EVALUATION OF STEREOVISION FOR EXTRACTING PLANT FEATURES

Z. Mohammed Amean1*, T. Low1, C. McCarthy2 and N. Hancock2
1 Faculty of Engineering and Surveying, University of Southern Queensland, Toowoomba, QLD
2 National Centre for Engineering in Agriculture, West Street, Toowoomba, QLD

ABSTRACT

Visual sensors produce an image that can be analysed by computational algorithms to extract useful information about image features. Plants have complex object structures in which images can help extract, however, not all plant features can be recognised through a single image. Stereovision enhances the single image features by adding the third dimension, depth, to obtain more accurate localisation of the plant's structures. Stereovision is a technique that produces a disparity map of a scene through the use of two or more images taken from different points of view. Depth information can be used to enhance the detection of fruit and plant parts; however research in using Stereovision for extracting plant structures is sparse.

In this paper, Stereovision is analysed in its ability to extract important features from two types of nursery plants taken in indoor and outdoor lighting conditions. From the colour images, colour and shape segmentation are evaluated on their ability to extract certain plant features, such as stems, branches and leaf. Depth images are also evaluated on their accuracy, coverage, and ability to improve image segmentation for colour images. The depth images have some gaps and missing data. The new algorithm develops the depth images by interpolating the gap data and smoothing depth images. Preliminary results show good plant feature can be extracted from depth images at indoor environment, while depth data from an outdoor environment contains more noise due to the variation in lighting conditions.

Keywords: 3D perception, depth perception, disparity map, feature extraction, image segmentation, mobile robot, plant part detection, precision agriculture.

1. INTRODUCTION

Fruit and plant parts detection face important challenges using image processing and segmentation techniques. The challenges are:

1. Extreme illumination conditions which occur when acquiring images of plant tree canopies from any camera orientation with respect to the sun.
2. Objects obscuring the fruit such as leaves, branches and other fruit or large fruit clusters.
3. The limited distance between the plants or trees in a crop row leading to limited working space for sensors.
4. Wind affects the images through blur and motion.

All these parameters affect and reduce the detection accuracy of machine vision and mobile robot. Feature extraction, image segmentation and feature matching are the main steps that researchers currently use to detect fruit.

Image segmentation is a pre-processing technique that partitions images into multiple-parts, each part being a set of pixels, and all pixels being covered by those parts (Raut et al. 2009). Image segmentation makes an image more meaningful and easier to analyse and extract useful data. Image segmentation assigns boundaries to objects such as lines and curves in images to represent an image by a set of segments (Valliammal & Geethalakshmi 2011). Plant parts are usually classified according to their colour, shape and structure of their leaves and flowers. There are multiple segmentation techniques that use colour, intensity, texture, shape and depth.
Colour image segmentation is a popular approach used to identify the region of interest according to its colour. Colour features have low reliability under different lighting condition such as the change in atmosphere, season and shadows (Valliammal & Geethalakshmi 2011). Plant images are complicated and have many details; therefore colour is not typically used as the optimal method to segment an

image. More than one feature has been used to segment a plant or tree image. RGB colour image contain information in the red (R), green (G) and blue (B) channels and the image can be segmented according to these channels. Transformation of the green channel to 'excess green' criterion *ExG* can enhance segmentation of green objects and perform effective discrimination of green plant from background (McCarthy 2009). Depth segmentation is a process of combination between 3D reconstruction and colour segmentation (Yang et al. 2007). Depth segmentation was used by (Yang et al. 2007) to recognise mature fruit and locate cluster position for green house automatic harvest application. A depth filter was applied to remove irrelevant information from the image. This method was applied in a real tomato greenhouse and was robust to strong sunlight and severe noise. This method was able to detect and locate the targets even with stereo occlusion. This method can be expanded to other types of fruit.

In order to save human labour and protect human health, a 3D leaf position measurement method was developed by (Xia et al. 2009) using a stereovision camera for the guidance of an automated pesticides spraying robot. The vertical binocular system was implemented by using one camera to move up and down on the vertical arm of the robot. Birchfield's stereo matching algorithm is used to calculate a disparity map which is reliable enough to detect the depth discontinuities of objects from a scene (Birchfield & Tomasi 1999). The disparity map represents the relative change in disparities and contour of the plant surface (Xia et al. 2009). This disparity map was segmented to several areas and all pixels in each area have the similar depth value. The depth of the leaf was calculated by the mean value of the segmented area of the disparity map. The geometry centre of leaf segmented area represents the 3D position of the leaf and the automatic spray system was positioned to the geometry centre of the detected leaves. Depth image can be used to perform 3D full reconstruction of plant. (Chéné et al. 2012) used depth image for phenotyping of entire plants. The developed algorithm segments depth images of plant from a single top view image for various biological applications. They used a Kinect Microsoft© RGB-depth camera which works at indoor conditions and cannot works at outdoor condition because IR radiation from sunlight degrades the camera measurements.

In this paper, Bumblebee2 stereovision camera has been used to obtain colour and depth/disparity images. This type of camera can work at indoor and outdoor illumination conditions. The developed method segments the colour images using RGB colour channels and edge detection technique. Disparity images have some missing information, so the developed algorithm tried to fill those areas of information and smoothed the result images. Disparity images have been used also to segment the plant parts using the depth and shape information from the disparity images. Both of colour and depth segmentation methods are evaluated on their ability to extracted plant features. Combined colour and depth segmentation has not been widely used in published research for detection of plant parts so this field of study need to be further developed.

1.1 Stereovision technique

Computer vision is also used to acquire a real time picture of the world in three dimensions using stereovision technique. Stereovision refers to the ability of producing 3D structural information (e.g. range and depth information) from two or more images taken from different views (Trucco & Verri 1998). The two images have a lot of similarities and small number of differences (Hamzah et al. 2010).The basic idea of stereovision is to monitor the same scene from slightly offset views and to find the relationship of pixel positions in the image according to the triangulation principle, from which 3D information of the object can be obtained (Xiong et al. 2009). According to the different views required, binocular systems have been frequently used to provide 3D information from 2D views (Xiong et al. 2009). In machine vision, 3D dimension model for an object is obtained by finding the similarities between the stereo images and using projective geometry to process these matches (Hamzah et al. 2010). Figure 1 depicts the image coordinate system with left and right deviations *dl* and *dr* for the left and right images of a stereo pair, respectively. The distance between the centres of the two parallel lenses is defined as the baseline *b*. *f* is defined as the focal length of the two lenses representing the distance between the image planes and the centres of each lens.

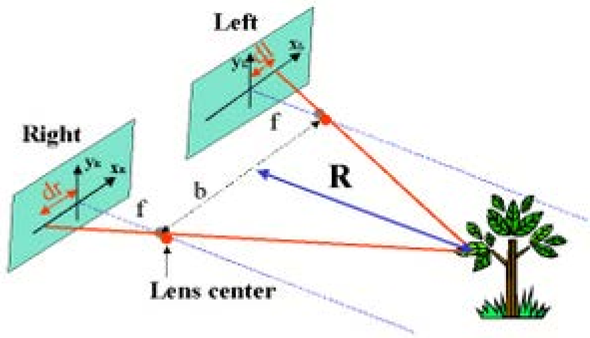


Figure 1. Tow images coordinate system (Rovira-Más et al. 2004).

The image taken for an object perceived through these two lenses will result in offsets. These offsets are proportional to the distance (*R*) between the object and the camera. Those offsets are used to calculate the object depth information. The correspondence of two images representing that the same point of the scene is called disparity matching. The disparity images can be defined as the horizontal position deviations of an object in the image captured by the left and right lenses of the stereovision camera (Rovira-Más et al. 2004).
The disparity (*D*) can be calculated, once the value of *dl* and *dr* are known, by equation (1).

$$D = dl - dr \qquad \text{……………………………(1)}$$

D is given in pixel units and the range *R* can be calculated by applying basic rules of geometry to the scene portrayed in the Figure 1 by applying equation (2).

$$R_{[mm]} = \frac{b_{[mm]} \cdot f_{[mm]}}{D_{[pixels]} \cdot w_{[mm/pixel]}} \qquad \text{…………………………(2)}$$

Where *b* = the baseline of the stereo camera (mm),
f= the focal length of the lenses,
D= the disparity (pixels), and
w= the size of the pixels (mm per pixel).

The range *R* represents the *Z* coordinate in a 3D frame. *X* and *Y* coordinates can be determined by triangulation in a similar manner. If the value of pixel x_L, y_L is known in the left image and this pixel represents the disparity value, the value of the *X* and *Y* coordinated can be calculated by equation (3) (Rovira-Más et al. 2004):

$$X = \frac{x_L \cdot w \cdot R}{f} \; ; \qquad Y = \frac{y_L \cdot w \cdot R}{f} \qquad \text{…………………………(3)}$$

Stereovision technique can be achieved by using two cameras separated by known distance or by using stereovision camera ‘one camera with two lenses separated by known distance.

1.1.1 Stereovision camera

In this paper we use Bumblebee2 stereovision camera with two lenses which provides a balance between 3D data quality, processing speed and size. Stereovision camera is ideal for multiple applications such as mobile robotics, people tracking and gesture recognition and other computer vision applications. Bumblebee2 is precalibrated for lens distortions and camera misalignments. The camera does not require in-field calibration. The calibration information is pre-loaded on the camera. A stereo rig is used to calibrate the cameras. The images have to be mapped to a pin-hole camera model. This image is called rectified. The camera synchronizes by itself which is practically useful for acquiring 3D data from multiple points of view. There are three important parameters related to the bumblebee2 stereovision camera:

1. Focal length which represents the distance from the lenses centre of the camera to the centre of the image plane and it is equal to 2.5 mm. The focal length determines the angle of coverage of the lens. The longer focal length narrows the angle of coverage and the shorter focal length wides the angle of coverage.
2. Base line which represents the distance between the two lenses and it is fixed by the manufacturer to 12 cm.
3. Image centre which represent the point in the image where the "optical axis" of the lens intersects the image plane.

The camera pixel resolution is 640x480 at 48 frames per second (FPS) or 1024x768 at 20 FPS. The camera pre-calibrated to within 0.1 pixel RMS error based on a stereo resolution of 640x480 and is valid for all camera models. Calibration accuracy will vary from camera to camera. There are many parameters related to this camera some of them are general for any camera and others are spatial for bumblebee2 stereovision camera which is called stereo and validation parameters. Those parameters are adjusted to obtain clear disparity image and depth calculation. Perhaps the most important parameter to set in order to insure the extraction of good stereo data is the disparity range. This sets the minimum and maximum distance over which measurements will be made. The larger the maximum disparity setting, the closer the minimum distance measured. The smaller the minimum disparity setting, the further the maximum distance measured.

2. THE EXPERIMENT

The evaluation of stereo vision camera is performed by capturing the images of the two nursery plant at different environment conditions. First the images have been taken at indoor environment from side and top view of the two nursery plants, then at outdoor environment with sun and cloud condition. Figure 2 shows set of stereovision and disparity images taken at three environment conditions for begonia nursery plant. The evaluation of disparity image is calculated according to the ability of the image to produce the depth information for each part of plant. The stereo vision camera has many stereo parameters, and parameter values are set as follows: *Stereo Mask* = 11; *Maximum Disparity* = 158; *Edge Mask* =7; *Surface Validation* = 1.2; *Texture Validation* = 1.1 and *Minimum disparity* parameter which ranged from 55 to 89. The parameter values need to be adjusted to produce good quality disparity image. The images are taken with constant stereo parameter values for all environmental condition except the *Minimum Disparity* parameter.

The values of *Minimum Disparity* parameter ranged from 55 to 89 as reported in Table 1 for images taken from different situation and different environmental conditions. This variation is caused by changing in the background and changing of distance between the camera and plant because each plant has different height. The images has been taken from constant distance between the camera and plant 100 cm for side view condition, for top view this distance was changed to 120 cm in order to capture complete view of plant. From this experiment, it is clear to recognize that Bumblbee2 stereo vision camera is robust to work at indoor and outdoor condition to obtain depth and range data of plant with some missed data for the disparity images at outdoor condition due to the illumination noise. The same set of data has been taken to the hibiscus nursery plant with the same environmental conditions.

Table 1. *Minimum Disparity* parameter values for the images taken from different views for two nursery plant (Hibiscus and begonia).

Parameter name	Minimum Disparity parameters values in:							
	Indoor side view for the two plants	Indoor top view for the two plants	Outdoor sun side view for the two plants	Outdoor sun top view hibiscus	Outdoor sun top view begonia	Outdoor cloud Side view for the two plants	Outdoor cloud Top view hibiscus	Outdoor cloud Top view begonia
Minimum Disparity	55	75	55	62	89	55	62	89

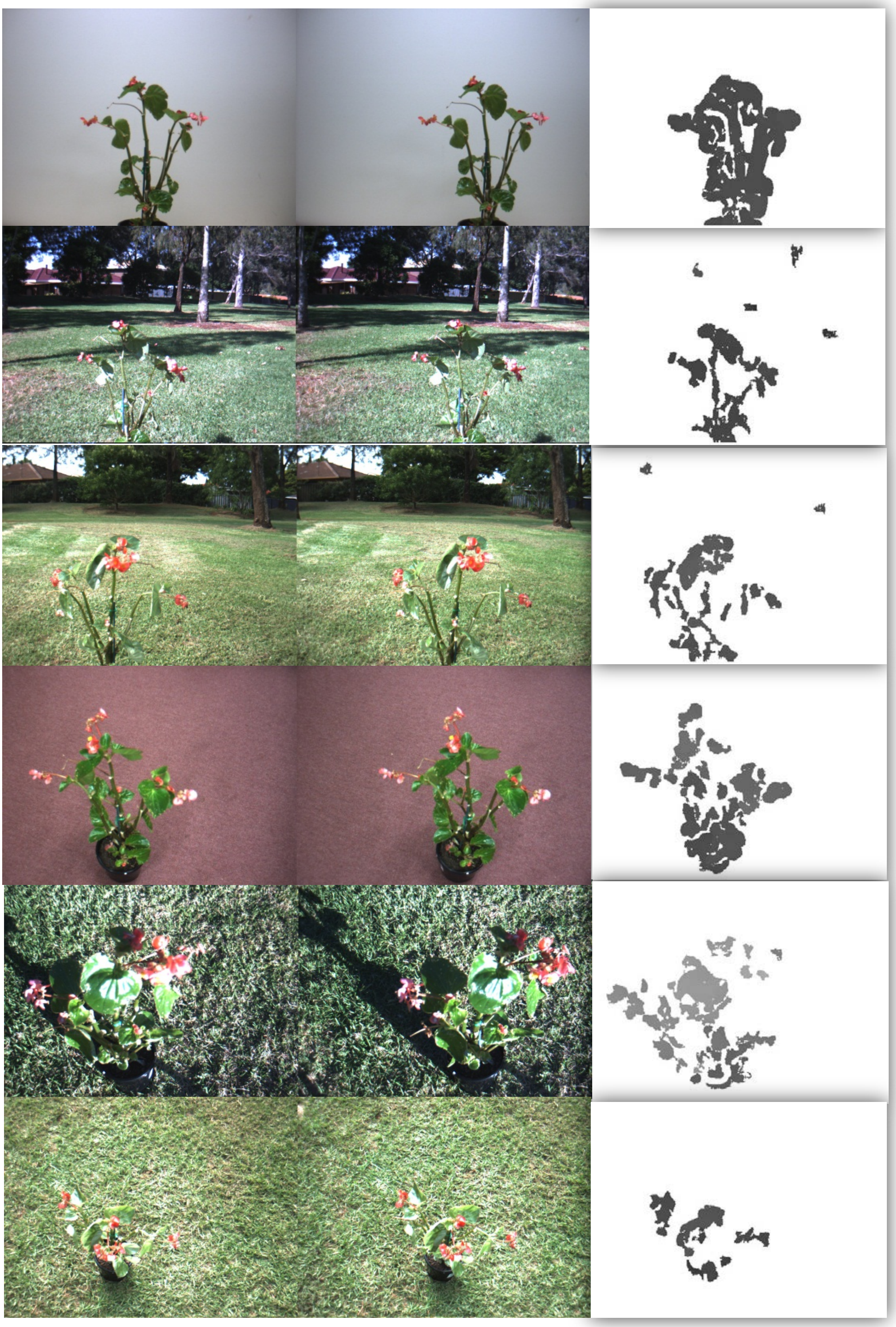


Figure 2. Left, right and disparity images for begonia nursery plant, row 1, 2, and 3, shows begonia side view for indoor, outdoor (sun) and outdoor (cloud) conditions respectively, row 3, 4, and5 shows begonia top view for indoor, outdoor (sun) and outdoor cloud conditions respectively.

3. PLANT SEGMENTATION ALGORITHM

Our algorithm presents the detection of plant parts according to the colour and depth information that captured using bumblbee2 stereo vision camera. The developed algorithm extracts the parts of plant such as stem, leaves and branches from the background and other non-interesting objects using the colour images information first, then segment the images according to the depth information and compare the result. The modified hue or 'excess green' criterion ExG or 2G-R-B (Woebbecke et al. 1995) will be used to enhance the green channel and to obtain the best separation of plant from the background, where R, G and B are the red, green and blue channel respectively. The ExG gray channel image converted to the binary image with sufficient threshold value with indoor environment condition images while Morphological opening operation was applied to minimize the illumination noise of the background and to subtract the background from the image for outdoor environment condition. The contrast of the images was increased by adjusting the intensity of the image foreground for the indoor and outdoor conditions.

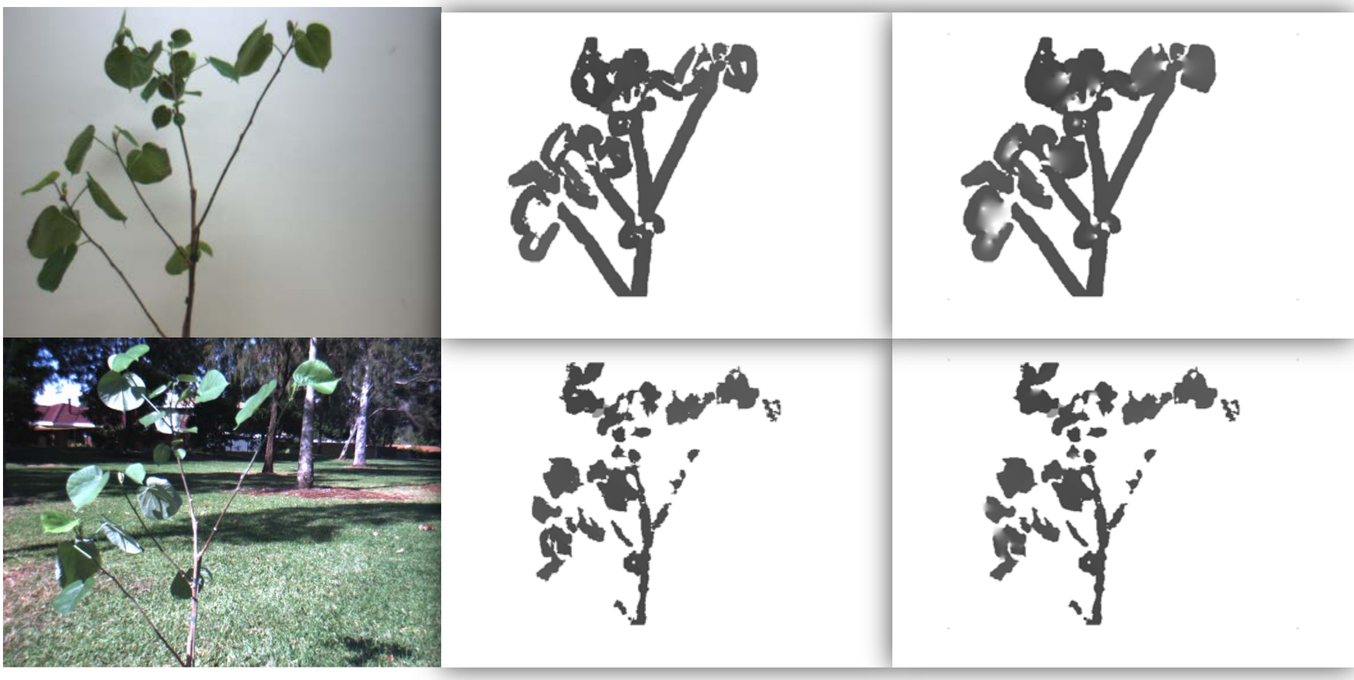
Thresholding was used on one component of the image RGB such as green or Excess green or on the grayscale of the images. The value of the threshold depends on contrast between the plant and the background. From preliminary tests, a low threshold value seems to be sufficient for discrimination of nursery plant in indoor environments, while plants in the field need a high threshold. An Otsu threshold method (Otsu 1975) was used to minimize the variance of the black and white pixels for binary images. Although plants have different parts with different tones and colour, edge detection was used to detect leaves, stem and branches edges. From the literature, Canny edge detection has been demonstrated to be more effective and a more accurate method for edge detection of plant parts (Canny 1986). Figure 3 shows the colour segmentation hibiscus nursery plant at indoor environment condition. Good colour segmentation and edge can be obtained for indoor environment image while poor result for outdoor conditions due to the illumination conditions and the similarity between plant and background. In this condition depth segmentation can provide more accurate results.

(a) (b) (c)

Figure 3. Colour segmentation of hibiscus nursery plant at indoor and outdoor conditions (a) colour image, (b) EXG channel, (c) Canny edge detection (sigma=1).

In addition to the raw colour images, stereovision camera also produces disparity image/map and 3D point cloud. Disparity image is a grayscale image and represents the output of stereo matching between the left and right images. Depth segmentation is a method to segment the disparity image according to the changes of the pixel intensity value, where changing intensity indicates changing depth. Parts of the plant with different intensity values in the disparity image (i.e. different depth) can be segmented according to this value. Result of depth segmentation can be combined with results of colour segmentation to enhance identification of plant parts. The disparity map points have gray level values that mean the darker value points has a farther distance to the camera than the lighter points.

Disparity/depth images as shown in figure 4 have some of missed parts, our developed algorithm fills Those missing parts with interpolated depth values and produced a modified disparity images. The modified disparity images have been smoothed using median filter to reduce the noise in the image. The disparity image will be useful for shape detection to count the plant parts as well as the depth information of each plant parts. High quality disparity image can be obtained on indoor environment While low quality disparity images on outdoor condition but, however it looks better than colour segmentation method to extract the important features of plant. Figure 4 represent the original colour images, disparity images and modified disparity images for indoor and outdoor environmental condition from the side view of the plant. This algorithm has been applied for the two types of nursery plans. Table 2 and Table 3 demonstrate a comparison between colour and depth segmentation for the two types of nursery plant images taken at different side of views and various environmental conditions. The comparison based on three types of images; colour images, Canny edge detection images, and modified disparity images.



(a) (b) (c)

Figure 4. (a) Colour image, (b) disparity image, (c) modified & smoothed disparity image. The picture was taken from constant distance between the camera and plant 100 cm and from the same side of view.

Table 1. Count of Hibiscus plant parts by using colour and depth segmentation.

No. of plant parts hibiscus plant at indoor side view,	Count by manual inspection.	Count from Edge detection.	Count from modified disparity image.
No. of stem	1	1	1
No. of branches	3	3	3
No. of leaves	23	20	16
at indoor top view			
No. of stem	1	1	1
No. of branches	3	3	1
No. of leaves	23	17	8
at outdoor side view/sun			
No. of stem	1	1	1
No. of branches	3	2	2
No. of leaves	23	Severe noise	13
at outdoor top view/sun			
No. of stem	1	Severe noise	0
No. of branches	3	Severe noise	2
No. of leaves	23	Severe noise	13
at outdoor side view/cloud			
No. of stem	1	0	1
No. of branches	3	1	1
No. of leaves	23	9	12
at outdoor top view/cloud			
No. of stem	1	0	0
No. of branches	3	0	0
No. of leaves	23	Severe noise	12

Table 2. Count of Begonia plant parts by using colour and depth segmentation.

No. of plant parts begonia plant at indoor side view,	Count by manual inspection.	Count from Edge detection.	Count from modified disparity image.
No. of stem	1	1	1
No. of branches	3	3	1
No. of leaves	15	12	7
at indoor top view			
No. of stem	1	1	0
No. of branches	3	1	0
No. of leaves	15	7	9
at outdoor side view/sun			
No. of stem	1	Severe noise	1
No. of branches	3	Severe noise	1
No. of leaves	15	Severe noise	5
at outdoor top view/sun			
No. of stem	1	Severe noise	0
No. of branches	3	Severe noise	0
No. of leaves	15	Severe noise	6
at outdoor side view/cloud			
No. of stem	1	Severe noise	1
No. of branches	3	Severe noise	3
No. of leaves	15	Severe noise	5
at outdoor top view/cloud			
No. of stem	1	Severe noise	0
No. of branches	3	Severe noise	0
No. of leaves	15	Severe noise	5

4. CONCLUSION AND PERSPECTIVE

The paper evaluates the performance of stereovision camera for obtaining the depth information and to extract the important features of two types of nursery plant at different lighting conditions. The result shows the accuracy of depth data produced disparity image is more accurate at indoor conditions and the result for side view is more accurate than top view. Depth information has been evaluated for different parts of plant and the result shows different percentage of performance. Range information that produced by stereovision camera has been compared with range information that obtained by using manual measurement and the results have high similarity. Indoor environment depth images were better than images taken at outdoor environment, which contains more noise due to the variation in lighting conditions. The developed disparity images look so promising to develop the extraction of plant parts if the depth information will combine with the colour information. The new algorithm needs to be developing to be robots to work at outdoor conditions and for complex plant structure.

Acknowledgements

The senior author would like to acknowledge the Ministry of Higher Education in Iraq who sponsored the PhD scholarship.

References:

Birchfield, S. and Tomasi, C. (1999). Depth discontinuities by pixel-to-pixel stereo, *International Journal of Computer Vision*, **35:** 269-293.

Canny, J. (1986). A computational approach to edge detection, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **8:** 679-698.

Chéné, Y., Rousseau, D., Lucidarme, P., Bertheloot, J., Caffier, V., Morel, P., Belin, É. and Chapeau-Blondeau, F. (2012). On the use of depth camera for 3D phenotyping of entire plants, *Computers and Electronics in Agriculture*, **82:** 122-127.

Hamzah, R., Rahim, R. and Rosly, H. (2010). Depth evaluation in selected region of disparity mapping for navigation of stereo vision mobile robot, *Proceeding of the 2010 IEEE Symposium on Industrial Electronics & Applications, Penang, Malaysia*, p.551-555.

McCarthy, C. 2009, 'Automatic non-destructive dimensional measurement of cotton plants in real-time by machine vision', PhD thesis, University of Southern Queensland.

Otsu, N. (1975). A threshold selection method from gray-level histograms, *Automatica*, **11:** 23-27.

Raut, S., Raghuvanshi, M., Dharaskar, R. and Raut, A. (2009). Image Segmentation--A State-Of-Art Survey for Prediction, *Proceeding of the International Conference on Advanced Computer Control, Singapore*, p.420-424.

Rovira-Más, F., Zhang, Q. and Reid, J. (2004). Automated agricultural equipment navigation using stereo disparity., *American Society of Agricultural Engineers Transactions of the ASAE*, **47:** 1289−1300.

Trucco, E. and Verri, A. 1998, *Introductory techniques for 3-D computer vision*, vol. 93, Prentice Hall, Upper Saddle River, New Jersey 07458.

Valliammal, N. and Geethalakshmi, S. (2011). Automatic Recognition System Using Preferential Image Segmentation For Leaf And Flower Images, *Computer Science and Engineering: an international Journal*, **1:** 13-25.

Woebbecke, D., Meyer, G., Von Bargen, K. and Mortensen, D. (1995). Color indices for weed identification under various soil, residue, and lighting conditions, *Transactions of the ASAE*, **38:** 259-269.

Xia, C., Li, Y., Chon, T. and Lee, J. (2009). A stereo vision based method for autonomous spray of pesticides to plant leaves, *Proceeding of the IEEE International Symposium on Industrial Electronics, Seoul, Korea*, p.909-914.

Xiong, J., Zou, X., Zou, H., Sun, Q., Chen, Y. and Deng, J. (2009). Research of positioning system for virtual manipulator based on visual error compensation, *Proceeding of the 2nd International Conference on Power Electronics and Intelligent Transportation System, Shenzhen, China*, p.316-319.

Yang, L., Dickinson, J., Wu, Q. and Lang, S. (2007). A fruit recognition method for automatic harvesting, *Proceeding of the 14th International Conference on Mechatronics and Machine Vision in Practice, Xiamen, China*, p.152-157.