Measuring Leaf Area in Soy Plants by HSI Color Model Filtering and Mathematical Morphology

M Benalcázar 1,3, J Padín 1,3, M Brun 1, J Pastore 1,3, V Ballarin 1, L Peirone 2,3 and G Pereyra 2,3

1 Laboratorio de Procesamiento Digital de Imágenes, Departamento de Electrónica, Universidad Nacional de Mar del Plata, Buenos Aires, Argentina.

2 Laboratorio de Fisiología Vegetal, Unidad Integrada Balcarce (Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata - Instituto Nacional de Tecnología Agropecuaria).

3 Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Mail: marco_benalcazar@hotmail.com

Abstract. There has been lately a significant progress in automating tasks for the agricultural sector. One of the advances is the development of robots, based on computer vision, applied to care and management of soy crops. In this task, digital image processing plays an important role, but must solve some important problems, like the ones associated to the variations in lighting conditions during image acquisition. Such variations influence directly on the brightness level of the images to be processed. In this paper we propose an algorithm to segment and measure automatically the leaf area of soy plants. This information is used by the specialists to evaluate and compare the growth of different soy genotypes. This algorithm, based on color filtering using the HSI model, detects green objects from the image background. The segmentation of leaves (foliage) was made applying Mathematical Morphology. The foliage area was estimated counting the pixels that belong to the segmented leaves. From several experiments, consisting in applying the algorithm to measure the foliage of about fifty plants of various genotypes of soy, at different growth stages, we obtained successful results, despite the high brightness variations and shadows in the processed images.

1. Introduction

An important aspect that emerges in agriculture is the identification of new genes to improve soy crops. This research should allow for new varieties of soy plants with a strong immunity to diseases, requiring fewer quantities of chemical products and water. This kind of research requires constant monitoring and evaluation of plants growth over the time. For this reason, it is very important the accurately measuring of physical characteristics of plants, called the phenotype, such as area and leaf angle distribution, which allows the researchers to match it with the genetic structure of the individuals, called the genotype. Artificial vision, or computer vision, is a technique from artificial intelligence, which allows for non-invasive automation of these tasks, with a significant increase in the capacity of information processing [1].

One of the most significant problems at applying machine vision for automatic phenotyping of plants is to find an automated and efficient method for segmentation of plants. The main issue is present
during the image acquisition, which is usually conducted in complex environments, exposed to continuous lighting changes. Different intensities and shadows in the images make too difficult the segmentation process because it causes deterioration of color and contrast, which are features used to separate the plant from the background of the image.

For the above reasons, it is fundamental to use the right digital image processing techniques, applying a profitable segmentation algorithm with parameters independent from brightness variations, which should reduce significantly the effects caused by lighting changes. In this way, it is possible to achieve optimal results on a large quantity and variety of images. In this paper we propose an automatic method for measure the foliage area after the leaves segmentation of soy plants, which showed a good performance on several light conditions.

2. Materials and Methods

2.1. Images

The images used for this work were obtained from an automated platform, which enables to photograph soy plants automatically. This platform is operating at the Unidad Integrada Balcarce that depends on the Facultad de Ciencias Agrarias, from the Universidad Nacional de Mar del Plata, and the Instituto Nacional de Tecnología Agropecuaria (INTA). The digital images have top and side views of soy plants with a size of 640x480 pixels, in the RGB color model and JPG format.

2.2. Automatic measurement method of leaf area

The proposed algorithm for automatic measurement of leaf area from digital images includes the following steps:

1. Color filtering
2. Binarization
3. Noise filtering
4. Foliage segmentation
5. Leaf area estimation

2.2.1. Color filtering

The color filtering is a fundamental and essential step previous the segmentation, because color is a powerful descriptor that simplifies object identification and extraction from a scene [2,3]. However, this process has different issues associated to the image acquisition, such as those caused by the changes in lighting conditions and when the background of the scene contains surfaces and objects with similar colors to the target [4]. Therefore, for color filtering it is important to select an appropriate color model that separates the brightness from chromaticity information. Chromaticity is used to distinguish the interest-object or target from the image background [5].

The color model selected to perform the filtering was the HSI model (hue, saturation and intensity). This model, different from others such as RGB (red, green and blue) that is oriented toward hardware [6], is very attractive for digital image processing because it represents a color using three components represented by the two solid cones shown in Figure 1. In this representation the hue (H) and saturation (S) are intimately related with the chromaticity information and are separated from the intensity (I). Hue is an attribute that describes a pure color and it is associated with the dominant wavelength in a mixture of light waves, their values are normalized in the range from 0 to 1. In this work, in order to simplify the image analysis and the filter definition, it will be ranged from 0 ° to 360 °. The saturation gives the degree to which a pure color is diluted by the white light, and their values are in the range from 0 to 1. The intensity represents the brightness information and their values are in the range from 0 to 1 [2].
The color filter proposed in this paper is intended to distinguish the soy plant from the background of the image, by defining a region in the HSI color space which represents the green color of plants. For the thresholds selection of the filter, the images were transformed from the RGB to the HSI model, using the conversion equations defined in [2]. Then, it was analyzed the color distribution of leaves at the HS plane (Figure 2). From this analysis it was determined that the hue for soy plants varies approximately in the range from 20° to 180°, whereas the saturation varies in the whole range of S (0 to 1). This region belongs to the range of yellow and green colors of the HS plane in the HSI model (Figure 3). The pixels intensity was approximately in the range from 0.2 to 1, whit the highest frequency in the range from 0.3 to 0.8. This variability in the intensity is due to images were taken in a greenhouse environment at different times of a day, with non-controlled daylight.
Figure 3. Color distribution in the HS plane of the HSI model with an intensity of 0.7

Depending on the analysis above, the filter between the plant and the rest of the scene is defined by the following equation:

\[
H_f = \begin{cases} 
H & \text{if } 20^\circ \leq H \leq 180^\circ \\
0^\circ & \text{otherwise}
\end{cases}
\]

\[
S_f = \begin{cases} 
S & \text{if } 20^\circ \leq H \leq 180^\circ \\
0 & \text{otherwise}
\end{cases}
\]

\[
I_f = \begin{cases} 
I_{\text{freq}} & [0.2 \ 0.9]
\end{cases}
\]

Where \(H_f\) is the filter output for hue and \(H\) is the original hue. \(S_f\) is the saturation of colors that passed the filter and \(S\) is the original saturation. \(I_f\) is the output for intensity and \(I_{\text{freq}}\) is the most frequent intensity in the range from 0.2 to 0.9 in the original image. The most frequent value, of the initial intensities in the mentioned range, was assigned to the filter output intensity to uniform the brightness across the whole image, and thus minimize the possible effects of shadows or excessive lighting at certain regions of the processed image.

After the image color filtering, it was converted back to the RGB model. The conversion was performed using the equations for this purpose defined in [2].

2.2.2. Binarization

Before foliage segmentation, it was performed a binarization of the resulting color filtered image, expressed in the RGB model. For this task the G (green) channel was used, because it contains the most useful information about the target. The image corresponding to this channel was treated as a grayscale image, to which was applied a contrast filter, via a linear transformation in the spatial domain, defined by the following equation:
\[ G_c = \frac{G}{G_{\text{max}}} \]  

(2)

Where \( G_c \) is the gray level of a pixel after the linear transformation, \( G \) is its initial gray level and \( G_{\text{max}} \) is the highest gray level among all the pixels in the G channel as it is shown in Figure 4.

For the selection of the binarization threshold the Otsu method was used [4]. This method selects a threshold that minimizes the intra-class variance (or the measure of separability of the resulting classes in gray levels) of black and white pixels. The resulting image from the binarization contains the interest-object represented by a 1 (white), and the background represented by 0 (black). This figure contains a certain amount of noise, both in the object of interest as in the background.

![Figure 4. Contrast linear transformation function](image)

2.2.3. Noise filtering

This step involved the removal of noise from the objects of interest, through the removal of objects smaller than 500 pixels. This value was selected heuristically based on the fact that all the size of branches and foliage of the plant, for all images analyzed, were above this value. The identification of objects was done by a connectivity 8 criterion. In a binary image, connectivity 8 defines a neighborhood between pixels in the horizontal, vertical and diagonals axis, as shown in Figure 5.

![Figure 5. Graphical representation of connectivity 8 in a binary image](image)
Having identified the objects in the image, we proceeded to count the pixels present in each of these, and then eliminate those with smaller size than the threshold, and so we obtained a new noise-free binary image.

2.2.4. Foliage segmentation

The next step was separate the stems from the leaves, to measure the area of the leaves. The foliage segmentation was done by Mathematical Morphology. The operation used was the opening with a disk-type structural element of radius 10 (in pixels). This size corresponds roughly to the radius of the stems to be deleted. Opening is an operation consisting of an erosion, which removes objects smaller than the structuring element, in this case stems, followed by a dilation with the same structuring element used for erosion, which recovers the approximate shape of the eroded objects, in this case the leaves. The mathematical detail of the operations mentioned above can be found in [8,9].

As the final stage of segmentation, it was applied a process of elimination of holes inside the leaves. A hole was defined as a small set of background pixels that is not connected with the edges of the image [10]. The holes appear in the images because in some cases plants have an identification tag on its leaves, which is erased during the color filtering. Holes also originate due to a non-uniform and intense illumination of some leaves.

Figure 6 shows graphically the procedure used for the segmentation. To observe the effect of the operations, we assigned to the pixels that define the plant and the structuring element a gray value of 0 (black) and the background a gray value of 1 (white). We also superimposed to each result the original edges. In this way, we can see that the opening removed from the plant the stems and some leaves.
details. This does not significantly affect the final result because it still retains the bulk of the leaves. It can also be noticed that some leaves contain small holes that were eliminated after the opening.

2.2.5. Leaf area estimation
In this last stage, the identification of leaves was done using a similar procedure to the one used for the identification and subsequent elimination of small objects described in section 2.2.3. Once identified the pixels belonging to each of the leaves, using the connectivity 8 criterion, we proceeded to count them. The number of pixels of each leaf estimates the foliage area of it and can be converted to area units by multiplying by a conversion factor that depends, among other factors, of the distance between the camera and the plant.

Because in some cases there was an overlap of the leaves, especially in those images that contain a top view of the plant, the leaves overlapping were considered as a single object.

3. Results

Figure 7. (a) Original image and resulting images from: (b) color filtering, (c) binarization of the G channel image resulting from the color filtering, (d) removal of objects smaller than 500 pixels and (e) final image with the area of each leaf and the total area in pixels.
In order to illustrate the steps used in the proposed algorithm, Figure 7 shows the results of the different stages of the algorithm applied to a side image of a soy plant, with high level of illumination. Figure (7a) presents the original image. Figure (7b) shows the result of applying the color filter in the HSI model. Figure (7c) shows the result of applying the binarization to the G channel of the resulting RGB image from the color filtering, with the selection of threshold by Otsu method. Figure (7d) shows the result of applying the noise filtering, through removing the objects smaller than 500 pixels and, finally, figure (7e) shows the result of the segmentation and subsequent estimation of area of each leaf. To improve the presentation of results, we superimposed the edges of the segmented leaves to the original image. Additionally, we added the value of the area in pixels in the center of each leaf and, the total value of area at the bottom right of image, which is the sum of all areas and corresponds to the foliage area estimation of the whole plant.

Figure 8. (a), (b) and (c) are results for images with top view and different levels of illumination. (d), (e), (f), (g) and (h) are results for images with side view and different levels of illumination.
Figure 8 shows some results obtained for several images of top and side views of soy plants, with different light levels and morphology. In images (a), (b) and (e) we can see that some leaves have not been detected due to excessive lighting in certain regions of the image, which deteriorates the color. This effect causes that during color filtering, some leaves are removed and considered as part of the background. In the case of images (c), (d), (f) (g) and (h) lighting is uniform, but varies from one image to another, however the algorithm works well in all these cases because the color filter is based on the hue and saturation and not at the level of illumination of the image. It should also be noticed that in some detected leaves, some details have been suppressed during the application of the morphological opening. This last fact is not significant because the estimated leaf area value is used only for comparative purposes.

Conclusions
This paper presents an algorithm to measure, automatically, the foliage area from an image of a soy plant. Because we used a color filter, based on the HSI model, together with the appropriate selection of the thresholds for hue, saturation and intensity, a correct detection of plants was achieved, regardless of the different levels or changes in brightness of the processed images.

Despite the wide variety of processed images tested, in many cases with significant differences, as a result of changes in lighting conditions during the acquisition stage, we were able to set general values for the parameters used in each stage of the algorithm, allowing us to process large groups of images, with a little processing time, and no human intervention, which, along with a proper presentation of the results, facilitates the analysis and comparison of growth for different soy genotypes.

References
[1] Biskup B., Scharr H., Schurr U., Rascher U.; 2007; A stereo imaging system for measuring structural parameters of plant canopies; Plant, Cell and Environment 30; pp. 1299-1308.
[2] Gonzalez R., Woods R.; 1992; Digital Image Processing, Addison - Wesley; pp. 221-248.
[3] Ordoñez C., Bouchet A., Pastore J., Blotta E.; 2011; Aplicación de filtros frecuenciales y cromáticos en tiempo real para la toma de biopsias dirigidas en endoscopias; V Congreso Latinoamericano de Ingeniería Biomédica (CLAIB2011).
[4] Chiho K., Bum-Jae Y., Mun-Ho J., Hagbae K.; 2008; Color segmentation robust to brightness variations by using B-spline curve modelin; Pattern Recognition 41; pp. 22-37.
[5] Qazi I., Alata O., Burie J., Moussa A., Maloigne C.; 2011; Choice of a pertinent color space for color texture characterization using parametric spectral analysis; Pattern Recognition 44; pp. 16-31.
[6] Bouchet A., Pastore J., Abras G., Jury G., Jury H., Ballarin V.; 2007; Barrett esophagus: guided biopsies through digital image processing; Journal of Physics; vol. 90; ISSN 1742-6596 (On-Line); ISSN 1742-6588 (Print).
[7] Otsu N.; 1979; A Threshold Selection Method from Gray-Level Histograms; IEEE Transactions on Systems, Man, and Cybernetics; vol. 9; pp. 62-66.
[8] Serra J.; 1989; Image Analisis and Mathematical Morphology; Academic Pres; vol. 1; pp. 43-56.
[9] Gonzalez M., Ballarin V., Celián A., Sánchez V., Rapacioli M., Flores V.; 2009; Analisis de axones en crecimiento mediante la función Granulométrica de Tamaños; XVII Congreso Argentino de Bioingeniería - VI Jornadas de Ingeniería Clínica (SABI); ISBN: 978-950-605-505-9.
[10] Soille P.; 2003; Morphological Image Analysis: Principles and Applications; Springer - Verlag; pp. 210.