Object Recognition on Cotton Harvesting Robot Using Human Visual System

Yong Wang, Xiaorong Zhu, Yongxing Jia, Changying Ji

Object Recognition on Cotton Harvesting Robot Using Human Visual System

Yong Wang[1]   Xiaorong Zhu [2]   Yongxing Jia [1]   Changying Ji [3*]

[1] College of Science, PLA University of Science and Technology, 210003, Nanjing, China

[2] College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, 210003, Nanjing, China

[3] College of Engineering, Nanjing Agricultural University, 210031, Nanjing, China

{Yong Wang:njwy1978@126.com; Changying Ji:chyji@njau.edu.cn}

Abstract. Object recognition is one of the hottest issues in the field of vision system for harvesting robot. How efficiently and accurately to remove the background and get the object in image is the key research. The attention mechanisms of human visual system (HVS) can be segmented an image into the region of interesting (ROI) which is considered important and the background which is less important, and recognized the object from ROI using the local information. In this paper, an algorithm based on the characteristic of HVS is proposed. In algorithm, the image was partitioned into many blocks of equal size. ROI was got through calculating the factor of weight of each sub-block image, and the object was extracted by segmenting the ROI. Experiment results show that the algorithm can be recognized the object efficiently and accurately. A new method for vision system of harvesting robot is provided.

Keywords: attention mechanisms; human visual system; cotton; object recognition

1   Introduction

Two main tasks of harvesting robot are to recognize and locate the object and to harvest the object without damaging. The first task recognize and locate the object is the key task. This task determined how the harvesting task can implement smoothly. The principal objective in the first task is to extract the object which was interesting from the image, in order to prepare for subsequent image processing steps. Nowadays, there are many methods for object reorganization, such as using the color difference between object and background. Some researchers use color difference (R-G) and color difference ratio (R-G)/(G-B) to recognize the apples[1-3]. The experiment results showed this method can eliminate the inferences of shade, backlighting and soil. The recognition rate reaches above 90%. Others use the spectral reflection ration of different parts of the crop[4,5]. These methods have a good effect for recognizing object. However, there is no considering the biological visual characteristics, especial the characteristics of human vision, to recognize the object. The human eyes were formed a unique visual characters in the long evolutionary history. The human eyes can recognize the target accurately and quickly in the complex environment. So many researchers are interesting in the biological vision. Liu Changqin[6] present a model of motion direction detection based on the characteristics of sustained-transient cell of biological vision, and simulated on the computer. The experiment results show this model can recognize the motion direction of object. Song Nong[7] developed an auto target recognition algorithm employing attention mechanisms with bottom-up and top-down control strategies. The experiment results show the objects for detection using the algorithm is effectively. Other researchers[8-13] present their own algorithms for image process based on the research of biological vision. Therefore, the bionic vision will be a hottest field for the future research[14-15]. This paper developed an algorithm for cotton recognition using the characteristic of attention mechanisms of HVS.

The rest of the paper is organized as follows. In Section II we introduce the materials of experiment throughout the paper. In Section III an algorithm using four factors of attention mechanisms of HVS can be found. We present some experiment results in Section IV. Finally, we conclude the paper in Section V.

2   Materials

This subject is Suman-12 which was in mature period Jiangpu farm of Nanjing Agricultural University. The color images in the outside were acquired by a digital color CCD camera WAT-231S in October. The image acquisition card is Matrox Meter-II/Standards. The color signals from WAT-231s were transferred by the image acquisition card as a 24-bit RGB color image and processed using Lenovo X200 Computer, the CPU is Intel Core2 Duo 2.4GHz, the memory is 2GB, and OS is Windows XP. The color images were taken at the time from 10:00 AM to 4:00 PM. These images were taken under sunny and cloudy weather. So the machine vision system is composed of a color CCD camera to acquire the cotton image, the image acquired card to transferred the analog signal to digital signal, and the PC to process the acquired image. The machine vision is shown as Figure 1.
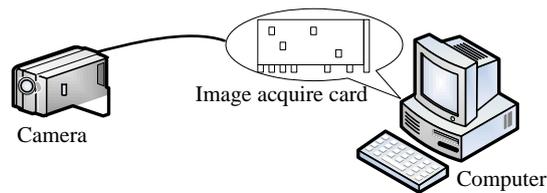

Image acquire card

Camera

Computer

Fig. 1. composition of machine vision

3   Principal of Human Visual System

We are interesting of the rapid change of image when we observed. It is so called the attention mechanisms. The rapid change area is called the region of interesting (ROI). To extract the ROI, the main factor affected human eyes are to be considered, such as the impact factor of brightness, contrast, sharpness and location. To calculate the effect of four factors mentioned above. First the image was blocked sub-image by equal space. Set the size of each block is m and n, and then calculated the impact factor of each block, obtained the weight factor ω of each blocks through normalization. Finally, the background was removed and ROI was remained based on the value of weight factor of ω .

3.1 Brightness Factor

Brightness factor describes the sensitivity level of human eyes to light. It is measured through Weber-Fechner law that is the relationship between the physical magnitude of stimuli and the percered intensity of the stimuli is the logarithm. This kind of relationship can be expressed as Equation (1):

$$s = k \, lg \frac{I}{I_0} \quad . \tag{1}$$

where k is the constant factor, I is the mean brightness of each block, I_0 is the absolute threshold, generally taken the maximum brightness value of each block.

3.2 Contrast Factor

Contrast factor describes the sensitivity level of human eyes to changed detail in image which can be expressed by variance. It can be expressed as Equation (2):

$$d = \sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} f^2(i, j) - \left[\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} f(i, j) \right]^2} \quad . \tag{2}$$

where $f(i, j)$ is the gray value of pixel(i, j).

3.3 Sharpness Factor

Sharpness factor describes the resolution of human vision. It can be expressed through the gradient. It can be written as Equation (3):

$$k = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \sqrt{\nabla^2 f_H(i, j) + \nabla^2 f_V(i, j)}}{m \times n} \quad . \tag{3}$$

where $f_H(i, j)$ and $f_V(i, j)$ is the horizontal and vertical gradient, respectively. Combining the horizontal and vertical gradient yields the spatial gradient.

3.4 Location Factor

Location factor describes the degree of interest in the different parts of image. It can be express as Equation (4):

$$r = 1 - (1 - B) \frac{\sqrt{(x_{i0} - x_c)^2 + (y_{i0} - y_c)^2}}{r_{max}} \qquad (4)$$

where B is the basic weight value, the range value is [0, 0.5], x_c, y_c is, respectively, the center coordinate of original image, x_{c0}, y_{c0} is the center coordinate of each block image, respectively, and r_{max} is the maximum distance from each point to the center coordinate of original image.

3.5   Weight Factor

The four factors of each sub-block were considered, ω_i was used to evaluate the importance of each sub-block. ω_i is expressed as Equation (5):

$$\omega_i = \sqrt{s^2 \times d^2 \times k^2 \times r^2} \qquad (5)$$

In order to decrease the influence of the absolute error of ω_i, the ω_i was progressed by normalized. The weight vector of ω is expressed as Equation (6):

$$\omega = \frac{\omega_i}{\sum_{i=1}^{N} \omega_i} \qquad (6)$$

where N is the number of block.

Now, the importance of each block can be expressed by the vector of ω. It is matched with the characteristic of human eyes. We call it as weight factor. If the value of ω is smaller than threshold, then the gray value was become zero, otherwise, the gray value is no changed. Then ROI can be extract through all the vector ω. The object of interesting can be got through segmenting ROI. The formula is expressed as Equation (7):

$$I = \begin{cases} I_i, & \omega > T \\ 0, & \omega < T \end{cases} \quad i = 1, 2, \cdots N \qquad (7)$$

where T is the threshold, N is the number of blocks.

Figure 2 shows the cotton recognition algorithm which was mentioned above. The process starts with the factor functions for segmentation. After segmenting the image, a low pass filter was conducted to remove noise. Then the features of object were

labeled, such as area, perimeter, etc, from the binary image. These features were used to locate the cotton in 3-dimension space.
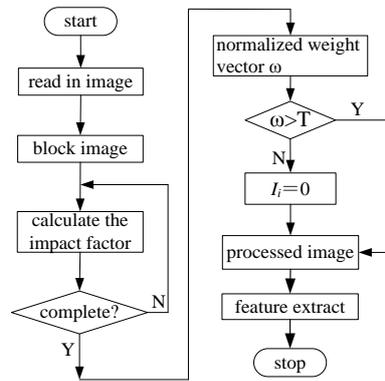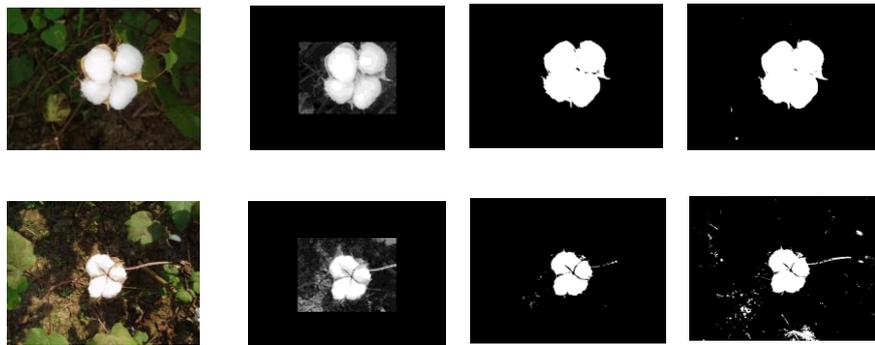

Fig. 2. cotton recognition algorithm

4   Experiments and Analysis

To check how effective the proposed method is extracting object in actual image, several experiments were conducted using sample color images under natural environment. Experiment results are shown in Figure 3.

(a)Original image   (b)ROI image (c)recognition with ROI(d)recognition without ROI

Fig. 3. experimental results

It can be seen that almost all the interesting object could be recognized by the proposed method. However, the number of sub-block images is changed by manually.

Under cloudy weather, the segmentation achieved good results when the number of sub-block images is 4×4. The object segmentation will be failure when the number of sub-block image is too many, for extraction of ROI is incomplete. Otherwise, the segmentation will be failure because too many background joining in ROI. Therefore, the number of sub-block will be changed interactive in order to achieve good results.

It can be seen from the experimental results. The segmentation results under cloudy weather are better than sunny weather. The brightness of leaves and ground closed to the cotton fruit under sunny weather that lead to mistake between leaves and fruit. There is spot noise after segmentation. It needs further process to remove the noise.

5   Conclusions

We proposed an algorithm to detect the object. The algorithm is based on the attention mechanism of human vision so that some experiments were conducted to investigate how effective and accurate it is. The experiment results show that the algorithm is effective and accurately at extracting the object present in complex scene. However, the accurate of detection is better under cloudy weather than sunny weather. Under the sunny weather, there is spot noise after the recognition of ROI. It needs further process to extract the object.

Reference

1. Si Yongsheng, Qiao Jun, Liu Gang, et al. Recognition and shape feature extraction of apples based on machine vision[J]. Transaction of the Chinese Society for Agricultural Machinery, 2009, 40(8):161-165.
2. Yang Xining, Duan Jianmin, Gao Deyi, et al. Research on Lane Detection Based on Improved Hough Transform[J].Computer Measurement and Control, 2010, 18(2):292-295.
3. Wang Yong, Shen Mingxia, Ji Changying. Using Color Data and Shape properties for Cotton Fruit Recognition[J]. Transaction of the Chinese Society for Agricultural Machinery, 2007, 38(2):77-79.
4. Yuan Ting, Zhang Junxiong, Li Wei, et al. Feature Acquisition of Cucumber Fruit in Unstructured Environment Using Machine Vision[J]. Transaction of the Chinese Society for Agricultural Machinery, 2009, 40(8):170-174.

5. Henten E. J., Hemming J., Tuijl B. J., et al. An autonomous robot for harvesting cucumbers in greenhouse[J]. Autonomous Robots, 2002, 13(3):241-258.

6. Liu Changqin, Wang Zhengzhi. A model of Motion Direction Detection based on Biological Vision[J]. Computer Simulation, 2006,23(8):181-183.

7. Sang Nong, Li Zhenlong, Zhang Tianxu. Applications of human visual attention mechanisms in object detection[J]. Infrared and Laser Engineering, 2004, 33(1):38-42.

8. Kang Mu, Li Yongliang. An Adaptive Image Enhancement Algorithm based on Human Visual properties [J]. Opto-Electronic Engineering, 2009, 36(7):71-77.

9. Kang Mu, Wang Baoshu. An Adaptive Color Image Enhancement Algorithm Based on Human Visual Properties[J]. Acta optical Sinica, 2009, 29(11):3018-3024.

10. Liu Libo, Zhou Guomin. Application of human visual system in extraction of plant leaf image[J]. Computer Engineering and Applications, 2009, 45(19):22-26.

11. Saghri J.A., Cheatham P.S., Habibi H. Image quality measure based on a human visual system model[J]. Optical Engineering, 1989, 28(7):813-818.

12. Cao Jian, Wang Wujun, Han Fei, et al. Research on Target Recognition Technology based on Local Feature[J]. Computer Engineering, 2010, 36(10): 203-205.

13. Yang Wei, Zhao Yan, Xu Dong. Method of image quality assessment based on human visual system and structural similarity[J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(1):1-4.

14. Liu Wei, Yuan Xiugan. Study on Biological Vision[J]. China Safety Science Journal，2000,10(6):51-56.

15. Yao Xingzhong, Lu Tongwei, Hu Hanping. Object Recognition Models Based on Primitive Visual Cortices: A Review[J]. Pattern Recognition and Artificial Intelligence，2001，22(4)：581-588.