CROP RECOGNITION UNDER WEEDY CONDITIONS BASED ON 3D IMAGING FOR ROBOTIC WEED CONTROL

Ji Li[a], Lie Tang[a*]

[a]Agricultural and Biosystem Engineering, Iowa State University, Ames, IA 50011, USA

*Corresponding author. Tel.: +1-515-294-9778; fax: +1-515-294-6633.

E-mail address: lietang@iastate.edu (L. Tang).

1. Abstract

A 3D time-of-flight (ToF) camera was applied to develop a crop plant recognition system for broccoli and green bean plants under weedy conditions. The developed system overcame the previously unsolved problems caused by occluded canopy and illumination variation. An efficient noise filter was developed to remove the sparse noise points in 3D point cloud space. Both 2D and 3D features including the gradient of amplitude and depth image, surface curvature, amplitude percentile index (API), normal direction, and neighbor point count in 3D space were extracted and found effective for recognizing these two types of plants. Separate segmentation algorithms were developed for each of the broccoli and green bean plant in accordance with their 3D geometry and 2D amplitude characteristics. Under the experimental condition where the crops were heavily infested by various types of weed plants, detection rates over 88.3% and 91.2% were achieved for broccoli and green bean plant leaves, respectively. Additionally, the crop plants were segmented out with nearly complete shape. Moreover, the algorithms were computationally optimized, resulting in an image processing speed of over 30 frames per second.

***Keywords*:** plant recognition; 3D point cloud; machine vision; robotic weed control

2. Introduction

An alternative approach to herbicide-based weed control that is clean, effective, and affordable will not only reduce the chemical usage for agricultural production systems, but also alleviate the concerns over herbicide-resistant weeds, environmental pollution, and human health issues (Jeschke, 2016).

Manual weeding is impractical because of labor costs and regulation. In contrast, automated weed control systems offer a great potential to reduce the economical and environmental costs while providing effective weed control efficacy (Thompson, Stafford, & Miller, 1991). There are two research areas for robotic weed control: one is about the weed removal mechanisms for robotic actuation; and the other is related to the control of the weeding tools. Four categories of weed removal mechanisms have been reported, including mechanical weeding (Mohler, 2001), precision chemical spraying (Franco, Pedersen, Papaharalampos, & Ørum, 2017), flaming (Datta & Knezevic, 2013), and high-voltage electrical discharge (Vigneault & Benoît, 2001).

Basically there are three approaches to guide and control the weeding tools. One approach is crop-row following based on machine vision (Åstrand & Baerveldt, 2005; Kise, Zhang, & Rovira Más, 2005; Søgaard & Olsen, 2003) and real-time kinematic global positioning system (RTK-GPS) for weeding tool guidance. The system has shown centimeter-level accuracy (Jørgensen, Søgaard, & Nielsen, 2002; Nagasaka, Umeda, Kanetai, Taniwaki, & Sasaki, 2004). However, its ability to identify individual crop plants was limited, thus not capable of intra-row (within or close to crop rows) weed

control (D. Slaughter, D. Giles, & D. Downey, 2008). Another approach is to use RTK-GPS during seeding or transplanting to generate a map of the crop planting locations. Then, during weeding, the RTK-GPS crop location map is used to actuate the weeding tools. This method can also facilitate the crop/weed detection process, as the plants detected at the places other than those of the recorded crop planting position can be regarded as weeds. Ehsani et al. (Ehsani, Upadhyaya, & Mattson, 2004) tested this kind of system for maize seed mapping, and reported that the average error between the seed map and the actual plant position after germination is about 34 mm. The errors were resulted from the RTK-GPS's error, the motion of the planting device relative to GPS antenna, seed bounce in the furrows, and the change of soil conditions (Griepentrog, Nørremark, Nielsen, & Blackmore, 2005). The third approach is via machine vision based plant species identification. Though this approach has been widely adopted for plant detection and analysis, practically proven solutions for robotic weeding with reliable performance have not been achieved yet (D. C. Slaughter, D. K. Giles, & D. Downey, 2008). Efforts to remove the roadblocks in machine vision based plant identification systems are needed.

With respect to the machine vision based plant species identification for robotic weeding, a large body of literature reporting image processing algorithms for extracting morphological and textural features from two dimensional (2D) color images can be found (Slaughter et al., 2008a,b). Statistical classifiers were also used to differentiate crops and weeds (Du, Wang, & Zhang, 2007; T. P. Kumar, Reddy, & Bora, 2017). More recently, deep learning architectures such as convolutional neural network (CNN) have been founded capable of creating and extracting features from raw representations of

input data without many human interactions. They were found effective in plant classification as well (Grinblat, Uzal, Larese, & Granitto, 2016; Lee, Chan, Mayo, & Remagnino, 2017). But as what were pointed out by Slaughter et al. (D. Slaughter et al., 2008), most of 2D image based plant recognition systems only worked under ideal conditions, where there were no substantial leaf occlusion and leaf damage problems. However, this ideal situation is rarely true in real field conditions.

Images containing invisible wavebands were reported to be effective in segmenting vegetation from soil, particularly images having near-infrared (NIR) band (L. Kumar, Schmidt, Dury, & Skidmore, 2002; Noh, Zhang, Han, Shin, & Reum, 2005). Hyperspectral imaging was investigated for plant species differentiation and was found to be effective under controlled illumination and by using thermally stabilized cameras (Chutia, Bhattacharyya, Sarma, Kalita, & Sudhakar, 2016; Vrindts & Baerdemaeker, 1997; Zhang, Slaughter, & Staab, 2012). Also, it is more robust to occlusion and less computationally intensive than shape-based pattern recognition algorithms. However, there are some major challenges in utilizing hyperspectral imaging for weed/crop detection: First, there are always plant species that are closely related in the field, thus requiring multi-season calibration and complex processes to train classifiers to discriminate crop plants from different weed species (Fennimore, Slaughter, Siemens, Leon, & Saber, 2016). Moreover, as a passive sensor, hyperspectral cameras are susceptible to the changes of environment factors including variation of sunlight, hence are not reliable for plant species classification (Du et al., 2007; Furbank & Tester, 2011; Jones & Vaughan, 2010).

Stereo vision is the method of reconstructing 3D surfaces from images taken from multiple 2D sensors by synthesizing objects from different views, and has been applied to plant recognition (Jin & Tang, 2009; Xiang, Jiang, & Ying, 2014). The main challenge in using stereo vision for field-based agricultural systems is the correspondence searching problems caused by the lack of leaf texture, the complexity of the canopy structure, occlusion, and variation in sunlight conditions (Weiss, Biber, Laible, Bohlmann, & Zell, 2010).

As active 3D imaging sensors such as LiDAR and ToF cameras are more resilient against the variable outdoor lighting conditions. These active 3D sensors can generate 3D point clouds with little to none computational load (Vázquez-Arellano, Griepentrog, Reiser, & Paraforos, 2016). An autonomous outdoor plant mapping system using LiDAR sensors was developed by (Weiss & Biber, 2011). A corn plant spacing sensing system based on a ToF camera was successfully developed by (Nakarmi & Tang, 2012), which achieved more than 98% corn plant detection accuracy. Gai, Tang, and Steward (2016) developed a ToF camera (Microsoft Kinect V2) based computer vision system for crop plant detection, where morphological and structural features of plant leaves and canopy were extracted from 3D point cloud data to build plant recognition classifiers for lettuce and broccoli. Reiser, Vázquez Arellano, Garrido Izard, Griepentrog, and Paraforos (2016) developed an algorithm to detect single plant in 3D point cloud acquired from a 2D LiDAR sensor, and a detection rate of up to 70.7 % was reported.

The overall objective of this research is to apply a 3D imaging sensor and develop corresponding machine vision algorithms to discriminate crop plants from weeds under challenging field conditions where weed infestation is severer than normal. Ultimately,

this research aims to provide a sound machine vision solution with satisfactory accuracy, reliability, and fast processing speed to meet the requirements of automated robotic weeding operations.

3. Sensor and Data Collection System

A 3D ToF camera (Swissranger SR4000, MESA Imaging AG, Zurich, Switzerland) was applied as the sensor in this study. Thanks to the powerful light source, this active 3D sensor has a degree of robustness to ambient sunlight, and can work outdoors when the sunlight is not too strong. A cover or umbrella allows this camera to work outdoors by reducing the sunlight intensity. The camera can provide 3D image with x, y, z coordinate locations for each pixel. It also can capture the amplitude image, which represents the intensity of the reflected light signal emitted from its light source. Its resolution is 144 × 176 pixels.

A data collection system was built by mounting the 3D sensor and a laptop on a modified golf cart. The height of the 3D camera was around 95 cm, and the angle between its view direction and vertical direction was around 18˚. In this study, the distance between camera and plant canopy was about 90 cm, and the camera's spatial resolution was around 4 mm × 4 mm per pixel. By pushing the golf cart and running the data capturing program, continuous amplitude image and point cloud data were collected and stored for the crop plant recognition study.

4. Experiment Design

The study objects of this crop plant recognition research were broccoli and green bean plants in the field with complex field conditions. Image data were collected in the broccoli and green bean fields shown in Figure 1. The broccoli field was full of various

types of weeds that are common in the state of Iowa, USA, including brome grass (*Bromus inermis Leyss*), Waterhemp (*Amaranthus rudis*), pigweed (*Amaranthus*), lambsquaters (*Chenopodium album*), cockspur grass (*Echinochloa crus-galli*), purslane (*Portulaca oleracea*), bindweed (*Convolvulus arvensis*), and clover (*Trifolium*). The broccoli plants were heavily infested by weeds with similar height. The green bean field was also full of weeds, mainly including brome grass (*Bromus inermis Leyss*), purslane (*Portulaca oleracea*), waterhemp (*Amaranthus rudis*), cockspur (*Echinochloa crus-galli*), and bindweed (*Convolvulus arvensis*). The majority of weed plants were brome grass plants. Some weed plants were higher than the green bean plants, and partially occluded the green bean canopies.

The image data were collected between 5 and 8 p.m. on sunny days (June 26, July 5, and July 18, 2012). The 3D ToF camera works the best when ambient light is weak and sun light is diffused. An umbrella was used to block direct sunlight from getting into the field of view of the camera.

Figure 1. Typical crop and weed infestation conditions of broccoli (left) and green bean (right) fields in this research.

In this study, the data collection system was pushed along the path between the crop rows at the speed of around 0.2 m/s, to collect consecutive 3D images and amplitude images of 206 broccoli plants and three rows of green bean plants. The length of one row of green bean plant was about 20 m. The interrow distance of both the broccoli and the green bean field is around 30 cm. The intrarow distance of the broccoli field was approximately 46 cm. Broccoli image sets were collected June 26, July 5, and July 18, 2012. And green bean image sets were collected on July 5, 2012. Each plant has multiple images, which were captured at different viewpoints when the system was pushed to pass by the plants during the data collection process. Moreover, this approach allowed a more comprehensive testing for the crop plant recognition at different viewpoints.

Table 1 displays the heights and growth stages of broccoli and green bean plants corresponding to each data collection date. There was little maturity difference observed in the broccoli plants between June 26th and July 5th, which was probably because of the high competition between weeds and crops.

Table 1. Height and growth stages of crop plants corresponding to each image capturing date

Data collection date	Height of broccoli (cm)	Growth stage of broccoli	Height of green bean plant (cm)	Growth stage of green bean
June 26, 2012	18–23	Early vegetative	N/A	Late vegetative or Pre-reproductive
July 5, 2012	18–23		Approx. 28	
July 18, 2012	18–30		N/A	

For each of the 206 broccoli plants and each of three data capturing dates, three image sets were randomly selected to evaluate the recognition performance of this system for the broccoli at different growing stages. The broccoli plant recognition accuracy of randomly selected images was recorded. Additionally, for practical applications, when

the mobile platform moved through the field, the platform had multiple chances to detect a crop by capturing and analyzing multiple sequential images of the plant. Therefore, the rate of recognizing a broccoli accurately from at least one of its three image sets collected at the same date was also evaluated. The false detection rate, which mistakenly recognized the background objects as broccoli plants, was also calculated.

Among the total 6967 image sets of three rows of green bean plant, 100 image sets were randomly selected to evaluate the green bean detection rate and segmentation accuracy by comparing to the manual counting result.

Algorithm Design

In this crop plant recognition research project, we developed 2D and 3D image processing algorithms to process the amplitude and depth images generated by the ToF camera. The algorithm of this research is introduced step by step below.

Initial Noise Filtering

Because of the strong ambient sunlight, some points of the data collected by the ToF camera did not have correct amplitude and 3D coordinate information. In that case, the strong ambient light saturated the corresponding pixels of image, leading to invalid amplitude and 3D data points. A threshold was applied to recognize the noise pixels from the image data where the amplitude was too big. Additionally, the pixels of the 3D image with a depth value over 1.5 m were treated as noise because the distance between the 3D sensor and the ground was only about 1 m in this study.

Curvature Estimation of Surface

Curvature is an important local feature that captures the 3D geometry of the local surface around a query point p_q. It is critical for this plant recognition study. One

example is that a query point p_q and its neighbor points belonging to the same green bean or broccoli leaf should have small curvature because the local surface of the plant leaves is relatively smooth and continuous.

To estimate the curvature of the local surface around a query point p_q, the first step is to search all of its neighbor points within a specific radius *r* in 3D space, and the selection of the radius for the neighborhood definition is important and application dependent. While the smaller maximum distance value will reduce the number of neighbor points and make the calculation results more susceptible to the local noise of the 3D image, a larger value is more likely to mistakenly select the points of other objects' surfaces as the neighbor points for curvature calculation. In this study, the maximum distance of neighbor points for the local surface feature analysis was set at 10 mm, and there were several reasons of this decision:

1) In this application, the surface curvature estimation at one point was based on the principal component analysis (PCA) to process the 3D coordinates of it and its neighborhood points, and PCA requires enough sample count. The distance between two closest points on a smooth leaf surface was around 4 mm. Therefore, inside the sphere with a radius of 10 mm and centered at the query point p_q, normally over 10 neighbor points could be found. This number of points was enough to perform PCA for curvature estimation.
2) As the size of green bean leaves was small and the broccoli leaves were not flat, a small radius would make the local surface estimation algorithm more accurate and would help to avoid taking the relatively far points into analysis.

3) The leaves of the crop plants and the weeds were close to each other, and occlusion of the canopy was common. A smaller radius could reduce the chance of taking the points from other leaves for local surface analysis.

The standard deviation of the 3D measurements of SR4000 used is around 4 mm, according to the datasheet. If the reflectance of the object is small, the standard deviation of measurement further increases. In this case, if the neighborhood definition radius is too small, the 3D measurement error can greatly affect the normal and curvature estimation accuracy of the local surface. Considering above problems, radius 10 mm was selected as the neighbor definition.

This study applied a PCA based method proposed by Pauly et al. (Pauly, Gross, & Kobbelt, 2002) to estimate the surface curvature. The surface curvature at a query point p_q is expressed by δ of Equation 1, where λ_1, λ_2, and λ_3 are the three eigenvalues of the covariance matrix of the 3D coordinates of $p_q$and its neighbor points set P_{10}^k. The range of δ is between 0 and 1/3, and the bigger δ indicates big curvature.

$$\delta = \frac{\min(\lambda_1,\ \lambda_2,\ \lambda_3)}{\lambda_1+\lambda_2+\lambda_3} \qquad (1)$$

Moreover, δ is not only related to curvature, it is also related to the noise level around the query point. The larger the noise level is, the higher δ becomes.

Figure 6f and 7f are the curvature images. As they indicate, the flat surface of scenery has low δ value, whereas the weeds and soil have relatively high δ value because of the discontinuity caused by the narrowness of the weed leaf and the roughness of earth surface. This indicates that δ value is effective to differentiate the crop plants from soil and weeds.

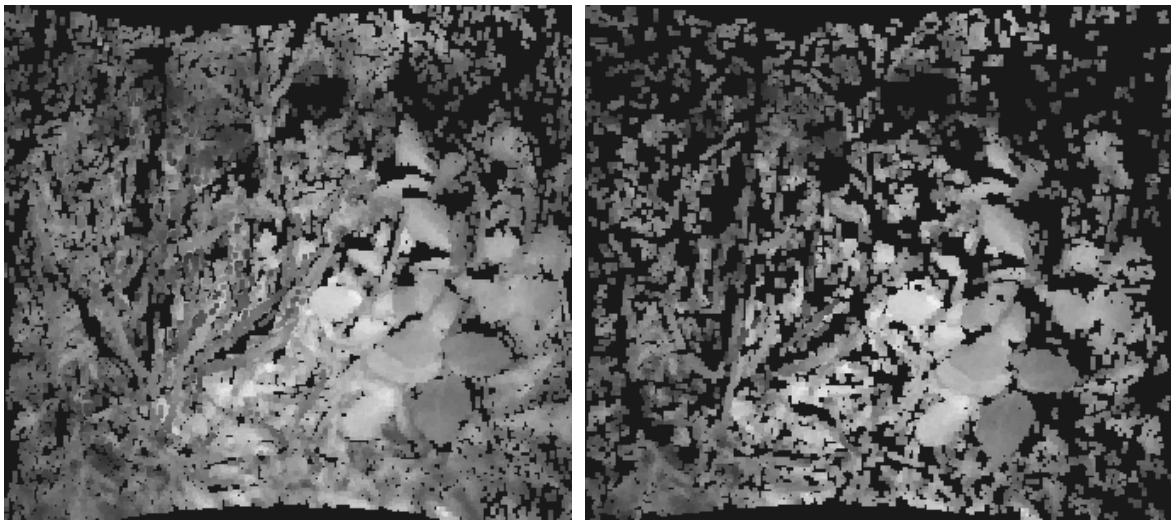
Neighbor Count Image

For every point in the 3D image data $\mathrm{p_q}$, the size of its neighbors $\mathrm{P_{10}^k}$ was also stored as the neighbor count image, N_C, which is shown by Figure 6e and 7e. The maximum N_C value of the whole data set collected in this research was 40. As N_C indicates, the pixels of broccoli and green bean leaves had significantly higher number of neighbor points in 3D space than most pixels of soil and weed background because of their relatively smooth 3D geometry. This observation was helpful for this crop plant recognition study.

Further Noise Filtering

Due to measurement error, the 3D image had outliers, namely sparse point noise, and they had relatively big distance to all other points. The above initial data cleansing method is simple, yet not sufficient to remove the sparse points. Since it can only be used as a preprocessing step, further noise filtering was conducted.

To perform further noise filtering, all the data points of the 3D image data processed by the initial noise filtering were organized into different regions in 3D space based on neighborhood analysis. The region separation is done by putting each point data p_q and its neighbor points P_{10}^k into the same region. The points in the regions with a member count less than four were removed as sparse point noise.



(a) (b)

Figure 2. Front view of 3D point cloud data after (a) initial noise removal, and (b) further noise removal process

The 3D image data before and after the further noise filtering is shown in Figure 2. As it indicates, many points were removed by the advanced noise filter. These filtered points were sparse point noise; most of which were weed plants and soil residue.

Gradient of Amplitude and Depth Image

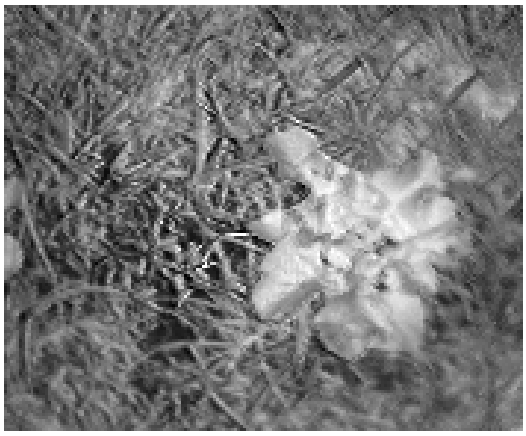
Gradient images of amplitude and depth information were produced by using a Sobel operator. The depth gradient image and amplitude gradient image are represented using G_Z *and* G_I, respectively, in this study. As Figure 6c, 6d, 7c and 7d indicate, the amplitude and depth gradient is high for the place where the change of amplitude value is high and the discontinuity of 3D geometry happens. They are critical features for the segmentation task of this research.

Percentile Index of Amplitude Image

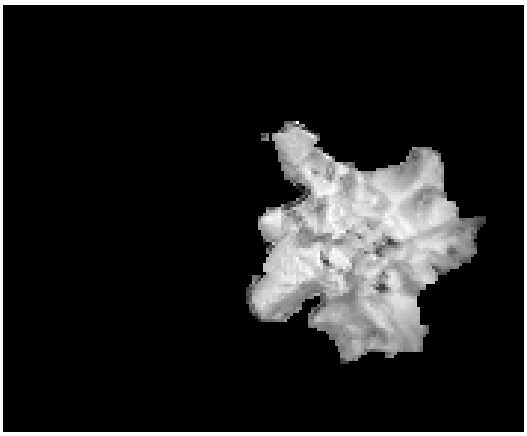
Based on the observation, the broccoli and green bean plants have higher intensity value than the weeds and soil background in the amplitude image. Figure 3d and e show the histograms of the background area and the broccoli area of the amplitude image

Figure 3a, respectively. As it can be seen, the amplitude range of broccoli is significantly higher than the soil and weed background. For this data set, although the majority of the background's amplitude value is smaller than 150, the major range of broccoli's amplitude value is between 125 and 255.

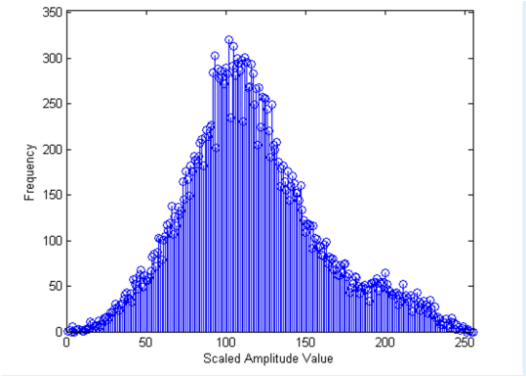
However, the commonly used histogram based threshold algorithms including the 2D maximum entropy thresholding algorithm (Kapur, Sahoo, & Wong, 1985), minimum error thresholding algorithm (Kittler & Illingworth, 1986), and Otsu's method (Otsu, 1975) could not work well to segment the broccoli and green bean out from the amplitude image. This is because that the histogram of the whole amplitude image has two obvious peaks separated by a trough, although the histograms of the background and the broccoli have significantly different ranges and peak values as Figure 3c indicates.



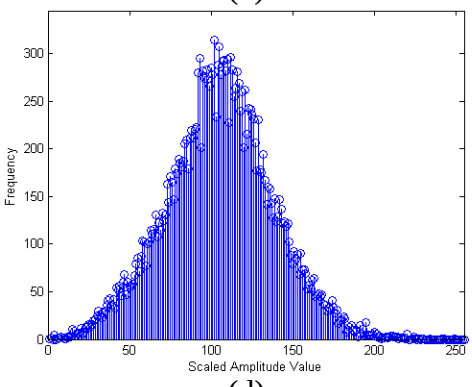
(a)



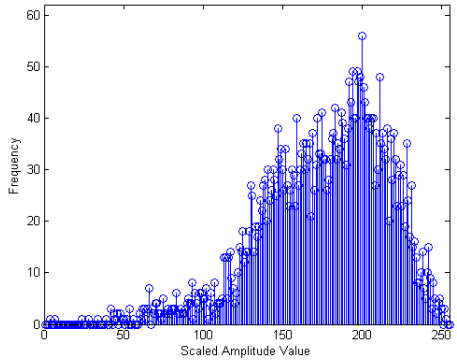
(b)



(c)



(d)



(e)

Figure 3. Broccoli amplitude image and histograms: (a) broccoli amplitude image, (b) broccoli segmented from amplitude image, (c) histogram of whole amplitude image, (d) histogram of background, and (e) histogram of broccoli

To solve this problem, this research created a feature called the API. While the data collection system was pushed through the field, the broccoli and green bean plants from right to the left side of view. Depending on the position of the crop plant, it partially or fully appeared in the image view. The size of crop plant in the view can greatly influence the histogram of the amplitude image. To solve it, API image was created, and the API value of a pixel was the percentile of its amplitude information within its column of pixels in the image, instead of the whole image. Therefore, when the data collection system went through in the field, the API value of the pixels of broccoli and green bean was less susceptible to their position in the view no matter whether the crop plant partially appeared at the left or the right boarder of the image or fully appeared in the center. Figure 6g and 7g are the API image of green bean and broccoli, indicate the pixels of green bean and broccoli have higher API value compared to the background pixels in the same column in the image.

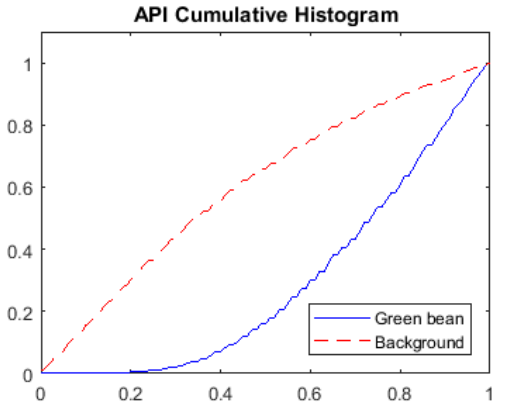
Statistics of Extracted Features

To develop a crop plant segmentation algorithm based on extracted 2D and 3D features, statistical analysis was done to investigate the distribution of each feature for broccoli, green bean, and background. Five randomly selected image sets of broccoli and green bean were used respectively for feature distribution analysis, and they were processed with manual segmentation to specify the area of crop plant and background.

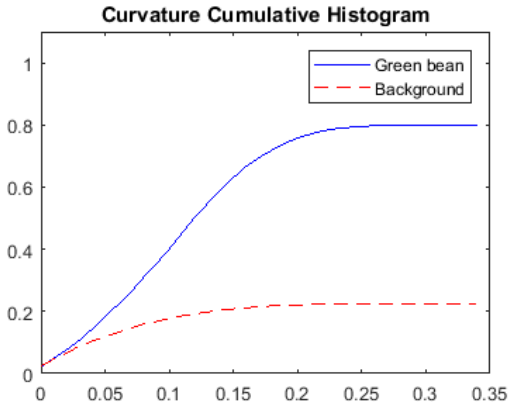
For broccoli and green bean plants, cumulative histogram of each extracted feature was generated at the crop plant and background areas of five sample image sets, respectively, as Figure 4 and 5 show. As Figure 4b and 5b indicate, the maximum cumulative histogram value of curvature feature is smaller than 1, this is because part of the crop plant and background area does not have effective curvature value because it does not have enough neighbor points for curvature estimation. For green bean image sets, 80% of the green bean and only 22.41% of the background area had valid curvature value. For broccoli image sets, 85.85% of the broccoli area and only 36.25% of the background had valid curvature value. Therefore, this is helpful to differentiate crop plant and background.

As the cumulative histogram indicates, the API and neighbor count features of broccoli and green bean leaves are mainly distributed in the higher value range, compared to the background, and the amplitude and depth gradient and curvature of these two crop plants are relatively smaller than the soil and weed background. Based on the feature distribution, this research carefully picked the threshold value for each feature in order to effectively remove the background area while keeping most of the crop plant area. The picked threshold and percentage of crop plant and background area within the range are given in Table 2 and 3. As these two tables indicate, the picked threshold values are

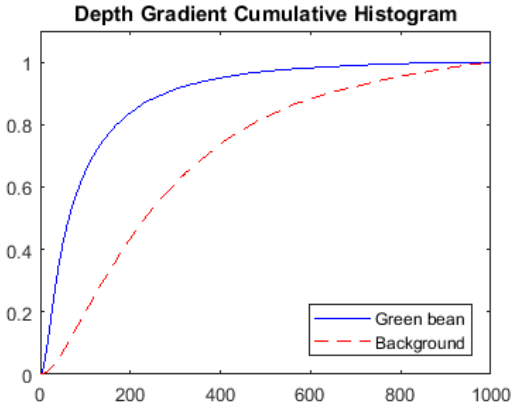
effective to keep most broccoli and green bean areas and remove some background area. An example is the selected curvature threshold value 0.25, which can keep 79.54% of the broccoli area while removing 77.61% of the background. The combination of the threshold of all features is powerful to segment the inner area of the broccoli and green bean leaves, which is introduced later.



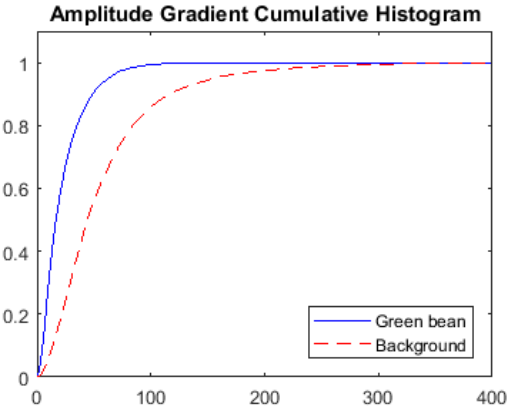
(a)



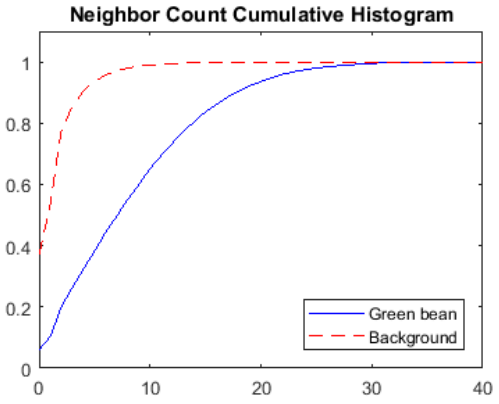
(b)



(c)

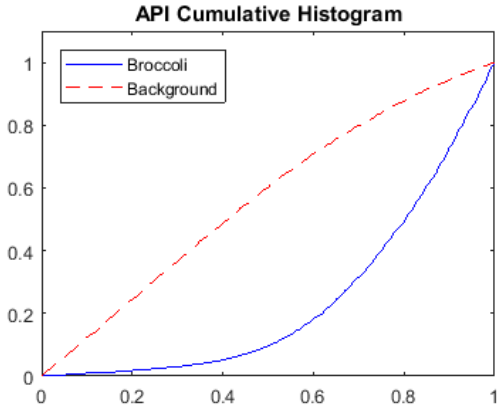


(d)

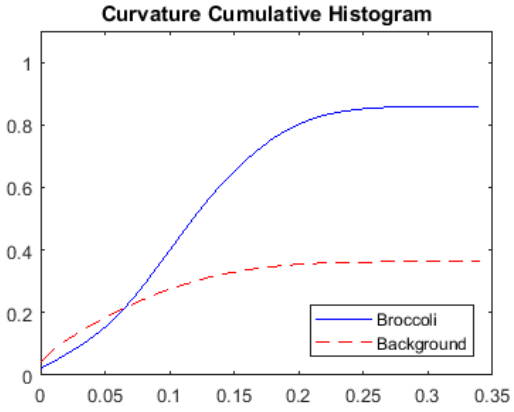


(e)

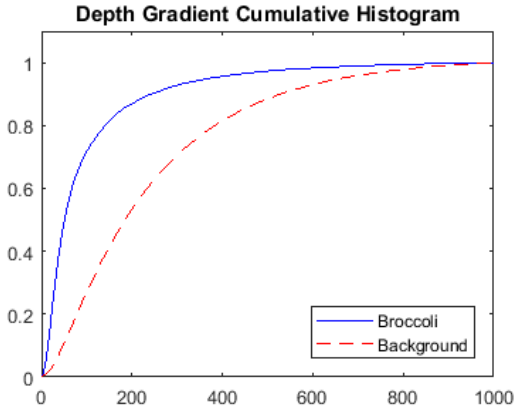
Figure 4. Cumulative histograms of various features of green bean and background: (a) cumulative histogram of API feature of green bean and background; (b) cumulative histogram of curvature feature of green bean and background, (c) cumulative histogram of depth gradient of green bean and background, (d) cumulative histogram of amplitude gradient of green bean and background, and (e) cumulative histogram of neighbor count feature of green bean and background



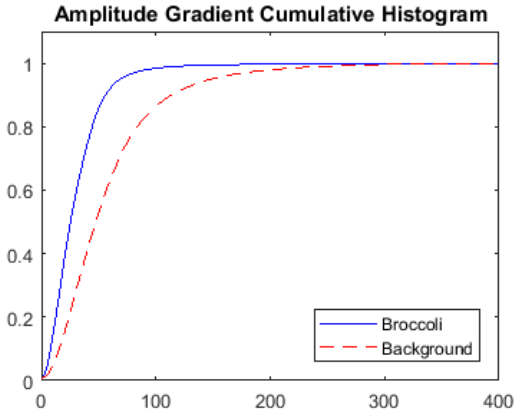
(a)



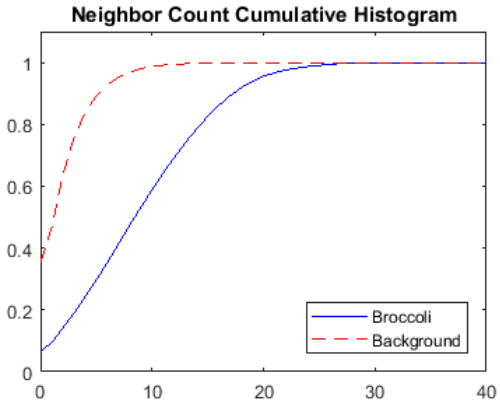
(b)



(c)



(d)



(e)

Figure 5. Cumulative histograms of various features of broccoli and background: (a) cumulative histogram of API feature of broccoli and background, (b) cumulative histogram of curvature feature of broccoli and background, (c) cumulative histogram of depth gradient of broccoli and background, (d) cumulative histogram of amplitude gradient of broccoli and background, and (e) cumulative histogram of neighbor count feature of broccoli and background

Table 2. Feature distribution table for green bean

	Feature Range	Percentage (%)		Feature Range	Percentage (%)	
		Green bean	Background		Green bean	Background
API	≥ 0.3	97.90	56.83			
Curvature	≤ 0.25	79.54	22.39	≤ 0.3	80.00	22.41
Depth gradient	≤ 80	.57.67	14.83	≤ 100	64.99	20.02
Amplitude gradient	≤ 25	.67.18	24.17	≤ 36	79.99	38.07
Neighbor count	≥ 5	67.56	9.50	≥ 7	55.64	4.23

Table 3. Feature distribution table for broccoli

	Feature Range	Percentage (%)		Feature Range	Percentage (%)	
		Broccoli	Background		Broccoli	Background
API	≥ 0.65	77.42	25.65	0.6	83.00	30.35
Curvature	≤ 0.25	85.16	36.17	≤ 0.3	85.81	36.25
Depth gradient	≤ 150	81.34	41.47	≤ 220	88.48	57.43
Amplitude gradient	≤ 65	93.96	67.50			
Neighbor count	≥ 5	75.81	16.21			

Crop Plant Segmentation Algorithm

After the noise filtering and feature extraction, the crop plant recognition algorithm segments the crop plants. The recognition algorithms of the broccoli and green bean plant share great similarity, but some differences exist because of their differences in 2D and 3D features described below.

5.1.1. Green bean Segmentation Algorithm

The raw green bean images used to introduce the green bean segmentation algorithm in this section are given by Figure 6a-b. This algorithm is based on various feature information calculated by the previously described steps, including curvature δ, neighbor count image N_C, amplitude and depth gradient image G_I and G_Z, and amplitude percentile image *API*; all of them are also provided in Figure 6c–g.

The green bean segmentation algorithm consists of multiple steps, which are introduced below.

1) By applying the feature images achieved in the previous step, two intermediate images, S1 and S2, were constructed based on the Equation 2 and 3, respectively. The threshold values of these two equations were picked based on the feature distribution discussion above. The main idea of this step was to apply these features to extract the green bean candidates areas based on several observations. First, the 3D geometry of green bean leaf was relatively flat and smooth. Therefore, the depth gradient value G_Z and curvature δ were small compared to the background. Second, the amplitude value of a green bean leaf was nearly uniform, so its amplitude gradient was small. Third, the

API value of green bean was not too small because its amplitude value is at least higher than soil and some weed background. Fourth, the points of the green bean leaf in the 3D point cloud data were relatively intense because of their flat 3D geometry, and thus, the neighbor point count value N_C was relatively high for the green bean points. Based on these criteria, image *S1* and *S2* are achieved as Figure 6h and i indicate.

$$S1 = \begin{cases} 1 & when\ G_z \le 80, G_I \le 25, N_c \ge 7, \delta \le 0.25, API \ge 0.3 \\ 0 & else \end{cases} \quad (2)$$

$$S2 = \begin{cases} 1 & when\ G_z \le 100, G_I \le 35, N_c \ge 5, \delta \le 0.3, API \ge 0.3 \\ 0 & else \end{cases} \quad (3)$$

2) The white pixels of the intermediate result image *S1* and *S2* can be separated into different regions in 2D space. In this study, the regions of size less than threshold size were considered as noise and were removed in image *S1* and *S2*, respectively. Because the amplitude image of a green bean leaf was smooth, the average amplitude gradient was small. Therefore, the regions with average amplitude gradient over 25 were also removed as background. The resulting images of *S1* and *S2* after this step are represented with *S11* and *S21*, respectively, which are shown by Figure 6 j and k.
3) Next, image *S11* was dilated and processed in 3D space. Each pixel of *S11* was checked to determine whether there was neighbor point in 3D space that was removed in *S11* but not in *S21* image. If there was a neighbor point, the qualified neighbor point was added into image *S11*. In addition, this operation was iterated to process new *S11* until there was no longer any new point to be added. After that, all of the points of the new *S11* image were separated into different regions in 3D space. For the region separation in 3D space, if the

distance between two points was no bigger than 10 mm, they were considered as connected points, belonging to the same region. Each region was a green bean leaf candidate. The regions with size smaller than threshold were removed because they were too small to be a green bean leaf. The final result of this step is represented with *S12*, as Figure 6l shows.

4) As Figure 6l indicates, the pixels extracted by the intermediate result image *S12* are only from the inner part of the green bean leaves instead of the complete results. Because the boundary area of the leaf has relatively high gradient value for both depth and amplitude, they were filtered out by the three steps mentioned above. In order to recover the boundary area of the green bean leaves, the intermediate image *S12* was expanded in 3D space. All of the removed neighbor points in 3D space of any point in image *S12* were added to create a new *S12* image. In addition, this process was iterated five times to update image *S12*, so that the boundary points, which are close to the extracted inner part of green bean leaf, could be recovered, as Figure 6m shows. The resulting image of this step is *S13*.
5) The intermediate result image *S13* still cannot fully recover the boundary part of the green bean leaves for some cases, so another algorithm was adopted to solve this problem. If a removed point q was a neighbor point of a valid point p in *S13* in 3D space, and it satisfied Equation 4–6, the removed point q was added to create a new S13 image, and this process was iterated five times.

$$\left|Z_q - Z_p\right| \leq 7\ mm, \qquad (4)$$

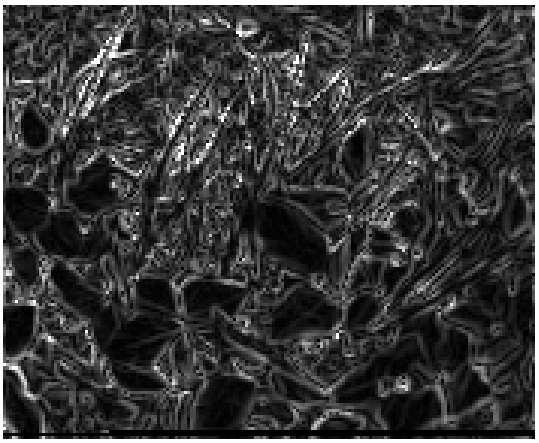
where Z_q and Z_p are the Z value of the 3D coordinates of point q and p, respectively. Equation 4 specifies that only the point p whose depth is similar to that of point q can be added, in order to ensure the newly added points and the existing points are at a single smooth surface.

$$G_{zq} \leq 130 \qquad (5)$$

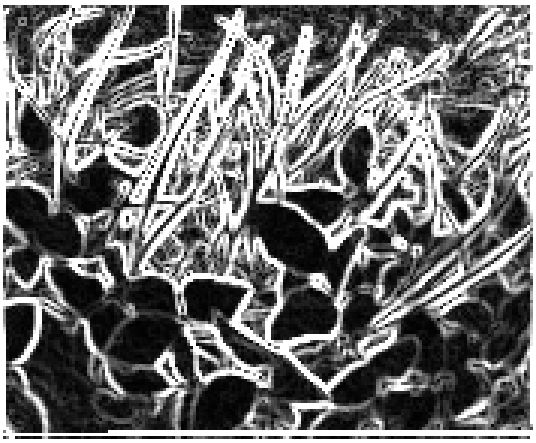
$$G_{Iq} \leq 70, \qquad (6)$$

where G_{zq} and G_{Iq} represent the depth gradient value and amplitude gradient value of point q. Equation 5 and 6 specify the upper limit of the corresponding gradient value for candidate green bean leaf boundary points to be recovered, and they are designed to avoid adding the points where the 3D geometry and amplitude image are not smooth. The threshold values in these two equations are higher than those in Equation 2 and 3 because of the nature differences between the inner part and the boundary area of the green bean leaves for both the 3D geometry and the 2D amplitude image. The resulting image of this step is the final green bean segmentation result, which is shown in Figure 6n.

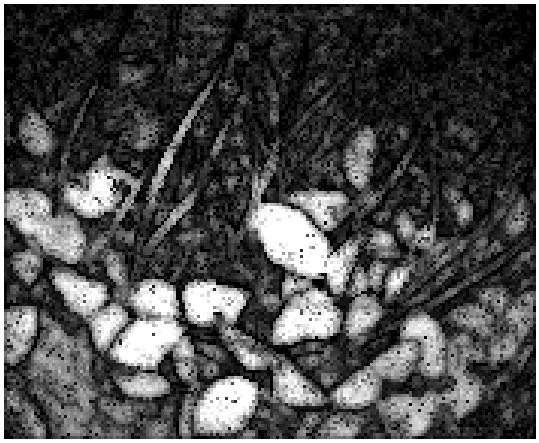
(a) (b)



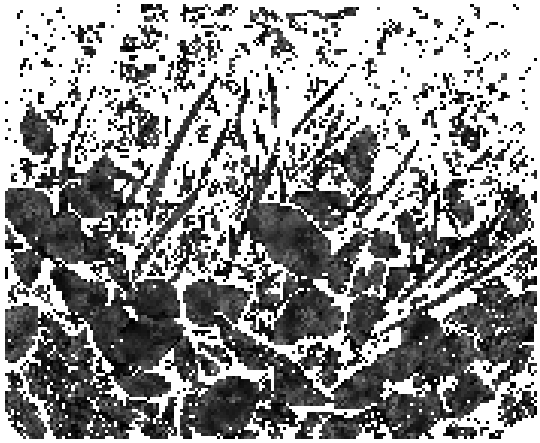
(c)



(d)

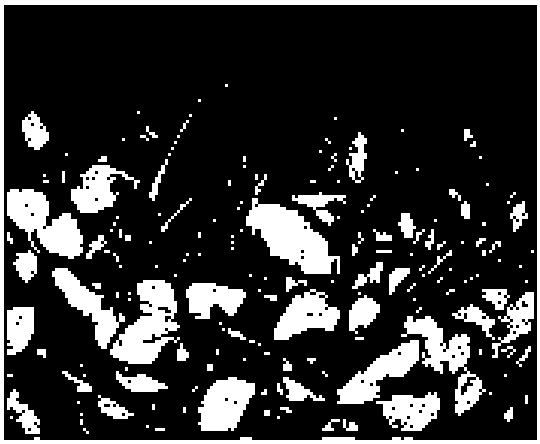


(e)

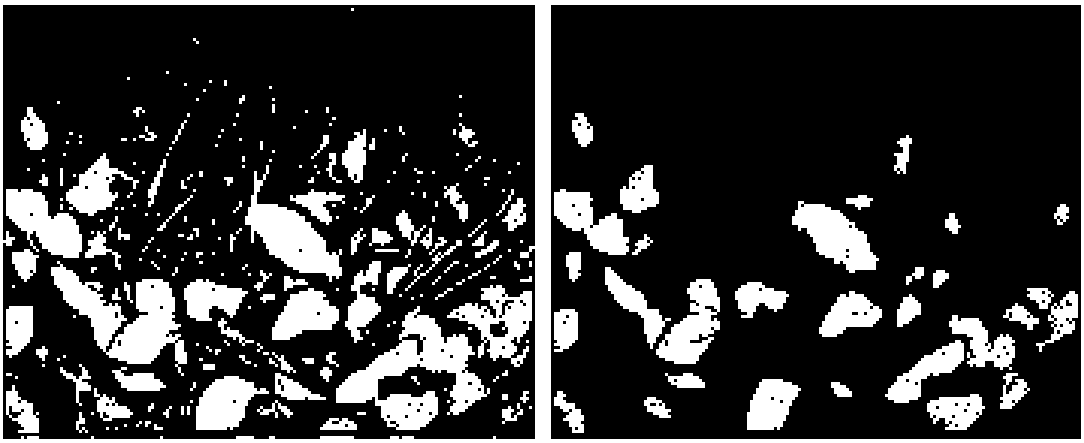


(f)

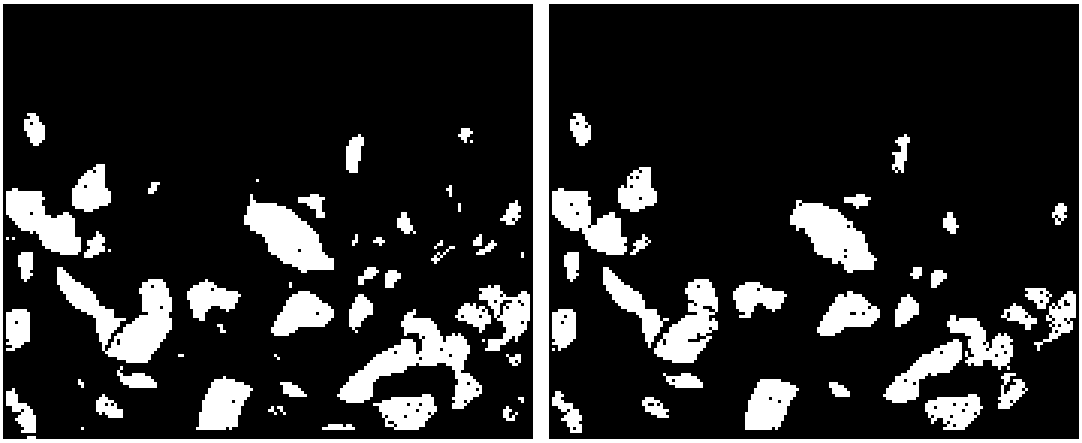
(g)



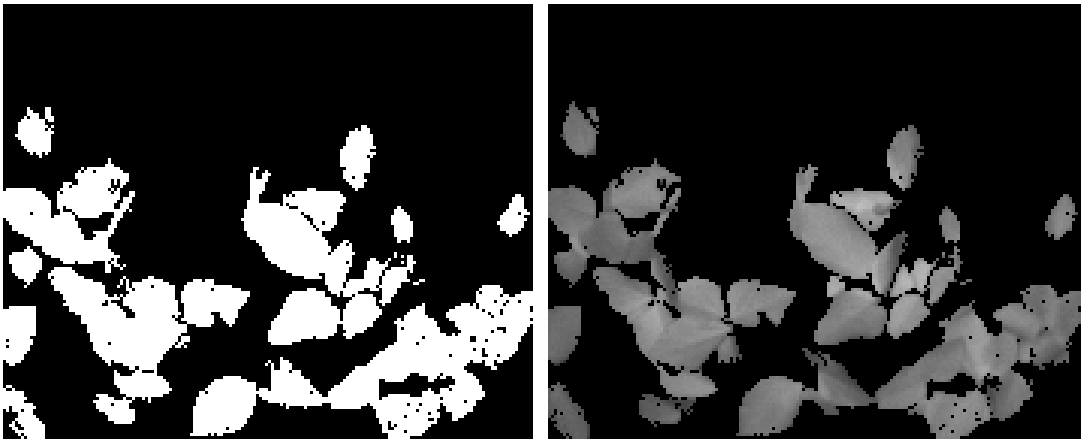
(h)



(i) (j)



(k) (l)



(m) (n)

Figure 6. Images of the green bean segmentation procedure: (a) amplitude image, (b) depth image, (c) amplitude gradient image G_I, (d) depth gradient image G_z, (e) neighbor

count image N_C, (f) curvature image δ, (g) amplitude percentile image API, (h–i) intermediate image *S1* and *S2*, (j–k) image *S11* and *S21*, (l) image *S12*, (m) Image *S13*, and (n) final green bean segmentation result image

5.1.2. Broccoli Segmentation Algorithm

The broccoli segmentation algorithms share great similarity with that of the green bean plant, but they still have some differences because of their various features. The detailed steps are introduced below:

1) Being similar to the first step of green bean segmentation, intermediate images *S1* and *S2* were constructed based on Equation 7 and 8. The threshold values of these two equations were picked based on the feature distribution discussion above. Compared to the first step of green bean segmentation, these two equations set a higher gradient threshold for both depth and amplitude, and they do not have any limitation in terms of the curvature of the 3D surface because of the inherent characteristics of the broccoli plant. As Figure 7a–g indicates, the 3D geometry of broccoli leaves is relatively complicated, instead of being simply flat, and this means that some broccoli area has high depth gradient and curvature value. Figure 7c also shows that the amplitude gradient of some parts of broccoli is relatively high compared to that of green bean leaves. Additionally, high *API* threshold value, which are 0.65 and 0.6 for generating *S1* and *S2*, respectively, were applied in Equation 7 and 8 for broccoli segmentation. This is because the relatively high *API* value resulted from the intensity of the broccoli plant's amplitude image, which is higher than that of the weeds and soil background. The resulting images *S1* and *S2* are shown in Figure 7h and i, respectively.

$$S1 = \begin{cases} 1, & when\ G_z \leq 150,\ N_c \geq 5, API \geq 0.65\ or \\ 0 & else \end{cases} \quad (7)$$

$$S2 = \begin{cases} 1, & when\ G_z \leq 220, G_I \leq 65,\ AN_c \geq 5, PI \geq 0.6\ or \\ 0 & else \end{cases} \quad (8)$$

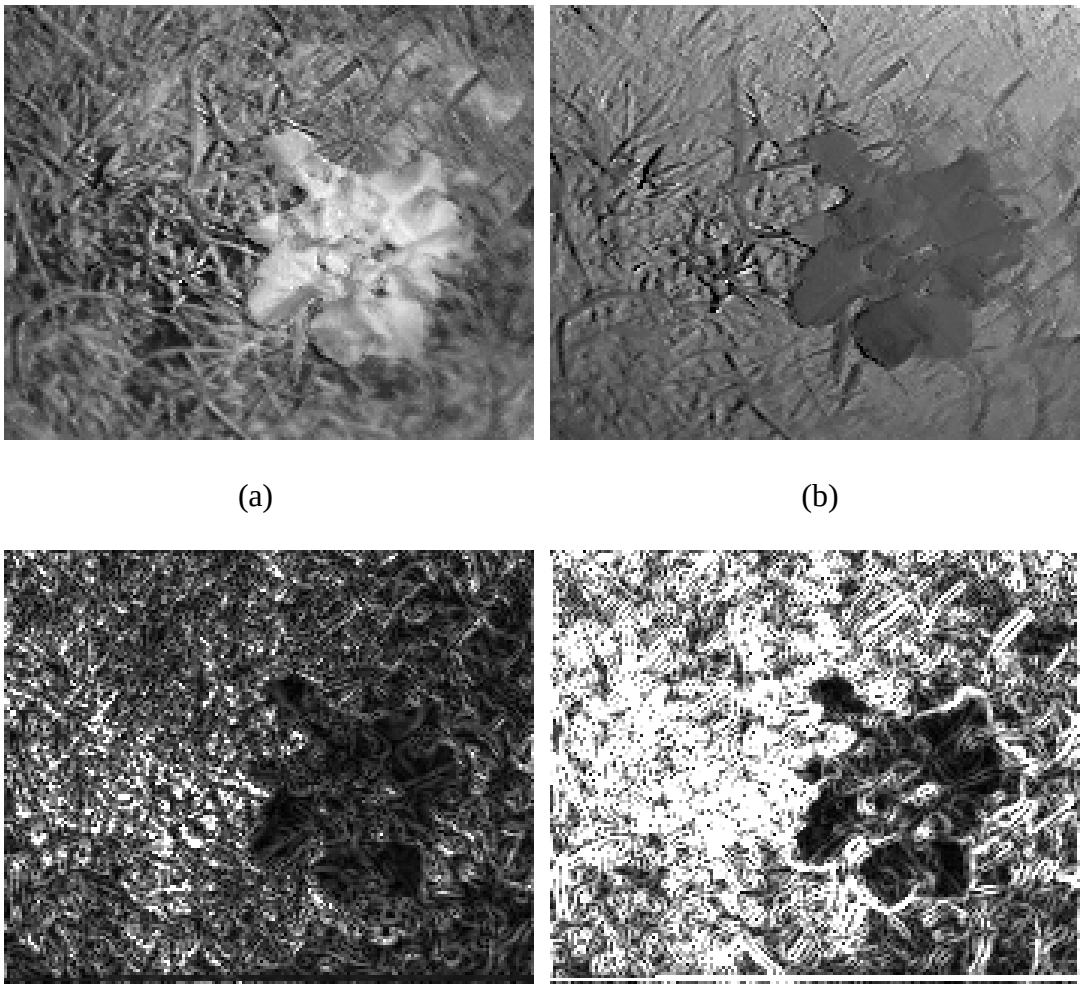
2) As Figure 7h and i show, the intermediate result images S1 and S2 have many small regions, which are noise after threshold. In order to get rid of the noise of S1, the opening operation, a 2D morphological processing method, was applied. The structuring element of the opening operator was a 3 × 3 square one. The result image of S1 after the opening operation is represented with S11, which is shown by Figure 7j. The figure indicates that the small regions are successfully removed. The intermediate image S11 can be separated into different regions in 2D space; regions with sizes smaller than 20 pixels were removed. The result image *S12* is shown by Figure 7k.
3) The intermediate result image *S12* did not contain a full broccoli plant. Most likely, the broccoli area extracted was only the inner part of a broccoli leaf. To recover the boundary area of the broccoli leaf and to get a more complete broccoli plant image, a method similar to the third step of the green bean segmentation algorithm was adopted. Each pixel of *S12* was checked to determine its neighbor point in 3D space, which was removed in *S12* but not in the *S2* image. If true, these qualified neighbor points were added into image *S12*. This operation was iterated to process new *S12* until there were no new points to be added. After this process, all of the points of the final *S12* image were separated into different regions in 3D space. Regions smaller than 30 points were removed because they were too small to be considered as parts of

a broccoli leaves. The final result of this step is represented with *S13*, as Figure 7l shows.

4) A method similar to the fourth step of the green bean segmentation algorithm was applied to further recover the boundary area of broccoli. Any background point with a distance to any foreground point of image *S13* that was smaller than 10 mm was added to generate a new *S13* image. This operation was iterated twice, and the final result is called image *S14*, which is shown in Figure 7m.
5) As Figure 7m shows, the broccoli image extracted in intermediate image *S14* is broken into unconnected pieces. This problem made it hard to analyze the broken pieces together as a whole plant. To solve this problem, a 2D dilation with a structural kernel of 3 × 3 square element was applied to process image *S14*. Being different from normal 2D dilation, only the pixels with *API* value over than 0.3 were added in this process. This is because the area with *API* value smaller than 0.3 is very likely to be soil and weed background, as discussed previously. This dilation process was iterated five times, and the result image is called *S15*. As Figure 7n indicates, the broken pieces of broccoli are merged together as a whole part. This procedure also took undesired soil or weed areas as part of the extracted broccoli image. The problem was solved in next step.
6) This study developed a method inspired by the flood fill algorithm to process intermediate image *S15* to remove the undesired background area from the extracted broccoli image. It is based on observation of the 3D geometry

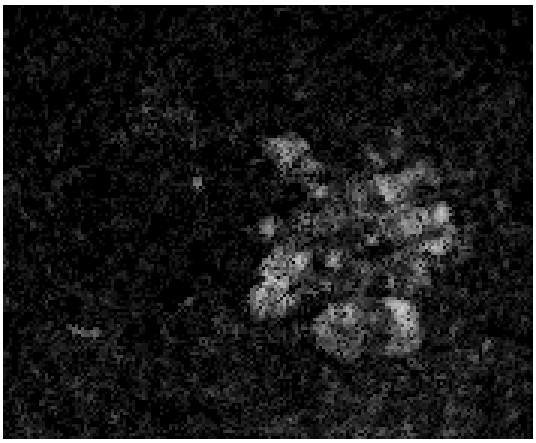
structure of the broccoli plant. First, the outer area of the broccoli plant is normally higher than the weeds and soil around. This makes the boundary area of broccoli like a “dam” that can prevent the “water” from flooding the inner part of broccoli for the flood fill algorithm. Second, the inner leaves are above the outer ones, and this means that there are several levels of “dam” from the outside to the inside area to protect the island (broccoli). Therefore, even if the outer leaves are covered by the “flood,” the inner leaves on top still have multiple chances to block the “water”. Because of the 3D structure of broccoli, the “flood” can easily cover the soil and weed background, but the broccoli image is relatively safe from “flood.” The difference between the algorithm in this study and the normal flood algorithm is that the “water” in this study flows to the neighbor pixel only when the height of the neighbor pixel is not more than 15 mm higher than that of the pixel where water it is. Moreover, the water continues to flood new areas unless the neighbor area is over 15 mm than where it is. In this study, all of the background pixels at the boarder of image S14 were used as the seed points of flood algorithm. The recursive algorithm will make the “water” continuously flood the new background area until no more change can be made. After this flood fill algorithm, broccoli and some other objects which were significantly higher than neighbor area were kept. By removing the point cloud region in 3D space whose size was smaller than 100 points, the objects other than broccoli can be reliably cleaned as Figure 7o shows. In Figure 7o, the white pixels represent the area covered by “flood,” and the black part is the extracted broccoli area,

which is safe from “water.” By rendering the amplitude value to the black area of Figure 7o, the final broccoli segmentation result is achieved as Figure 7p shows.



(a) (b)

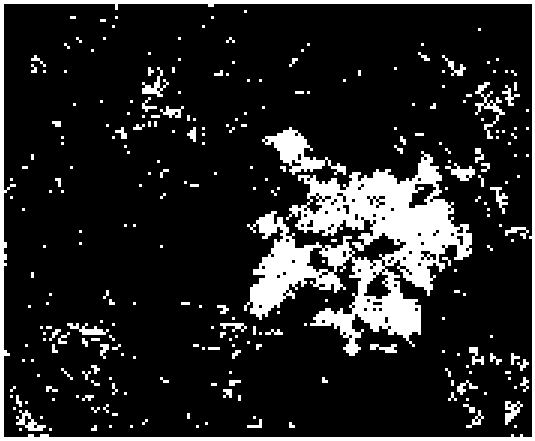
(c) (d)



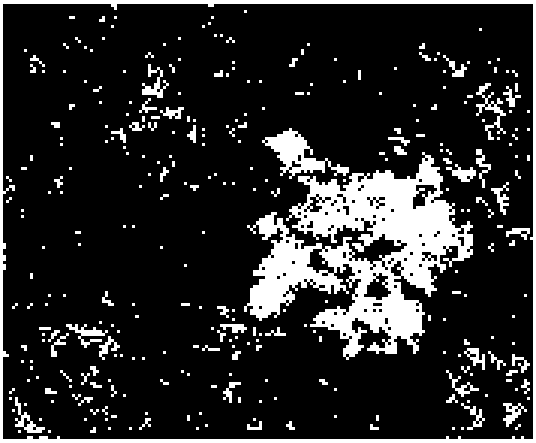
(e)

(f)

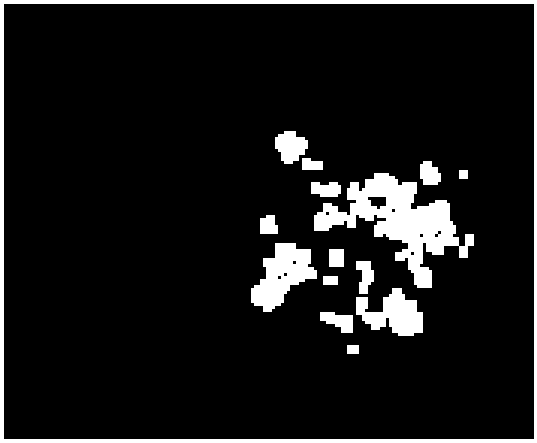
(g)



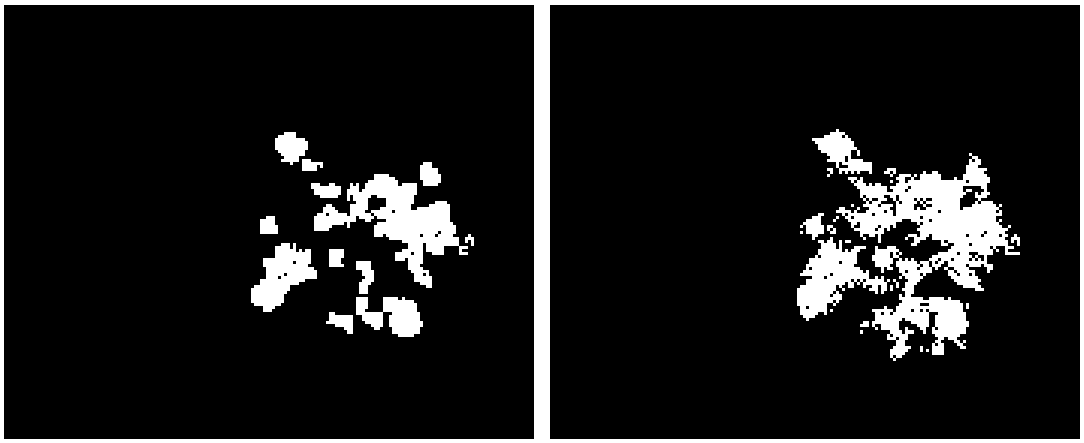
(h)



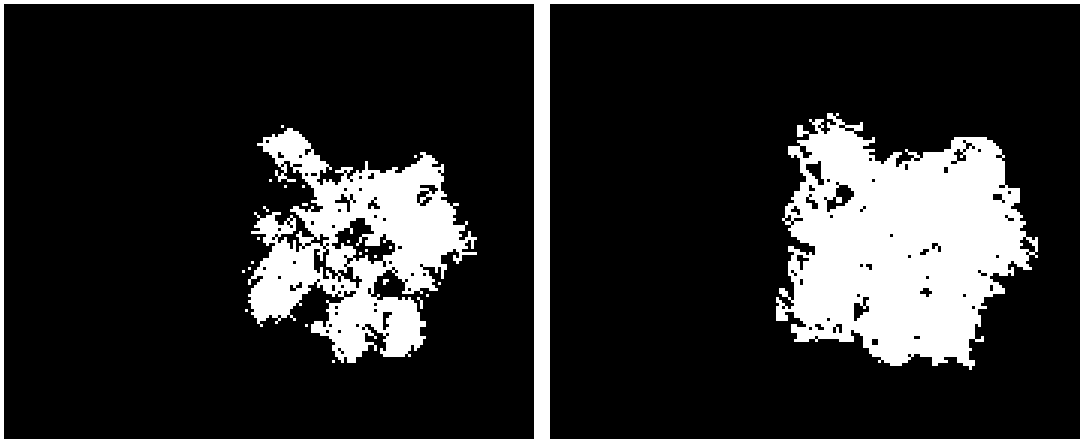
(i)



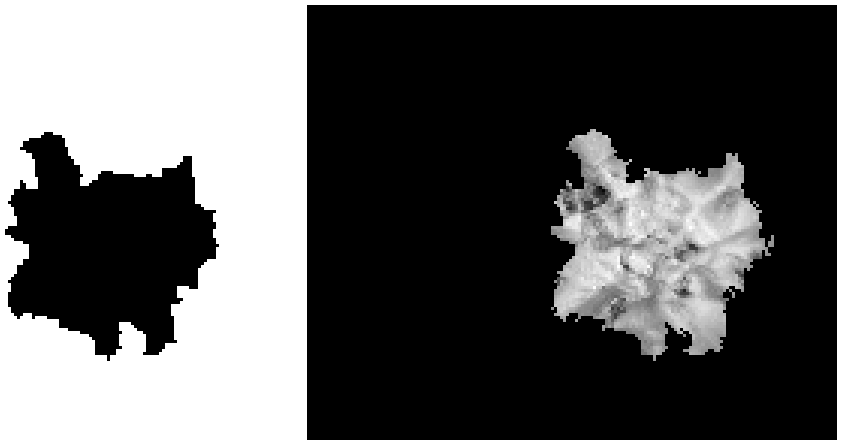
(j)



(k) (l)



(m) (n)



(o) (p)

Figure 7. Images of broccoli segmentation procedure: (a) amplitude image, (b) depth image, (c) amplitude gradient image G_I, (d) depth gradient image G_Z, (e) neighbor count

image N_C, (f) curvature image δ, (g) amplitude percentile image *API*, (h–i) intermediate images S1 and S2, (j) image S11, (k) image S12, (l) image S13, (m) image S14, (n) image S15, (o) result image of flood fill, and (p) final broccoli segmentation result image

6. Results & Discussion

Broccoli Recognition Result

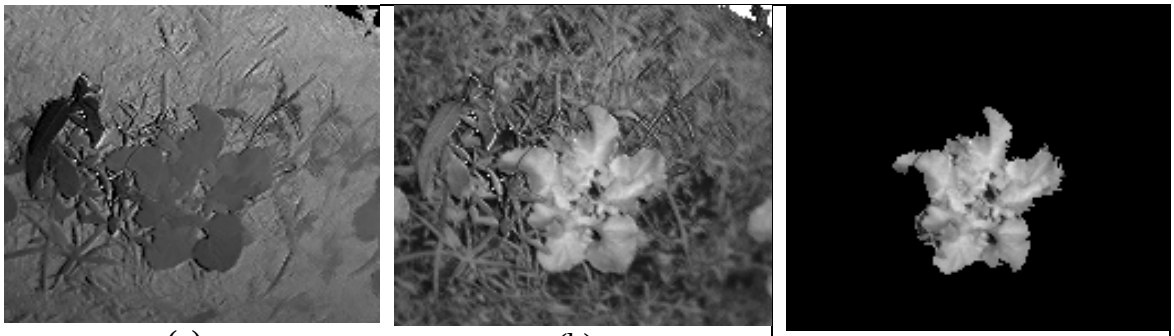
The broccoli plant recognition accuracy results are given in Table 4. As it indicates, from one image, the broccoli plant detection rate is over 84.3%, and the rate of the broccoli plant being recognized from at least one of its three images is over 88.3%. The false detection rate is less than 1.5%. As Table 4 shows, the recognition accuracy of the later image sets is slightly higher than those of earlier image sets, which indicates that broccoli plants at a higher growth stage are relatively easy for this system to recognize.

The major reasons for the broccoli plants being missed by this system were the plants' small size and the low resolution of the ToF camera. The diameter of all of the missed plants was less than 12 cm. As the resolution of the ToF camera used in this research was only 144 × 176 pixels, and the camera was over 0.6 m away from the top of the plants, the missed plants appeared too small in the captured image. The occlusion problem caused by weeds also led to miss detection.

The broccolis plants were extracted with relatively accurate shape as Figure 7 and 9 indicate. The broccoli plants can be segmented out with relatively precise and complete shape. Because of the small resolution of input images and occlusion caused by weeds, some small parts of broccoli plants may be missed by this system. However, the major areas of the broccoli plants were accurately kept in the segmentation result for all samples. Additionally, the segmentation result was clean, and in less than 1.5% of the randomly selected images had the background mistakenly recognized as a broccoli plant.

Table 4. Broccoli recognition accuracy rate

Data collection date	Height of broccoli (cm)	Broccoli recognition rate from one image (%)	Detection rate of one plant from 3 images (%)	False detection rate (%)
June 26, 2012	18–23	84.3	88.3	1.45
July 5, 2012	18–23	86.1	90.3	0.65
July 18, 2012	18–30	88.0	92.7	0.97



(a) (b) (c)

Figure 8. Broccoli segmentation result: (a) Depth image, (b) amplitude image from ToF camera, and (c) broccoli segmentation result

Green Bean Recognition Result

For the randomly selected 100 images of green bean plants, there were about 3938 green bean leaves according to the manual counting result, and 3593 of them were accurately detected by the system. Therefore, the detection rate of green bean leaves is 91.2%. Most of the leaves missed by the system were either too small or occluded. The main idea of the green bean detection algorithm is to search for the flat surface with relatively uniform amplitude value. Small or partially occluded leaves did not have enough pixels in the low resolution images captured by the ToF camera and can hardly be recognized as flat surface in this system; therefore, they were missed. Moreover, there was no weed or soil area recognized as green bean leaves for all of these 100 images.

The green bean leaves were extracted with relatively accurate shape by the segmentation algorithm of this research (Figures 6, 9). The big green bean leaves that are not occluded by weeds can always be extracted without any significant loss of area. However, for the small green bean leaves or the ones occluded, if they are not missed, their loss of fragment is more obvious in the segmentation result.

(a) (b) (c)

Figure 9. Green bean segmentation result: (a) depth image, (b) amplitude image from the ToF camera, and (c) green bean segmentation result

Processing Speed

Automated weeding application requires satisfactory processing speed for the crop plant detection task. The algorithm of this research is highly optimized. To test the processing speed of the system, it was run on a 3.4 GHz Intel Xeon CPU to process 1091 broccoli images and 2184 green bean images. According to the experiment result, the processing time range of one broccoli image was between 16.82 ms and 48.53 ms; the mean value was 26.82 ms; and the standard deviation was 6.6 ms. For green bean data, the processing time of one image ranged from 24.11 ms to 59.01 ms; the mean value was 32.68 ms; and the standard deviation was 5.34 ms. The average processing speed for broccoli and green bean were 37.29 and 30.60 frames per second (fps), respectively.

7. Conclusions

This research developed a green bean and broccoli plant detection system based on the use of a 3D ToF camera for automated weeding application. From the result of this research, it can be concluded that the 3D imaging based crop plant recognition exhibited promising potentials for automated robotic weeding application. First, the sparse noise filter of this research was effective and efficient. Second, the 2D and 3D features, including the gradient of amplitude and depth, surface curvature, *API*, normal direction, and neighbor point count in 3D space, were effective to discriminate broccoli and green bean plants from weeds and soil. Third, according to the 3D geometry and 2D amplitude characteristics of broccoli and green bean plants, a segmentation algorithm was developed for each crop. The detection rate of this system reached 88.3% and 91.2% for broccoli and green bean under weedy conditions, respectively. Crop plants that were not too small-sized in the images were extracted and recovered with their nearly complete canopies. Fourth, both 2D and 3D machine vision algorithms developed in this project were highly optimized, and the image processing speed of this system was over 30 fps for both types of crop plants.

It was found that using a low spatial resolution ToF camera is a limitation to achieving higher crop plant detection rate and segmentation accuracy. It led to a higher likelihood of missing small-sized broccoli plants and green bean leaves or those with earlier growing stage since they did not have enough pixels to allow accurately analyzing their 2D characteristics and 3D geometry. Moreover, the precision of captured 3D image data is relatively low. The noise of 3D image brought challenges to the 3D geometry analysis. A 3D image sensor with higher spatial resolution and lower noise level can help to improve the accuracy of the system.

Moreover, the 3D characters extracted in this research needed relatively flat leaf surfaces. The detection rate of broccoli is worse than that of green bean even though the leaves of broccoli are much larger. This is because that green bean leaves are much smoother, leading to more consistent 3D feature extraction.

In addition, the research found that strong sunlight could increase the noise level of 3D image captured by the ToF camera. The ToF camera requires a longer exposure time under strong sunlight, resulting in more motion blur issue in the depth image. A design which can effectively block sunlight will be very helpful to improve the 3D image quality, and therefore improving plant recognition accuracy.

8. Reference

Åstrand, B., & Baerveldt, A.-J. (2005). A vision based row-following system for agricultural field machinery. Mechatronics, 15(2), 251-269.

Chutia, D., Bhattacharyya, D., Sarma, K. K., Kalita, R., & Sudhakar, S. (2016). Hyperspectral remote sensing classifications: a perspective survey. Transactions in GIS, 20(4), 463-490.

Datta, A., & Knezevic, S. Z. (2013). Flaming as an alternative weed control method for conventional and organic agronomic crop production systems: a review. Advances in agronomy.

Du, J.-X., Wang, X.-F., & Zhang, G.-J. (2007). Leaf shape based plant species recognition. Applied Mathematics and Computation, 185(2), 883-893.

Ehsani, M., Upadhyaya, S. K., & Mattson, M. L. (2004). Seed location mapping using RTK GPS. Transactions of the ASAE, 47(3), 909-914.

Fennimore, S. A., Slaughter, D. C., Siemens, M. C., Leon, R. G., & Saber, M. N. (2016). Technology for Automation of Weed Control in Specialty Crops. Weed Technology, 30(4), 823-837.

Franco, C., Pedersen, S. M., Papaharalampos, H., & Ørum, J. E. (2017). The value of precision for image-based decision support in weed management. Precision Agriculture, 18(3), 366-382.

Furbank, R. T., & Tester, M. (2011). Phenomics - technologies to relieve the phenotyping bottleneck. Trends in Plant Science, 16(12), 635-644.

Gai, J., Tang, L., & Steward, B. (2016). Plant Localization and Discrimination using 2D+ 3D Computer Vision for Robotic Intra-row Weed Control. Paper presented at the 2016 ASABE Annual International Meeting.

Griepentrog, H.-W., Nørremark, M., Nielsen, H., & Blackmore, B. (2005). Seed mapping of sugar beet. Precision Agriculture, 6(2), 157-165.

Grinblat, G. L., Uzal, L. C., Larese, M. G., & Granitto, P. M. (2016). Deep learning for plant identification using vein morphological patterns. Computers and Electronics in Agriculture, 127, 418-424.

Jeschke, P. (2016). Progress of modern agricultural chemistry and future prospects. Pest management science, 72(3), 433-455.

Jin, J., & Tang, L. (2009). Corn Plant Sensing Using Real-Time Stereo Vision. Journal of Field Robotics, 26(6-7), 591-608.

Jones, H. G., & Vaughan, R. A. (2010). Remote sensing of vegetation: principles, techniques, and applications: Oxford University Press, Oxford, UK.

Jørgensen, M., Søgaard, H. T., & Nielsen, P. S. (2002). Præcision ved automatisk styring af radsensere. Grøn Viden-Markbrug(268).

Kapur, J., Sahoo, P. K., & Wong, A. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. Computer Vision, Graphics, and Image Processing, 29(3), 273-285.

Kise, M., Zhang, Q., & Rovira Más, F. (2005). A stereovision-based crop row detection method for tractor-automated guidance. Biosystems Engineering, 90(4), 357-367.

Kittler, J., & Illingworth, J. (1986). Minimum error thresholding. Pattern recognition, 19(1), 41-47.

Kumar, L., Schmidt, K., Dury, S., & Skidmore, A. (2002). Imaging spectrometry and vegetation science Imaging spectrometry (pp. 111-155): Springer.

Kumar, T. P., Reddy, M. V. P., & Bora, P. K. (2017). Leaf Identification Using Shape and Texture Features. Paper presented at the Proceedings of International Conference on Computer Vision and Image Processing.

Lee, S. H., Chan, C. S., Mayo, S. J., & Remagnino, P. (2017). How deep learning extracts and learns leaf features for plant classification. Pattern Recognition, 71, 1-13.

Mohler, C. L. (2001). Mechanical management of weeds. Ecological management of agricultural weeds, 139-209.

Nagasaka, Y., Umeda, N., Kanetai, Y., Taniwaki, K., & Sasaki, Y. (2004). Autonomous guidance for rice transplanting using global positioning and gyroscopes. Computers and Electronics in Agriculture, 43(3), 223-234.

Nakarmi, A. D., & Tang, L. (2012). Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Computers and Electronics in Agriculture, 82, 23-31.

Noh, H., Zhang, Q., Han, S., Shin, B., & Reum, D. (2005). Dynamic calibration and image segmentation methods for multispectral imaging crop nitrogen deficiency sensors. Transactions of the ASAE, 48(1), 393-401.

Otsu, N. (1975). A threshold selection method from gray-level histograms. Automatica, 11(285-296), 23-27.

Pauly, M., Gross, M., & Kobbelt, L. P. (2002). Efficient simplification of point-sampled surfaces. Paper presented at the Proceedings of the conference on Visualization'02.

Reiser, D., Vázquez Arellano, M., Garrido Izard, M., Griepentrog, H. W., & Paraforos, D. S. (2016). Using Assembled 2D LiDAR Data for Single Plant Detection.

Slaughter, D., Giles, D., & Downey, D. (2008). Autonomous robotic weed control systems: A review. Computers and electronics in agriculture, 61(1), 63-78.

Slaughter, D. C., Giles, D. K., & Downey, D. (2008). Autonomous robotic weed control systems: A review. Computers and Electronics in Agriculture, 61(1), 63-78.

Søgaard, H. T., & Olsen, H. J. (2003). Determination of crop rows by image analysis without segmentation. Computers and electronics in agriculture, 38(2), 141-158.

Thompson, J., Stafford, J., & Miller, P. (1991). Potential for automatic weed detection and selective herbicide application. Crop protection, 10(4), 254-259.

Vázquez-Arellano, M., Griepentrog, H. W., Reiser, D., & Paraforos, D. S. (2016). 3-D Imaging Systems for Agricultural Applications—A Review. Sensors, 16(5), 618.

Vigneault, C., & Benoît, D. L. (2001). Electrical weed control: theory and applications Physical control methods in plant protection (pp. 174-188): Springer.

Vrindts, E., & Baerdemaeker, J. d. (1997). Optical discrimination of crop, weed and soil for on-line weed detection.

Weiss, U., & Biber, P. (2011). Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. Robotics and autonomous systems, 59(5), 265-273.

Weiss, U., Biber, P., Laible, S., Bohlmann, K., & Zell, A. (2010). Plant species classification using a 3d lidar sensor and machine learning. Paper presented at the Ninth International Conference on Machine Learning and Applications (ICMLA).

Xiang, R., Jiang, H., & Ying, Y. (2014). Recognition of clustered tomatoes based on binocular stereo vision. Computers and electronics in agriculture, 106, 75-90.

Zhang, Y., Slaughter, D. C., & Staab, E. S. (2012). Robust hyperspectral vision-based classification for multi-season weed mapping. ISPRS journal of photogrammetry and remote sensing, 69, 65-73.