

Apple Detection for Harvesting Robot Using Computer Vision

^{*1}Abhipray Paturkar, ²Gourab Sen Gupta, ³Donald Bailey

School of Engineering and Advanced Technology, Massey University, New Zealand

Email: A.Paturkar@massey.ac.nz

Received: 20th September 2018, Accepted: 11th October 2018, Published: 31st October 2018

Abstract

Development of fruit detection has always been a challenge for a harvesting robot. Detection of fruit in an unregulated environment with change in illuminations affects the harvesting robot's efficiency. This paper presents the improvement of robotic apple harvesting by enhancing apple detection and visibility by handling occlusion in outdoor conditions. For detection of the fruit, we developed an image processing method that is robust to change in illumination and partial occlusion of the apples; we used a stereo camera pair for image acquisition. To detect the fruits efficiently, global thresholding, based on chromaticity approach, is used. The results demonstrated that fruit detection achieved accuracy of 95% in sunny condition and handled partial occlusion of the fruit. This study will be helpful for detecting apples for harvesting robots.

Keywords

Computer Vision, Stereo Vision, Harvesting Robot, Apple Detection, Occlusion Handling

Introduction

Advancement in precision agriculture has improved the quality of the products, productivity and yield of the field and at the same time reduced production, and labor costs [1]. Regardless of these advancements, many of the agricultural tasks are still being done manually, which are tedious, difficult and monotonous operations but are still complicated for robots to do.

An important part in successful harvesting robots is their capability to process the sensory data, specifically, their ability to evaluate and understand video input. Practically, combination of video input and smart computer vision algorithms might help to perform many other operations and improve harvesting robots considerably. Nonetheless, the issues related to computer vision in outdoor conditions are numerous such as immensely unstructured conditions with large degree of ambiguity; fruits with different shapes, colours, textures, reflectance properties and sizes and serious occlusions, are the part of the issues that computer vision systems might face. These computer vision systems are yet to reach full potential to implement solutions at practical level [1]. These systems have been used for numerous applications in agricultural fields e.g. estimation of yield, obstacle avoidance and autonomous navigation, spraying, planting, and quality evaluation. Still, possibly the most popular application has been detection of fruit, which has the goals:

1. To detect the fruits by differentiating them from the background
2. To locate them accurately

These detected fruits are then used to help the interplay of the fruit with the end effectors and manipulators for more physical processing and harvesting operations. In spite of enough attention given to these tasks in the last three decades, not a single harvesting robot has been implemented at a commercial level. The failure of harvesting robots at commercial level is because of lack of success of computer vision system to handle issues due to unstructured environment. It is obvious that efficient computer vision system is important for achieving excellent detection rates of fruits in real-time.

Generally, a stereo camera pair provides more information than a single camera. It gives depth information based on triangulation principle and the disparity between two pixels in corresponding images. The use of stereo camera pair in harvesting robots has increased greatly in the last few years [2-8]. Xiang et al. [9] scrutinized detection of clustered tomatoes on trees using stereo vision. In this vision system, depth data of clustered tomatoes was achieved. In clusters, majority of the back tomatoes were occluded by front tomatoes and front tomatoes were preventing the end effector from reaching the tomatoes at the back due to false detection of tomatoes in the cluster. Noise present in the depth map was an issue as many pixels with strange depth values due to false matching leads to incorrect detection. Even though this approach is useful for identification of clustered tomatoes, the accuracy of this approach is not enough if there is serious occlusion.

Si et al. [10] presented a system which is capable of automatically locating apples on trees using a stereo vision system. Shape and colour based study are two main methods for fruit identification. Analysis based on colour is inefficient if there is a shadow present in the scene. Edge detection should be done ahead of feature extraction; once an apple is recognised then the branches and the leaves are identified to remove unwanted noise pixels. Feature extraction of the apple commonly depends on the contour. A Hough transform is mostly used for this but it is computationally expensive. The accuracy of the presented vision system was not adequate and thus there is a further space for enhancement.

Furthermore, there are some issues that these researchers did not solve efficiently, which are divided into two types: occlusion handling and change in illumination [1]. Resolving these issues is very essential for the success of harvesting robots. The goals of this research are:

1. To detect the apples under change in illumination scenario
2. To handle the occlusion

Materials and Methods

Image Acquisition: The stereo camera used for this experiment is Kinect Xbox360 a gaming webcam capable of capturing standard video. It has an RGB camera that stores three channel data with a 1280x960 resolution making it possible to capture colour images.

The images were captured of the apples hanging on the tree in outdoor conditions. The different lighting conditions considered were:

1. Sunny condition
2. Apple in the shade
3. Cloudy condition

90 images were captured (30 in each lighting condition).

Test Computer: The computer used is a HP personal computer with Intel Core i5 processor with a clock speed of 3.20 GHz with 8GB memory. It is running on a 64-bit Windows 10 Operating System. Camera is connected to the computer via USB 2.0.

Development of Apple Detection Algorithm:

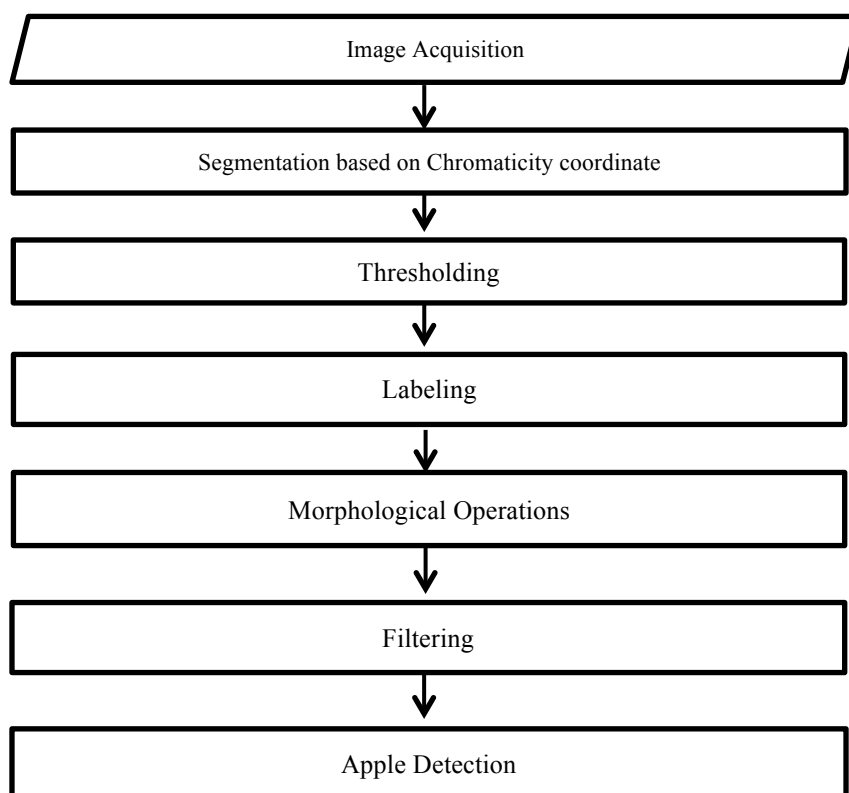


Figure 1: Apple Detection Algorithm

Interpreting the colour models of images is a fundamental process of the detection part. Generally, three colour models are used, chromaticity, RGB model, difference in colour and luminance approach [11]. In these models, the leaf, the branch, and the fruit have specific colours. This shows that

the fruits can be discriminated from the background using these colour features. The results suggested that the difference in colour and luminance, and RGB model approaches were affected by the change in illumination condition. The chromaticity approach was barely affected by change in illumination [11]. Thus in our method, the global thresholding, based on chromaticity approach, is used for detection of fruit. Figure 1 illustrates the detailed fruit detection algorithm. Segmentation differentiates the apples from the complex background. In our method, a global thresholding is based on chromaticity coordinate " r " is used.

$$r = \frac{R}{R+G+B} \quad (1)$$

Where, R , G , and B are red, green, and blue pixel values respectively.

Equation (2) is used to get the threshold value based on [13]:

$$T = \mu_r + c * \sigma_r \quad (2)$$

Where,

μ_r is the mean value of r of an image

c is a constant ranging from 1 to 3

σ_r is standard deviation of r of an image

Labeling of the image helps to differentiate the apple pixels from the binary image. Labeling method basically scans through the image searching for the pixels which are connected to each other and are assigned as a region. Once all the regions have been found in an image, the regions which are not associated to apples are filtered out using a median filter. These filtered regions are used to detect and locate the apples.

Occlusion Handling

Stereo camera has helped immensely to tackle the issue of occlusion [12] [13]. Based on the theory proposed by [14], we implemented an occlusion handling method. To get the occluded area, we defined two different occlusion maps : occlusion map 1 showing regions in image 1, which are not clearly visible in the image 2 and occlusion map 2 showing regions in image 2 which are not clearly visible in image 1, where the disparity between these two images is the important part.

Results and Discussion

Successful segmentation of apples and background greatly depends on the illumination conditions. Change in illumination could happen in practical scenarios because of moving clouds, sunlight intensity, and the position of the sun. This basically is responsible for the change in appearance of the apples in the images captured by the camera.

Apple Detection

Figure [2] shows the results of image processing algorithm for apple detection under sunny condition. Figure 2[a] is a sample image of apples captured under sunny condition. The background of the apples consists of leaves and branches. Figure 2[b] is the result of segmentation based on chromaticity approach in which apples are brighter than the background. Figure 2[b] was then gone through thresholding, morphological operations, and filtering operations to get figure 2[c]. The apple portion was effectively segmented from the background area (more than 98%). Figure 2[d] illustrates the final detection output of the apples.

Performance of the algorithm under different lighting conditions is shown in Table 1. Segmentation under sunny condition had the highest accuracy (95%) and lowest accuracy was under cloudy condition (85%). When the apple was under shade the accuracy was around 90%. The overall error rate was very less (2%) under all the conditions. These results show that this approach is effective. After comparison with the difference in colour approach [11] and segmentation algorithm based on luminance [15], our approach performed very well under sunny condition. The reason behind that is the chromaticity approach which was barely manipulated by the sunny lighting.

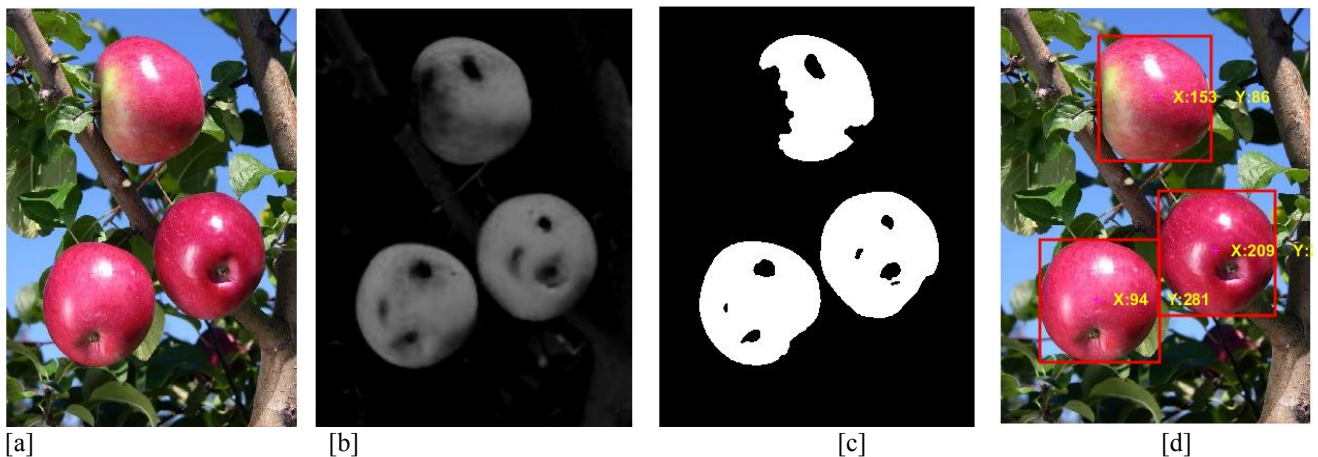


Figure 2: Results of Image Processing of Developed Algorithm in Sunny Condition

In the literature, researchers used shape and texture for feature extraction. The common shape features detailed in the literature are aspect ratio, roundness, convexity, ratios of length and width, perimeter dimensions, elongation, central moment, and principal axis moment shape [16]. They have achieved good results by using these features but the main problem with shape features is that in practical conditions, the apples can be overlapping or occluded. This means that complete fruit edge information or fruit orientation may be missing, which impacts the performance of the system. A combination of colour and shape feature should be explored to further improve detection accuracy.

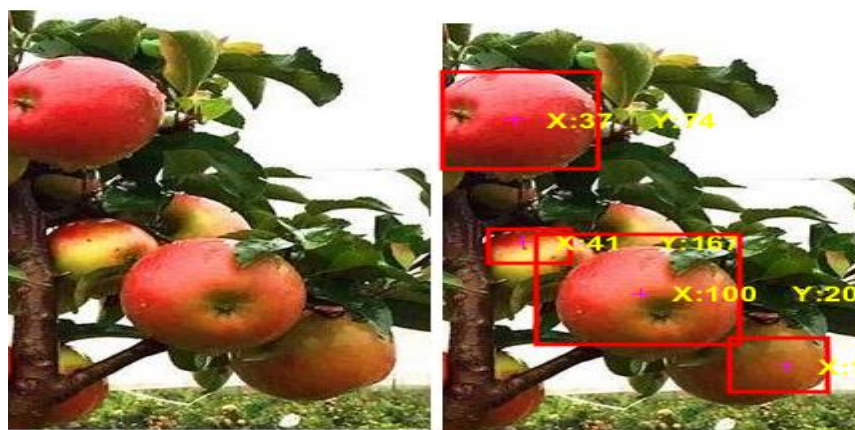
Image Processing Approach	Accuracy in different lighting conditions (%)		
	Sunny	Shade	Cloudy
Difference in colour	85	80	73
Luminance based segmentation	90	83	80
Chromaticity Approach	95	90	85

Table I Performance of the Algorithm

Occlusion Handling:

In practical scenarios, laborers can handle the occlusion by changing their viewing angle. But in a harvesting robot, to handle occlusion, the system should make use of 3D vision to make the system more robust.

Figure [3] shows the result of occlusion handling method under sunny lighting condition. This occlusion handling method has performed differently in different lighting condition. So, this method performed very poorly under cloudy condition and was not efficient to handle occlusion by considering cluster of apples as a single apple. When the apples were in the shade condition this method performed comparatively better than the cloudy condition. It performed extremely well in the sunny condition compared to the rest of the conditions and the reason behind that is the robust feature detection method. Figure 3[a] is an input image with occluded apples and figure 3[b] is a resultant image.



[a] [b]
Figure 3: Result of Occlusion Handling in Sunny Condition

From the literature survey, we analyzed that because of the inability of vision systems to detect apples in changing lighting conditions and inefficient occlusion handling, the harvesting robots are yet to be implemented in practical conditions. As there is advancement in the sensing technology, many commercial 3D sensors such as LiDAR, ToF, and polarization cameras have appeared. These sensors are reported to perform extremely well in the natural lighting scenarios. One of the renowned institute in Austria [17] recently developed a stereo system which constantly rotates around the tree or plant to develop a real-time 360° view. This camera could be very useful for harvesting robots to handle occlusion in real-time.

Every 3D vision system has some merits and demerits. Nonetheless, there is no 3D vision system which is completely immune to change in lighting conditions and adverse climate conditions (snow, rain, wind).

Conclusion

In this paper, the improvement of apple detection for harvesting robots was presented and detailed experiments were performed. The vision system consists of a stereo camera and a PC. There are two main contributions; an apple detection algorithm that can deal with change in lighting condition and can handle partial occlusion. The chromaticity approach with red coefficient was effective for segmentation and detection. Total 90 images, captured under varying lighting conditions, were tested. 85% to 95% of the apples were correctly detected under different lighting conditions. Partial occlusion of apples was handled. The developed apple detection algorithm could be used in harvesting robots.

This approach provides an effective solution to the detection of apples and could be used for other fruits or objects.

References

- [1] A. Paturkar, G. S. Gupta, and D. Bailey, "Overview of image-based 3D vision systems for agricultural applications," in 2017 International Conference on Image and Vision Computing New Zealand (IVCNZ), 2017.
- [2] N. Kondo, M. Monta, and T. Fujiura, "Fruit harvesting robots in Japan," *Advances in Space Research*, vol. 18, pp. 181-184, 1996.
- [3] D. Kong, D. Zhao, Y. Zhang, J. Wang, and H. Zhang, "Research of apple harvesting robot based on least square support vector machine," in *Electrical and Control Engineering (ICECE)*, 2010 International Conference on, 2010, pp. 1590-1593.
- [4] P. P. Ling, R. Ehsani, K. Ting, Y. Chi, N. Ramalingam, M. H. Klingman, *et al.*, "Sensing and end-effector for a robotic tomato harvester," in 2004 ASAE Annual Meeting, 2004, p. 1.
- [5] S. Kitamura, K. Oka, K. Ikutomo, Y. Kimura, and Y. Taniguchi, "A distinction method for fruit of sweet pepper using reflection of LED light," in *SICE Annual Conference*, 2008, 2008, pp. 491-494.
- [6] Y. Kang, C. Lee, and Y. Ho, "An efficient rectification algorithm for multi-view images in parallel camera array," in *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video*, 2008, 2008, pp. 61-64.
- [7] T. Yuan, W. Li, Q. Feng, and J. Zhang, "Spectral imaging for greenhouse cucumber fruit detection based on binocular stereovision," in 2010 Pittsburgh, Pennsylvania, June 20-June 23, 2010, 2010, p. 1.

- [8] H. Jiang, Y. Peng, C. Shen, and Y. Ying, "Study of area-based stereovision method for locating tomato in greenhouse," in *2008 Providence*, Rhode Island, June 29–July 2, 2008, 2008, p. 1.
- [9] R. Xiang, H. Jiang, and Y. Ying, "Recognition of clustered tomatoes based on binocular stereo vision," *Computers and Electronics in Agriculture*, vol. 106, pp. 75-90, 2014.
- [10] Y. Si, G. Liu, and J. Feng, "Location of apples in trees using stereoscopic vision," *Computers and Electronics in Agriculture*, vol. 112, pp. 68-74, 2015.
- [11] D. M. Bulanon, T. Kataoka, Y. Ota, and T. Hiroma, "A colour model for recognition of apples by a robotic harvesting system," *Journal of the Japanese Society of Agricultural Machinery*, vol. 64, pp. 123-133, 2002.
- [12] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision*, vol. 47, pp. 7-42, 2002.
- [13] Y. Zhao, L. Gong, Y. Huang, and C. Liu, "A review of key techniques of vision-based control for harvesting robot," *Computers and Electronics in Agriculture*, vol. 127, pp. 311-323, 2016.
- [14] J. Weng, N. Ahuja, and T. S. Huang, "Two-view Matching," in *ICCV*, 1988, p. 64.
- [15] D. M. Bulanon, T. Kataoka, Y. Ota, and T. Hiroma, "AE—automation and emerging technologies: a segmentation algorithm for the automatic recognition of Fuji apples at harvest," *Biosystems Engineering*, vol. 83, pp. 405-412, 2002.
- [16] O. Bawden, J. Kulk, R. Russell, C. McCool, A. English, F. Dayoub, *et al.*, "Robot for weed species plant-specific management," *Journal of Field Robotics*, vol. 34, pp. 1179-1199, 2017.
- [17] A. N. Belbachir, S. Schraml, M. Mayerhofer, and M. Hofstätter, "A novel HDR depth camera for real-time 3D 360° panoramic vision," in *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2014 IEEE Conference on, 2014, pp. 425-432.