



# Segmentation of foreground apple targets by fusing visual attention mechanism and growth rules of seed points

Weifeng Qu, Wenjing Shang, Yanhua Shao, Dandan Wang, Xiuli Yu and Huaibo Song

Northwest A&F University, College of Mechanical and Electronic Engineering, Yangling, Shaanxi 712100, China.

## Abstract

Accurate segmentation of apple targets is one of the most important problems to be solved in the vision system of apple picking robots. This work aimed to solve the difficulties that background targets often bring to foreground targets segmentation, by fusing the visual attention mechanism and the growth rule of seed points. Background targets could be eliminated by extracting the ROI (region of interest) of apple targets; the ROI was roughly segmented on the HSV color space, and then each of the pixels was used as a seed growing point. The growth rule of the seed points was adopted to obtain the whole area of apple targets from seed growing points. The proposed method was tested with 20 images captured in a natural scene, including 54 foreground apple targets and approximately 84 background apple targets. Experimental results showed that the proposed method can remove background targets and focus on foreground targets, while the k-means algorithm and the chromatic aberration algorithm cannot. Additionally, its average segmentation error rate was 13.23%, which is 2.71% higher than that of the k-means algorithm and 2.95% lower than that of the chromatic aberration algorithm. In conclusion, the proposed method contributes to the vision system of apple-picking robots to locate foreground apple targets quickly and accurately under a natural scene.

**Additional key words:** picking robots; *Malus domestica*; ROI; Itti model.

**Abbreviations used:** GBVS (graph-based visual saliency); HSV (Hue-saturation-value); RGB (red-green-blue); ROI (region of interest);

**Citation:** Qu, W. F.; Shang, W. J.; Shao, Y. H.; Wang, D. D.; Yu, X. L.; Song, H. B. (2015). Segmentation of foreground apple targets by fusing visual attention mechanism and growth rule of seed points. Spanish Journal of Agricultural Research, Volume 13, Issue 3, e0214, 9 pages. <http://dx.doi.org/10.5424/sjar/2015133-7047>.

**Received:** 02 Nov 2014. **Accepted:** 20 Jul 2015

**Copyright © 2015 INIA.** This is an open access article distributed under the Creative Commons Attribution License (CC by 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Funding:** This work was supported by the National High Technology Research and Development Program of China, 863 Program (No. 2013AA10230402); Natural Science Basic Research Plan in Shaanxi Province of China (No. 2014JQ3094); and the “Fundamental Research Funds for the Central Universities” of China (No. QN2011031).

**Competing interests:** The authors have declared that no competing interests exist.

**Correspondence** should be addressed to Huaibo Song: [songyangfeifei@163.com](mailto:songyangfeifei@163.com), or Weifeng Qu: [quweifeng@nwsuaf.edu.cn](mailto:quweifeng@nwsuaf.edu.cn) (shared corresponding authors)

## Introduction

Apple (*Malus domestica* Borkh.) picking is an extremely important aspect of the apple production chain. Harvest mechanization not only can achieve labor saving but also helps to improve the rate of high-quality fruit by reducing the damages caused by artificial factors. At present, however, the picking operation is still an artificial, seasonal and labor-intensive work. Along with the development of studies related to apple picking robots, the ability to identify apple targets quickly and accurately is affecting the working efficiency of picking robots (Tu *et al.*, 2010). In fact, the existing apple-picking robots pick one fruit at a time, and background targets not only cause difficulties for foreground target extraction but also seriously affect the efficiency of picking robots.

For the identification of fruits, in-depth studies on fruits, such as oranges, apples, tomatoes, strawberries and pineapples, have been performed by many scholars from different angles (Xu *et al.*, 2005; Zhou *et al.*, 2007; Jiang *et al.*, 2008; Li *et al.*, 2010; Peng *et al.*, 2011; Linker *et al.*, 2012). Zhang *et al.* (2008) used a neural network for apple image segmentation by selecting the color feature ( $R/B$  value) and texture features as input nodes. Wang *et al.* (2009) extracted the color features of apple color images and used the support vector machine method to identify apple targets. Si *et al.* (2009) identified red apple targets under different circumstances, such as suitable light and backlight, by combining the  $R-G$  and color ratio ( $R-G/G-B$ ). Tu *et al.* (2010) proposed an apple identification method based on an illumination invariant graph and

identified apple targets under different circumstances such as low light, weak light, bright light and strong light. The fruit recognition algorithm was developed by Zhao *et al.* (2011) to detect and locate apples in trees automatically by using a support vector machine with radial basis function. Ji *et al.* (2012) proposed an automatic recognition vision system guided for an apple harvesting robot. Lv *et al.* (2012) developed a recognition method for apple fruits in three different states: non-obscured, overlapped and severely obscured by branches and leaves. Based on the theory of convex hull, Song *et al.* (2012, 2013) identified and located shadowed apples by segmenting and reconstructing overlapping apple targets. Cheng & Shi (2013) segmented the apple image according to the sample color space they established for the vision system of an apple-picking robot. Wang *et al.* (2015) proposed the method of using the k-means clustering algorithm and convex hull theory to recognize and localize occluded apples. Si *et al.* (2015) proposed an apple location method in trees using stereoscopic vision.

Faced of a complex natural scene, the human visual attention system can quickly focus on a few salient visual objects and address them first. This process is called visual attention, and these salient visual objects are called ROI or “region of interest” (Jung & Kim, 2012). Accurate extraction of the ROI of images can improve the efficiency and accuracy of image processing and analysis. Image segmentation algorithms based on ROI cannot only solve the self-adaptive problems caused by excessive identification parameters but also eliminate the influence of small background targets, while the conventional method cannot. On the basis of the Itti model, Zhang & Wang (2005) and Zhang *et al.* (2009) satisfactorily used the detection method of saliency area or ROI of an image.

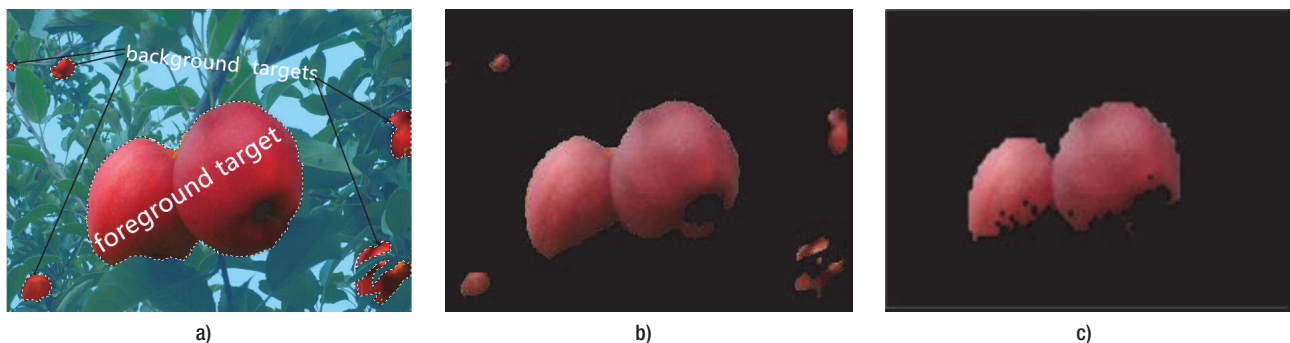
Although these methods have achieved good effects, excessive identification parameters, especially the segmentation threshold, have been introduced, making

it difficult to select parameters adaptively. Meanwhile, they almost overlooked the fact that background targets often cause difficulties for foreground target segmentation and that conventional identification algorithms cannot eliminate background targets. Moreover, the visual attention mechanism can quickly focus on a few salient visual objects and address them first. Therefore, the objectives of this study were to i) apply the visual attention mechanism and growth rule of seed points to detect the foreground apple objectives and overcome the effects caused by background apple objects, and to ii) evaluate the performance of the presented method and the commonly used k-means algorithm and the chromatic aberration algorithm for the detection of apples in natural conditions.

## Material and methods

A personal computer with a 2.60 GHz processor and 4.0 GB of RAM was used as the hardware part of the computer vision system, and all algorithms were developed in Matlab version R2013a software. A digital camera (Fuji film A900, CMOS color camera) was selected and the shooting distance was approximately 1.5 m. All of the images were acquired under natural daylight conditions in the RGB color model. The image frames were 1600×1200 pixels in the jpeg format.

Our research was focused on “Fuji” apples, which is the most popular apple species in China. The images used in the experiment were collected from September 2011 to September 2014 at the standard rootstocks density planting orchard of Northwest A&F University, Yangling, Shaanxi, China. In this study, 20 apple images, including 54 foreground apple targets and approximately 84 background apple targets, were selected to test the performance of the presented algorithm, chromatic aberration algorithm and k-means algorithm. Fig. 1 shows an example of background targets and foreground targets and their segmentation results.



**Figure 1.** Example of background targets and foreground targets and its segmentation results. a. Original image. b. Segmentation result of k-means. c. Segmentation result of the proposed method.

## Extracting ROI of apple images

Extracting ROI is one of the most important procedures for extracting apple targets. Itti *et al.* (1998), Zhang & Wang (2005) and Harel *et al.* (2006) proposed many ROI detection algorithms. According to the different methods of extracting ROI, the existing ROI detection algorithms could be divided into three categories: those based on interaction, those based on transformation and those based on visual features. Among them, the third type is currently the most popular. In particular, the Itti model and the graph-based visual saliency model are the most popular visual attention mechanism models for extracting ROI, and they are suitable for identifying a seed point for the growth rule of seed points.

### Itti visual attention mechanism model

The Itti model is one of the most classic visual attention models. Various characteristics of the input image, such as brightness, color, and direction, are extracted in this model, and the conspicuity maps of these characteristics are formed by the Gaussian pyramid and central-surround operator (Wang *et al.*, 2011). Then, a saliency map is obtained by normalization and fusion. On this basis, the most salient area wins out and the ROI is obtained through the winner-take-all mechanism of a neural network to attract attention focus in the internal saliency map. Finally, attention is turned to the next salient area by the returning inhibition mechanism to curb the current salient area (Rumelhart & Zipser, 1985).

The Itti model's concrete implementation steps are as follows:

(1) Visual preprocessing. Starting from the color values (red, green and blue) of the input image, an intensity image  $I=(r+g+b)/3$  is obtained. Color channels  $R=r-(g+b)/2$  for red,  $G=g-(r+b)/2$  for green, and  $B=b-(r+g)/2$  for blue are generated by each pixel in the pyramid. The detection of local orientation at each point in the image is achieved by over-complete steerable filters.

(2) Center-surround differences. Compute center-surround differences to determine contrast, by taking the difference between a fine (center) and a coarse scale (surround) for a given feature. This operation across spatial scales is performed by interpolation to the fine scale and then by point-by-point subtraction.

(3) Normalization. The values in the map are normalized to a fixed range  $[0...M]$ , to eliminate modality-dependent amplitude differences. The location of the map's global maximum  $M$  is found, and the average  $m$  of all its other local maxima is computed; the map is then multiplied by  $(M-m)^2$ .

(4) Conspicuity maps. The feature maps are combined into three conspicuity maps at the scale of 4. This is obtained through across-scale addition by reducing each map to the lowest resolution (scale 4) and by point-by-point addition.

(5) Saliency map. The three conspicuity maps are normalized and summed into the final input to the saliency map.

### Graph-based visual saliency model

The GBVS (graph-based visual saliency) model was proposed by Harel *et al.* (2006). It was based on the Itti model; the Markov chain of a 2D image was built by using the characteristics of the Markov random field, and the saliency map was obtained by calculating its equilibrium distribution (Dandapat *et al.*, 2004).

The GBVS model's concrete implementation steps are as follows:

(1) To obtain multi-scale brightness information, the input gray image  $I$  is filtered by the Gaussian pyramid low-pass filter. Each order of the Gaussian pyramid is a two-dimensional Gaussian low-pass filter, as shown in Eq. [1]:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad [1]$$

where  $(x, y)$  is the position of a pixel in the image and  $\sigma$  is the scale factor. The so-called pyramid is constantly sampling the original image by 1/2 and Gaussian low-pass filtering, in which  $\sigma$  decreases continuously with decreasing image size. Then the sets of the filtering results of different scales, which express brightness channels, are obtained.

(2) To obtain multi-scale direction information, the original gray image  $I$  is filtered by a set of Gabor pyramid filters. The process is shown in Eq. [2]:

$$H(x, y, \sigma, \theta) = \frac{1}{\sigma^2} \exp\left(-\pi \frac{x^2 + y^2}{\sigma^2}\right) \left[ \exp(i2\pi f(x \cos \theta + y \sin \theta)) - \exp\left(-\frac{\theta^2}{2}\right) \right] \quad [2]$$

where  $\sigma$  is the scale factor,  $f$  is the sine wave frequency,  $\theta$  is the direction angle and usually  $\theta = [0, \pi/4, \pi/2, 3\pi/4]$ , namely, the image is filtered in four directions. Thus, like the brightness information, four groups of filtering results of different scales, which express direction channels, are obtained.

(3) The Markov equilibrium distribution of images of different scales and different characteristics is cal-

culated. According to the differences between the pixels and Euclidean distances, the Markov chain of the filtering result of each scale within each channel is established, and then the Markov equilibrium distribution can be calculated.

(4) The saliency map is obtained. The comprehensive saliency map, whose size is equal to the original image, is obtained by adding the Markov equilibrium distributions of a channel respectively, and normalizing the previous result.

After comparing of the segmentation results between the Itti model and GBVS model, Fig. 2 was obtained, which showed the segmentation effect by both the Itti and GBVS models for apple images: (i) when extracting the ROI of the image of a single apple, both the Itti model and the GBVS model had good effects; (ii) when extracting the ROI of the image of a few apples, the

Itti model could include more targets, while the GBVS model was concentrated in a small area.

Considering that apple images captured under the natural scene contain many targets, the Itti model was selected to extract the ROI of apple images. Moreover, the extracted target area was still not complete. Therefore, in this study, the ROI was segmented again, and the segmentation result was regarded as a seed point to obtain the whole apple area by the growth rule of seed points.

### Apple images segmented based on the growth rule of seed points

The growth rule of seed points is that pixels of a similar quality are made to constitute an area (Deng & Manjunath, 2001). First, a seed point will be selected as



**Figure 2.** Original digital images (a1, a2) captured in natural daylight condition, ROIs (b1, b2) of the two original images extracted by the Itti model, and ROIs (c1, c2) of the two original images extracted by GBVS model.

the starting point of growth in each area before segmentation. Then, the pixels will be sought at the seed area whose qualitative similarity meets the rules of growth, and these pixels will merge with the seed area. Thus, these merged pixels become a new seed, and continue seeking and merging until there is no pixel to be merged. Qualitative similarity is mostly color similarity in common image segmentation, but it has to meet not only the color similarity but also visual attention conditions in the segmentation of the ROI. We introduced the growth rule of seed points to achieve the whole apple area. At first, the ROI of each original apple image was extracted using the Itti model; then, the seed point was extracted and converted into binary. Finally, the whole apple areas were obtained via the growth rule of seed points.

### Obtaining seed areas

The apple target saliency map obtained from the apple target via the Itti model was not complete. On the basis of the previous results, an apple image was segmented preliminarily in the HSV color space, and then the segmentation result was taken as a seed point for growth.

Given that there is an obvious color difference between a ripe apple and the background, as the color of the former is red, we used a super red image ( $k=2$ ) to remove the background and cause the apple target to stand out, which was achieved by Eq. [3]:

$$Sub = k * (R - G) \quad [3]$$

where *Sub* represents the result; *R-G* represents red and green components of the image, respectively. When parameter *k* is 2, the processing effect is best, as is shown in Fig. 3a. Then, the image was changed into a binary image (Fig. 3b), which is regarded as a seed area. Because the seed image is not completed, it should be further processed by the growth rule of seed points to get the whole apple area.

### Growth rule of seed points

Suppose that the interested object *O* grows from a seed point *R*. The seed point *R* that is identified as the point of interest is called the marked point, while the point that does not belong to any area is called the unmarked point. Thus, the area growth procedure of object *O* is described as follows: the unmarked point next to *O* will be merged into object *O* after the testing of qualitative similarity criteria. The point set belonging to *O* can be defined as Eq. [4]:

$$H = \{(x, y) \notin O \mid N(x, y) \cap O \neq \emptyset\} \quad [4]$$

where  $N(x, y)$  is a  $3 \times 3$  small area whose center is point  $(x, y)$ .

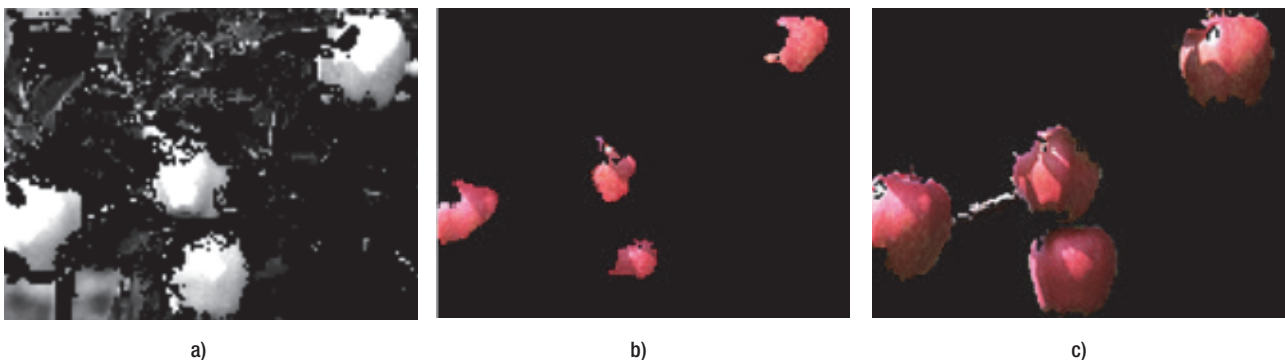
Supposing pixel *P*,  $P \in H$ , and its visual attention degree *S*, the corresponding relative position indication is *PSD*. Qualitative similarity testing that decides whether *P* merges into *O* is defined by Eq. [5]:

$$ST(P \rightarrow O) = \begin{cases} true & \text{if } S > T_s \text{ and } PSD < T \\ false & \text{otherwise} \end{cases} \quad [5]$$

After testing, if pixel *P* meets the above conditions, *P* will be merged into object *O*. Both  $T_s$  and  $T_{PSD}$  are thresholds. The essence of qualitative similarity testing is that the point, which has a high attention degree and does not belong to an edge, belongs to the interested object.

### Extract the whole apple area by the growth rule of seed points

The HSV color model has linearity and flexibility, which is aligned with human eye perception features. HSV is a transformation of an RGB color space, and its components and colorimeter are relative to the RGB color space from which it is derived.



**Figure 3.** The results of image processing. a: enhanced image of ROIs; b: seed image; c: final result.

Because  $H$  could make the apple target more striking, it becomes the qualitative similarity condition to control the seed image growth in the HSV color space. Fig. 3c shows that combining the visual attention mechanism with the growth rule of seed points in a certain color space can realize accurate segmentation of the apple targets.

## Results

To verify the effectiveness of the proposed method, 20 images were selected to conduct the experiment. The main procedures of the experiment were as follows:

- *Step 1.* Load an original image and extract ROI of the image via the Itti model.
- *Step 2.* Remove the background and enhance the ROI by means of Eq. [3].
- *Step 3.* Binarize the result of step 2, and extract its seed points.
- *Step 4.* Transform the result of step 3 from the RGB color space into the HSV color space, then process the seed image by the given growth rule of seed points, and finally obtain the whole apple area.
- *Step 5.* Using the k-means algorithm, segment apple targets from the 20 original images captured from a natural scene.
- *Step 6.* Using chromatic aberration algorithms, segment apple targets from the 20 original images captured from a natural scene.

Fig. 4 shows six images and their corresponding segmentation results obtained after the experiment using the k-means algorithm, the chromatic aberration algorithm and the proposed method.

To verify the efficiency and accuracy of the proposed method, the area of 20 original apple images and its corresponding processing results obtained using the k-means algorithm, chromatic aberration algorithm and the proposed method were calculated, respectively. Moreover, the segmentation effectiveness criterion was proposed, namely, the segmentation error rate of an apple, which is designated as  $\sigma$  and defined by Eq. [6]:

$$\sigma = |S - S_i| / S \times 100\% \quad (i = 1, 2, 3) \quad [6]$$

where  $S$  is the real apple area and  $S_1$ ,  $S_2$  and  $S_3$  are the apple areas segmented by the k-means algorithm, the chromatic aberration algorithm and the proposed method, respectively.

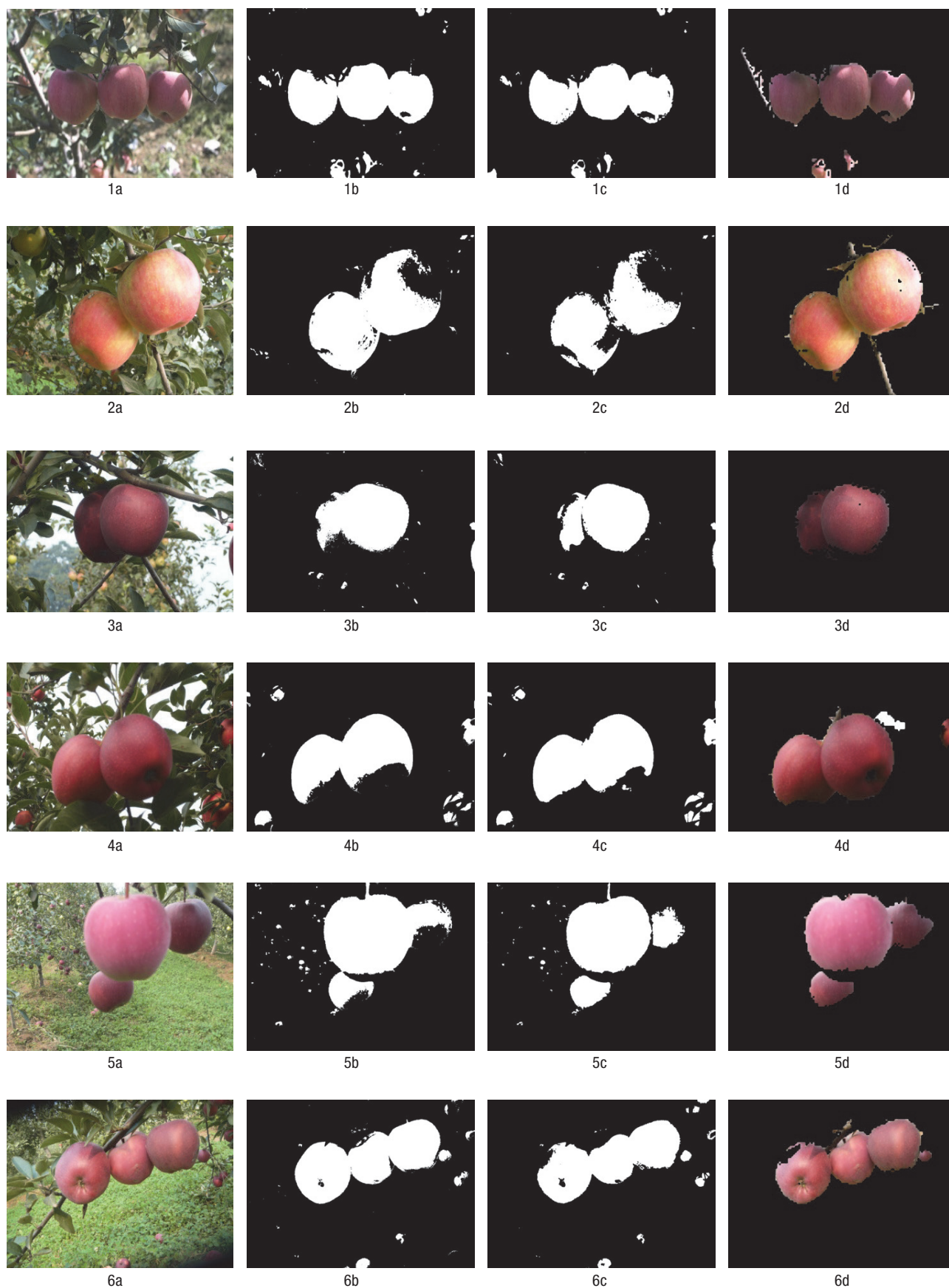
From Figs. 4-2b, 4-3b, and 4-4b, it can be seen that the k-means algorithm was highly affected by the strong light and shadows. From Figs. 4-2c, 4-3c and 4-5c, it can be seen that the chromatic aberration algorithm was highly affected by the strong light and shadows. From Figs. 4-1d and 4-2d, it can be seen that the proposed algorithm was influenced by the branches.

Additionally, by comparing Figs. 4 1b-1d, 3b-3d, 4b-4d, 5b-5d and 6b-6d, we can see that there were many small background apple targets in the segmented images made by the k-means and chromatic aberration algorithms, while there were none in the images segmented by the proposed method, which suggests that the k-means and chromatic aberration algorithms could not remove small background apple targets and that the vision attention system model only focuses on the ROI of the target and could remove small background targets. Thus, the proposed method is conducive to the vision system of an apple-picking robot to locate apple targets quickly and accurately.

All of the calculation results of the 20 images and their corresponding processing results among the k-means algorithm, the chromatic aberration algorithm and the proposed method are shown in Table 1. After analyzing the segmentation error rate in Table 1, we found that the highest segmentation error rate and the average segmentation error rate were 42.97% and 10.52%, 42.68% and 16.81%, and 27.71% and 13.23%, respectively, for the chromatic aberration algorithm, the k-means algorithm and the proposed method by us. It could be concluded that the highest segmentation error rate of the proposed method is the lowest among these three methods and that the average segmentation error rate of the proposed method was 2.71% higher than that of the k-means algorithm and 2.95% lower than that of the chromatic aberration algorithm. Meanwhile, the proposed method requires less time than the k-means algorithm to process images, which conduces to picking robots to pick apples quickly.

## Discussion

For the identification of apples, many in-depth studies have been performed (Zhang *et al.*, 2008; Si *et al.*, 2009; Tu *et al.*, 2010; Si *et al.*, 2015), however, which merely pursue the identification numbers including all the apples in the machine vision, while those studies do not consider the fact that apple-picking robots pick one fruit at a time, and overlook the fact that background targets not only can cause difficulties for foreground target extraction but also seriously affect the efficiency of foreground apple's localization for picking robots. Meanwhile, some studies introduce the neural network



**Figure 4.** Six examples of original images captured in natural scenes. a: original images of apple targets captured from the natural scene; b: results of the k-means algorithm; c: results of the chromatic aberration algorithm; d: results of the proposed method.

**Table 1.** Calculation results of the 20 images and the corresponding processing results among the k-means algorithm, chromatic aberration algorithm and the proposed method

Apple test number	Real apple area, $S$	Segmentation area, $S_i$			Segmentation error rate $\sigma$ (%)		
		k-means method	Chromatic aberration	Text method	k-means method	Chromatic aberration	Text method
1	150122	140689	135526	155385	6.28	9.72	3.51
2	103640	100449	98181	120781	3.08	5.27	16.54
3	272465	247559	232639	301761	9.14	14.62	10.75
4	157819	89999	154572	180130	42.97	2.06	14.14
5	104227	96295	92862	88626	7.61	10.90	14.97
6	216041	207762	204976	192099	3.83	5.12	11.08
7	191978	170075	148065	244404	11.41	22.87	27.31
8	198745	161835	137991	231220	18.57	30.57	16.34
9	299653	278743	235823	235644	6.98	21.30	21.36
10	269652	254006	220357	237119	5.80	18.28	12.06
11	130368	105785	98431	116540	18.86	24.50	10.61
12	249391	227827	211069	283954	8.65	15.37	13.86
13	320049	303947	264267	231355	5.03	17.43	27.71
14	179672	157303	102995	140811	12.45	42.68	21.63
15	192223	161006	156274	190953	16.24	18.70	0.66
16	198011	178839	152358	177829	9.68	23.06	10.19
17	273715	260991	252951	280402	4.65	07.59	2.44
18	216541	203046	189003	200381	6.23	12.72	7.46
19	154626	146176	136714	153099	5.46	11.58	0.99
20	148462	137476	134629	179524	7.40	9.32	20.92
Average					10.52	16.18	13.23

algorithm (Zhang *et al.*, 2008) and support vector machine (SVM) algorithm (Wang *et al.*, 2009), which not only increase the identification of parameters but also increase the time consumption to process these images.

The results from this study demonstrate that fusion of the visual attention mechanism and the growth rule of seed points applied to apple images in natural conditions is an effective tool for detecting foreground apple objects. This study further confirms that the presented method could ignore background apples accurately and sufficiently, while the commonly used methods could not.

By considering the colors, shapes, feature points, textures and so on, the commonly used image segmentation methods always use a threshold to realize the detection of apples, and image clustering algorithms need to know the clustering numbers also. For practical applications, none or a few input parameters is preferable; by applying the visual attention mechanism and growth rule of seed points, the presented method is a non-paramour input required algorithm and can run without human intervention.

Fusing the visual attention mechanism and growth rule of seed points could reduce the influence of the artificial threshold setting without introducing many parameters, which could also help to remove background apple targets, overcome the influences of small area targets in the background and improve the running speed of the algorithm simultaneously. It is conducive to the vision system of apple-picking robots to locate foreground apple targets quickly and accurately within the natural scenes.

## References

- Cheng X, Shi X, 2013. Target extraction study on the vision system of apple picking robot. Proc Chinese Intelligent Automation Conference. Springer Berlin Heidelberg, pp: 45-52. [http://dx.doi.org/10.1007/978-3-642-38466-0\\_6](http://dx.doi.org/10.1007/978-3-642-38466-0_6)
- Dandapat S, Chutatape O, Krishnan SM, 2004. Perceptual model based data embedding in medical images. Image Processing, Int Conf IEEE ICIP'04, Vol 4, pp: 2315-2318. <http://dx.doi.org/10.1109/icip.2004.1421563>



- Deng Y, Manjunath BS, 2001. Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(8): 800-810. <http://dx.doi.org/10.1109/34.946985>
- Harel J, Koch C, Perona P, 2006. Graph-based visual saliency. *Advances in Neural Information Processing Systems*, Vancouver, BC, Canada. pp: 545-552. Available in: [http://machinelearning.wustl.edu/mlpapers/paper\\_files/NIPS2006\\_897.pdf](http://machinelearning.wustl.edu/mlpapers/paper_files/NIPS2006_897.pdf).
- Itti L, Koch C, Niebur E, 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence* (11): 1254-1259. <http://dx.doi.org/10.1109/34.730558>
- Ji W, Zhao D, Cheng F, Xu B, Zhang Y, Wang J, 2012. Automatic recognition vision system guided for apple harvesting robot. *Comput Electr Eng* 38(5): 1186-1195. <http://dx.doi.org/10.1016/j.compeleceng.2011.11.005>
- Jiang H, Peng Y, Chen C, Ying Y, 2008. Recognizing and locating ripe tomatoes based on binocular stereovision technology. *T CSAE* 24(8): 279-283.
- Jung C, Kim C, 2012. A unified spectral-domain approach for saliency detection and its application to automatic object segmentation. *IEEE Transactions on Image Processing* 21(3): 1272-1283. <http://dx.doi.org/10.1109/TIP.2011.2164420>
- Li B, Wang N, Wang M, Li L, 2010. In-field pineapple recognition based on monocular vision. *T CSAE* 26(10): 345-349.
- Linker R, Cohen O, Naor A, 2012. Determination of the number of green apples in RGB images recorded in orchards. *Comput Electron Agr* 81: 45-57. <http://dx.doi.org/10.1016/j.compag.2011.11.007>
- Lv J, Ji W, Chen F, Zhao D, Xu B, 2012. Research on the recognition method for obscured apple in natural environment. *IEEE Control Conference (CCC)*, 31st Chinese. pp: 3932-3937. Available in: [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6390613&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs\\_all.jsp%3Farnumber%3D6390613](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6390613&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D6390613).
- Peng H, Wen Y, Wu L, 2011. Citrus edge detection based on self-adaptive canny operator. *Comput Eng Appl* 47(9): 163-166.
- Rumelhart DE, Zipser D, 1985. Feature discovery by competitive learning. *Cognitive Science* 9(1): 75-112. [http://dx.doi.org/10.1207/s15516709cog0901\\_5](http://dx.doi.org/10.1207/s15516709cog0901_5)
- Si Y, Qiao J, Liu G, Liu Z, Gao R, 2009. Recognition and shape features extraction of apples based on machine vision. *T CSAE* 40: 161-165, 73.
- Si Y, Liu G, Feng J, 2015. Location of apples in trees using stereoscopic vision. *Comput Electron Agr* 112: 68-74. <http://dx.doi.org/10.1016/j.compag.2015.01.010>
- Song H, He D, Pan J, 2012. Recognition and localization methods of occluded apples based on convex hull theory. *T CSAE* 28(22): 174-180.
- Song H, Zhang C, Pan J, Yin, X, Zhuang Y, 2013. Segmentation and reconstruction of overlapped apple images based on convex hull. *T CSAE* 29(3): 163-168.
- Tu J, Liu C, Li Y, Zhou J, Yuan J, 2010. Apple recognition method based on illumination invariant graph. *T CSAE* 26 (Suppl. 2): 26-31.
- Wang J, Zhao D, Ji W, Zhang C, 2009. Apple fruit recognition based on support vector machine using in harvesting robot. *T CSAM* 40(1): 148-151.
- Wang X, Wang B, Zhang L, 2011. Airport detection in remote sensing images based on visual attention. *Neural Information Processing*. Springer Berlin Heidelberg, pp: 475-484. [http://dx.doi.org/10.1007/978-3-642-24965-5\\_54](http://dx.doi.org/10.1007/978-3-642-24965-5_54)
- Wang D, Song H, Tie Z, Zhang W, He D, 2015. Recognition and localization of occluded apples using K-means clustering algorithm and convex hull theory: a comparison. *Multimedia Tools and Applications*, 1-22. <http://dx.doi.org/10.1007/s11042-014-2429-9>
- Xu H, Ye Z, Ying Y, 2005. Identification of citrus fruit in a tree canopy using color information. *T CSAE* 5: 023.
- Zhang P, Wang R, 2005. A survey of detecting regions of interest in a static image. *J Image Graph* 10(2): 142-148. Available in: <http://www.oalib.com/paper/1639359>.
- Zhang Y, Li M, Qiao J, Liu G, 2008. Segmentation algorithm for apple recognition using image features and artificial neural network. *Acta Optica Sinica* 28(11): 2104-2108. <http://dx.doi.org/10.3788/AOS20082811.2104>
- Zhang J, Shen L, Gao J, 2009. Region of interest detection based on visual attention model and evolutionary programming. *J Electron Inf Technol* 31(7): 1646-1652.
- Zhao D, Lv J, Ji W, Zhang Y, Chen Y, 2011. Design and control of an apple harvesting robot. *Biosyst Eng* 110(2): 112-122. <http://dx.doi.org/10.1016/j.biosystemseng.2011.07.005>
- Zhou T, Zhang T, Yang L, Zhao J, 2007. Comparison of two algorithms based on mathematical morphology for segmentation of touching strawberry fruits. *T CSAE* 23(9): 164-168.