

# Improved vegetation segmentation with ground shadow removal using an HDR camera

Hyun K. Suh<sup>1</sup> · Jan Willem Hofstee<sup>1</sup> · Eldert J. van Henten<sup>1</sup>

Published online: 3 March 2017

© The Author(s) 2017. This article is published with open access at Springerlink.com

**Abstract** A vision-based weed control robot for agricultural field application requires robust vegetation segmentation. The output of vegetation segmentation is the fundamental element in the subsequent process of weed and crop discrimination as well as weed control. There are two challenging issues for robust vegetation segmentation under agricultural field conditions: (1) to overcome strongly varying natural illumination; (2) to avoid the influence of shadows under direct sunlight conditions. A way to resolve the issue of varying natural illumination is to use high dynamic range (HDR) camera technology. HDR cameras, however, do not resolve the shadow issue. In many cases, shadows tend to be classified during the segmentation as part of the foreground, i.e., vegetation regions. This study proposes an algorithm for ground shadow detection and removal, which is based on color space conversion and a multilevel threshold, and assesses the advantage of using this algorithm in vegetation segmentation under natural illumination conditions in an agricultural field. Applying shadow removal improved the performance of vegetation segmentation with an average improvement of 20, 4.4, and 13.5% in precision, specificity and modified accuracy, respectively. The average processing time for vegetation segmentation with shadow removal was 0.46 s, which is acceptable for real-time application (<1 s required). The proposed ground shadow detection and removal method enhances the performance of vegetation segmentation under natural illumination conditions in the field and is feasible for real-time field applications.

**Keywords** Image processing · Vegetation segmentation · High dynamic range · Shadow detection and remove · Weed control

---

✉ Hyun K. Suh  
davidsh79@gmail.com

<sup>1</sup> Farm Technology Group, Wageningen University, P.O. Box 16, 6700 AA Wageningen, The Netherlands

## Introduction

This work was part of the EU-funded project SmartBot, a project with the research goal to develop a small-sized vision-based robot for control of volunteer potato (weed) in a sugar beet field. Such a vision-based weed control robot for agricultural field application requires robust vegetation segmentation, i.e. a vegetation segmentation that has good performance under a wide range of circumstances. The output of vegetation segmentation is the fundamental element in the subsequent process of weed and crop discrimination as well as weed control (Meyer and Camargo Neto 2008; Steward et al. 2004). There are two challenging issues for robust vegetation segmentation in agricultural field conditions: (1) to overcome the strongly varying natural illumination (Jeon et al. 2011; Wang et al. 2012); (2) to avoid the influence of shadows under direct sunlight conditions (Guo et al. 2013; Zheng et al. 2009).

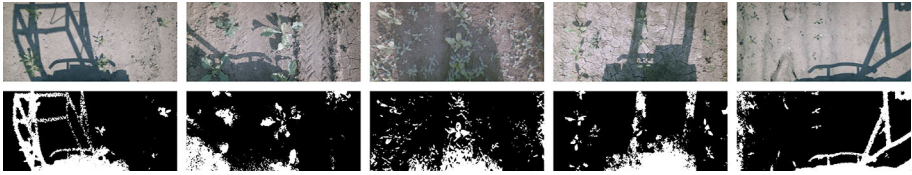
Illumination conditions constantly change in an agricultural field environment depending on the sky and weather conditions. These illumination variations greatly affect Red–Green–Blue pixel values of acquired field images and lead to the inconsistent color representation of plants (Sojodishijani et al. 2010; Teixidó et al. 2012). In addition, shadows often create extreme illumination contrast, causing substantial luminance differences within a single image scene. These extreme intensity differences make vegetation segmentation a very challenging task.

Researchers addressed the above two problems by using a hood covering both the scene and the vision acquisition device. By doing so, any ambient visible light was blocked (Ahmed et al. 2012; Åstrand and Baerveldt 2002; Haug et al. 2014; Lee et al. 1999). Constant illumination under the cover was then obtained using artificial lighting (Nieuwenhuizen et al. 2010; Polder et al. 2014).

Such a solution was not feasible within the framework of the Smartbot project because a small-sized mobile robotic platform was to be used. An extra structure for the cover was not viable due to the reduced carrying capacity of the platform. Moreover, using additional energy for artificial lighting would be another critical issue, considering the mobile platform was battery operated. Therefore, a solution was needed that uses the ambient light while overcoming the drawbacks mentioned earlier.

A way to resolve the issue of varying natural illumination and substantial intensity differences within a single image scene is to use high dynamic range (HDR) camera technology as has been indicated by a number of studies (Graham 2011; Hrabar et al. 2009; Irie et al. 2012; Mann et al. 2012; Slaughter et al. 2008). Under direct sunlight conditions, the dynamic range of the scene is much larger than a traditional non-HDR camera covers, especially if an image scene contains shadows (Dworak et al. 2013). Having a larger dynamic range, an HDR camera enables the capture of stable and reliable images even under strong and direct solar radiation or under faint starlight (Reinhard et al. 2010).

HDR cameras, however, do not resolve the shadow issue. When the scene in the field is not covered by a hood or similar structure, shadows are inevitable. In many cases in vegetation segmentation, shadows tend to be classified as part of the foreground, i.e., as vegetation regions (Fig. 1). Therefore, shadows need to be detected and preferably removed for better segmentation performance. However, shadow detection is extremely challenging especially in an agricultural field environment because shadows change dramatically throughout the day depending on position and intensity of the sun. Besides, shadows have no regular shape, size, or texture, and can even be distorted on an uneven ground surface. In recent years, many shadow detection and removal algorithms have been



**Fig. 1** Example of shadow images (*top*), and vegetation segmentation output with excess green (ExG) segmentation (*bottom*). Shadows are partially segmented as vegetation

proposed in the computer vision research area using a feature-based or a brightness based compensation (Sanin et al. 2012). However, these shadow detection and removal algorithms are difficult to implement and require a significant amount of computation time, which is an important issue for real-time field applications. Moreover, these algorithms provide poor shadow removal for outdoor scenes (Al-Najdawi et al. 2012). Therefore, a simple and effective shadow detection and removal algorithm is needed for real-time weed detection and control application in an agricultural field environment.

This paper proposes an algorithm for ground shadow detection and removal, and assesses the effectiveness of using this algorithm in vegetation segmentation under natural illumination conditions in an agricultural field. The proper quantitative measure to evaluate the performance of vegetation segmentation are discussed.

## Materials and methods

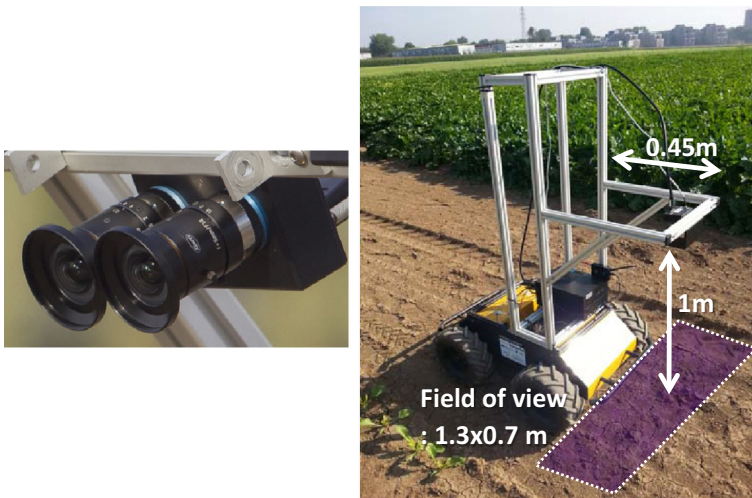
### High dynamic range (HDR) camera

A common definition of the dynamic range of an image is the ratio of the maximum and minimum illuminance in a given scene. More precisely, Bloch (2007) defines dynamic range as the logarithmic ratio between the largest and the smallest readable signal (an image is treated as a signal from the camera hardware):

$$\text{Dynamic Range (dB)} = 20 \times \log_{10} \frac{\text{Max Signal}}{\text{Min Signal}} \quad (1)$$

The illumination difference in a real-life image scene can easily exceed a dynamic range of 80 dB. In outdoor field conditions, the dynamic range can exceed 120 dB (Radonjić et al. 2011). Human eyes have a dynamic range of around 200 dB, while a conventional imaging device such as a non-HDR CCD digital camera typically has a dynamic range of around 60 dB (Bandoh et al. 2010; Ohta 2007). Under direct sunlight conditions, the dynamic range of the scene can be much higher than a traditional non-HDR camera can cover, especially when the image scene contains sharp dark shadows. Thus, a conventional imaging device is not feasible for machine vision applications in a natural agricultural environment, because strong direct solar radiation and shadows frequently cause extreme lighting intensity changes. Piron et al. (2010) used an exposure fusion method to generate a high dynamic range scene in plant images and reported that high dynamic range acquisition supported obtaining a quality image of the scene with a strong signal to noise ratio. In the past few years, HDR cameras have become commercially available at an affordable price.

In this study, a HDR camera (NSC1005c, New Imaging Technologies, Paris, France) having a dynamic range of 140 dB and a bit depth of 36 bits per pixel was used (Fig. 2).



**Fig. 2** Field images were acquired with an HDR camera (*left*) which was mounted at a height of 1 m viewing perpendicular to the ground surface, resulting in a field of view:  $1.3 \times 0.7$  m (*right*)

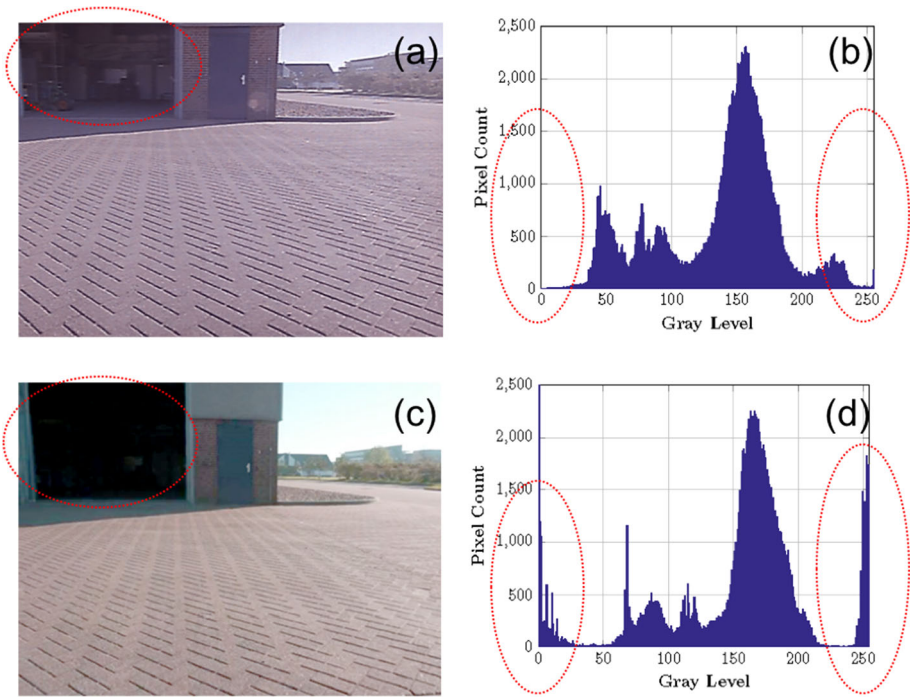
This camera has two identical CMOS sensors providing the stereo images (left and right), but only the left sensor image was used in this study.

Example images of a similar scene made with the HDR and traditional non-HDR CCD camera are shown in Fig. 3. The HDR camera captures the objects even in the dark shadow region (Fig. 3a) where as a traditional non-HDR CCD camera (Sony NEX-5R) captures no objects but produces black pixels (Fig. 3c). The histogram of the HDR image is well balanced across the darkest and lightest margins (Fig. 3b) while the histogram of a traditional non-HDR CCD camera image is imbalanced with peaks both on the left and right edges due to clipping (Fig. 3d).

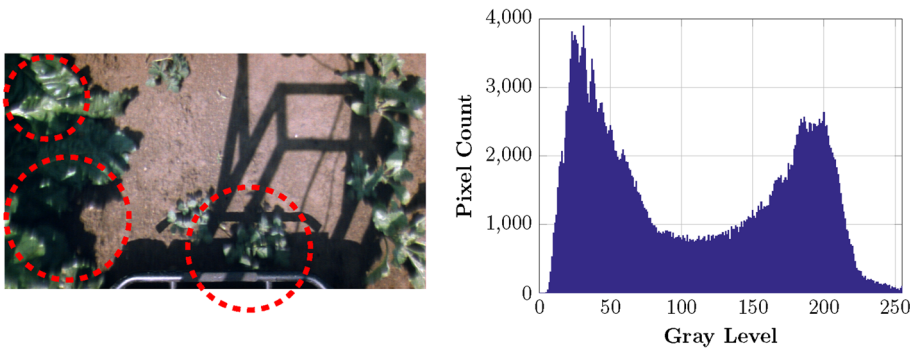
An example field image that was acquired with an HDR camera under very bright sunny conditions is shown in Fig. 4. Some pixels in the green leaves were bright due to specular reflection; while some pixels in the shadow region were very dark. The extreme lighting intensity difference with a high dynamic range is often found in a field image scene. In such a condition in the field, a conventional non-HDR imaging device would not be able to adequately capture the objects in both the bright regions as well as in the dark shadow regions but an HDR camera does adequately capture these objects under these lighting conditions.

### Algorithm—ground shadow detection and removal

As was shown in Fig. 1, shadows in agricultural field images are often classified as part of vegetation when applying a commonly used vegetation segmentation method based on the excessive green index ( $2g-r-b$ ), ExG (Woebbecke et al. 1995). To further process the shadows, a ground shadow detection algorithm was developed using color space conversion. Color pixel values in RGB space can be highly influenced by the illumination conditions because illumination and color parts are not separated in this color representation (Florczyk 2005). Using a different color space (or conversion of the RGB image to another color space) that uses a color representation separating color and illumination, pixel values are less influenced by the illumination conditions with the shadow detection



**Fig. 3** An example outdoor image scene on a sunny day: (a) HDR camera image with (b) image histogram (c) traditional non-HDR CCD camera (Sony NEX-5R, ISO 100, 1/80, f/11, dynamic range optimizer activated) image with (d) image histogram. The *red ellipses* indicate that (b) the histogram of the HDR image is well balanced across the darkest and lightest margins, but (d) the histogram of a traditional non-HDR CCD camera image is imbalanced with peaks both on the left and right edges due to clipping



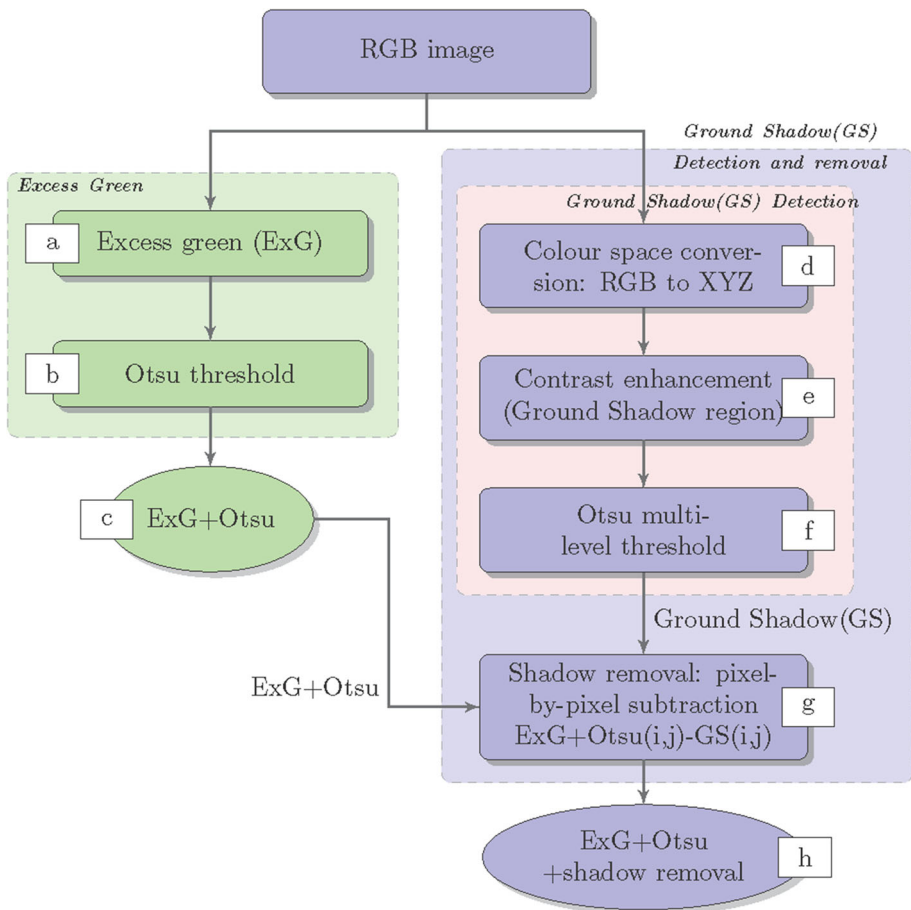
**Fig. 4** HDR camera image in a sugar beet field (*left*) and its image histogram (*right*). The *red circles* indicate that some pixels in the green leaves are bright due to specular reflection; while some pixels in the shadow region are very dark

procedure. Many studies have shown that a color space conversion approach is simple to implement and computationally inexpensive; thus, is very useful for real-time field applications (Sanin et al. 2012).

In this study, the XYZ color space was chosen because the normalized form of this color space separates luminance from color (or rather from chromaticity). Also this color space is based on how a human would perceive light (Pascale 2003). The XYZ system provides a standard way to describe colors and contains all real colors (Corke 2011). Besides, this particular color space has been shown to be robust under illumination variations (Lati et al. 2013).

The procedure used for ground shadow detection and removal is shown in Fig. 5. Two main processes are shown: (1) ExG with Otsu (1979) threshold (Fig. 5 steps a–c), and (2) ground shadow detection and removal (Fig. 5 steps d–h).

The left column in Fig. 5, steps (a)–(c), shows the conventional vegetation segmentation procedure. ExG, one of the most commonly used methods, was used in this study to compare the performance of vegetation segmentation before and after shadow removal because ExG showed good performance in most cases in our preliminary studies. The Otsu threshold was used because the Otsu method showed good performance in a preliminary study.



**Fig. 5** Flow diagram of ground shadow detection and removal algorithm

The right column in Fig. 5, step (d)–(h), shows the ground shadow detection and removal procedure. The three individual steps (d)–(f) are referred to as a ground shadow detection, and pixel-by-pixel subtraction in step (g) is referred to as ground shadow removal. The detected ground shadow region was subtracted from ExG with Otsu threshold (ExG + Otsu) which resulted from step (c). Then, the shadow-removed image (Fig. 5h) was compared with ExG + Otsu (Fig. 5c) to evaluate the performance improvement when using vegetation segmentation after shadow detection and removal.

The details of the algorithms are described below

#### Step a: excess green (ExG)

The excess green index ( $ExG = 2g - r - b$ ) was applied to the RGB image (Woebbecke et al. 1995). The normalized r, g and b components, in the range [0,1], were obtained with (Gée et al. 2008):

$$r = \frac{R_n}{R_n + G_n + B_n}, g = \frac{G_n}{R_n + G_n + B_n}, b = \frac{B_n}{R_n + G_n + B_n} \quad (2)$$

where  $R_n$ ,  $G_n$  and  $B_n$  are the normalized RGB co-ordinates ranging from 0 to 1. They were obtained as follows:

$$R_n = \frac{R}{R_{max}}, G_n = \frac{G}{R_{max}}, B_n = \frac{B}{B_{max}} \quad (3)$$

where  $R_{max} = G_{max} = B_{max} = 255$ .

#### Step b: Otsu threshold

The Otsu threshold method was applied to obtain an optimum threshold value. The pixels of the image were divided into two classes:  $C_0$  for  $[0, \dots, t]$  and  $C_1$  for  $[t + 1, \dots, L]$ , where  $t$  was the threshold value ( $0 \leq t < L$ ) and  $L$  was the number of distinct intensity levels. An optimum threshold value  $t^*$  was chosen by maximizing the between-class variances,  $\sigma_B^2$  (Otsu 1979):

$$t^* = \arg \max_{0 \leq t < L} \{\sigma_B^2(t)\} \quad (4)$$

#### Step c: ExG + Otsu

Using the optimum threshold value  $t^*$  (Eq. 4), vegetation pixels were classified:

$$\begin{cases} \text{Background region; if } ExG(i,j) < t^* \\ \text{Vegetation region; if } ExG(i,j) \geq t^* \end{cases} \quad (5)$$

where  $ExG(i,j)$  was the ExG value of the pixel (i,j).

#### Step d: color space conversion

The first step involved color space conversion. The RGB values were converted to the 1931 International Commission on Illumination (CIE) XYZ space using the following matrix (Lati et al. 2013):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6}$$

where R, G and B were pixel values in RGB color space in the range [0, 255]. X,Y and Z were pixel values in XYZ color space. Finally, XYZ values were then normalized using the following equation:

$$x = \frac{X}{X + Y + Z}, y = \frac{Y}{X + Y + Z}, z = \frac{z}{X + Y + Z} \tag{7}$$

*Step e: contrast enhancement of the ground shadow region*

The contrast of the ground shadow region was enhanced from the rest of the image. This contrast enhancement was achieved by dividing the product of the chromaticity values *x* and *y* by *z*:

$$\text{Contrasted Ground Shadow (CGS}(i,j)) = \frac{x(i,j) \cdot y(i,j)}{z(i,j)} \tag{8}$$

where *CGS*(*i,j*) was the contrasted pixel (*i,j*) of the ground shadow, and *x*(*i,j*), *y*(*i,j*), *z*(*i,j*) were normalized values of the pixel (*i,j*) in the XYZ color space.

*Step f: Otsu multi-level threshold*

The Otsu multi-level threshold method was applied to the image based on the observation that the shadow image contained three components—ground shadow, plant material and soil background. The previous steps (d) and (e) made the ground shadow region more distinct from other components. Thus, the Otsu multi-level threshold enabled to separate the ground shadow region, which had the lowest intensity level, from plant material and soil background. The lowest intensity level was selected as the ground shadow region, but plant material and soil background regions were not separated because they were not clearly distinct from each other.

All pixels of the image obtained in the previous step (e) were divided into the following three classes: *C*<sub>0</sub> for [0,...,*t*<sub>1</sub>], *C*<sub>1</sub>, for [*t*<sub>1</sub> + 1,...,*t*<sub>2</sub>], and *C*<sub>2</sub> for [*t*<sub>2</sub> + 1,...,*L*], where *t*<sub>1</sub> and *t*<sub>2</sub> were threshold values (0 ≤ *t*<sub>1</sub> < *t*<sub>2</sub> < *L*), and *L* was the total number of distinct intensity levels. An optimal set of threshold values *t*<sub>1</sub><sup>\*</sup> and *t*<sub>2</sub><sup>\*</sup> was chosen by maximizing the between-class variances, σ<sub>B</sub><sup>2</sup> (Otsu 1979):

$$\{t_1^*, t_2^*\} = \arg \max_{0 \leq t_1 < t_2 < L} \{\sigma_B^2(t_1, t_2)\} \tag{9}$$

For the ground shadow detection, the optimal threshold value *t*<sub>1</sub><sup>\*</sup> was used. The threshold value *t*<sub>2</sub><sup>\*</sup> was ignored since it did not have any added value in this ground shadow detection process. Consequently, the ground shadow pixels were classified in two classes as follows:

$$\begin{cases} \text{Ground shadow (GS)} & \text{if } CGS(i,j) < t_1^* \\ \text{Non Ground shadow (NGS)} & \text{if } CGS(i,j) \geq t_1^* \end{cases} \tag{10}$$



### *Step g: Ground shadow removal by subtraction*

Once the ground shadow region was identified, the shadow-removed image was generated by a pixel-by-pixel subtraction from the ExG + Otsu. The shadow-removed pixel values were simply the values of ExG minus the corresponding pixel values from the ground shadow region image (Eq. 11).

$$F(i,j) = \text{ExG}(i,j) - \text{GS}(i,j) \quad (11)$$

where  $F(i,j)$  was the shadow-removed pixel  $(i,j)$ , and  $\text{GS}(i,j)$  was the detected ground shadow pixel  $(i,j)$ .

### **Field image collection**

For crop image acquisition, the HDR camera was mounted at a height of 1 m viewing perpendicular to the ground surface on a custom-made frame carried by a mobile platform (Husky A200, Clear path, Canada), as was shown in Fig. 2. The camera was equipped with two identical Kowa 5 mm lenses (LM5JC10 M, Kowa, Japan) with a fixed aperture. The camera was set to operate in automatic acquisition mode with automatic point and shoot, having an image resolution of  $1280 \times 580$  pixels per image of left and right sensors. The ground-covered area was  $1.3 \times 0.7$  m, corresponding to three sugar beet crop rows. The acquisition program was implemented in LabVIEW (National Instruments, Austin, USA) to acquire five images per second. Field images were taken while the mobile platform was manually controlled with a joystick and driven along crop rows using a controlled traveling speed of 0.5 m/s.

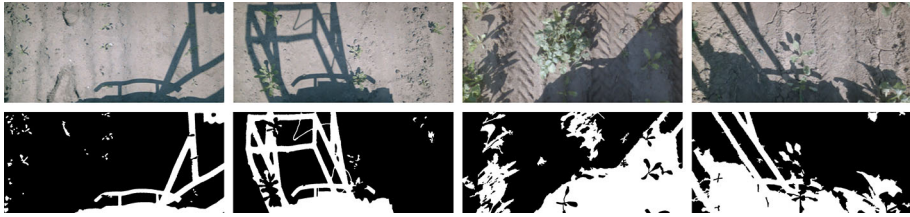
Sugar beet was seeded three times (Spring, Summer, and Fall) in 2013 and 2014 in two different soil types (sandy and clay soil) on the Unifarm experimental sites in Wageningen, The Netherlands. Crop images were acquired under various illumination and weather conditions on several days in June, August and October of 2013 as well as in May, July and September of 2014.

### **Image dataset**

The following image datasets were chosen for this study: (1) Set 1: only containing images with shadow to purely test and evaluate the performance of the shadow detection algorithm against human generated ground truth, and (2) Set 2: containing a mix of images with and without shadows to assess the effectiveness of shadow removal on segmentation.

Set 1 consisted of 30 field images that all contained shadows ranging from shallow to dark with various shadow shapes (Fig. 6). The images in this set were acquired on several days under various weather conditions at different growth stages of the crop. Ground truth images for shadow regions was manually generated using Photoshop CC (Adobe Systems Inc., San Jose, USA).

For Set 2, a total of 110 field images was selected from all acquired field images. During the selection of this set, a wide range of natural conditions was considered, including different stages of plant growth, various illumination conditions from a cloudy dark to sunny bright day conditions and extreme illumination scenes caused by strong direct solar radiation and shadows. Half of the images in this set contained shadows (from shallow to dark shadows) while the other half contained no shadows. Vegetation regions were



**Fig. 6** Example images in Set 1 (*top*) and its human-labeled ground truth of ground shadow regions (*bottom*)

manually labeled for ground truth using Photoshop CC. All images were processed with the Image Processing Toolbox in Matlab (The MathWorks Inc., Natick, USA).

### Quantitative performance measures of vegetation segmentation

The segmentation results were compared and evaluated at pixel level with human-labeled ground truth images. In this study, a set of quantitative measures based on the confusion matrix (Table 1) was used to assess the performance of the vegetation segmentation. Positive prediction value (precision), true-positive rate (recall or sensitivity), true-negative rate (specificity) and modified accuracy (MA) were used. Each of these has a different goal to measure, thus assessing above measures altogether helps to evaluate the performance of vegetation segmentation in a balanced way. The details of the measures are described below (Benzeth et al. 2008; Metz 1978; Prati et al. 2003):

$$Precision (Positive predict value) = \frac{TP}{TP + FP} \tag{12}$$

$$Recall (True positive rate or sensitivity) = \frac{TP}{TP + FN} \tag{13}$$

$$Specificity (True negative rate) = \frac{TN}{TN + FP} \tag{14}$$

where TP is true-positive; FP is false-positive; TN is true-negative; FN is false-negative.

Precision indicates how many of the positively segmented pixels are relevant. Precision refers to the ability to minimize the number of false-positives. Recall indicates how well a segmentation performs in detecting the vegetation and thus relates to the ability to correctly detect vegetation pixels that belong to the vegetation region (true-positive). Specificity, on the other hand, specifies how well the segmentation algorithm performs in avoiding false-positive error, which also indicates the ability to correctly

**Table 1** Confusion matrix

	Algorithm	
	Vegetation	Background
Ground truth		
Vegetation	TP	FN
Background	FP	TN

*TP* true-positive, *TN* true-negative, *FP* false-positive, and *FN* false-negative

detect non-vegetation pixels that belong to non-vegetation regions (true-negative). A single measure above does not fully reflect the performance of vegetation segmentation because each one can have a biased value under certain conditions. For example, if a segmentation produces vegetation pixels only when there are strong green components, precision will have a higher value (close to 1) in a poorly segmented image. Moreover, if a segmentation always identifies all the pixels as vegetation, recall will attain large values.

Accuracy is commonly used as a single representative performance indicator in the literature. However, this measure has a drawback if there is a significant imbalance between vegetation and background (Bac et al. 2013; Rosin and Ioannidis 2003). An alternative way to measure the performance would be balanced accuracy, i.e. the average of sensitivity and specificity. However, this measure can also provide a biased value if segmentation output has a large number of false-positives in case an image contains only a small amount of vegetation. Therefore, the amount of vegetation (foreground area) needs to be considered to reflect better the performance of vegetation segmentation. Sezgin and Sankur (2004) used the relative foreground area error (RAE) and combined this indicator with accuracy (misclassification error). Many studies have used this combined approach since then (Guan and Yan 2013; Nacereddine et al. 2005; Navarro et al. 2010; Shaikh et al. 2011). Inspired by these studies, the modified accuracy (MA) was defined in this study. This performance indicator uses a harmonic mean of relative vegetation area error (RVAE) and balanced accuracy (BA). Both measures have values between 0 and 1, where 0 represents very poor segmentation, and 1 represents perfect segmentation. The harmonic mean indicates if there is a large imbalance between these two measures, thus providing a better indication of the performance. The equations are described below:

$$\text{Balanced accuracy (BA)} = \frac{\text{Recall} + \text{Specificity}}{2} \quad (15)$$

$$\text{Relative vegetation area error (RVAE)} = \begin{cases} 1 - \frac{A_{GT} - A_{SEG}}{A_{GT}} & \text{if } A_{SEG} < A_{GT} \\ 1 - \frac{A_{SEG} - A_{GT}}{A_{SEG}} & \text{if } A_{GT} \leq A_{SEG} \end{cases} \quad (16)$$

$$\text{Modified accuracy (MA)} = \frac{2 \cdot \text{BA} \cdot \text{RVAE}}{\text{BA} + \text{RVAE}} \quad (17)$$

where  $A_{GT}$  is the vegetation area in ground truth (TP + FN);  $A_{SEG}$  is the vegetation area in segmented image (TP + FP).

In addition, receiver operating characteristic (ROC) and precision-recall curves were used. These curves have been used in several studies on image processing (Bai et al. 2014; Bulanon et al. 2009) helping to visually assess the segmentation performance. One performance indicator that is often used in these curves is the Area Under Curve (AUC), a measure represented with a single scalar value ranging from 0 to 1. AUC indicates how reliably the segmentation can be performed. An AUC value of 1 indicates a perfect segmentation (Mery and Pedreschi 2005).

Finally, processing time was measured to indicate how fast the algorithm performed. The processing time was measured on a PC equipped with an Intel(R) Core(TM) i7-377T 2.5 GHz processor and 8 GB RAM running 64bit Windows 7.

## Results

### Ground shadow detection

The performance measures of the ground shadow detection with Set 1 is shown in Table 2. Ground shadow detection performance was generally successful (modified accuracy  $\geq 0.9$ ) under natural lighting conditions. An average processing time of 0.33 s is satisfactory for real-time application as well. Example images in Set 1 and their ground shadow detection output are shown in Fig. 7. From the original field images (Fig. 7a), enhanced contrast of the ground shadow region is shown in the second column (Fig. 7b). The third column shows the detected ground shadow region (Fig. 7c), and its ground truth images are displayed in the fourth column (Fig. 7d). The last column (Fig. 7e) displays the difference image between detected shadow and ground truth.

### Ground shadow removal and vegetation segmentation performance

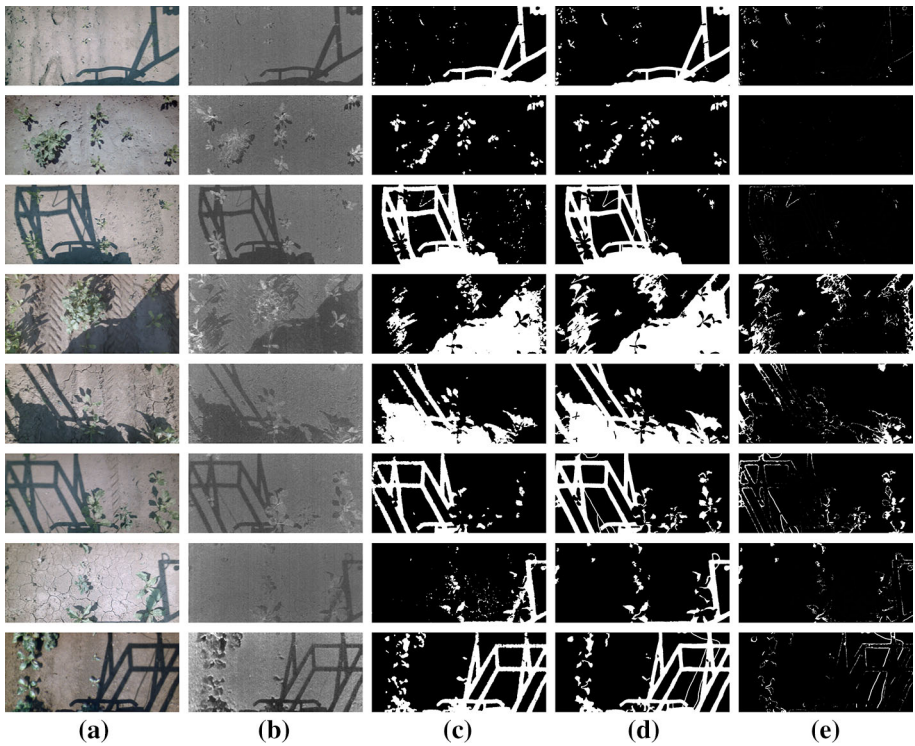
The images with shadow removal (ExG + Otsu + shadow removal, Fig. 5h) were compared with those without shadow removal (ExG + Otsu, Fig. 5c) to assess the performance improvement in vegetation segmentation when using ground shadow removal. The quantitative performance comparison is shown in Fig. 8. When it comes to precision, sensitivity and modified accuracy, the figure indicates that the vegetation segmentation with shadow removal has a higher performance than the segmentation without shadow removal. The average values of precision, specificity and modified accuracy for vegetation segmentation with shadow removal were 0.67, 0.96 and 0.71, respectively, indicating 20, 4.4 and 13.5% improvement over indicator values achieved without shadow removal. A *T*-test revealed that these improvements are significantly different ( $P < 0.001$ ). Only recall indicated that there were some losses of true-positive pixels (vegetation pixels) due to the shadow subtraction process. This minor loss was mainly observed when the shadow removal was applied to images without shadows (see the last column in Fig. 9). The combined measure, modified accuracy, indicated considerable improvement (13.5%) in vegetation segmentation. The average processing time for vegetation segmentation without shadow removal was 0.12 s, and was 0.46 s for segmentation with shadow removal. This is acceptable for a real-time application ( $< 1$  s).

Figure 9 shows the results of the segmentation process with ground shadow removal applied to Set 2, including images without and with shadows. When there was a shadow in the image scene, the ground shadow detection algorithm was, in general, successful in

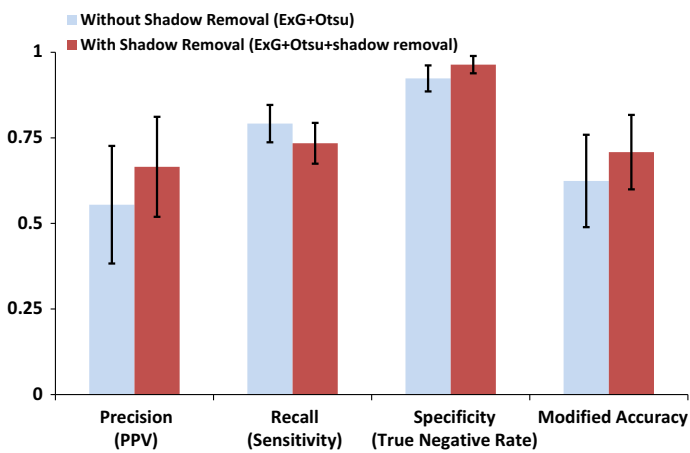
**Table 2** Quantitative performance measure of ground shadow detection in Set 1 (30 images)

	Precision (PPV)	Recall (sensitivity)	Specificity (TNR)	Modified accuracy	Processing time, s
Mean	0.94	0.87	0.99	0.92	0.33
Median	0.96	0.87	0.99	0.92	0.33
SD	0.07	0.04	0.01	0.03	0.01

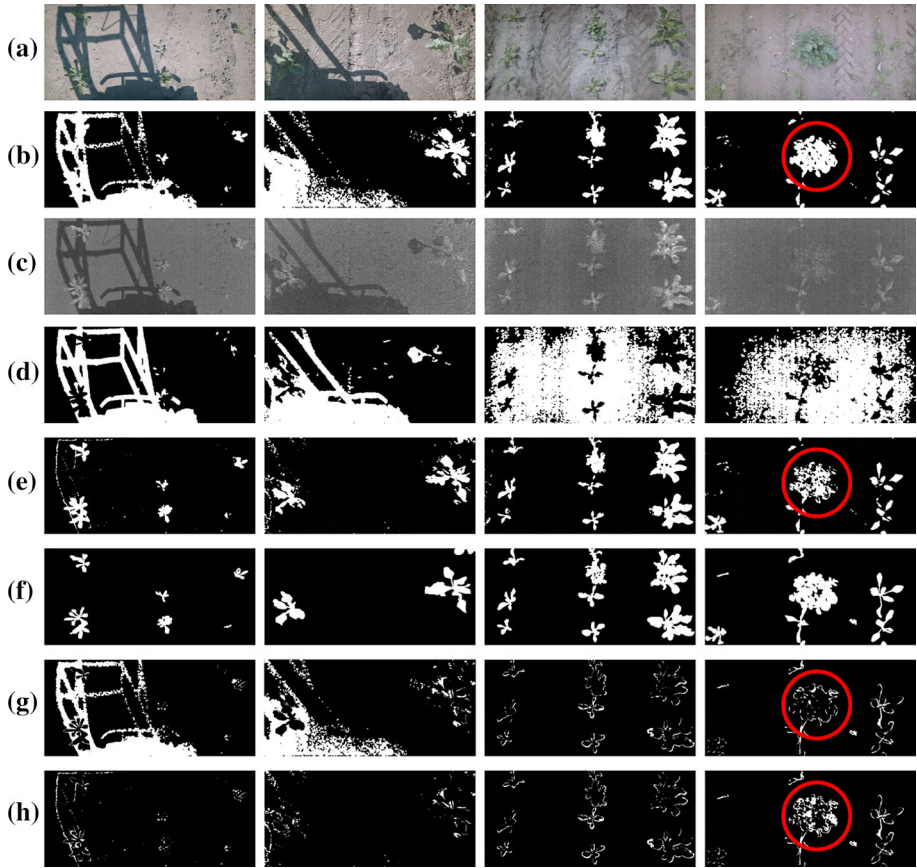
The mean, median, and standard deviation (SD) of the performance measures are indicated *PPV* positive predictive value, *TNR* true negative rate



**Fig. 7** Example images in Set 1 using the ground shadow detection process: (a) original image, (b) contrasted ground shadow region, (c) ground shadow detected, (d) ground truth, and (e) difference between (c) and (d)



**Fig. 8** The performance of vegetation segmentation: without shadow removal (ExG + Otsu) versus with shadow removal (ExG + Otsu + shadow removal)

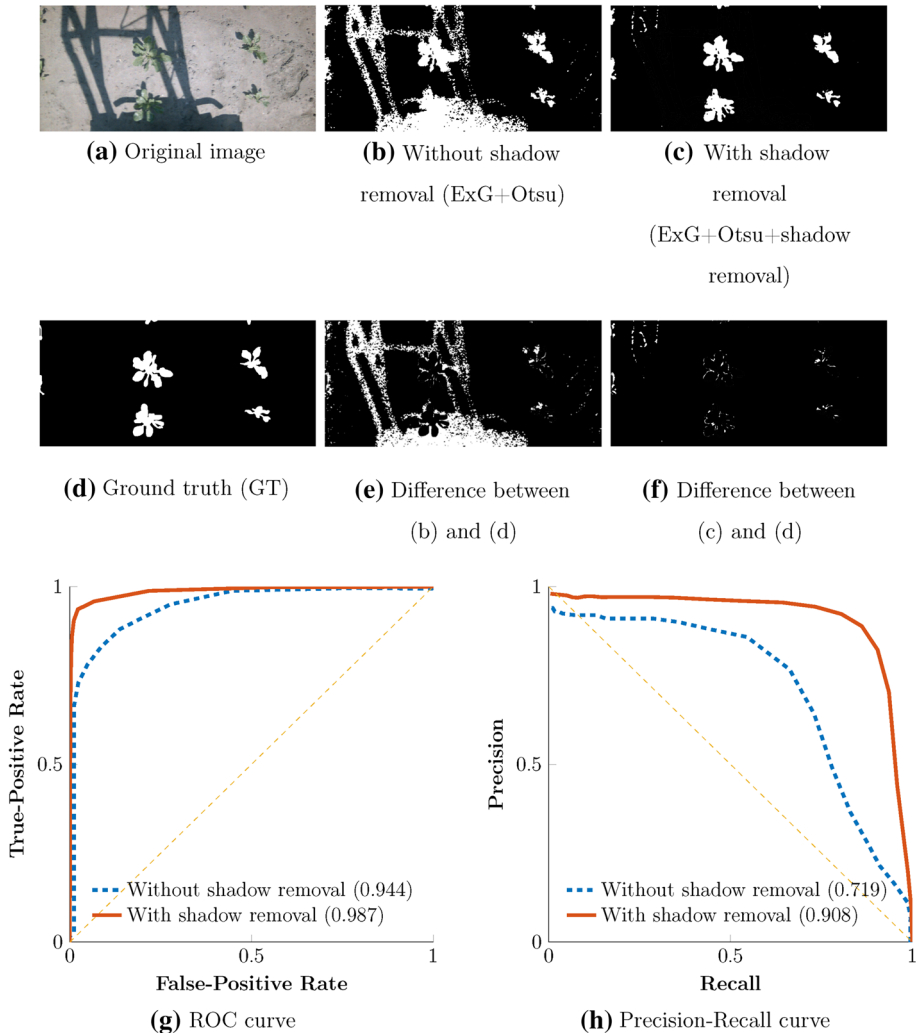


**Fig. 9** Example of images with vegetation segmentation and ground shadow removal process: (a) original image, (b) vegetation segmentation without shadow removal (ExG + Otsu), (c) contrasted ground shadow region, (d) ground shadow detected, (e) vegetation segmentation with shadow removal (ExG + Otsu + shadow removal), (f) vegetation ground truth, (g) difference between (b) and (f), and (h) difference between (e) and (f). Indicated with circles in the last column are some vegetation pixels lost during shadow removal

detecting the ground shadow region; but when there was no shadow in the image scene, almost the entire soil background was classified as a ground shadow region (Fig. 9d). In both cases, however, green-related pixels (plant materials) were not included in the ground shadow region, leading to no significant loss of vegetation pixels in the shadow removal process (Fig. 9e). The last column in Fig. 9 contains some examples of vegetation pixel loss with shadow removal indicated by circles.

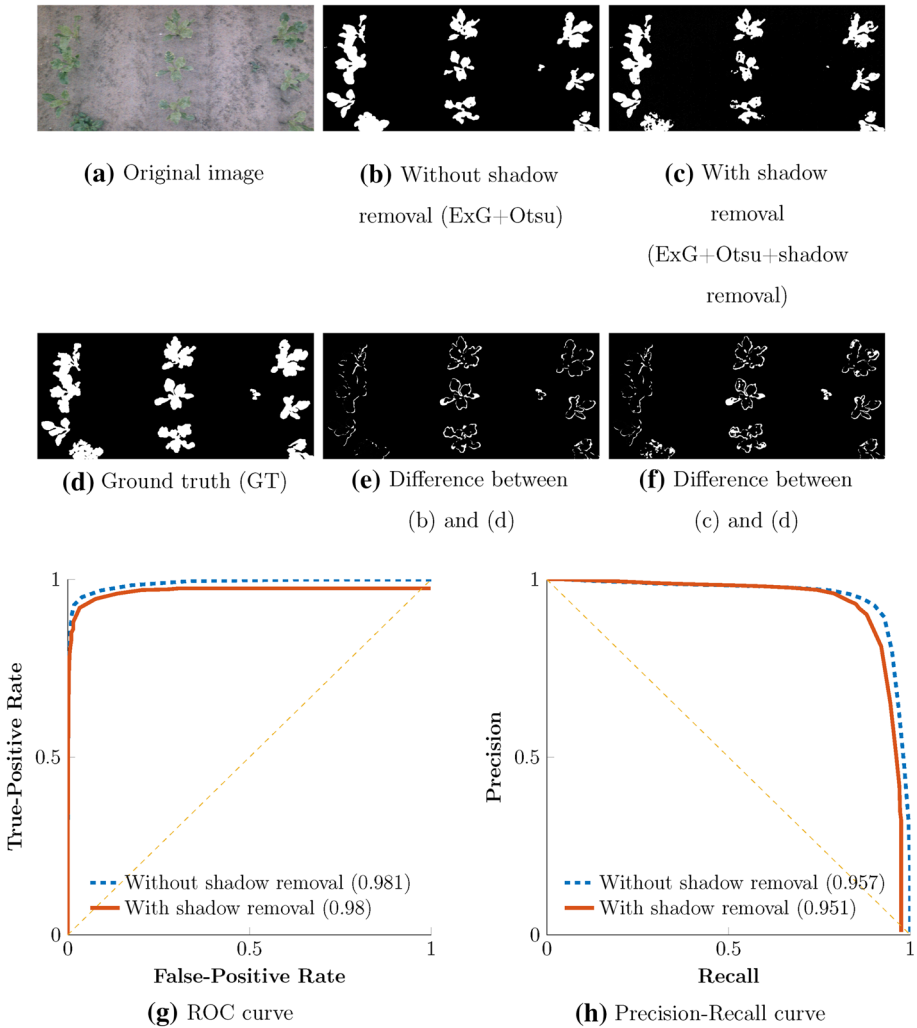
The ROC and precision-recall curves with a shadow image before and after the ground shadow removal are shown in Fig. 10. The AUC before and after shadow removal in ROC analysis were 0.944 and 0.987 respectively, and those in the precision-recall analysis were 0.729 and 0.908 respectively. Both curves showed that after ground shadow removal the performance improved and the vegetation segmentation (with a shadow image) succeeded ( $AUC \geq 0.9$ ).

The ROC and precision-recall curves are shown in Fig. 11 for segmentation with ground shadow removal applied to an image which contained no shadows. Then, the AUC



**Fig. 10** Segmentation of a shadow image with ground shadow removal and its performance analysis: (a) original image, (b) vegetation segmentation without shadow removal (ExG + Otsu), (c) vegetation segmentation with shadow removal (ExG + Otsu + shadow removal), (d) ground truth, (e) difference between (b) and (d), (f) difference between (c) and (d), (g) and (h) ROC and precision-recall curves for vegetation segmentation with and without shadow removal, the area under curve (AUC) in parenthesis

values before and after shadow removal in ROC analysis were 0.981 and 0.980 respectively, and those in the precision-recall analysis were 0.957 and 0.951 respectively. Both curves showed that the performance was not considerably different before and after ground shadow removal. The vegetation segmentation in an image without shadows was successful even after ground shadow removal. The ground shadow removal led to better performance of vegetation segmentation when applied to an image containing a shadow and did not negatively affect the result when applied to an image without shadows (Fig. 10 and 11).



**Fig. 11** Segmentation of a non-shadow image with and without ground shadow removal and its performance analysis: (a) original image, (b) vegetation segmentation without shadow removal (ExG + Otsu), (c) vegetation segmentation with shadow removal (ExG + Otsu + shadow removal), (d) ground truth, (e) difference between (b) and (d), (f) difference between (c) and (d), (g) and (h) ROC and precision-recall curves for vegetation segmentation with and without shadow removal, the area under curve (AUC) in parenthesis

## Discussion

### High dynamic range (HDR) camera

The HDR camera enabled capture of quality images in the high dynamic range scene of the field. During the field image acquisition, no image saturation caused by strong direct solar radiation was observed, and plants under sharp dark shadows were still clearly noticeable.



Two other studies (Dworak et al. 2013; Piron et al. 2010) also reported that high dynamic range acquisition enabled a strong signal to noise ratio for all pixels of the image as well as a better Normalized Difference Vegetation Index (NDVI). In this study, however, an HDR and a conventional non-HDR cameras were not simultaneously used in parallel in the field. Thus, a quantitative comparison between these two cameras under agricultural field conditions could not be made. However, the added value of using a HDR camera is expected in the agricultural field under natural light conditions.

## Shadow detection and removal

Although the proposed algorithm effectively detects and removes ground shadows, and thus improves the performance of vegetation segmentation, the algorithm itself alone does not extract any green material. The algorithm has to be combined with vegetation extraction methods (vegetation index), such as ExG, NDVI and CIVE. However, the shadow detection and removal algorithm is not limited to any specific vegetation extraction method because the algorithm is a separate procedure that can work as an add-on process.

The algorithm is based on color space conversion and chromaticity difference. This approach is simple, easy-to-implement and computationally inexpensive. Sanin et al. (2012) reported that the color space conversion approach needed the least computation time among the reviewed methods. However, the color space conversion approach requires the selection of an optimal threshold value that relies on the assumption that the image scene consists of a fixed number of components. In this study, a hypothesis was made that the field image scene can be divided into three classes: vegetation (green plants), background (soil) and ground shadow. Although hardly any other materials than these three were found in field images, a crop image scene may contain, according to Yang et al. (2015), various kinds of straw, straw ash or non-green plants. If an image scene contains a significant amount of the above mentioned or other materials, the algorithm may have limited performance.

There were few losses of true-positive pixels (vegetation pixels) during the shadow removal process. Although this loss was not critical, there are two ways to improve this procedure: post-image processing, and selective application of shadow removal. Post-image processing such as a hole filling or erode/dilate operation can recover some true-positive pixels that were lost during the removal process. Alternatively, applying shadow removal only when an image contains a shadow can improve the performance. In this study, shadow removal was applied to all images (images with and without shadows). However, technically there is no need to apply shadow removal when the image contains no shadows. This selective approach, however, would require a procedure that detects the presence of shadows in a given image scene. An alternative might be to use an illuminant-invariant image based on physical models of illumination and colors (Álvarez and López 2011; Finlayson et al. 2006).

The processing time for vegetation segmentation with shadow removal was 0.46 s, and this should be acceptable in a real-time application (<1 s required). There is a way to further reduce the processing time. If the processing time is highly critical for certain applications, a faster processor with multiple/parallel processing implementations might be an alternative approach to reduce the processing time.

## Conclusions

In this study, a ground shadow detection and removal method based on color space conversion and multi-level threshold was proposed. This method is to be used in a real-time automated weed detection and control system that has to operate under natural light conditions. Then vegetation segmentation is challenging due to shadows.

Applying shadow removal improved the performance of vegetation segmentation with an average improvement of 20, 4.4 and 13.5% in precision, specificity and modified accuracy, respectively, compared with no shadow removal. The average processing time for vegetation segmentation with shadow removal was 0.46 s, which is acceptable for the real-time application (<1 s required).

The proposed method for ground shadow detection and removal enhances the performance of vegetation segmentation under natural illumination conditions in the field, and is feasible for real-time field applications and does not reduce segmentation performance when shadows are not present.

**Acknowledgements** The work presented in this study was part of the Agrobot part of the Smartbot project and funded by Interreg IVa, European Fund for the Regional Development of the European Union and Product Board for Arable Farming. We thank Gerard Derks in Unifarm for arranging and managing the experimental fields. We also thank Wim-Peter Dirks for his contribution to creating the ground truth.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Ahmed, F., Al-Mamun, H. A., Bari, A. S. M. H., Hossain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection*, *40*, 98–104.
- Al-Najdawi, N., Bez, H. E., Singhai, J., & Edirisinghe, E. A. (2012). A survey of cast shadow detection algorithms. *Pattern Recognition Letters*, *33*(6), 752–764.
- Álvarez, J. M., & López, A. M. (2011). Road detection based on illuminant invariance. *IEEE Transactions on Intelligent Transportation Systems*, *12*(1), 184–193.
- Åstrand, B., & Baerveldt, A. J. (2002). An agricultural mobile robot with vision-based perception for mechanical weed control. *Autonomous Robots*, *13*(1), 21–35.
- 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.
- Bai, X., Cao, Z., Wang, Y., Yu, Z., Hu, Z., Zhang, X., et al. (2014). Vegetation segmentation robust to illumination variations based on clustering and morphology modelling. *Biosystems Engineering*, *125*, 80–97.
- Bandoh, Y., Qiu, G., Okuda, M., Daly, S., Aach, T., & Au, O. C. (2010). Recent advances in high dynamic range imaging technology. In *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP 2010)* (pp. 3125–3128). Hong Kong: IEEE.
- Benezeth, Y., Jodoin, P. M., Emile, B., Laurent, H., & Rosenberger, C. (2008). Review and evaluation of commonly-implemented background subtraction algorithms. In *Proceedings of the 19th International Conference on Pattern Recognition (ICPR 2008)* (pp. 1–4). New York, USA: IEEE.
- Bloch, C. (2007). *The HDRI Handbook: High Dynamic Range imaging for photographers and CG artists*. Santa Barbara, USA: Rocky Nook.
- Bulanon, D. M., Burks, T. F., & Alchanatis, V. (2009). Image fusion of visible and thermal images for fruit detection. *Biosystems Engineering*, *103*(1), 12–22.
- Corke, P. (2011). Light and Color. *Robotics, Vision and Control* (pp. 223–250). Berlin Heidelberg, Germany: Springer.

- Dworak, V., Selbeck, J., Dammer, K., Hoffmann, M., Zarezadeh, A. A., & Bobda, C. (2013). Strategy for the development of a smart NDVI camera system for outdoor plant detection and agricultural embedded systems. *Sensors*, *13*(2), 1523–1538.
- Finlayson, G. D., Hordley, S. D., Lu, C., & Drew, M. S. (2006). On the removal of shadows from images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(1), 59–68.
- Florczyk, S. (2005). *Robot Vision: Video-based indoor exploration with autonomous and mobile robots*. Weinheim, Germany: Wiley-VCH.
- Gée, C., Bossu, J., Jones, G., & Truchetet, F. (2008). Crop/weed discrimination in perspective agronomic images. *Computers and Electronics in Agriculture*, *60*(1), 49–59.
- Graham, D. J. (2011). Visual perception: Lightness in a high-dynamic-range world. *Current Biology*, *21*(22), R914–R916.
- Guan, P. P., & Yan, H. (2013). A hierarchical multilevel image thresholding method based on the maximum fuzzy entropy principle. *Image Processing: Concepts, Methodologies, Tools, and Applications* (pp. 274–302). IGI Global: Hershey, USA.
- Guo, W., Rage, U. K., & Ninomiya, S. (2013). Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Computers and Electronics in Agriculture*, *96*, 58–66.
- Haug, S., Biber, P., & Michaels, A. (2014). Plant stem detection and position estimation using machine vision. In *Proceedings of the International Workshop on Recent Advances in Agricultural Robotics (RAAR2014)*. Padova, Italy. CD-ROM.
- Hrabar, S., Corke, P., & Bosse, M. (2009). High dynamic range stereo vision for outdoor mobile robotics. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2009)* (pp. 430–435). Kobe, Japan: IEEE.
- Irie, K., Yoshida, T., & Tomono, M. (2012). Outdoor localization using stereo vision under various illumination conditions. *Advanced Robotics*, *26*(3–4), 327–348.
- Jeon, H. Y., Tian, L. F., & Zhu, H. (2011). Robust crop and weed segmentation under uncontrolled outdoor illumination. *Sensors*, *11*(6), 6270–6283.
- Lati, R. N., Filin, S., & Eizenberg, H. (2013). Estimating plant growth parameters using an energy minimization-based stereovision model. *Computers and Electronics in Agriculture*, *98*(1), 260–271.
- Lee, W. S., Slaughter, D. C., & Giles, D. K. (1999). Robotic weed control system for tomatoes. *Precision Agriculture*, *1*(1), 95–113.
- Mann, S., Lo, R. C. H., Ovtcharov, K., Gu, S., Dai, D., Ngan, C., et al. (2012). Realtime HDR (High Dynamic Range) video for eyetap wearable computers, FPGA-based seeing aids, and glasseyes (EyeTaps). In *Proceedings of the 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp.1–6). Montreal, Canada: IEEE.
- Mery, D., & Pedreschi, F. (2005). Segmentation of colour food images using a robust algorithm. *Journal of Food Engineering*, *66*, 353–360.
- Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, *8*(4), 283–298.
- Meyer, G. E., & Camargo Neto, J. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, *63*(2), 282–293.
- Nacereddine, N., Hamami, L., Tridi, M., & Oucief, N. (2005). Non-parametric histogram-based thresholding methods for weld defect detection in radiography. *World Academy of Science, Engineering and Technology*, *1*(9), 1237–1241.
- Navarro, P., Iborra, A., Fernández, C., Sánchez, P., & Suardíaz, J. (2010). A sensor system for detection of hull surface defects. *Sensors*, *10*(8), 7067–7081.
- Nieuwenhuizen, A. T., Hofstee, J. W., & Van Henten, E. J. (2010). Performance evaluation of an automated detection and control system for volunteer potatoes in sugar beet fields. *Biosystems Engineering*, *107*(1), 46–53.
- Ohta, J. (2007). *Smart CMOS Image Sensors and Applications*. Boca Raton, USA: CRC Press.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66.
- Pascale, D. (2003). *A Review of RGB Color Spaces, from xyY to R'G'B'*. Technical report, The BabelColor Company, Montreal, Canada.
- Piron, A., Heijden, F., & Destain, M. F. (2010). Weed detection in 3D images. *Precision Agriculture*, *12*, 607–622.
- Polder, G., Van der Heijden, G. W. A. M., Van Doorn, J., & Baltissen, T. A. H. M. C. (2014). Automatic detection of tulip breaking virus (TBV) in tulip fields using machine vision. *Biosystems Engineering*, *117*, 35–42.
- Prati, A., Mikic, I., Trivedi, M. M. M., & Cucchiara, R. (2003). Detecting moving shadows: algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(7), 918–923.

- Radonjić, A., Allred, S. R., Gilchrist, A. L., & Brainard, D. H. (2011). The dynamic range of human lightness perception. *Current Biology*, 21(22), 1931–1936.
- Reinhard, E., Ward, G., Pattanaik, S., Debevec, P., Heidrich, W., & Myszkowski, K. (2010). *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. Burlington, USA: Morgan Kaufmann.
- Rosin, P. L., & Ioannidis, E. (2003). Evaluation of global image thresholding for change detection. *Pattern Recognition Letters*, 24(14), 2345–2356.
- Sanin, A., Sanderson, C., & Lovell, B. C. (2012). Shadow detection: A survey and comparative evaluation of recent methods. *Pattern Recognition*, 45(4), 1684–1695.
- Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146–165.
- Shaikh, S. H., Maiti, A., & Chaki, N. (2011). Image binarization using iterative partitioning: A global thresholding approach. *Proceedings of the International Conference on Recent Trends in Information Systems (ReTIS)* (pp. 281–286). IEEE: Kolkata, India.
- Slaughter, D. C. C., Giles, D. K. K., & Downey, D. (2008). Autonomous robotic weed control systems: A review. *Computers and Electronics in Agriculture*, 61(1), 63–78.
- Sojodishijani, O., Ramli, A. R. R., Rostami, V., Samsudin, K., & Saripan, M. I. I. (2010). Just-in-time outdoor color discrimination using adaptive similarity-based classifier. *IEICE Electronics Express*, 7(5), 339–345.
- Steward, B. L., Tian, L. F., Nettleton, D. S., & Tang, L. (2004). Reduced-dimension clustering for vegetation segmentation. *Transactions of the ASAE*, 47(2), 609–616.
- Teixidó, M., Font, D., Pallejà, T., Tresanchez, M., Nogués, M., & Palacín, J. (2012). Definition of linear color models in the RGB vector color space to detect red peaches in orchard images taken under natural illumination. *Sensors*, 12(6), 7701–7718.
- Wang, Q., Wang, H., Xie, L., & Zhang, Q. (2012). Outdoor color rating of sweet cherries using computer vision. *Computers and Electronics in Agriculture*, 87, 113–120.
- Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, 38(1), 259–269.
- Yang, W., Wang, S., Zhao, X., Zhang, J., & Feng, J. (2015). Greenness identification based on HSV decision tree. *Information Processing in Agriculture*, 2(3–4), 149–160.
- Zheng, L., Zhang, J., & Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture*, 65(1), 93–98.