Optimisation Of Natural Images Processing Using Different Evolutionary Algorithms

Xavier P. Burgos-Artizzu Angela Ribeiro Alberto Tellaeche Gonzalo Pajares

Abstract—The development of image processing methods to discriminate between weed, crop and soil is an important step for Precision Agriculture, the main goal of which is the site-specific management of crops. The main challenge in terms of image analysis is to achieve an appropriate discrimination in outdoor field images under varying conditions of lighting, soil background texture and crop damage. This work presents several developed computer-vision-based methods for the estimation of percentages of weed, crop and soil in digital images of a crop field. These methods are interchangeable among them, having each one of them a set of input parameters that need to be adjusted differently for each image. Two different evolutionary methods (standard genetic algorithm and NSGA-II) have been used to adjust these parameters and find the best method combinations. The proposed approach can reach a correlation with real data of up to 97% for a set of images acquired from different fields and under different conditions.

I. INTRODUCTION

Precision Agriculture (AP) is a concept that addresses the in-field variability of factors that influence crop growth. It seeks to avoid applying same management practices to a crop regardless of site conditions, [1]. The most significant benefits of site-specific management are cost reduction in the production of the crop and reduction in environmental pollution, [2]. In order to accomplish an AP practice, it is essential to generate a treatment map showing the degree of weed coverage present in each area of the field. To obtain this information, automatic computer-vision-based weed detection methods are widely used.

Research work in this area is difficult to classify and compare due to the variations that each different crop presents. However, almost all existing weed detection methods process the image in two steps: 1) segmentation of vegetation against background and 2) detect which of the vegetation pixels belong to weeds. Some examples of the existing research works can be seen in [3], [4]. Unfortunately, none of these studies have resulted so far in commercialization of the technologies developed. The major problems associated with this concern the high computing and economic costs involved, as well as the difficulties of correctly representing all the possible situations present in real and outdoor conditions.

This paper presents a computer-based image processing system that process images of a crop field, computing the percentages of weed, crop and soil present in the image. The proposed image processing is divided into different steps and for each one of these steps different methods are available. In its turn, each proposed method needs the adjustment of a series of attributes, which determine how the method works. For each particular image, the best set of parameter values, as well as the best combination of methods to be used at each step will be different. Therefore, to fully exploit the potential of the proposed system and check its correct operation, an adjustment of method combination and their parameters values needs to be made.

Such adjustment has to face with an near infinite search space, since each attribute can take an infinite range of values, and also with the lack of knowledge on how each attribute value change affects the global processing. Amongst the various optimization algorithms existent, evolutionary computation methods are known for their ability to deal with vast search spaces as well as for their flexibility, [5], [6]. Moreover, these methods are often used in image processing, because of their ability to solve problems with a high combinatorial complexity relatively fast, [7], [8], [9]. In this work we present the results of using two different evolutionary methods 1) standard genetic algorithm, and 2) NSGA-II.

The rest of this paper is structured as follows. Sect. II describes the proposed image processing. Sect. III presents the results obtained by the proposed image processing on a set of digital images with different natural conditions, after optimization with each evolutionary computation method. Sect. IV outlines the conclusions obtained from this study.

II. IMAGE PROCESSING

A. Image Acquisition

All digital images used for this study were taken in barley fields sited in Madrid (Spain) on different days and therefore under varying lighting conditions. The most common weed found is *Avena sterilis*. This weed specie, at the time of
herbicide applications, is indistinguishable from the barley crops in colour, shape and texture, so that the only possible method of discrimination left is its position. Images are taken between two consecutive crop rows always with the same focal length and as overhead views. The area covered by each image is 0.5m², covering two crop rows and the space in between. Moreover, soil samples were collected from the same plot and at the same time as when the digital images, and a bio-mass laboratory measurement was made (amount of weed), so that each image has its associated bio-mass value.

B. Main Objectives

Given any sample image, the system has to be able to compute the Weed Pressure (WP), WP = c_w/(100 - c_s), expressed as a function of the percentages of weed cover (c_w), crop cover (c_c), and soil (c_s). The WP can be used to measure the risk the weed represents and therefore the dosage of herbicide required, [10]. Moreover, since the value of the WP represents the proportion of weed present in the area covered by the processed image, it can be compared with the bio-mass value associated to the image in order to evaluate the goodness of the performed processing.

In this context, the main challenge in terms of image analysis is to achieve an appropriate discrimination among weed, crop and soil in outdoor field images under varying conditions of lighting, soil background texture and crop damage. The images in the first row of Fig. 1 are illustrative examples of such conditions. From an initial standpoint one can see that it is very difficult to distinguish between crop and weed due to similarities in shape or colour, so that the only possible discriminatory characteristic left is the spatial position each one occupies. Formally, the vegetation cover grows between two crop rows is considered as weed cover. Therefore, the initial problem is converted into another one: to detect the crop rows and identify as weed only the vegetation growing in between.

This new perspective of the problem reduces its complexity but still presents many difficulties. For example, sometimes weed grows leaned against the crop, being hard to determine where the crop ends and where the weed begins, as in Fig. 1(B). Another problem is shown in Fig. 1(D), where the crop rows are intermittent, or practically non-existent, due to sowing errors. Moreover, the state of crop growth can vary from image to image, as for example in Fig. 1(C), where the crop rows are so thick that weed cannot be seen. Also, it is essential to consider lighting conditions if appropriate discrimination between weed, crop and soil has to be achieved. The images shown in Fig. 1 show different illumination levels corresponding to different weather conditions (sunny or cloudy days).

C. Proposed Image Processing

The image processing system distinguishes three different components from each input image and then uses their total sizes (in pixels) to compute the relative percentages. In the proposed approach, once images are taken, the steps involved are: a) to transform the colour image into a black&white image where vegetation and non-vegetation are separated and the vegetation cover appears in white (segmentation), b) to detect and eliminate from vegetation pixels those that are part of crop rows (elimination of the crop rows), and c) to determine which white zones remaining in the image are considered as weed and which not (weed extraction). At each stage a different component is extracted: in a) soil, in b) crop and in c) weed. For each step different methods are proposed with the aim of achieving a robust discrimination.

1) Image segmentation: As said before, the objective of the first step is to convert the input digital red-green-blue (RGB) image into a black&white image, where the vegetation parts of the input image (weed and crop) will be represented with white pixels and the rest (soil) as black. This is a crucial step in the image analysis. In fact, if the segmentation is incomplete or erroneous, the entire process will suffer from it. The illumination conditions on which the image was taken play a decisive role in the success of this step. Two methods of segmentation have been developed: S1 and S2.

In the first method, S1, the segmentation is performed by a lineal combination of three colour (RGB) planes that describe each original image pixel into a one-dimensional gray level (monochrome) image, where vegetation pixels present high intensity values and non-vegetation pixels low intensity values: Monochrome = r * input_red + g * input_green + b * input_blue. The values of real coefficients r, g, b of the lineal combination will determine how the monochrome image is constructed and are the key to obtain the best discrimination between vegetation and the rest, [10], [11]. After the monochrome image has been produced, S1 applies a threshold operation, setting the contrast breakpoint between pixels containing vegetation and pixels containing non-vegetation, transforming the monochrome image into a binary black&white image.

The second segmentation method, S2, converts directly each RGB image pixel to black or white depending directly
S1 \( (r = -1, g = 2, b = -1, \text{thresh} = 237) \)

S2 \( (\text{min intensity} = 50, \text{max intensity} = 200) \)

E1 \( (p_{\text{elimination}} = 65\%) \)

E2 \( (\text{length} = 15, \text{min} = 5) \)
on its colour. Vegetation pixels show always a strong green component, a weak blue component and are neither too bright nor too dark. Formally, an input RGB pixel will be considered as a vegetation pixel if 1) $\text{input\_green} > \text{input\_red} > \text{input\_blue}$, 2) $\text{input\_green} \geq \text{min\_intensity}$ and 3) $\text{input\_blue} \leq \text{max\_intensity}$, where $\text{min\_intensity}$ and $\text{max\_intensity}$ are input parameters used to discard any pixel belonging to shadows (low intensity) or light reflections (high intensity) that might appear in the input image.

Results of both methods on four different images (for a fixed value of their parameters) can be seen in the second and third row of Fig. 1.

2) Crop Row Elimination: The goal of the second step is to detect and eliminate from vegetation pixels those that are part of crop rows, leaving intact those that represent weed. This stage will have to overcome many difficulties, as different states of crop and weed growth, sow errors, or infestation. Three methods for crop row elimination have been developed: $E1$, $E2$ and $E3$.

The first and simplest elimination method, $E1$, exploits the geometrical distribution and shape of crop in the sample images (as vertical columns) due to the fact that all images are taken as overhead views. Simply, all columns in the binary image whose number of vegetation pixels is greater than or equal to $(\text{height} \times p_{\text{elimination}})$ are eliminated (set to black). The simplicity of $E1$ causes it to be unable to adapt to difficult conditions, as for example when the crop rows are not homogeneous or when the weed is mixed with the
crop, figures 1 B and C. However, this simplicity also makes E1 the fastest method.

The second elimination method, E2, performs a more complex and exhaustive analysis of the image. E2 first finds the neuralgic column centres of each crop row as the first and last image columns where the number of vegetation pixels is closest to the total height of the image, or number of rows. Both centres are then used as starting points for an independent and separate analysis of each image row. Lets suppose the algorithm is currently exploring pixel \( p = input(i, j) \) where \( i \) is a fixed image row and \( j \) is the neuralgic column centre of one of the crop rows. To eliminate all pixels belonging to the crop at image row \( i \), E2 moves to the left and to the right eliminating all white pixels encountered, until left and right edges of the crop are reached. The edges are identifiable as where the transition from groups of white pixels (crop) to groups of black pixels (space between crop rows) occurs. However, since the binarized image result of segmentation is not perfect, groups of black pixels can exist inside the crop region, causing the illusion that the edge of the crop has been reached. The decision on whether the edge has been reached or not must therefore take into account this, allowing the exploration to go on even when some black pixels are encountered. This is achieved by E2 using an exploration window of length pixels and the threshold value \( \text{min} \) as outlined in figure 2, where length and \( \text{min} \) are its input parameters.

\[
\text{IF} \quad (\text{WhitePixels}(input(i, j \ldots j + \text{length})) > \text{min}) \\
\text{THEN} \quad input(i, j \ldots j + \text{length}) = 0; \\
\quad j = j + \text{length}; \\
\text{ELSE} \quad i = i + 1; \quad \text{//jump to next row}
\]

Fig. 2. Outline of E2 code. While the number of white pixels encountered is above \( \text{min} \), the algorithm eliminates all pixels inside the exploration window and continues. If not, the algorithm considers that a crop edge has been reached and jumps to the next image row.

A third approach is proposed that combines the achievements of both previous methods. E3 has the same basic structure than E2, but redefines the decision on whether the crop edge has been reached or not, using border pixel data. E3 uses Sobel borders to identify black-to-white and white-to-black transitions, so that crop edges and any other form of sudden pixel value change are clearly marked. The workings of E3 are explained in table I. For each crop row, E3 will perform an independent exploration at each image row starting from the neuralgic column centre and moving to the right and to the left until crop edges are reached. Let us suppose the method starts, at the pixel \( p = input(i, j) \), where \( i \) is a fixed image row and \( j \) is the neuralgic column centre of the crop row. As long as all explored pixels are white, the method is sure to be inside the crop and therefore keeps eliminating all of them. However, if a border pixel is found, it can only mean two things: the exploration has reached either a crop edge or a group of black pixels inside the crop. The main difference between the two is the distance to the next border or white pixel: in the case of a crop edge, the next change will be far away (a weed edge or another crop edge), while in the case of black pixels inside the crop, the next change will be much closer. Therefore, distance can be used as a good discriminative. E3 uses distance thresholds:

\[
\theta_1 \leq \theta_2
\]

Below \( \theta_1 \) is always considered too short a distance for a crop edge, while above \( \theta_2 \) is the exact opposite and is always considered a long distance and therefore a proof of having reached a crop edge. A distance between \( \theta_1 \) and \( \theta_2 \) is considered as not definitive in either way, so that the decision is taken looking at the ratio of vegetation pixels present at the same image columns in the rest of the image rows. If a crop edge has been reached, all other image rows will have lots of black pixels so the proportion of white pixels will be low, and viceversa. E3 allows the manual adjustment of the proportion threshold that discriminates between the two by its parameter \( \text{min} \cdot \text{proportion} \).

Results of the three crop row elimination methods (for a fixed value of their parameters) can be seen from fourth to sixth row of Fig. 1, using the result of method S1 as input.

3) Weed extraction: The goal of the previous step was to completely eliminate crop from segmented image, leaving an image where all remnant white pixels represent weed. However, natural images can present a great number of different situations to which is very difficult to adapt, resulting that not always the crop rows can be totally detected and fully eliminated. Therefore, a third step is necessary, where weed needs to be filtered from other remaining white pixels. More precisely, this stage must remove all white pixels outside the area delimited by the crop rows that were eliminated in the previous stage, since weed only grows between crop rows, and then differentiate from the rest of white pixels those that represent weed from noise. Two different methods have been developed: F1 and F2.

Most of the weed species and in particular the *Avena sterilis* grows in small associations, called patches, so that usually where a weed plant grows others develop in its...
vici. This association is the fundamental characteristic taken into account to determine which remaining white pixels belong to weed and which are just noise. The simplest way to discriminate between them is to group each pixel with its neighbours (region extraction) and to eliminate all regions whose size in number of pixels is too small to be considered weed. This is precisely what \( F_1 \) does. It groups all white pixels in different regions, where each region contains all pixels that are connected between them by a 3 x 3 neighbourhood. Then, \( F_1 \) finds the region of largest size and eliminates all pixels belonging to regions whose size is smaller than largest * percent, where percent is an input parameter.

The second approach, \( F_2 \), is a further amelioration of \( F_1 \). Before the regions are computed, \( F_2 \) performs a morphological opening of the image using a 3x3 neighbourhood binary operator, and then the same region extraction method than \( F_1 \) is used. Finally, \( F_2 \) enhances the remaining white regions by a morphological dilation, in order to get more accurate results.

Results of both weed extraction methods (for a fixed value of their parameters) can be seen in seventh and eighth row of Fig. 1, using the result of method \( E_3 \) as input.

### III. Adjustment Using Evolutionary Algorithms

The adjustment algorithm has to find the best possible combination of methods and input parameter values for a given set of images, i.e., reach the highest statistical correlation between the system computed weed\_pressure values and real bio-mass values associated to each image. The test set contains 136 images, divided in two groups: 65 being collected in sunny days and 71 in cloudy days, each group being adjusted separately.

#### A. Adjustment Using a Standard Genetic Algorithm

First, a standard genetic algorithm (GA) was used. Each method combination was evaluated separately, in a different running of the GA, so that the GA individuals only had to encode the value of the input parameter values needed by the methods used in the current running. The genome length ranged from 6 (S1-E1-F1) to 9 (S2-E3-F2) genes, being all genes a real value, generated randomly. Each individual, at each generation, had to process all images of the test set using the fixed combination of methods and the parameter values encoded in its genome, obtaining for each individual processing more than 60 images at each algorithm iteration, the population size had to be restricted to only 30 individuals. To overcome the limitations that such small population causes, such as poor variability and small covered search space, a very high mutation probability was used (30%), combined with elitism to prevent loosing good solutions. The rest of parameters of the GA was fixed in a classical fashion: Gaussian mutation and Scattered selection functions, with a selection probability of 70% [6], [5]. Stopping criterion was to obtain a Pearson correlation coefficient

<table>
<thead>
<tr>
<th>Type of current Pixel ( p = \text{input}(i, j) )</th>
<th>Distance ( d ) (in pixels) until next pixel not black ( n = \text{input}(i, j + d) )</th>
<th>( d \leq \theta_1 )</th>
<th>( \theta_1 &lt; d \leq \theta_2 )</th>
<th>( d &gt; \theta_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>Erase all pixels from ( p ) to ( n ) and jump to ( n ) ( (p = \text{input}(i, j + d)) )</td>
<td>Erase all pixels from ( p ) to ( n ) and jump to ( n ) ( (p = \text{input}(i, j + d)) )</td>
<td>go to next row ( (i = i + 1) )</td>
<td></td>
</tr>
<tr>
<td>Border</td>
<td>Erase all pixels from ( p ) to ( n ) and jump to ( n ) ( (p = \text{input}(i, j + d)) )</td>
<td>IF ( \text{Whites}(1 \ldots N, j \ldots j + d) &gt; \text{min_proportion} ) THEN Erase all pixels from ( p ) to ( n ) and jump to ( n ) ( (p = \text{input}(i, j + d)) ) ELSE go to next row ( (i = i + 1) )</td>
<td>go to next row ( (i = i + 1) )</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>Jump to ( n ) ( (p = \text{input}(i, j + d)) )</td>
<td>Jump to ( n ) ( (p = \text{input}(i, j + d)) )</td>
<td>go to next row ( (i = i + 1) )</td>
<td></td>
</tr>
</tbody>
</table>
higher than 99% or to have been in a stall for the last 100 generations. In average, 500 generations were necessary, representing in our actual working conditions (Pentium IV, 3.2 Ghz, 1Gb RAM) more than 9 days of constant running (an average of approximately 0.9s per image processed).

Table II shows the results of the genetic algorithm over the two different sets of images: Cloudy (71 images) and Sunny (65 images). The proposed methods show very satisfactory results, reaching a mean correlation coefficient with biomass of 87.1%, with peaks of 96%. As predicted, the system adjusts more easily the images taken on cloudy days (91%), than those of sunny days (83%), due to the difficulties introduced by differences in illumination. Best combinations are S1-E3-F1 and S1-E3-F2, being capable of adjusting both sets of images with correlation coefficients of 93% and 93.7% respectively.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Cloudy</th>
<th>Sunny</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-E1-F1</td>
<td>0.915</td>
<td>0.728</td>
</tr>
<tr>
<td>S1-E1-F2</td>
<td>0.909</td>
<td>0.890</td>
</tr>
<tr>
<td>S1-E2-F1</td>
<td>0.924</td>
<td>0.881</td>
</tr>
<tr>
<td>S1-E2-F2</td>
<td>0.957</td>
<td>0.837</td>
</tr>
<tr>
<td>S1-E3-F1</td>
<td>0.952</td>
<td>0.908</td>
</tr>
<tr>
<td>S1-E3-F2</td>
<td>0.961</td>
<td>0.901</td>
</tr>
<tr>
<td>S2-E1-F1</td>
<td>0.904</td>
<td>0.704</td>
</tr>
<tr>
<td>S2-E1-F2</td>
<td>0.772</td>
<td>0.788</td>
</tr>
<tr>
<td>S2-E2-F1</td>
<td>0.918</td>
<td>0.834</td>
</tr>
<tr>
<td>S2-E2-F2</td>
<td>0.912</td>
<td>0.811</td>
</tr>
<tr>
<td>S2-E3-F1</td>
<td>0.903</td>
<td>0.824</td>
</tr>
<tr>
<td>S2-E3-F2</td>
<td>0.924</td>
<td>0.867</td>
</tr>
</tbody>
</table>

### Table II
**SYSTEM RESULTS USING A STANDARD GENETIC ALGORITHM.**

**B. Adjustment Using NSGA-II**

Since each parameter can take values in an infinite domination and that the time required for the processing of a great number of images is considerable, the adjustment using GA, even if it gave satisfactory results, was an arduous task. Therefore, the application of a more direct and fast method of optimisation was considered. For this purpose, the Non-dominated Sorting Genetic Algorithm (NSGA-II) proposed by [12] has proved to work very well under conditions similar to ours. Moreover, NSGA-II is multi-objective, so it permitted us to include into the adjustment another crucial factor not taken into account by GA: time.

Instead of fixing the method combination, this time it was included into the genome. The first 3 genes were used to encode which method was used at each step, and then the value of all input parameters for all methods had to be included into the genome, since the method combination can change in run-time due to crossovers and mutations. Genome length was therefore of 20 genes, the first 3 being integers, and the rest real values.

Each individual, at each generation, had to process all images of the test set using the configuration encoded into its genome (method combination and parameters value), obtaining for each image a corresponding weed pressure value. As was done for the GA, this weed pressure values are then compared with bio-mass values using Pearson correlation coefficient, being the maximization of this value the first objective function of NSGA-II. The second objective function, as mentioned earlier, is to minimize the CPU time used: simply, each individual record how much time the processing of all images took. Since NSGA-II makes a much better use of memory the population size could be incremented up to 80 individuals. For more information on the experimental setup of NSGA-II algorithm, please see [12].

Table III shows the main results of NSGA-II on the same two test sets. Since NSGA-II is multi-objective, best results can be seen from different points of view, which in this case are: best individual in terms of the first objective (maximum correlation), fastest individual (minimum CPU time) and best overall individual.

As in GA, best combinations are S1-E3-F1 and S1-E3-F2, while combinations S1-E2-F1 and S1-E2-F2 reach coefficients a bit inferior but are almost half a second faster. As expected, fastest method combination is S1-F1-E1, but achieves a much poor adjustment. The NSGA-II achieved a mean correlation coefficient with bio-mass of 89.4%, with peaks of 97%, slightly better than GA. All results obtained with the NSGA-II were achieved in 300 iterations (almost half the iterations needed by the GA) which in our case meant a reduction in running time of almost 86 hours (Pentium IV, 3.2Ghz, 1Gb RAM). Moreover, with the GA each running meant only the parameter adjustment of a single method combination, while with NSGA-II all method combinations are being adjusted simultaneously, so the improvement is considerable.

### IV. CONCLUSIONS

In this work different methods for natural image processing are proposed in order to reach an optimal discrimination among weed, crop, and soil. For each step of the image processing, different interchangeable methods have...
been developed, achieving a robust discrimination that works under all possible situations. Moreover, all the methods can be adjusted by their input parameters, allowing to refine one step further the processing to each of the circumstances found.

An adjustment on the values of the parameters and method combination has been performed using a genetic algorithm and NSGA-II. This adjustment has been performed on two different test sets of images. The results of the computed best values have been compared with bio-mass (taken simultaneously and in the same points as images), obtaining mean correlation coefficients of 89.7% in average and of 97% in the best cases, values that prove the goodness of the proposed methods.

This work is part of a research project whose main goal is the development of a spraying system integrated by a computer-vision-based system, able to detect weeds in real-time, and a sprayer that operates where weed is identified [13]. The proposed methods show very good results, both in the correct detection of weeds and in small computational complexity, so that they will be used as starting point to the development of a real-time vision system (25fps), [14].

**Acknowledgement**

The Spanish Ministry of Education and Science (MEC) and the CAM provides full and continuing support for this research work through projects AGL2005-06180-C03-03 and PRICIT-CAM-P-DPI-000176-0505(ROBOCITY2030). The first author currently holds a Ph.D. scholarship given by the Spanish Council for Scientific Research (CSIC).

**References**


