition algorithm that may be applied in varying lighting conditions, while taking into account fruit occlusion with overlapping; (2) to design a circle detection algorithm of apples; (3) to develop a stereo matching algorithm for the location of fruit. It was found that these algorithms provided acceptable performance under actual outdoor conditions.

2. Materials and methods

2.1. System components

The binocular stereovision system consisted of the following parts: two CMOS color cameras, a camera-link interface and gigabit nets interface conversion box, a camera guide and a tripod, as shown in Fig. 1. The cameras (model MVC1000SAC-GE30), which were produced by Beijing microview company, had a digital video output of 1280 by 1024 effective pixels and a 6 mm fixed focal length. Two cameras were mounted on the slide bar in parallel with the adjustable length between the centers of the two camera lens (baseline length), which was fixed at the length of 200 mm in this study.

The images used for further processing were retrieved from a commercial apple orchard in Changping district of Beijing in apple maturation period (in September and October). And the images were processed on a notebook computer (Intel Core 2 Duo 2.4 GHz, RAM 0.97G, Windows XP). The image acquisition and processing algorithms were implemented in the VC++ 6.0 programming environment. The variety of apple used in this study was *Fuji*.

Fig. 1. The binocular stereovision system.

2.2. Apple recognition algorithm

Shape based analysis and color based analysis are the two main methods for fruit recognition in images. To avoid the effects of illumination, Kelman and Linker, 2014 presented a method for localization of mature apples in tree images using convexity. Many other studies (Yaqin and Hua, 2004; Bulanon et al., 2002) focused on approaches based on color analysis. In the color based studies, some color indicators were proposed to emphasize the fruits in the image. Subsequently, the fruits were determined with a certain threshold value. Most of these indicators were empirical parameters, and did not function well in some conditions. The indicator $2R–G–B$ (Yaqin and Hua, 2004) was vulnerable to the impact of shadow, and the results were not ideal in cases such as those where backlighting was present or the background contained soil. The red color of mature Fuji apple fruit is the most obvious feature in comparison with background. In this study, color based analysis was adopted.

To find a more effective color indicator, line profiles of apple images in RGB color space were plotted and analyzed. A careful observation of Fig. 2 revealed that pixels between M and N on the white line belonged to fruit, and the other pixels belonged to the background, including leaves, stems and sky. Observing the profile curves, it was clear that in the fruit, R values were greater than G values, while the opposite was true in the background. Attempting to use the $R - G$ color difference image segmentation method, the $R - G$ difference was binary processed (if $R - G > 0$, $R - G = 255$, or $R - G = 0$), and the profile curve $R - G$ was obtained. Apple fruit (the part between M and N with the value of 255) could be recognized completely by $R - G$, while in the background area there were some error recognitions. Therefore, the results would show much noise when only using the $R - G$ color difference method for image segmentation. Some other color characteristics need to be combined with $R - G$ color difference in order to recognize the fruits. In the fruit area, the difference between R and G was clearly greater than that between G and B, while in the part where error recognitions occurred with only using $R - G$ (background area where the $R - G$ value is 255), the difference between R and G was not so significant in comparison to that between G and B. This can be used as the second color characteristic, combined with the first color characteristic of R greater than G in the fruit area to recognize the fruits. Specifically, the value of color difference ratios $(R - G)/(G - B)$ in the part where error recognitions occurred with only using $R - G$ was different from that in the fruit area. The cutoff value was 1 (for the convenience of display, the value in the figure

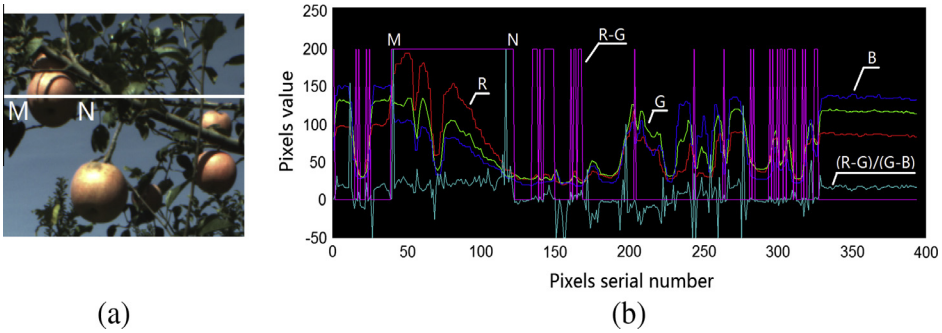

Fig. 2. Result of line profile map. (a) Raw image and (b) line profile map.

was 10 higher than the original value). If a pixel met $R - G > 0$ and $(R - G)/(G - B) > 1$, the pixel was identified as an apple, else the pixel was background.

2.3. Image cleaning

Pretreatments such as noise elimination and edge extraction should be performed before feature extraction. After the apples were recognized, a small group of leaves, branches and soil were incorrectly identified to form the noise points. These noise points not belonging to apples must be filtered. Some highly occluded apples should be removed because it was difficult to extract their features, which in turn influences processing speed. The image regional labeling was applied after the noise points were removed, and the area of each region was calculated, then the regions with areas of less than 1/8 of the largest area were eliminated. Morphological opening operation and closing operation with 3×3 square structuring elements were applied once. Most of the noises and highly occluded apples in the image were filtrated after the above treatment. The holes often appeared within the fruit image due to reflected light. To eliminate these holes for facilitating further image processing, area filling was applied with the seed filling algorithm. Then, eight connected boundaries were traced and the fruit contour image was obtained, which was beneficial to reducing the computational cost of the extraction process. Finally, eight-connected region mark algorithm was used to mark the different fruit contour regions.

2.4. Circle detection based on Random Ring Method (RRM)

The feature extraction of point, line and so on was the foundation of feature matching and location. The feature extraction of apple fruit was based on the fruit contour image. Hough transform is a very conventional approach to this problem, but is also regarded as time consuming. Moreover, it was empirically found that Hough transform was unable to work well with clusters of fruits, or when a substantial part of the fruit surface was occluded (Plebe and Grasso, 2001). Apples were circular in the images. Any three non-collinear pixels on the apple contour can determine a circle O_A. Assuming that the center and radius of the circle O_A are (O_x, O_y) and R_0 respectively. The radius of the apples in the images were within a certain range(30–60 pixels) due to the camera shooting distance limitations. The distance in this range was defined as the constraint radius. If R_0 exceed the above range, then discarded. Otherwise, the distances D_i between the points on the contour and the center (O_x, O_y) were calculated. Theoretically, these distances should be equal to the radius R_0. The apple contours are generally irregular circles and pixels on the contour are not exactly on one circle. A width R_w was set and the number N_{max} of the contour points within $R_0 \pm R_w$ was counted. The higher the number N_{max} was, the better the circle fitting was. Circle fitting was performed several times. The circle corresponding to the largest number was selected as the final result. Fig. 3 was the flow chart of the proposed algorithm. In short, this algorithm was circle fitting by counting the points within a random ring. It can be defined as random ring method (RRM). As for the apples in a cluster, there was possibility that the RRM will find a radius larger than 60 pixels through 3 random points. In that case, the RRM algorithm would discard it because that radius was out range. And then the RRM will detect the circle through anther 3 random points till the RRM algorithm finish.

2.5. Stereo matching based on area and cross-correlation function of vertical projection

The main difficulty of the stereo vision method was obtaining the points in both images corresponding to the same point in the actual scene, known as the correspondence problem. The process of searching the corresponding point in the other image is referred to as stereo matching. Some researchers (Batlle et al., 1998) have attempted to solve this problem by using structured lighting. However, this method was not suitable for use in outdoor environments. Tarrío et al. (2006) used a matrix of optical spot laser diodes to project light, whenever it was necessary to distinguish individual fruit growing in bunches. This method was proven effective, but also involved some mechanical movements with potential influence on the speed. Stereo matching methods may be further classified into the area-based method and feature-based method. The feature-based method was comparatively less sensitive to the illumination and therefore adopted in this study. Different fruits possessed different shape features in the binary images due to occlusion or overlapping, but the same fruit had similar shape in both the left and right images. Thus, an area was selected as having matching features. The focal length of the cameras f and the baseline length B were the two key parameters of the binocular vision system. According to the camera's shooting distance, f was selected as 6 mm. If the baseline is short, both the area and the position of apple circle in left and right image will be more consistent, in other words, a short baseline makes the matching of features easy, while a short baseline will make distance errors large due to narrow triangulation (Kanade et al., 1992). Considering the accuracy and matching difficulty, the baseline B was adjusted to 200 mm. Given any point on an object, the 'epipolar plane' was defined as the plane passing through the object point and both foci of the camera lens. This plane intersects the two image planes, defining an 'epipolar line' in each (Arnold and Binford, 1980a,b). The epipolar constraint states that given a point in one view (say the left image), its corresponding point in the other view lies on the epipolar line (Song et al., 2006). The fruits were located nearly at the same height range in the left and right images, according to the epipolar

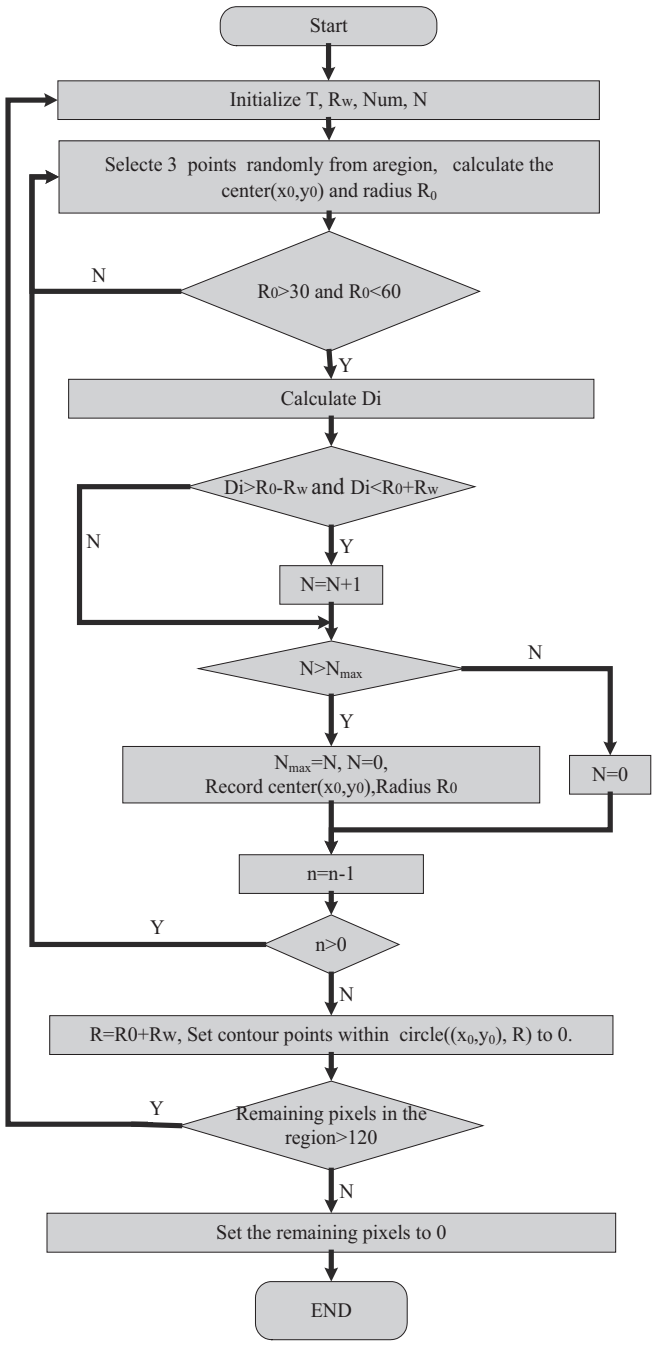

Fig. 3. Flow chart of random ring method.

constraint. When searching a fruit in right image corresponding to a fruit in the left image, the search area was not the entire right image, but a rectangular region with the same height as the fruit being matched in the left image. In the search area, there may have been one or more fruits, and the fruit with minimum area difference was selected as the corresponding fruit. Through the epipolar constraint mentioned above, the possibility of mismatch was reduced significantly. However, in the search area, there still may have been some fruits with similar areas to the fruit being matched. As shown in Fig. 4c and d, there were two fruits, P′ and Q′, which both had similar areas to fruit P in Fig. 4c. In order to obtain the only correct match, other constraints were required for further discrimination. The ordering constraint stated that if an object is to the left of another in one stereo mage, it is also to the left in the other image (Kolmogorov and Zabih, 2001). According to the ordering constraint, the correct corresponding fruit P′ could be determined. However, in certain conditions, some fruits had no corresponding objects in the other image, such as fruits *G* and K in Fig. 4c and d. In these conditions, directly using the ordering constraint method to perform the correspondence between two images showed a high degree of difficulty.

Fig. 4c was equivalent to the image obtained by shifting Fig. 4d to the right by a certain distance. A reference line could be found as shown in Fig. 4c. The fruits on the right side of the baseline had the same sequence as the fruits in Fig. 4d. Thus, the ordering constraint could be used to perform matching between the fruits with similar areas in the two images. To obtain the reference line, vertical projection was applied to the binary image for accumulating the values of the pixels in the vertical direction:

$$S(j) = \sum_{i=1}^{h} I(i,j) \quad j = 1,2 \ldots M \tag{2}$$

where $I(i, j)$ was the value of the pixel at (i, j) (value is 0 or 1). The above calculation was applied to two images, and two curves $Sl(j)$ and $Sr(j)$ were obtained as shown in Fig. 4c and d. There was little difference between the two images, so the two curves were similar. The reference line could be obtained by using the cross-correlation function (CCF). The CCF of two curves was shown as below:

$$R(n) = \frac{1}{M}\sum_{j=1}^{M} Sr(j) * Sl(j+n) \quad n = 1,2 \ldots M \tag{3}$$

It was assumed that R reached the maximum value when $j = j_{max}$, then the line $j = j_{max}$ was the reference line (Dotted line in Fig. 4c). In most cases, the fruits on the left of the reference line had no corresponding objects in the other image, and these fruits were ignored.

3. Results and discussions

Recognition, shape features extraction, stereo matching and distance measurement experiments were performed. It should be noted that these experiments were not separated from each other. For instance, the results of the recognition may have affected the accuracy of shape feature extraction. The first three experiments were laboratory tests and the last one was performed in an outdoor environment.

3.1. Recognition of apples by segmentation

Fig. 5 showed the recognition results of a typical image composed of leaves, branches, soil, shadows and sky. The results of the algorithm $2R–G–B$ and $Y–R$ were also displayed for comparison purposes. It could be observed from Fig. 5d that the fruit portion was properly segmented from the background. Fig. 5b and c showed that the fruits were segmented, but some parts of the background, such as leaves, branches and, in particular, soil, were also segmented as fruit. Fig. 5b and c also revealed that the fruits located in shadowed areas were not successfully recognized. Compared with $2R–G–B$ and $Y–R$ algorithm, this method performed more ideally under natural lighting conditions.

To evaluate the effect of the recognition algorithm proposed, 160 images captured in natural outdoor conditions were tested. Among the 160 images, 50 were captured in sunny day front lighting conditions, 50 in sunny day back lighting conditions, and the rest of images were in cloudy day conditions. Comparing the fruits in the images before and after segmentation, the fruit success recognition rate could be calculated. The images were processed using the segmentation algorithm mentioned above, recording the

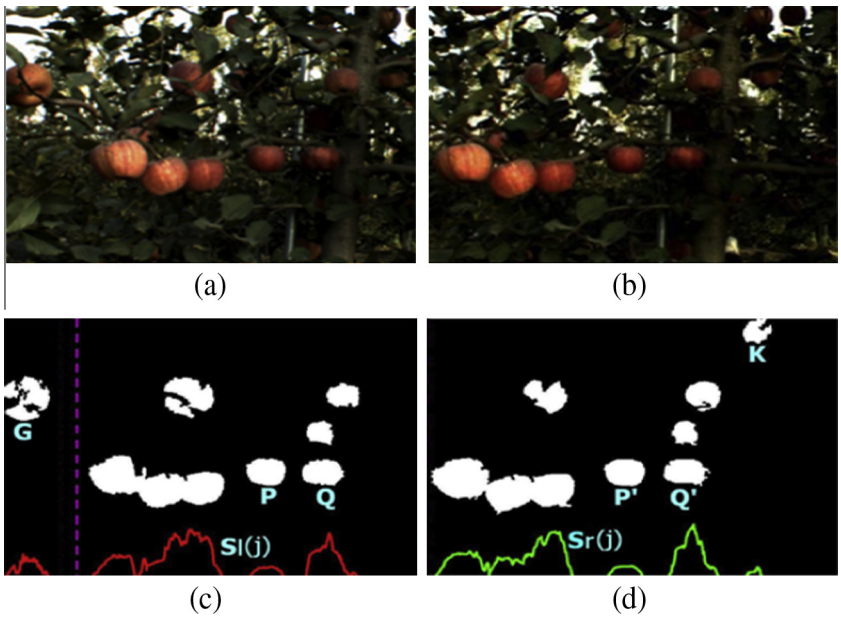

Fig. 4. Image match. (a) Left raw image, (b) right raw image, (c) left binary image and (d) right binary image.

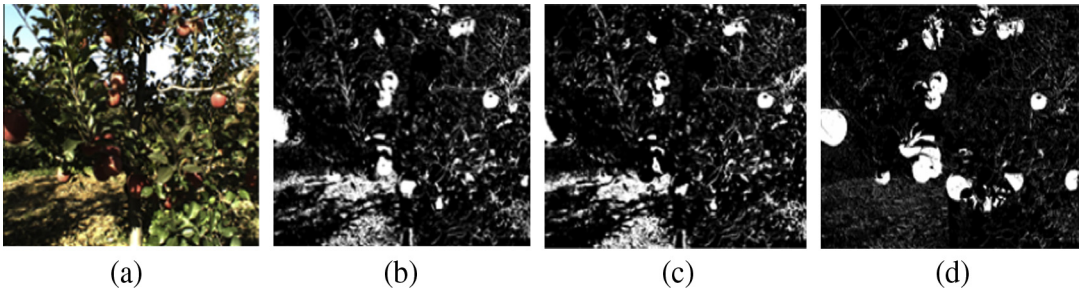

Fig. 5. Comparison of segmentation results. (a) Original image, (b) *Y–R*, (c) 2R–G–B and (d) proposed method.

results. Then the same images were processed manually, labeling apples, backgrounds pixels using image editing software (photoshop cs3). The automatic and manual identifications were then compared to evaluate the segmentation algorithm's results. Table 1 showed the performance of recognition algorithm in segmenting the fruits under different lighting conditions. As can be seen in Table 1, the proposed image segmentation method yields very satisfactory results. It achieves on average 92.5%, 89.5% and 97.5% fruit hits on front lighting, back light lighting and cloudy day conditions respectively. Segmenting under cloudy day conditions achieved the highest fruit hits (97.9%) while the lowest fruit hits occurred under back light conditions (89.5%). As expected, according to the color characteristics, it was difficult to fully distinguish the fruit and the background with similar color. This resulted in higher levels of false positives (Background was mistakenly recognized as fruit) under both front lighting conditions and backlighting conditions, 5.3% and 6.7%, respectively. Fruit false positives are preferable to fruit false negatives (Fruit was mistakenly recognized as background) because these false positive patches were very small in area and easy to be eliminated according to the area. So this result validates the proposed image segmentation method.

Table 1
Performance of fruit segmentation algorithm.

	Hits (%)	False positives (%)	False negatives (%)
Front lighting	92.5	5.3	2.2
Back lighting	89.5	6.7	3.8
Cloudy day	97.9	1.2	0.9

3.2. Circle fitting

In the RRM algorithm, the width R_w of the ring and the number of circle fitting times should be set. R_w was set to 5 according to the experience. In order to take into account the effect and speed of the circle fitting, different apple contours were selected to circle fitting. The fitting times n were set as follows: 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 400, 600, 800, 1000 and 1500. For each fitting times n, the radius was calculated and the calculations were repeated 100 times. The variances of each group of 100 radiuses were calculated. The circle fitting got satisfactory results when the variance was small. Fig. 6 was the variance curve of the apple contour. As can be seen in Fig. 6, when the fitting times was greater than 400 the variance changed slowly and close to 0. So in this study, n was set to 400.

To evaluate the performance of RRM circle fitting algorithm, the RRM and Hough transform algorithm were applied in Fig. 7. The cvHoughCircles in OpenCV library was the function to identify circles based on function to identify circular shape based on Hough transform algorithm. The cvHoughCircles was used to identify the circles in Fig. 7. The 2 parameters min_radius and max_radius were set to 30 and 60 respectively. There were five apples in Fig. 7.

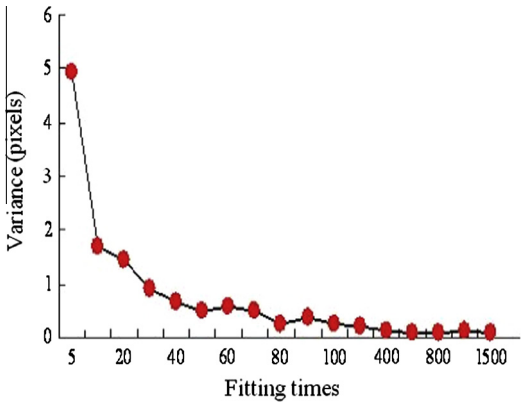

Fig. 6. The variance curve of the apple contour.

Most of Apple B was occluded by apple A. Apple D and E were occluded by leaves. Contours occluded of apple D and E were both about half of a circle. Fig. 7b was the result of the Hough transform. The circle fitting effect of Apple C and E were satisfactory. The contour of apple B was occluded more than half of the contour, so circle fitting was not successful for apple B. The Fitting radius of apple A was smaller than that of real apple. The main reason was that the contour of apple A was an non-standard nearly circular. The Fitting radius of apple D was larger than that of real apple due to the contour occlusion by leaves. Fig. 7c was the result of the RRM. The fitting result of RRM for apple A was significantly better than that of HT. This showed that RRM was able to detect irregular circular contour. Apple B was not able to be identified by RRM due to serious occlusion. For apple C, the circle fitted by RRM covered more points of the contour than the circle fitted by HT. So the fitting result of RRM for apple C was better than that of HT. For apple D, the radius of the circle fitted by RRM was closer to the real radius than that of HT. The circle fitted by RRM for apple E was similar with that of HT. The RRM identified circles by counting the points within random rings, so it was able to get better result when the apple contour was not standard circle. In summary, RRM was competent for circle fitting even if the contours were not standard circle or part of the contour was occluded.

3.3. Matching and distance estimation

The matching experiment was performed on 45 pairs of images. There were 323 pairs of apples viewable in both the left and right images. More than 95% of the apples were correctly matched. A total of 15 apple pairs were incorrectly matched. The main reason for these incorrect matches was that an apple was separated in the left image and connected with other apples in the right image, resulting in differences in order and area. To examine the influences of occlusion and overlapping as well as location accuracy, a distance estimation experiment was performed under indoor and outdoor conditions. In the empirical test under indoor conditions, an artificial apple tree was used as background, and a real apple was hung in front of the tree. Under the outdoor conditions, a real apple tree was used as background. Under indoor conditions, the apple was placed in three situations: no occlusion and no overlapping, with occlusion, and with overlapping. Under the outdoor conditions, the variation in illumination was the main difference from the indoor conditions. The distances between the apple and the camera were measured under five conditions, respectively: indoor with no occlusion and no overlapping, with occlusion, with overlapping, with outdoor front lighting, and with outdoor backlighting. For every 100 mm, a value was measured in the range of 400–1500 mm. At each point, the measurement was conducted three times and the average was taken as the measurement value. The stereo vision system and the laser rangefinder type PD4 (produced by Hilti, Germany) were both used to measure the distance. The measurement error of PD4 was less than 2 mm within the range of 50,000 mm, and this value was taken as the true distance. The errors between PD4 and the stereo vision system were recorded. Fig. 8 shows the errors under the above five conditions.

In the range of 400–1500 mm, all the errors were less than 20 mm. T-test was used to compare these errors and the result is shown in Table 2.

According to Table 2, there is no significant difference between the errors of ‘indoor no occlusion’ and ‘outdoor front lighting’. And there is no significant difference between the errors of ‘outdoor

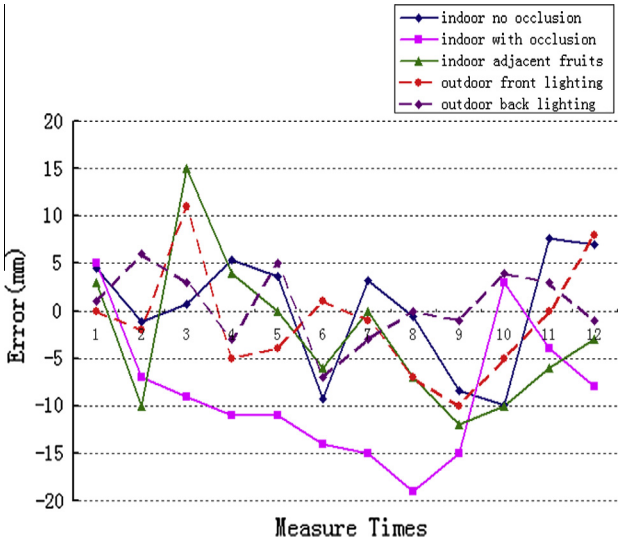

Fig. 8. Distance estimation result.

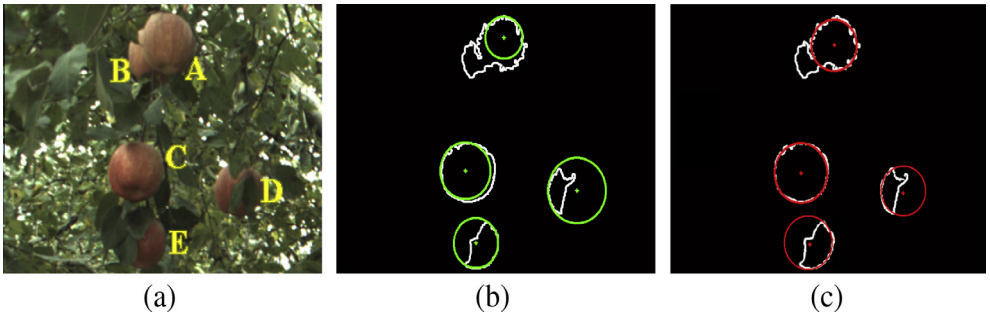

Fig. 7. Circle fitting of fruits. (a) Original image. (b) Circle fitting by HT. (c) Circle fitting by RRM.

Table 2
The *t*-test results of measuring errors.

Sample pair	*t*	t_{table}[a]	Significance
'indoor no occlusion' and 'outdoor front lighting'	0.12	2.07	No
'outdoor front lighting' and 'outdoor back lighting'	−0.86	2.07	No
'indoor no occlusion' and 'indoor with occlusion'	3.24	2.07	Yes
'indoor no occlusion' and 'indoor adjacent fruits'	1.01	2.07	No

[a] The value of t_{table} was at the significance level of 95%.

front lighting' and 'outdoor back lighting'. This implied that the proposed apple recognition algorithm was robust to the changing lighting conditions. The difference between the errors of 'indoor no occlusion' and 'indoor with occlusion' was significant, which indicated that, compared with illumination changes, occlusion had a larger influence on the proposed circle detection algorithm. There is no significant difference between the errors of 'indoor no occlusion' and 'indoor adjacent fruits'. And this implied that the RRM was effective when the fruits were adjacent.

The stereovision system was tested in real orchard. The test showed that the vision system can provide real-time location for harvesting. The average image processing cycle (from image acquisition to fruit location) was 2350 ms, which accounted for a small proportion of the entire harvesting cycle (about 45000 ms). However, the test in orchard also revealed the weakness of the proposed approach. The robot worked in the outdoors under natural sunlight conditions. The cameras need to capture images at different positions and from different directions among the fruit trees. It is impossible for the cameras to capture good quality images all the time. The proposed method did not work well under some challenging lighting conditions duo to the poor quality images captured, such as overexposed or underexposed images. To counteract the influence of these challenging light conditions, in the future research some other sensors, such as depth cameras, should be used to build a combined vision system and to combine the output of both color cameras and depth cameras.

4. Conclusions

A mechanical vision system for the recognition and location of apples was developed, and brief tests were performed. The vision system was composed of two cameras and a notebook PC. Three algorithms to recognize and locate the apples were introduced. Based on the test results, the following conclusions were achieved:

(1) An apple recognition algorithm with color difference $R - G$ and color difference ratio $(R - G)/(G - B)$ was stable and less affected by light conditions. A total of 160 images captured under different lighting conditions were tested, and over 96% of the images were correctly segmented.
(2) RRM was able to circle fitting even when the contours were not standard circle or when part of the contour was occluded. However, occluded and clustered fruits showed a certain degree of influence on the proposed RRM.
(3) The matching algorithms based on area features and epipolar geometry were effective in matching the apple pairs in the left and right images. More than 95% of the apples were correctly matched.
(4) In the range of 400–1500 mm, all the distance estimation errors were less than 20 mm.

Acknowledgements

This project was supported by a grant from the National Natural Science Foundation of China (No. 31371532), and by the Foundation of Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education (Jiangsu University).

References

Arnold, R.D., Binford, T.O., 1980a. Geometric constraints in stereo vision. Proc. SPIE 238, 281–292.
Arnold, R.D., Binford T.O., 1980. Geometric constraints in stereo vision. In: Image Proceeding for Missile Guidance, pp. 281–292.
Batlle, J., Mouaddib, E., Salvi, J., 1998. Recent progress in coded structured light as a technique to solve the correspondence problem: a survey. Pattern Recogn. 31 (7), 963–982.
Bulanon, D.M., Kataokab, T., Ota, Y., et al., 2002. A segmentation algorithm for the automatic recognition of Fuji apples at harvest. Biosyst. Eng. 83 (4), 405–412.
Bulanon, D.M., Kataoka, T., Okamoto, H., Hata, S., 2004. Determining the 3-D location of the apple fruit during harvest. Proc. Automat. Technol. Off-Road Equip., 91–97
Fujiura, T.S., Ueda, K.J., Chung, S.H., et al., 2000. Vision system for cucumber-harvesting robot. In: Bio-robotics II, 3rd IFAC International Workshop on Bio-Robotics, Information Technology and Intelligent Control for Bioproduction Systems, Osaka, Japan, November 25–26.
Hannan Michael, W., Burks, Thomas F., 2004. Current developments in automated citrus harvesting. In: ASAE/CSAE Meeting, 2004, Paper No. 043087.
Harrell, R.C., Adsit, P.D., Munilla, R.D., Slaughter, D.C., 1990. Robotica 8, 269–278.
Hayashi, Shigehiko, Ganno, Katsunobu, Ishii, Yukitsugu, et al., 2002. Robotic harvesting system for eggplants. JARQ 36 (3), 163–168.
Jiménez, A.R., Ceres, R., Pons, J.L., 2000. A vision system based on a laser range-finder applied to robotic fruit harvesting. Mach. Vis. Appl. 11 (6), 321–329.
Kanade, T., Okutomi, M., Nakahara, T., 1992. A multiple-baseline stereo method. In: Proceedings of the 1992 DARPA Image Understanding Workshop, pp. 409–426.
Kawamura, N. et al., 1984. Study on agricultural robot (Part 1) microcomputer-controlled manipulator system for fruit harvesting. J. Jpn. Soc. Agric. Mach. 46 (3), 353–358 (In Japanese with English summary).
Kelman, Eliyahu (Efim), Linker, Raphael, 2014. Vision-based localisation of mature apples in tree images using convexity. Biosyst. Eng. 118, 174–185.
Kolmogorov, V., Zabih, R., 2001. Computing visual correspondence with occlusions via graph cuts. In: Proc. Int'l Conf. Computer Vision, vol. II, pp. 508–515.
Kondo, Naoshi, Ting, K.C., 1998. Robotics for Bioproduction Systems.
Kondo, N., Shibano, Y., Mohri, K., Fujiura, T., Monta, M., 1992. Request to cultivation method from tomato harvesting robot. In: International Symposium on Transplant Production Systems, vol. 319. Yokohama, Japan, 21–26 July.
Kondo, Naoshi, Yamamoto, Kazuya, Yata, Koki, Kurita, Mitsutaka, 2008. A machine vision for tomato cluster harvesting robot. In: An ASABE Meeting Presentation, Paper Number: 084044.
Li, P., Lee, S., Hsu, Y.Y., 2011. Review on fruit harvesting method for potential use of automatic fruit harvesting systems. Proc. Eng. 23, 351–366.
Plebe, Alessio, Grasso, Giorgio, 2001. Localization of spherical fruits for robotic harvesting. Mach. Vis. Appl. 13, 70–79.
Safren, O., Alchanatis, V., Ostrovsky, V., Levi, O., 2007. Detection of green apples in hyperspectral images of apple-tree foliage using machine vision. Trans. ASABE 50 (6), 2303–2313.
Slaughter, David C., Harrell, Roy C., 1987. Color vision in robotic fruit harvesting. Trans. ASABE 30 (4), 1144–1148.
Song, B., Bursalioglu, O., Roy-Chowdhury, A., Tuncel, E., 2006. Towards a multiterminal video compression algorithm using epipolar geometry. In: Proc. IEEE Int. Conf. Acoust., Speech Signal Processing, Toulouse, France, pp. II-49–II-52.
Subrata, I.D.M., Fujiura, T.S., Nakao, S.J., et al., 1997. 3-D vision sensor for cherry tomato harvesting robot. Jpn. Agric. Res. Q. 31 (4).
Sun, M., Takahashi, T., Fukuchi, H., Zhang, S., 1994. Existing State and Problem of Development for Robotics of Fruit Harvesting. Tohoku Branch Report of the Japanese Society of Agricultural Machinery, No. 41, pp. 23–28 (in Japanese).
Tanigaki, K., Fujiura, T., Akase, A., Imagawa, J., 2008. Cherry-harvesting robot. Comput. Electron. Agric. 63 (1), 65–72.
Tarrío, Paula, Ana, M., Bernardos, José, R., 2006. A harvesting robot for small fruit in bunches based on 3-D stereoscopic vision. In: Computers in Agriculture and Natural Resources, 4th World Congress Conference, ASABE Publication No. 701P0606.
Ting, Yuan, Wei, Li, Yuzhi, Tan, et al., 2009. Information acquisition for cucumber harvesting robot in greenhouse. Trans. Chin. Soc. Agric. Machinery 40 (10), 151–155.
Van Henten, E.J., Hemming, J., Van Tuijl, B.A.J., et al., 2002. An autonomous robot for harvesting cucumbers in greenhouses. Autonomous Robots 13, 241–258.
Van Willigenburg, L.G., Hol, C.W.J., Van Henten, E.J., 2004. On-line near minimum-time path planning and control of an industrial robot picking fruits. Comput. Electron. Agric. 44, 223–237.
Yaqin, Wang, Hua, Gao, 2004. Study on the segmentation and orientation of fruit image under natural environment. Comput. Eng. 30 (13), 128–129, in Chinese.