2950 Niles Road, St. Joseph, MI 49085-9659, USA
269.429.0300 fax 269.429.3852 hq@asabe.org www.asabe.org

An ASABE Meeting Presentation

Paper Number: 2189355

A Robo-vision Algorithm for Automatic Harvesting of Green Bell Pepper

Ali Moghimi, PhD Candidate

Bioproducts and Biosystems Engineering, University of Minnesota, moghi005@umn.edu

Biosystems Engineering, Ferdowsi University of Mashhad

Mohammad Hosein Aghkhani

Biosystems Engineering, Ferdowsi University of Mashhad

Mahmood Reza Golzarian

Biosystems Engineering, Ferdowsi University of Mashhad

Abbas Rohani

Biosystems Engineering, Ferdowsi University of Mashhad

Ce Yang

Bioproducts and Biosystems Engineering, University of Minnesota

Written for presentation at the

2015 ASABE Annual International Meeting

Sponsored by ASABE

New Orleans, Louisiana

July 26 – 29, 2015

Abstract. *One of the main concerns of greenhouse growers is the cost for labor-intensive tasks including planting, monitoring, spraying and most importantly harvesting. Within the last two decades, there have been great efforts for developing automatic harvesting robots, but they are not commercialized yet. There is a need to conduct further research about different aspects of robots. Machine vision is one major aspect of a harvesting robot, and generally is inseparable part of robot automation. The main objective of this study was to develop a vision system that is simple, low-cost but effective with a reasonable accuracy for detecting bell pepper in greenhouse. Green bell pepper was chosen not only for its nutrient importance but also for its challenging segmentation due to color similarity between samples of interest and leaves. To overcome this challenge, images were firstly segmented into objects. In the next step, texture characteristic as one of the object-based features was utilized to segment objects into smooth and rough classes. Categorized smooth objects were then classified into plant and non-plant regions using adjusted thresholds of color indices of hue, saturation and Excessive Green Index (EGI). This approach produced promising classification results on images taken under natural light for ultimate purpose of automatic harvesting. The algorithm could recognize 94 out of 108 (detection accuracy of 87%) bell peppers located within workspace of robot.*

Keywords. *Bell Pepper, Computer Vision, Edge detection, Harvesting Robot, Object-based features, Texture*

Introduction

In recent years, greenhouse cultivation has received more attention due to resource constraints and increasing demands for off-season crops. However, one of the main concerns of greenhouse growers is the cost for labor-intensive tasks including planting, monitoring, spraying and most importantly harvesting. Greenhouse fruits such as bell peppers do not ripe simultaneously and this makes cost of manual harvesting rise even higher and make the whole process even more time-consuming and repetitive. Furthermore, ambient temperature and humidity inside greenhouse is rather high and usually with poor ventilation, which poses potential health issues of harvesting labor (Sánchez-Hermosilla et al., 2010).

Adopting automation technologies in greenhouse is a promising solution to overcome these challenges. In some aspects such as monitoring of temperature and moisture, automation has shown promise and it is considerably successful, while in harvesting of crops automation has not yet played the expected role, although within last two decades, there have been great efforts for developing automatic robots for harvesting crops such as cherry (Tanigaki et al., 2008), cucumber (Van Henten et al., 2003), apple (De-An et al., 2011), strawberry (Hayashi et al., 2010), Gerbera Jamesonii (Rath and Kawollek, 2009). Nevertheless, most of these robots have remained in research stage and not commercialized due to high complexity and cost, less reliability and usually being single purpose robots.

Software optimization, efficiency enhancement and mechanical structure improvement were considered the three major aspects by De-An et al. (2011) in order to commercialize and improve the practicability of agricultural robots. Many studies have been conducted on various aspects of robot to make it more practical, efficient and accurate with low cost. For a harvesting robot, machine vision is mostly utilized for fruit detection (Tanigaki et al., 2008) besides automatic navigation (Xue et al., 2012), position and orientation identification (Mehta et al, 2008). However, the main challenge of machine vision system for a harvesting robot is to accurately recognize the fruit and its position on plants by different image processing techniques. In this regard, machine vision system will meet some challenges including variation in ambient light condition, low fruit detectability especially when they are located in the shadow of other fruits and leaves and also color similarity of green crops and leaves or stems as background. In order to overcome these challenges and enhance the efficiency and the competency of robot vision system, different camera positions (Hemming et al., 2012), stereo vision system (Kitamura et al., 2008), other types of camera like as multispectral camera (Bac et al., 2013), and a variety of algorithms have been implemented in different research.

Kitamura et al. (2008) developed a prototype of harvesting robot for bell peppers that had an image processing system using a parallel stereovision consisting of two color CCD cameras and LED light to improve the distinction ability. They mentioned that the distinction ability was lacking; and to overcome this deficiency, LED light was used. It was reported that five frames out of the 21 total frames were not recognized in spite of using LED lights and two cameras.

Hemming et al. (2012) compared the fruit detectability (FD) of fourteen different camera positions in order to detect yellow and red bell pepper in green house. They reported that the maximum of FD was 66% with the criterion of a minimum 50% fruit visibility. They could reach to the maximum FD of 86% for yellow and red bell pepper with a combination of five positions. Multiple viewpoint positions have been also utilized by other researchers. Images of cucumber plants were taken with a linear displacement of 0.33 m so that every plant was visible in three subsequent images (Van Henten et al. 2002). Tanigaki et al. (2008) conducted a research on a cherry harvesting robot and investigated different positions around the crop to increase visibility. Images were taken from four different positions around the trunk of the plant. They reported 59% of the fruits were visible when all images were used. Bulanon et al. (2009) investigated multiple positions around a citrus tree to determine the positions that were needed to get the highest fruit visibility. A combination of up to 6 views resulted in a significant higher visibility.

Multispectral imaging systems have been used to increase the accuracy of target detection by taking the advantage of having near infrared and red-edge bands besides regular RGB bands. Bac et al. (2013) utilized multi-spectral system covering 6 bands and also 16 halogen lamps to classify objects into five classes including fruit, top and bottom of a leaf, stem and petiole for bell pepper harvesting. Even using multi-spectral system covering 6 bands and also 16 halogen lamps did not yield to high classification performance. True positive bell pepper detection rate was 54.5% with standard deviation of 9.9. They stated the results were not high enough to develop an accurate path planning for collusion-free motion of a robot’s arm. They suggested considering object-based features in algorithm and also removing temporal variation of ambient light.

Besides the problems that some of these methods failed to improve the efficiency, the cost of robot was driven up due to using expensive thermal or multispectral camera, increasing the time of processing and registration images captured by two cameras of a stereo vision system, and therefore making robot more complicated. Moreover, although adding artificial light can help in overcoming variation in light condition, it needs bulky and heavy power source which occupies additional space and consume more energy of robot to carry them. Furthermore, it may yield to misclassification because of variation in objects illumination according to their distance and orientation with respect to the position of lamps.

Taking all of the above-mentioned points into consideration, we believe that focusing on software side solution is more reasonable than adding hardware. We can take advantages of development in computer programming and high speed processor to reduce the cost and make the solution more affordable. The main goal of this research was to develop a vision system that is simple, low-cost but effective with a reasonable accuracy in green bell pepper detection. For this purpose, we need to add object-based features such as texture and objects area to color processing in order to enhance detection of green area of interest out of green background.

Material and Methods

Image acquisition

Imaging has been done in a commercial greenhouse with the area of 1000m^2 located in North West of Mashhad, Iran. All of the images have been captured under natural light condition between 11:00 and 13:00 local time. To ensure that there is a slight variation of light condition during experiment, a camera was fixed at the constant distance from specific plant to take images every 15 minutes (nine images in total). Statistical analysis of the crop images showed that the standard deviations of HSI components were 0.003, 0.004 and 0.004 which, showed that light condition altered slightly during experiments.

The imaging system was composed of a RGB Nikon camera (Coolpix p510, Nikon Inc, Japan) and an adjustable tripod camera stand. Distance between the camera lenses and crop was approximately 40 cm.

Image Processing

Generally, the proposed algorithm for bell pepper detection has three main steps: image segmentation, object-based analysis and color features analysis. MATLAB (version R2015a, MathWorks, Natick, MA, USA) has been used for developing algorithm of image analysis.

Segmentation of image into objects

Since the edges in the view would be segmented, smoothing helps to remove noise that causes redundant and excessive edges. Size of median filter mask was chosen to be 3x3. Using larger mask size will lead to smoother and less noisy image but result in blurring of edge and removing of sharp details in the images, which plays important roles in the further step. After smoothing, edge segmentation was done on the images in grayscale level. It should be mentioned that edges of an object are pixels that considerable intensity change is occurring around them, so the intensity gradient of edge pixels is locally maximum. Therefore, it can be concluded that in the areas of high edge diversity, intensity changes more. Although edge segmentation significantly reduced the amount of data in an image, it preserved the structural properties to be used for further image processing (Canny, 1986).

Object-based Image Analysis

When human is asked to determine objects of an image, either consciously or unconsciously, they break the image and interpret the information by assigning attributes to the objects. These characteristics can be reduced to a set of eight elements of image interpretation include: shape, size, tone, shadow, pattern, texture, site and association (Campbell and Wynne 2011). Therefore, it is a wise decision to utilize these elements in algorithms of a robot vision system trying to emulate human’s interpretation as much as possible. However, most research on vision system of harvesting robot failed to exploit the characteristic of adjacent pixels and consider Object- Based Image Analysis (OBIA) when assigning classes or detecting a target; on the contrary they primarily relied on pixel-based analysis. The growing number of peer-reviewed articles is a sufficient proof that shows the advantages of OBIA over per-pixel analyses (Blaschke, 2010).

However, OBIA has been mostly used for remote sensing image analysis so far (Benz et al., 2004; Dronova et al., 2015). In current study, it is intended to consider OBIA in algorithm of green bell pepper detection. For this purpose, the next step is to generate closed objects by morphological dilation operation which is adding pixels to the edges of objects. Then, generated objects were classified according to texture characteristics which is one of the object-based feature for image interpretation. Texture of an image is about the apparent roughness or smoothness of a region in the image (Campbell and Wynne 2011).

Although there is high degree of color similarity and resemblance between bell peppers and green background, bell peppers usually have a smooth and waxy surface that can distinguish them from leaves and stems. This feature results in smooth texture of bell pepper objects (BPO) in an image compared to the rough texture of leaves and stem. In other words, BPO would have large number of pixels. As a result, an area threshold was defined in the number of pixels to segment the objects into two classes, smooth objects and rough objects.

Color Feature Analysis

Only smooth objects remained after applying area threshold. These smooth objects included soil, sky, bell peppers and also some leaf area which had smooth texture. So the next step is to detect bell peppers out of these smooth objects. For this purpose, some color components and index should been utilized. Threshold was defined for hue component to remove soil and sky objects. In addition, Saturation threshold was used to remove leaves although some leaves were still detected as bell pepper. In order to segment bell pepper from leaves which have similar hue and saturation, excessive green index (EGI) as Equation 1 was implemented

$$\mathrm{EGI} = 2g - r - b \tag{1}$$

In the original equation of EGI, non-normalized color components were used (Ohta et al., 1980). However, in this study we used normalized color components in EGI equation. Having known the summation of three normalized components of RGB space is equal to one, then the equation 1 can be simplified to equation 2

$$\mathrm{EGI} = 3g - 1 \tag{2}$$

To summarize all the process, a step-by-step model of algorithm is shown in Figure 1.

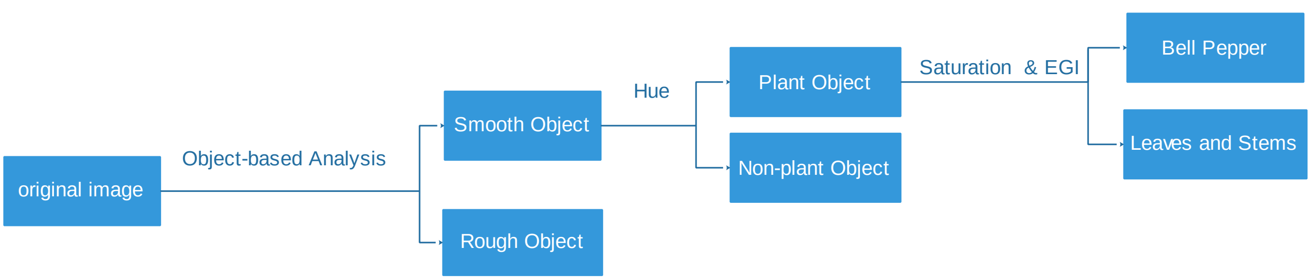

Figure 1. Step-by-step model of algorithm to detect green bell pepper

Results

Figure 2 shows how closed objects are formed. The first image depicts the edge detection of an image which is subsequently transformed to negative image for further process. Then closed objects are generated by morphological dilation operation which is adding pixels to the edges of objects in negative image. As expected, bell pepper almost has a smooth texture because of the less variation of intensity and consequently less edge density while leaves and stems mostly have rough surface due to the high density of leaves and their veins, along with stems. In other words, the edge density is much higher around objects presenting leaves and stems especially when they are at long distance from the camera. In a particular part of an image, the higher the edge density is, the smaller the objects' area is in that part. In general, edge density was sparse in objects belonging to bell pepper, soil, sky and also those leaves which are broad and close to camera.

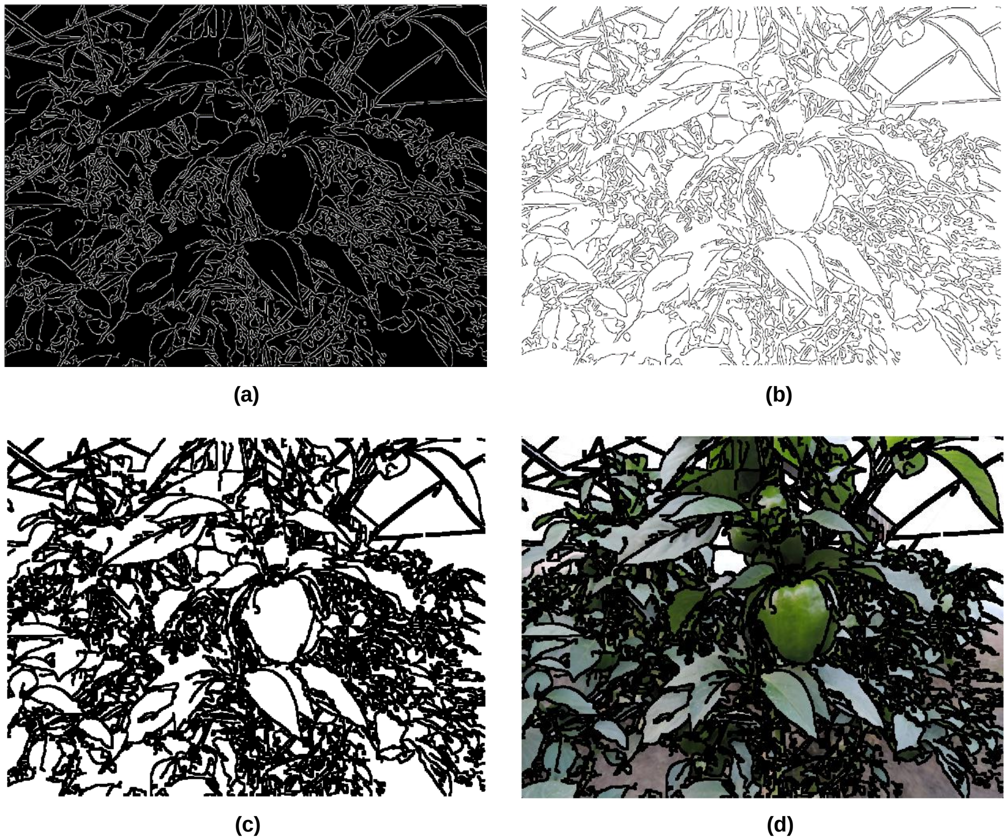

Figure 2. Generating of objects in an image, (a) edge detection, (b) negative of edge detection, (c) closed object formed by applying morphological operation, (d) segmentation of the original image into objects

However, most of these smooth objects can be removed by defining an efficient threshold for objects area. Figure 3 illustrates the result of applying area threshold on objects. BPO was among objects that could pass area threshold filter in most cases. Actually, its smooth texture makes the object consist of a large number of pixels. Nevertheless, some bell peppers might be missed in this step and be recognized as background. It happens usually for two main reasons. The first reason is when bell peppers are located at a far distance from camera and then have small area therefore they are recognized as background, but this should not be a big concern since they are out of the work space of a robot's arm and can be detected in the following images taken while robot is moving along the rows and changing its current position. The next reason is that some bell peppers might be highly shaded by another objects, then they have less area than defined threshold. Figure 2 (b) shows a bell pepper with less degree of visibility covered by stems and leaves so that there is high density of edges and therefore it could not pass area threshold. At this moment further research is needed to overcome this challenge. If threshold defined for area is reduced, a lot of other objects including leaves can pass the filter. This makes the process even more challenging since they have similar color with bell pepper. In this step of algorithm, 9 bell peppers have been missed out of 108 bell peppers which were supposed to be detected by algorithm in 50 images.

Figure 3. (a) Objects passing the area threshold (b) not-detected bell pepper due to being covered by leaves and stems

These problems seem to be unavoidable regardless of using pixel-based or objects-based analysis unless we utilize more than one camera imaging from different angles. However, it should be mentioned that color analyzing of objects generated by morphological operation can increase the speed of processing compared to the pixel-by-pixel analysis because there will be just few objects left to process. The elapsed time to detect and dilate edges and finally apply area threshold on 50 images was approximately 0.2 second in average using Intel(R) Core(TM) i5 CPU 2.4 GHz Quad core processor with 4 GB memory.

As previously mentioned, the next step was applying hue threshold in order to remove soil and sky objects. This step was done with 100 percent accuracy due to major color difference between plant objects (bell peppers and leaves) and non-plant objects (soil and sky).

The most difficult step was bell pepper recognition out of plant objects. This is where saturation and EGI threshold were used. Figure 4 presents how bell pepper is detected by applying hue, saturation and EGI.

There were 108 bell peppers in 50 images taken from plants in greenhouse. According to Table 1, 9 bell peppers have been missed in the first step which was segmentation of image into smooth and rough objects and then remove the rough objects as background. It means that the accuracy of this step to let objects presenting bell pepper pass to the next step is 91.6% (i.e. error=8.4%).

In the second step, 5 bell peppers were detected as background. Then, we have almost 94.9% accuracy since it is 94 out of 99 bell peppers which could pass from the previous filter. Therefore, the algorithm could detect 94 out of 108 bell pepper which means the total accuracy of the presented algorithm for detecting green bell pepper is 87.03 %. This high accuracy seems to be reasonable comparing to the detection rate of 80.8% achieved using parallel stereo vision and also LED light (Kitamura et al., 2008). However, further research is still needed to reduce the amount of leaves detected as target (False Positive Error). Approximately, one leaf was detected as target per each bell pepper. This is high rate of false positive, but would be decreased using machine learning techniques.

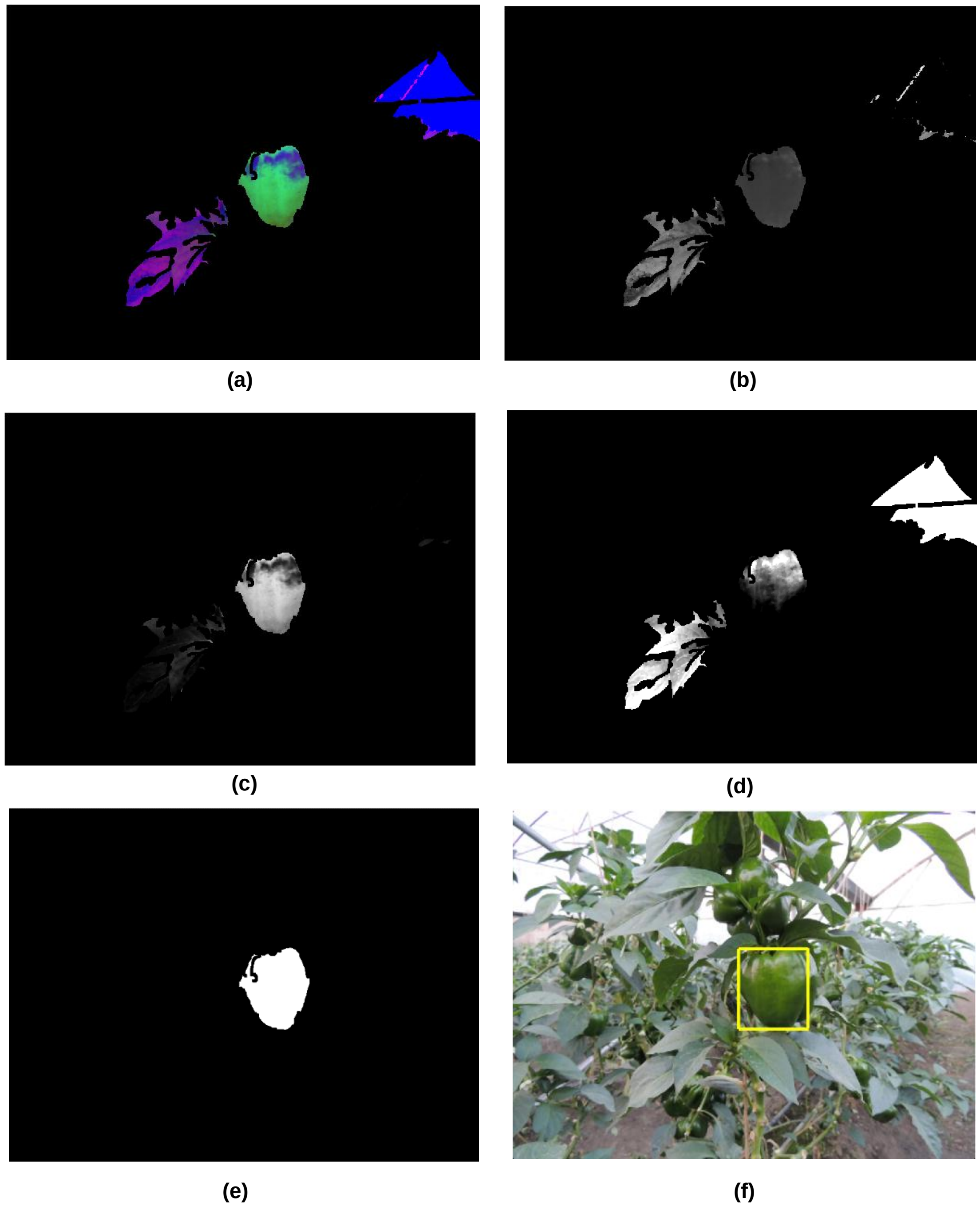

Figure 4. Bell pepper detection by color feature analysis of smooth objects, (a) HSI space, (b) hue, (c) saturation, (d) EGI, (e) object left after color analysis, (f) labeling the target

Table 1. Accuracy of each step of algorithm

	No. of Bell pepper	No. of Detected Bell Pepper	Accuracy
Object-based Analysis	108	99	91.66%
Color features processing	99	94	94.94%
Algorithm	108	94	87.03%

Conclusion

In this research, the main challenge was the similar color of green bell pepper and background. Therefore, just using color processing would not assist a lot in bell pepper detection. To overcome this challenge, Images were segmented into “objects” by detection and dilation of edges. Each of these objects was composed of several pixels which have similar color feature. Then texture, as one of the elements of image interpretation, helped to more effectively classify these objects into rough and smooth objects. Bell peppers were among smooth objects since they tend to have less density of edges due to its waxy and smooth surface. Some of the smooth objects were detected as background and removed by applying the area threshold. So, the algorithm only has to deal with a few objects rather than a lot of pixels which increases the speed of image processing. Finally, some color features have been used to recognize bell pepper out of smooth objects remained from previous steps. The result of this study suggested that objects-based image analysis can be utilized for vision system of robot while it has been widely used for remote sensing image analysis. In general, the algorithm developed in this study was a cost-effective solution for vision system of a harvesting robot with a satisfactory accuracy detection.

Acknowledgements

The authors would like to express their gratitude towards Department of Biosystems Engineering, Ferdowsi University of Mashhad for supporting this research. Also, we greatly appreciate Department of Bioproducts and Biosystms Engineering, University of Minnesota for their grate assist in this study.

References

Bac, C. W., Hemming, J., & van Henten, E. J. (2013). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and Electronics in Agriculture, 96*, 148-162. doi: 10.1016/j.compag.2013.05.004

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing, 58*(3-4), 239-258. doi: 10.1016/j.isprsjprs.2003.10.002

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*(1), 2-16. doi: 10.1016/j.isprsjprs.2009.06.004

Bulanon, D. M., Burks, T. F. & Alchanatis, V. (2009). Fruit visibility analysis for robotic citrus harvesting. *Trans. ASABE 52*, 277-283.

Campbell, J. B. & Wynne, R. H. (2011). Introduction to Remote Sensing, Fifth Edition. New York, NY: The Guilford Press.

Canny, J. (1986). A Computational Approach to Edge Detection. *The IEEE Transactions on Pattern Analysis and Machine Intelligence 8*(6), 679-698.

De-An, Z., Jidong, L., Wei, J., Ying, Z., & Yu, C. (2011). Design and control of an apple harvesting robot. *Biosystems Engineering, 110*(2), 112-122. doi: 10.1016/j.biosystemseng.2011.07.005

Dronova, I., Gong, P., Wang, L., & Zhong, L. (2015). Mapping dynamic cover types in a large seasonally flooded wetland using extended principal component analysis and object-based classification. *Remote Sensing of Environment, 158*, 193-206. doi: 10.1016/j.rse.2014.10.027

Hayashi, S., Shigematsu, K., Yamamoto, S., Kobayashi, K., Kohno, Y., Kamata, J., & Kurita,

M. (2010). Evaluation of a strawberry-harvesting robot in a field test. *Biosystems Engineering, 105*(2), 160-171. doi: 10.1016/j.biosystemseng.2009.09.011

Hemming, J., Ruizendaal, J.L., Hofstee, J.W., & van Henten, E.J. (2012). *Fruit detectability analysis for different camera positions in sweet-pepper*. In *Proc. International conference of agricultural engineering, CIGR-AgEng,* Valencia, Spain.

Kitamura, S., Oka, K., Ikutomo, K., Kimura, Y., & Taniguchi, Y. (2008). *Distinction Method for Fruit of Sweet Pepper Using Reflection of LED Light*. In *Proc. SICE Annual Conference,* The University Electro-Communications, Japan.

Mehta, S. S., Burks, T. F., & Dixon, W. E. (2008). Vision-based localization of a wheeled mobile robot for greenhouse applications: A daisy-chaining approach. *Computers and Electronics in Agriculture, 63*(1), 28-37. doi: 10.1016/j.compag.2008.01.013

Ohta, Y., Kanade, T., & Sakai, T. (1980). Color information for region segmentation. *Computer Graphics and Image Processing, 13*, 222-241.

Rath, T., & Kawollek, M. (2009). Robotic harvesting of Gerbera Jamesonii based on detection and three-dimensional modeling of cut flower pedicels. *Computers and Electronics in Agriculture, 66*(1), 85-92. doi: 10.1016/j.compag.2008.12.006

Sánchez-Hermosilla, J., Rodriguez, F., Gonzalez, R., Guzman, J. L., & Berenguel, M. (2010). A Mechatronic Description of an Autonomous Mobile Robot for Agricultural Tasks in Greenhouses, Mobile Robots Navigation, Alejandra Barrera (Ed.), ISBN: 978-953-307-076-6, InTech, DOI: 10.5772/9003.

Tanigaki, K., Fujiura, T., Akase, A., & Imagawa, J. (2008). Cherry-harvesting robot. *Computers and Electronics in Agriculture, 63*(1), 65-72. doi: 10.1016/j.compag.2008.01.018

Van Henten, E. J., Hemming, J., Vav Tuijl, B. A. J., Kornet, J.G., Meuleman, J., Bontsema J., & Van Os, E.A. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses. *Autonomous Robots, 13*, 241-258.

Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003). Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering, 86*(3), 305-313. doi: 10.1016/j.biosystemseng.2003.08.002

Xue, J., Zhang, L., & Grift, T. E. (2012). Variable field-of-view machine vision based row guidance of an agricultural robot. *Computers and Electronics in Agriculture, 84*, 85-91. doi: 10.1016/j.compag.2012.02.009