Development and testing of image processing algorithm to estimate weed infestation level in corn fields

Wesley E. Santiago1*, Neucimar J. Leite2, Bárbara J. Teruel2, Manoj Karkee3, Carlos AM. Azania4, Renan Vitorino5

1Federal University of the Valleys Jequitinhonha and Mucuri, Minas Gerais, Brazil
2University of Campinas, São Paulo, Brazil
3Center for Precision and Automated Agricultural Systems and Department of Biological Systems Engineering, Washington State University, Prosser, Washington, USA
4IAC - National Pride in the Lands of São Paulo, Brazil

*Corresponding author: esdras.wes@gmail.com

Abstract

The challenges of modern agriculture have led to the development of localized management tools which allow the rationalization of the use and application of pesticides, a reduction in production costs and the optimization of agricultural processes. This study was carried out to develop an algorithm capable of orienting weed control in the management of a corn crop, using digital image analysis to identify the level of weed infestation in the field. The seeds of six species of weed were sown in an experimental plot of corn, and daily images were captured for 40 days for the evaluation of the level of weed infestation (low, intermediate or high). The algorithm developed was able to target information about the plants and soil accurately and discriminate the residual information as referring to either the culture or weeds. The proposed algorithm has achieved 90% accuracy in identifying the level of infestation from images already evaluated by experts. The results can thus be used as part of weed control strategy, with the incorporation of the geographic coordinates of the image making possible the construction of a map of the level of weed infestation in the different areas where the crop is growing.

Keywords: precision agriculture, specialized system, herbicide.

Abbreviations: Pdp_Number of pixels belonging to the category weed, Lum_Image width in pixels , Cim_Image length in pixels. RMSE_Root Mean Squared Error, P_in Input value, P_est Estimated Value.

Introduction

Weed control is a crucial part of all agriculture systems, since competition with weeds can reduce the yield (Silva et al., 2009) and increase the costs, at times even compromising the feasibility of harvest. Since weeds require the same elements for growth as cultivated plants (water, sunlight, nutrients and space), the absence of weeds on the farm fields is crucial in the production of commercial crops such as sugarcane, soybeans, and corn (Kuva et al., 2001; Cury et al., 2012; Vasconcelos et al., 2012; Wandscheer et al., 2013). The level of competition of weeds with agriculture crops depends on the specific community of weeds involved, as well as factors linked to the crop, the environment and the length of time of co-habitation (Ferreira et al., 2010). Competition is initiated when the requirements for one or more growing factors is greater than the supply (Rizzardi et al., 2001). A study by Ribas et al. (2013) concluded that when facing competition from weeds, the corn yield was reduced by 17%. Similar results have been reported for beans and soybeans by Manabe et al. (2015) and Tavares et al. (2012), respectively. In an attempt to avoid losses resulting from the presence of weeds in cultivated areas, efficient management measures should be adopted. Currently, the most frequently adopted approach is the use of herbicides. This practice has proved to be quite effective. However, in the long run, weed management in agricultural fields must be reconsidered, since new herbicides will be needed to meet new challenges, and their development will require time and financial resources. Moreover, herbicides can be a mixed blessing, and in recent years, many cases of environmental damages and deleterious effects on human health caused by agrochemicals have been reported in the scientific literature (Koller et al., 2012; Helander et al., 2012; Hernandez et al., 2013). The challenges of increasing production without damaging the environment have led to the development of tools for site specific management which allow the rationalization of the use and application of chemical products, as well as a reduction in production costs and the optimization of the agricultural processes involved (Nyko et al., 2013). Studies involving the chemical management of weeds have shown how certain technologies of site-specific application can effectively reduce the volume of herbicides needed in up to 90% (Gerhards et al., 2002; Shiratsuchi et al., 2002). Because of the demand for such a reduction, interest in the use of image processing techniques in agriculture has increased, and the positive results of some of the studies have led to the inclusion of these techniques as part of the basic arsenal of strategies and tools available for the control of weeds in agricultural management. Smart systems which can distinguish weeds from crops have been reported; these make it possible for equipment to be programmed to automatically
target only weeds, with only the weeds being sprayed (Bakker et al., 2010; Burgos-Artizzu et al., 2011). However, for these systems to work properly, it is necessary to develop computational tools which can discriminate weeds from crops and bare ground (Gerhards et al., 2002). The present paper was thus designed to develop an algorithm for the identification of the level of weed infestation in fields of corn using image processing techniques.

Results and Discussion

Weed infestation level based on expert evaluation

The values for weed coverage obtained by the algorithm were compared to those attributed by experts. The results (Table 1) suggest that for a given level of weed infestation, the adoption of control measures similar to those adopted for the conventional method can be adopted without the need of a more robust rating system. The experts varied only slightly in relation to level of tolerance when analyzing weed coverage (Table 1). Expert 3, for example, considered that infestation was low for 34 of the images, whereas for Experts 1 and 2, 24 and 31 of the images were so evaluated. For all of the images evaluated as having a low infestation, the average weed coverage was only 4.8%. Seven of the images were considered to show a high level of weed infestation by Expert 1, and six by Expert 2; the third expert evaluated only one of the images as revealing high weed coverage. For these images evaluated as having high weed cover, the weed coverage exceeded 81%. Since it is only when infestation surpasses 7.4% that immediate intervention is necessary, this breakdown of a field into areas of low, intermediate and high levels of infestation makes it possible to construct herbicide treatment maps identifying where intervention is needed (Gerhards et al., 2002; Longchamps et al., 2012), with the insertion of such maps in equipment performing localized applications promoting economy in herbicide use (Shirasuchi et al., 2003). Evaluating a system of images to help in making decisions about intervention in the control of weeds, Hamouz et al. (2014) identified a potential reduction of up to 90% in the volume of herbicides needed in comparison with what was recommended by a more conventional method. For Rizzardi and Fleck (2004), densities of 4 or more plants per m² of the common weed Bidens pilosa L. can limit the development of soybeans. Hamouz et al. (2013) organized the management of weeds for a specific site on a 4 x 1.5-meter grid using a threshold of weed density of 0.02 to 0.20% (0.10 to 15.51 plants m⁻²). They found this procedure led to a reduction in herbicides needed from 15.6 to 100%, depending on the type of herbicide and threshold of application used.

Performance analysis from algorithm for automatic estimate of weed cover

Fig. 2 presents the results for a single image as evaluated by the neuro-fuzzy classifier. The weed cover is evaluated in relation to each of the levels of infestation to determine what the estimated level is. Table 2 summarizes the performance of the fuzzy logic classifier. The rating precision is showed for both image groups, although the RMSE refers only to training image groups. The correlation between the level of weed infestation estimated by algorithm and that attested to by the experts was high, with a RMSE value of 0.44 and 90% accuracy, which represents a good fit of the data for all three categories considered. A similar level of accuracy in the identification of a specific weed (Rumex obtusifolius) in an area of pasture was obtained by Hiremath et al. (2012), but here the problem consists of a complex system of various kinds of weeds and the similar leaf shapes of the crop. This is an extremely challenging problem, especially since various environmental factors affect the images, such as differences in lighting, position, scale and overlapping of leaves, all of which can contribute noise and hinder the rating process, as discussed in the literature review presented by Peteinatos et al. (2014).

Despite the challenge of plant discrimination under field conditions, Burgos-Artizzu et al. (2011) report that acquiring images in real time, or from prior mapping of a cultivated area, encourages more efficient weed control than does the conventional method, because in the former decisions are specific for the area in the image, rather than for the whole crop. Thus, there are possibilities for the development of a weed management program using the proposed algorithm as a strategy on which to base decisions in relation to spraying and or herbicide dosages. From the perspective of agronomy, the results observed (Table 3) are very promising for reducing the use of herbicides because, with this tool, clusters of weeds or even individual plants have a high probability of being identified, thus allowing the farmer to tailor the treatment geospatially. In other words, it can be used for the processing of “offline” images obtained from representative areas of crops or fed into mechanical weed removal tools, providing input to be utilized in determining pulverization with a direct-injection sprayer. For Feyertaerts and Gool (2001), systems of weed cover identification from images with accuracy rates of at least 80% can promote a reduction in the volume of herbicide applied in up to 90%. In the paper presented by Gerhards et al. (2002), this reduction reached 98% when a prescription map and a pulverization mechanism with a direct-injection sprayer were used.

Materials and Methods

This study was conducted on a single site of approximately 70 m² during the crop growing season of 2014. The study site was located in the proximity of 22°53’ S, 47°05’ W in the municipality of Campinas, in the state of São Paulo in Brazil, it had been used continuously for the production of corn for the previous three growing seasons. The Köppen classification for the climate of Campinas is Cwa, i.e. a subtropical/tropical climate; the average annual temperature is 22.3 °C, the relative humidity is 62%, and the total annual rainfall of 1425 mm. The experimental plot is located at an altitude of 620 m. The soil of the study site was classified as udox (Embrapa, 2006). DKB 310 PRO hybrid corn was planted at a density of 55000 seeds ha⁻¹ in March of 2014, with a row spacing of 0.9 m. The most common weeds which contaminate the grain crops cultivated in the state of São Paulo were selected for the study: Urochloa plantaginea, Urochloa decumbens, Panicum maximum, Euphorbia heterophylla, Ipomoea hederifolia and Ipomoea quamoclit. Approximately 2 kg of a mixture of the seeds from all of these weeds was sown by hand, aiming for an approximate density of 60 seeds m⁻².

Image acquisition

Daily images of the experimental plot were captured from the tenth to the fiftieth day after sowing using an RGB digital camera (Nikon Coolpix P520) set for automatic focusing, exposure, shutter speed and lens opening. The camera was affixed to a tripod at a height of 1.5m. Each
Table 1. Number of images classified by the experts as belonging to each category (low, intermediate or high infestation) and threshold resulting from the analysis of each of the experts. Average percentage of weed coverage for each image category is in parentheses. Each expert reviewed a total of 50 images.

<table>
<thead>
<tr>
<th>Expert</th>
<th>Low</th>
<th>Intermediate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>24 (4.7%)</td>
<td>19 (7.3%)</td>
<td>7 (9.8%)</td>
</tr>
<tr>
<td>A2</td>
<td>31 (3.8%)</td>
<td>13 (6.1%)</td>
<td>6 (8.1%)</td>
</tr>
<tr>
<td>A3</td>
<td>34 (5.9%)</td>
<td>15 (7.4%)</td>
<td>1 (9.6%)</td>
</tr>
<tr>
<td>Average</td>
<td>29.66 (4.8%)</td>
<td>15.66 (6.9%)</td>
<td>4.66 (9.2%)</td>
</tr>
</tbody>
</table>

Numbers refer to the number of images evaluated as revealing this level of infestation by the expert; values in parentheses show the percentage of coverage required for inclusion in that level by that expert.

Fig 1. Image with regions of soil removed, highlighting the information about the plants.

Table 2. Overall performance of the expert system.

<table>
<thead>
<tr>
<th>Performance analysis</th>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (%)</td>
<td>66.67</td>
<td>90.00</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.44</td>
<td>----</td>
</tr>
</tbody>
</table>

Fig 2. Visualization of the application of fuzzy rules to an image.
image covered an effective area of 1.15 by 1.80m (approximate resolution of 38 pixels cm\(^{-2}\)). All images captured two rows of crops and the intervening space between them, whether or not it was occupied by weeds.

**Vegetation segmentation**

The images acquired were analyzed with a customized Matlab program to separate the pixels indicating vegetation from the others using the absolute green method described by Nejati et al. (2008), where the value of *Pixel Color Distance* is obtained by calculating the Euclidean distance applied to the normalized values of the red and green channels of each pixel. This distance is given by:

\[
PCD = \sqrt{pixel(r)^2 + [pixel(g) - 1]^2}
\]

Where, PCD is the distance to absolute green of the pixels, \(pixel(r)\) the value for the pixel in the red channel and \(pixel(g)\) the value for the green channel. The PCD represents a new value for the pixel in a monochromatic image; this was calculated for all the pixels of the image. This Otsu method (Otsu, 1979) involves the iterative analysis of the histogram of the new image to automatically determine a threshold value; this is then used to determine whether each pixel refers to a plant or background (Fig. 1). A total of 70 pictures were taken, and these were evaluated for visual quality; for the development of the algorithm, 50 of the images were included.

**Extraction of information about weed coverage**

For each image, weed coverage was extracted using machine vision. Segmented binary images from which background had been removed so that only vegetation remained (Fig. 1) were processed by computing the area of the objects (connected pixels) present in the image. Objects covering an area larger than a threshold were considered to be the crop, while those smaller than this threshold were considered to be weeds. The threshold was obtained by using the “Kmeans” algorithm (Jain et al. 1999), which divided the total area into the two parts.

After identifying the pixels belonging to the two categories of crop and weed, a template with information about the position of the pixels of the category crop is overlain on the original image: the level of weed infestation is obtained by dividing the remaining pixels in the image by the total area of the image in pixels, as shown in the equation below:

\[
I_{in}(\%) = \left(\frac{ppd}{I_{im} \times C_{im}}\right) \times 100
\]

Where, \(I_{in}\) is the level of weed infestation or estimate of weed coverage, \(ppd\) the number of pixels belonging to the category weed, \(I_{im}\) the image width in pixels, and \(C_{im}\) the image length in pixels.

**Input from Crop Experts**

After the scanning of the experimental plots with the RGB camera, the whole data set of images was sent to three crop experts for the evaluation of weed infestation. They were asked to determine whether the level of infestation of each image was low (1), intermediate (2) or high (3), the three levels used to indicate weed infestation in the definition of management strategies for an area (Longchamps et al. 2012). The three crop experts were specialized in the production of corn and worked in three different regions in the state of São Paulo. All three are accredited agronomists with vast experience in crop production and weed management.

**Automatic estimate of area of weed coverage**

Using an Intel Core 2 CPU computer, 2.13 GHz and 2 Gb Ram, running the Windows operating system and the Toolbox software Image Processing Matlab 9.0 R2011 (Mathworks), an algorithm was developed to estimate the ground area covered by weeds and determine whether a herbicide was necessary or not on the basis of the level of infestation observed.

The model for making decisions was developed by means of a supervised learning technique, i.e., the group of 50 images already analyzed by the experts, was randomly divided into two subsets, with 30 images for training and 20 for testing. An index of the ground area covered by weeds was obtained from each of the images of the training subset; these indices were then associated with the level of infestation attributed by the experts (1 low; 2 intermediate; 3 high); these indices were used to obtain a model of classification based on neural networks and fuzzy logic.

Fuzzy logic is a generalization of classic set theory, where various elements which are uncertain in the analysis of patterns are flexibilized and evaluated according to their level of relevance (Aguirar et al. 1999). Artificial neural networks are designed to simulate the behavior of the nervous system of living beings during the machine learning process, with the strategy consisting of the capture of key elements capable of promoting various interactions between the “neurons”, as happens in the human nervous system. This kind of classifier is adaptable, tolerant of errors, and robust, thus reaching optimal solutions (Peña et al. 2014; Eddy et al. 2014). The integration of fuzzy systems and neural networks in a neuro-fuzzy classifier generates a hybrid system with a high potential for pattern recognition, because of the flexibility, speed and adaptability introduced into the new system (Ghosh et al. 2009; Pradhan 2013; Khoshevisan et al. 2014). The adaptive neuro-fuzzy classifier relied on linguistic barriers capable of altering the primary meaning of the member functions to a secondary meaning. In order to improve the result of the fuzzy rules and rating precision, an additional layer defining the linguistic barriers was added to the proposed network.
These barriers were trained with other network parameters by a training algorithm in a conjugated gradient scale and then tuned to fuzzy sets, thus improving the indices for distinction of overlapping categories. Additional information about the neuro-fuzzy classifier adopted can be obtained in Cetíşli (2010).

The algorithm proposed was evaluated by accuracy analysis, with accuracy referring to the percentage of the total number of images in the testing set classified in agreement with the classification of the experts. In addition to accuracy, Root Mean Square Error analysis (RMSE) also was used as a parameter to describe the performance of algorithm:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=0}^{N} (P_{iA} - P_{iD})^2}$$

Where, RMSE: Root Mean Square Error

$P_{iA}$: Input value

$P_{iD}$: Estimate Value

Conclusions

The proposed algorithm was capable of identifying plants and satisfactorily discriminating weeds from crops. It was also able to estimate the area of the image occupied by weeds and properly classify the image in relation to level of infestation. The accuracy achieved shows that the algorithm can be used in complex problem solving, such as the task of weed recognition in agricultural areas, as well as providing essential subsidies for the formulation of strategies for making decisions in weed management. This paper provides a preliminary study for the development of a smart system for decision-making in weed management, with the information provided by image-processing software being used to make a map of the level of weed infestation which can serve as a guide for localized application; it could also be used to trace the distribution of weed seeds in an area. Its greatest contribution resides in its potential to innovate the management of weeds in Brazil, since the adoption of the proposed software makes more efficient management decisions possible, not only economically, but also in relation to environmental sustainability.

Acknowledgments

The authors would like to acknowledge the grants provided by the Coordination for the Improvement of Higher Education Personnel (CAPES – Process 11249-13-7) and the Foundation for the Support of Research of the State of São Paulo (FAPESP – Processes 2012/19958-7 and 2012/2036-6) for carrying out this research. Thanks also go to Vanusa Sobrinho and Linda Gentry El-Dash for the translation and revision of the text.

References


Embrapa - Empresa Brasileira de Pesquisa Agropecuária.


1236


