Algorithm of Weed Detection in Crops by Computational Vision

A. J. Irías Tejeda[1]
Faculty of engineering and architecture
Universidad Tecnológica Centroamericana (UNITEC)
Tegucigalpa, Francisco Morazán, Honduras
alvaroirias@unitec.edu

R. Castro Castro[2]
Faculty of engineering and architecture
Universidad Tecnológica Centroamericana (UNITEC)
Tegucigalpa, Francisco Morazán, Honduras
rigocastro@unitec.edu

Abstract—**This research has been based on the use of precision agriculture tools for the management of weeds in crops. It has focused on the creation of an image-processing algorithm to detect the existence of weeds in a specific site of crops. The main objective has been to obtain a formula so that a weed detection system can be developed through binary classifications. The initial step of image processing is the detection of green plants in order to eliminate all the soil in the image, reducing information that is not necessary. Then, it has focused on the vegetation by segmentation and eliminating unwanted information through medium and morphological filters. Finally, a labeling of objects has been made in the image so that weed detection can be done using a threshold based on the area of detection. This algorithm establishes an accurate monitoring of weeds and can be implemented in automated systems for the eradication of weeds in crops, either through the use of automated sprayers for specific site or a weed-cutting mechanism. In addition, it increases the performance of operational processes in crop management, reducing the time spent searching for weeds throughout a plot of land and focusing weed removal tasks on specific sites for effective control.**

Keywords— Image processing, weed detection, crop monitoring, morphological filters, precision agriculture.

I. Introduction

According to [1], a study carried out by professors of the Faculty of Agronomy of the University of Buenos Aires (FAUBA), has estimated that at least some 1,300 million dollars are destined every year for the control of weeds. Most people who practice agricultural activity for export, are people who use conventional methods for land treatment, irrigation and crop management, maintaining the quality of the product at an outstanding level. Despite the high quality of the national product, the outstanding quality standards developed countries have not yet achieved. The disparity lies in the use of new technologies to improve and optimize the processes of soil study and crop management.

The Precision Agriculture (PA) has opened the doors so that technology can be incorporated into the farming processes and improve the effectiveness of production in the crops. This new concept has led to developed countries highly productive in agriculture, opting for the use of new tools to improve their technological management in the agricultural enterprise. In this way, agricultural practices are determined to replace the usual inputs based on average values, as in traditional agriculture, for a more precise agriculture, with localized management, which studies the changes in yield in an entire area. Consequently, greater benefits are obtained such as the optimization of the use of inputs, determination of the availability of nutrients, organic matter, water, etc. on the land, reducing production costs and improving the quality of crops.

The control of weeds is of vital importance in agriculture, these are unwanted by the farmer since they are causing several problems in the crop. Among its negative effects is the contamination of production, shelter of insects and diseases, facilitates the growth of other pests and increases irrigation costs. The monitoring of these weeds allows us to detect the presence and/or abundance of weeds, gather information that allows decision-making during the campaign, provide data to build the record of the sites on which long-term actions can be designed term, detect the entry of invasive species, not yet present in the lot and provide bases for precision agriculture and specific site management of inputs [2].

The idea of making an algorithm that by means of image processing detects the weeds that are located in a specific area of plantation then arises. This methodology is advantageous because it offers a technological tool for farmers throughout the process of sowing, growing and harvesting crops. In addition, it increases the performance of operational processes in crop management, reducing the time spent searching for weeds throughout a plot of land and focusing weed removal tasks on specific sites for effective control.

II. MATERIALS AND METHODOLOGY

For the acquisition of the test images applied to this research, was used a Nikon D3200 camera with AF-P DK NIKKOR 18-55 mm focal lens. Image processing was carried out using the Matlab R2015a software, implementing the Image Processing Toolbox. The algorithm was tested on a Dell Latitude E5450 computer with Intel Core i5-5200 processor, 8 GB of RAM, Windows 10 Home and 64-bit operating system.For the development and implementation of the algorithm for image processing the following stages were followed:

A. Acquisition of the Image

The ability to acquire own images and use them for the tests of the algorithm is advantageous, because in this way, the algorithm will adapt to the characteristics of the images obtained. It has been proposed to use a semi-professional camera of 24.2 megapixels with the ability to take photographs at 1080 pixels resolution, enough to capture good quality images. The images were acquired at a height of 1.20 meters from the surface of the ground, this height was set to obtain a good resolution of the crop and the weeds on the surface. Each acquired image has a resolution of 4512x3000 pixels, which results in it covering an area of 180x120 cm above the crop. The position of the camera was positioned vertically to avoid shadows and ensure uniform illumination. The proposed algorithm considers the processing of images that contain uniform illumination; images with better illumination are processed with greater precision [10].

Fig. 1. Image acquired for algorithm tests.

B. Detection of green plants

Previous studies have based their criteria for selection on an index that stands out green component of source image as the NDVI (Normalized Difference Vegetation Index) and SAVI (Soil Adjusted Vegetation Index) [7]. In this project, the source image is converted to grayscale intensity whereby the hue and saturation information is eliminated while retaining the luminance, this operation can be performed through the function *rgb2gray* of Matlab 2015.

Taking advantage of the RGB components of the image, all the components in the XY space that correspond to the green value in the image, are subtracted from the grayscale image to separate the vegetation from the other components [3]. The operation that makes this stage possible is shown in equation (1), the results are shown in Fig. 2.

$$I_{Plant}(X_{Pixel},Y_{Pixel}) = I_{Green}(X_{Pixel},Y_{Pixel},G) - I_{Gray}(X_{Pixel},Y_{Pixel}) \quad (1)$$

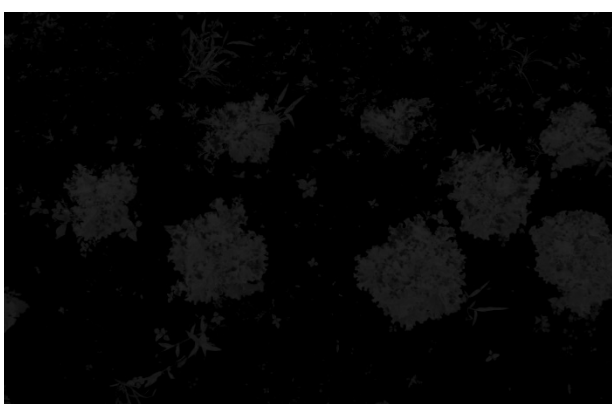

Fig. 2. Subtraction of green components.

C. Medium Filter and Threshold Segmentation

The median filter is used regularly to reduce noise in images with subtraction of components like the one that has been used in this process [4]. This filter works by replacing the central pixel of a region called neighborhood, in this case a neighborhood of 3x3 pixels, creating a mask over the image. The value of the center of the mask is replaced with the calculation of the median of the values of the neighborhood pixels [8]. This operation can be executed through the function *medfilt2* of Matlab 2015.

Once the medium filter is applied, the image must be segmented. For the segmentation it is recommended to use the Otsu method described in [5], one of the most famous methods and used for this type of applications. Segmentation is accomplished by using the following expression (2):

$$I_{bin}(x,y) = \begin{cases} 0, I_{Median}(x,y) < t \\ 1, I_{Median}(x,y) \geq t \end{cases} \quad (2)$$

The selection of an appropriate threshold is carried out with image histogram, taking the valium t calculated by the operation *graythresh* of Matlab 2015, the value of t is based on the average intensity value of the light and dark areas of the image. The results of the threshold segmentation are shown in Fig. 3.

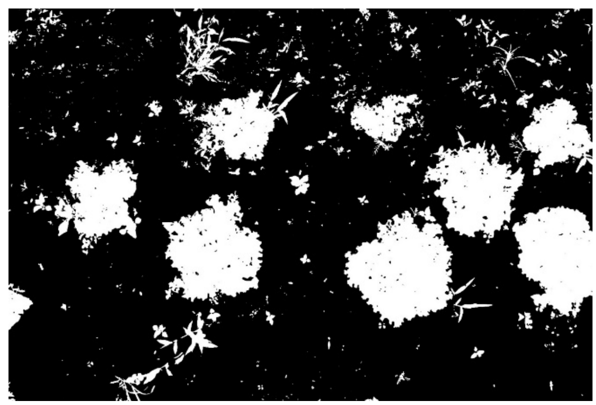

Fig. 3. Threshold segmentation.

D. Morphological filters

The classification of the labels is based on the area of each object, it is convenient to fill the holes in the objects of the image. Therefore, a filter based on morphological reconstruction should be applied in order to fill the holes and obtain a more effective area [6]. The algorithm calculates a marked image stemming from source image borders using (3).

$$I_{mark}(x,y) = \begin{cases} 1 - I_{source}(x,y), \\ (x, y \text{ is on the border of } I_{source}) \\ 0, \text{otherwise} \end{cases} \quad (3)$$

The methodology iterates until all holes of the objects in the image are filled. This operation can be done using the morphological operators *imfill* and *bwmorph* of Matlab 2015.

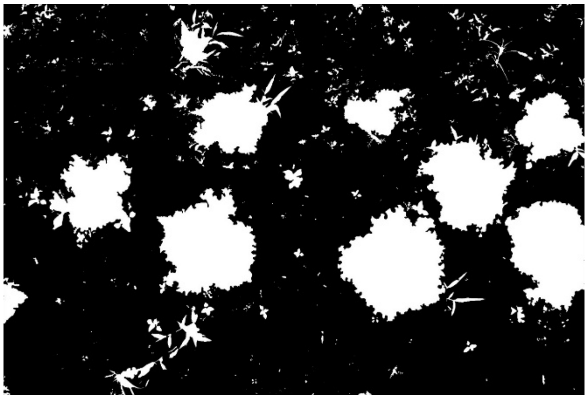

Fig. 4. Image with filled holes.

E. Labeling and classification labels

To identify the objects in the image as plants it is necessary to label them, the labels are rectangles that border the areas of all the identified objects, this is a crucial step since those labeled areas can be extracted characteristics [9]. The region labeling stage evaluates each pixel with a 4 neighbor-connectivity, using a heuristic stated on pixel values according to predecessor labels at north and west position. The computational tool to perform this step is the *blabel* operator of Matlab 2015.

Once objects on the scene are labeled, the next step is to extract area features from each element to discriminate weed and crop. The algorithm defines an area counting the number of pixels in the object region; then, the value is stored for all items. The extraction of characteristics of areas in images is done through of the *regionprops* function of Matlab 2015.

The algorithm performs two stages of discarding to identify weeds, the first consists in eliminating those areas that are too small to be considered considerable weeds.

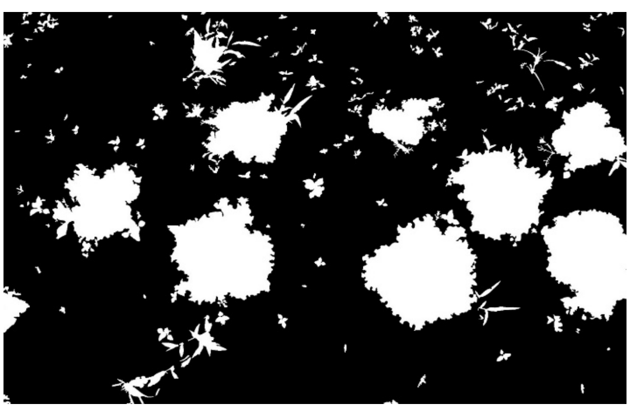

Fig. 5. Discard excessively small areas.

After the first discard stage, a threshold based on the classification of weeds is defined by the difference between the average values of the largest areas corresponding to crop plants. The average value of the crop is calculated, this value is taken as the threshold for crop and weed classification. The value of the threshold is compared with the size of all the objects in the image. If the area of the analyzed object is smaller than the threshold value, is identified as weed. Fig. 6 demonstrates the results obtained.

Fig. 6. Weed detection and labeling.

The summary of the procedure to develop the algorithm for weed detection is presented in Fig. 7.

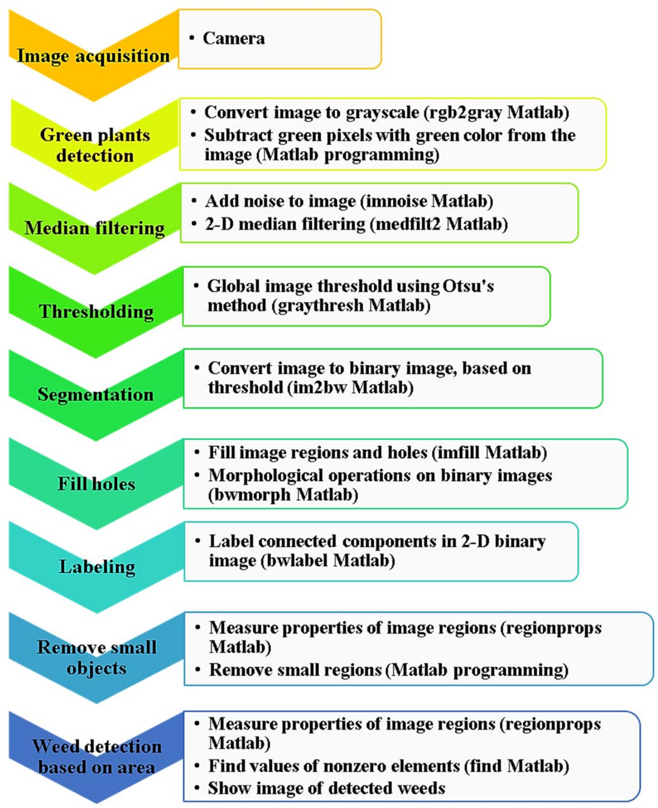

Fig. 7. Weed detection algorithm process diagram with description of hardware and software tools to execute each step.

III. ANALYSIS AND RESULTS

The results of the algorithm are defined by the average of its specificity and sensitivity. To check both indexes it is necessary to tabulate the following variables:

- True Positive (TP): Number of weeds detected correctly.

- True Negative (TN): Number of crop plants detected correctly.
- False Positive (FP): Number of crop plants detected as weeds.
- False Negative (FN): Number of weeds detected as crop plants.

The calculations for the indexes are the following:

$$\text{Sensitivity} = TP / (TP + FN) \quad (4)$$

$$\text{Specificity} = TN / (TN + FP) \quad (5)$$

$$\text{Positive Predictive Value (PPV)} = TP / (TP + FP) \quad (6)$$

$$\text{Negative Predictive Value (NPV)} = TN / (TN + FN) \quad (7)$$

Subsequently, the analysis and results of a sample image are shown. The calculations of indexes and variable counts were executed.

Fig. 8. Sampling image for algorithm validation.

TABLE 1. RESULTS OF SENSITIVITY AND SPECIFICITY IN LETTUCE CROP

Weed detection analysis			
Plants detected as weeds	True Positive (TP)	False Positive (FP)	PPV
	271	1	0.99
Plants detected as a vegetable	False Negative (FN)	True Negative (TN)	NPV
	11	12	0.52
Results	Sensitivity	Specificity	
	0.96	0.92	

To validate the image processing algorithm, the algorithm was tested on 36 crop images of multiple vegetables, this in order to be able to average the appropriate specificity and sensitivity of the algorithm. Table 2 shows the results of the corresponding sample.

TABLE 2. GENERAL RESULTS OF THE WEED DETECTION ALGORITHM

Weed detection analysis				
Sensitivity	Specificity	PPV	NPV	Processing Time
97%	79%	99%	47%	2.98 seconds

As shown in Table 2, the sensitivity result of the algorithm is 97%, which indicates a good performance of the algorithm proposed to detect weeds correctly. On the other hand, the specificity value of 79% represents the ability of the system to correctly detect the vegetables. Comparing the specificity and sensitivity indexes, the specificity is of lower index because several of the acquired images were from sections of crops that the crop plants were the same size as the weed. As a result, the algorithm detects the vegetable as a weed and increases the FP variable. For this reason, the high value of FP affects the effectiveness of the algorithm's specificity.

The index of positive predictive value is 99% which indicates that in almost all cases, the weed was correctly identified. In contrast, the negative predictive value is 47% which indicates that the vegetables detected were not always vegetables, this is due to cases where the weeds were almost the same size as the crop plants.

IV. CONCLUSIONS

This research has proposed a practical way to detect weeds by image processing based on the characteristic of the area of each object in an image. Although research has been limited in that the size of the weed is smaller than that of the crop, high indices of sensitivity, specificity and positive predictive value have been achieved, contrary to the negative predictive value, which is lower than 50%.

The proposed algorithm has the advantage of detecting weeds present between the plants in the crop lines. It also detects effectively as crop plants even those that are outside the crop lines, which is an objective difficult to achieve with other methods using computational vision [11]. However, the algorithm loses effectiveness when the sizes of the weeds are similar to the sizes of the plants of the crop, since the characteristic that is taken as variable of classification is the size of the plants. This problem can be solved by adding another characteristic as a classification method.

The use of low level characteristics such as the color of the plants and the area is an advantage given that the specific characteristics of the weeds as texture or shape are not relevant, providing versatility for the application of the algorithm in different crops of vegetables. This advantage is important, due to the great variety and types of weeds that exist in crops. A specific database of weeds is not necessary to be able to train the algorithm and identify weeds, as an automatic learning algorithm would do.

It was concluded that the proposed algorithm using low level characteristics and a threshold based on the area, have an improvement field in the specificity indexes and NPV, but the results are good enough to use the algorithm in practical applications of precision agriculture.

References

[1] Ruralnet, «Ruralnet», 2015. [Online]. Available: https://ruralnet.com.ar/nueva-tecnologia-la-deteccion-malezas/. [Last access: 18 September 2018].

[2] E. S. Leguizamón, «El menejo de malezas: desafíos y oportunidades», Universidad Nacional de Rosario, Rosario, Argentina, 2007.

[3] M. A. Molina-Villa y L. E. Solaque-Guzmán, «Machine vision system for weed detection using image filtering in vegetables crops», *Facultad de Ingeniería Universidad de Antioquia,* vol. 80, pp. 124-130, 2016.

[4] E. M. García, Visión artificial, Catalunya: Universitat Oberta de Catalunya, 2012.

[5] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE transactions on systems, man, and cybernetics,* vol. 9, no. 1, pp. 62-66, 1979.

[6] P. Soille, Morphological image analysis: principles and applications, New York: Springer Science & Business Media, 2013.

[7] K. R. Thorp and L. F. Tian, "A review on remote sensing of weeds in agriculture", *Precision Agriculture,* p. 477–508, 2004.

[8] B. I. Justusson, "Median filtering: Statistical properties", in Two-Dimensional Digital Signal Prcessing II, Berlin, Springer, 1981, pp. 161-196.

[9] R. Kumar, K. Ramareddy and B. Rao, "A simple region descriptor based on object area per scan line," *International Journal of Computer Applications,* vol. 3, no. 7, pp. 24-27, 2010.

[10] J. L. Tang, X. Q. Chen, R. H. Miao and D. Wang, "Weed detection using image processing under different illumination for site-specific areas spraying," *Computers and Electronics in Agriculture,* vol. 122, pp. 103-111, 2016.

[11] A. Paikekari, V. Ghule, R. Meshram and V. B. Raskar, "Weed detection using image processing", International Research Journal of Engineering and Technology (IRJET), vol. 3, no. 3, pp. 1220-1222, 2016.