

# Recognition and Quantification of Area Damaged by *Oligonychus Perseae* in Avocado Leaves

Gloria Díaz<sup>1,\*</sup>, Eduardo Romero<sup>1</sup>, Juan R. Boyero<sup>2</sup>, and Norberto Malpica<sup>3</sup>

<sup>1</sup> Universidad Nacional de Colombia, Colombia

<sup>2</sup> Centro de Investigación y Formación Agraria Cortijo de la Cruz, Spain

<sup>3</sup> Universidad Rey Juan Carlos, Spain

**Abstract.** The measure of leaf damage is a basic tool in plant epidemiology research. Measuring the area of a great number of leaves is subjective and time consuming. We investigate the use of machine learning approaches for the objective segmentation and quantification of leaf area damaged by mites in avocado leaves. After extraction of the leaf veins, pixels are labeled with a look-up table generated using a Support Vector Machine with a polynomial kernel of degree 3, on the chrominance components of YCrCb color space. Spatial information is included in the segmentation process by rating the degree of membership to a certain class and the homogeneity of the classified region. Results are presented on real images with different degrees of damage.

**Keywords:** Leaf damage, segmentation, quantification, machine learning.

## 1 Introduction

The perseae mite, *Oligonychus perseae*, is a serious pest of avocado harvesting and every year it results in high economical losses for the productive sector. The presence of the mite in avocado crops is easily recognizable by the damage produced, as nearly circular necrotic regions of brownish color in the underside of leaves, distributed mainly along the central vein and other main veins (view 1). Manual delineation of leaves can be tedious and time consuming, specially when a high number of leaves needs to be analyzed. In order to study the susceptibility of various avocado crops to the *O. perseae* in the south region of Spain, we wanted to quantify the mite feeding damage, using an image analysis technique to count brown spots on leaves and calculate the percentage of damaged leaf area. Several

---

\* Gloria Díaz is supported by a grant from the Colombian Institute for the Advancement of Science and Technology (COLCIENCIAS), Grant no. 109-2005.

methods have already been proposed for this task. Kerguelen et al.[1] compared three methods for spot detection and obtained the best results with a simple color thresholding and heuristic constant thresholding. This procedure is not robust when different leaves have to be analyzed, or when spots of different age are considered, as it was the case for the present study. Wijekoon et al. [2] evaluate a method for quantification of fungal disease using Scion software. The method is interactive, and several parameters must be adjusted for each leaf. Bakr [3] describes a software for leaf damage quantification, based on color classification of image regions. We have tested the software on our own leaves and the automatic segmentation gives very poor results (not reported here).

In the present paper, a semi-automatic approach for determining the damaged area caused by the *Oligonychus perseae* mite in avocado leaves is designed, implemented and validated. This is achieved in three main steps: First avocado leaves are digitized and segmented from the background using a simple Otsu thresholding [4]. A preprocessing step is applied in order to standardize the leaf color distributions and to enhance the image contrast. Then, images are segmented using a two stage color pixel classification approach: the principal veins are first extracted and the remaining pixels are then classified into damaged area or healthy leaf. Finally, the resulting segmentation is improved using a function which combines the degree of membership to the labeled class and the homogeneity of the neighborhood region.



**Fig. 1.** Different damage levels of avocado leaves caused by *Oligonychus perseae*, from high (left) to low (right) damage levels. Notice how the intermediate and low levels present a very blurry regions in which, it is very difficult, even for a human observer, to establish the presence of the disease.

## 2 Materials and Methods

### 2.1 Image Acquisition

Thirteen leaves of avocado (*Persea americana*), of the Hass variety, damaged by *Oligonychus Perseae*, were harvested from a plantation in the province of Malaga, in Spain. Leaves were randomly chosen from different adult trees and scanned using an Epson Stylus DX 8400 scanner, with a resolution of 400 dpi and saved in bitmap format. Necrotic areas were manually delineated using Photoshop (Adobe Systems, San Jose, USA).

### 2.2 Image Preprocessing

The inherent variability of many factors, such as biological variability and acquisition conditions, results in different image luminance and color distribution in digitized leaves, which must be reduced for improving the pixel classification performance. Two major problems are addressed at this stage: contrast enhancement and color normalization between leaves. Contrast was enhanced applying a classical histogram equalization approach, whereas a color normalization approach was based on the grey world normalization assumption [5,6], which assumes that changes in the illuminating spectrum may be modeled by three constant multiplicative factors applied to the red, green, and blue color components. So, an image of unknown illumination  $I^u$  can be simply transformed to a known luminance space  $I^k$  by multiplying pixel values with a diagonal matrix  $I_{rgb}^k = MI_{rgb}^u$ , with  $M$  defined as 1.

$$M = \begin{pmatrix} m_{11} & & \\ & m_{22} & \\ & & m_{33} \end{pmatrix} m_{11} = \frac{\mu^{I_r^k}}{\mu^{I_r^u}}, m_{22} = \frac{\mu^{I_g^k}}{\mu^{I_g^u}}, m_{33} = \frac{\mu^{I_b^k}}{\mu^{I_b^u}} \quad (1)$$

In this work, the color distributions of all leaves were normalized to the color distribution of one specific leaf, which was used for selecting the training points in the classification step. Figure 2, shows one example of the normalization step applied on a leaf section.



**Fig. 2.** Pre-processed image results. From left to right: original image, color distribution normalized and contrast enhanced.

### 2.3 Damaged Area Segmentation

Segmentation of the damaged area is carried out using an efficient pixel classification strategy. Given an input image  $I$ , the classification of a target pixel  $I(x, y)$  is performed according to a color feature vector  $X$  describing it. In order to reduce the time needed to perform a single pixel classification, the approach presented herein is based on a classification process that find the set of boundaries that optimally separates a particular color-space representation into target classes, labeling each color component as belonging to any class. The labeled color space is used as a look-up-table (LUT) for deciding the class of each image pixel. As the selection of a color space for pixel classification is application dependent [7], we evaluated the classification performance of several color spaces and classification algorithms in order to select the most suitable for our pixel-based classification task. Four supervised classification techniques (KNN, Naive Bayes, Support Vector Machine and Multi Layer Perceptron neural networks), and four color spaces (RGB, normalized RGB, HSV and YCrCb) were assessed.

Classification models were created using a collection of the labeled pixels as training data set, which were manually extracted by an expert from a representative leaf with different damage levels (the same which was used as reference in the color normalization step). The training dataset was filtered for obtaining unique instances. Finally, less than 1000 pixels by label were used for training each classification model.

The main problem of the proposed strategy is that some colors corresponding to different damage levels are strongly mixed up with colors corresponding to leaf veins. Initially, we tried to create a learning model able to split the color spaces in two classes: damaged area and healthy leaf, but its performance was poor. We also trained some multi-class classification models, which tried to distinguish between damaged areas, veins and healthy leaf, but its performance was also poor. So, a first supervised learning model was generated for detecting the principal veins and a second one was used for classifying remaining pixels as damaged area or leaf background. Finally, a correction factor that takes the spatial relationship of the pixels into account was applied.

From the evaluated classification schemes, a SVM technique with a 3 degree polynomial kernel and the CbCr color transformation presented the best performance in the two classification stages. SVM are learning systems that project the input vectors into a high dimensional feature space, using the hypothesis space of linear functions, induced by a kernel function chosen a priori. In the feature space, the learning algorithm produces an optimal separating hyperplane between two classes, maximizing a convex quadratic form, subjected to linear constraints. The optimal hyperplane found with SVM corresponds to the maximal distance to the closest training data and is represented as a linear combination of training points called support vectors. In this work, a version of SVM that uses a sequential minimal optimization algorithm was used [8]. The classification model produces a decision class according to the distance from the instance to the optimal hyperplane. On the other hand, the *YCrCb* is a family of color spaces commonly used to represent digital video. Luminance information

is stored as a single component (Y), and chrominance corresponds to the two color-difference components (Cb and Cr). We have used the YCbCr transformation specified in the ITU-R BT.601 standard for computer-display oriented applications, but the luminance component was not considered.

**Segmentation of main veins.** The first stage in the segmentation approach is to extract the main leaf veins. Approaches for leaf vein extraction have been proposed previously [9,10,11,12,13]. These approaches are mainly based on the assumption that both veins and remaining leaf tend to be uniform in color and contrast [14]. Soille applied morphological filters to extract leaf veins [9]. Fu and Chi [10] proposed a two stage vein extraction that performed a preliminar segmentation based on the intensity histogram of the leaf image, which was then enhanced by a pixel classification based on edge, contrast and texture properties of the pixels. Similarly, a rough segmentation, based on the intensity histogram was used for obtaining the veins of leaf by Li et al. [15], and an active contour technique based on cellular neural network, was used to extract the veins in the obtained rough regions of leaf pixels. Li et al. [11] applied independent component analysis (ICA) to learn a set of linear basis functions that were used as patterns for vein extraction in patches of grey level leaf images. These approaches fail in our problem because damaged areas are commonly located along the veins and their color and contrast are strongly mixed. However, the color pixel classification strategy proposed was able to correctly detect the majority of main veins of the evaluated leaves.

Figure 3 displays the veins obtained when the images of Figure 1 were segmented using the LUT corresponding to veins and leaf partition of the  $CrCb$  color space.

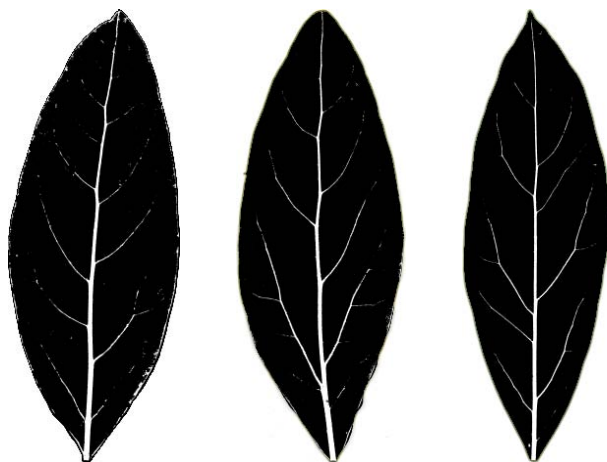


Fig. 3. Vein extraction results of leaves from Figure 1

**Segmentation of damaged area.** After eliminating the veins from the original images, remaining pixels were classified using the LUT generated for partitioning the color space between damaged area and healthy leaf. Figure 4 displays results of the segmentation process applied on the digitized leaves shown in Figure 1. Red color corresponds to the damaged area contour marked by the expert and white regions correspond to damaged areas found by the approach.



**Fig. 4.** Segmentation detail of damaged area of the leaves in Figure 1

**Segmentation refinement.** The aim of this stage was to reclassify pixels incorrectly classified, based on the use of contextual information, not taken into account by the previous phases. Pixels are reclassified using the function defined by equation 2.

$$f(p) = w_{LUT} l(x) + w_N \overline{l_{N(p)}} \quad (2)$$

where  $l(x)$  stands for the label assigned to the color component  $x$  by the LUT and  $\overline{l_{N(p)}}$  corresponds to the mean label value computed from a neighborhood  $N$  around pixel  $p$ .  $w_{LUT}$  and  $w_N$  are weighting functions for the label assigned by the LUT to the actual pixel and for the labels of a neighborhood  $N$  around the pixel.  $w_{LUT}$  represents a value of membership degree to the corresponding label, given by the normalized distance of the color component to the SVM hyperplane, whereas  $w_N$  represents a value of color homogeneity in a neighborhood  $N$  that is given by the difference of the color of the actual pixel with respect to  $N$ . After applying this postprocessing, some small artifacts can still be present, which are removed using a morphological opening.

## 2.4 Experimental Results

The proposed approach was applied to estimate the percentage of damaged area in 13 avocado leaves with different levels of damage. Damaged areas were marked by a botanist, expert in the analysis of avocado leaves. Segmentation performance was evaluated by computing average accuracy, precision and recall of damaged area segmented. Results are shown in table 1. As the measure commonly used by the botanists is the area of damaged leaf, the mean difference between the percentage of damaged area computed from the manual and automatic segmentation images is also reported in the last row of the table. The



**Fig. 5.** Segmentation of damaged area. Detail of segmentation errors caused by poor contrast between damaged area and healthy leaf. Note that similar regions are assigned to different labels.

**Table 1.** Average performance of proposed approach

	Pixel-based classification	Improved Segmentation
Accuracy	$0.924 \pm 0.022$	$0.926 \pm 0.022$
Precision	$0.703 \pm 0.157$	$0.679 \pm 0.131$
Recall	$0.638 \pm 0.189$	$0.799 \pm 0.171$
Estimation difference	$1.2 \pm 2.5$	$2.7 \pm 2.8$

results show that our approach is able to detect different levels of damage in avocado leaves, produced by the *Oligonychus perseae* mite. The postprocessing approach improves the sensitivity of the method, although it reduces the general accuracy. Inaccuracies arose mainly due to early damage levels that are difficult to estimate even visually as shown in Figure 5. It is worth noting that differences in computed leaf damage are negligible when assigning a discrete damage level.

### 3 Conclusions

A simple semi-automatic approach for segmenting and quantifying damaged area in avocado leaves was proposed. The approach is based on an initial pixel based classification according to the chrominance feature from the *YCrCb* color space, which is improved by taking into account local context information and the membership degree of the color value to a specific class (healthy leaf or damaged area). Classification time is reduced through the construction of a lookup table (LUT) in which the classes for the whole *YCrCb* chrominance space are assigned by a learning model in an off-line process. So, a minimal user intervention is needed, to select a sample of pixels for training the color space classifier. The proposed approach was tested on 13 leaves with different color distributions and several degrees of damage and results are very promising. However, an evaluation with a higher number of leaves is warranted.

## References

1. Kerguelen, V., Hoddle, M.S.: Measuring mite feeding damage on avocado leaves with automated image analysis software. *Florida Entomologist* 82, 119–122 (1999)
2. Wijekoon, C., Goodwin, P., Hsiang, T.: Quantifying fungal infection of plant leaves by digital image analysis using scion image software. *Journal of Microbiological Methods* 74, 94–101 (2008)
3. Bakr, E.M.: A new software for measuring leaf area, and area damaged by *tetranychus urticae* koch. *Journal of Applied Entomology* 129, 173–175 (2005)
4. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man And Cybernetics* 9, 62–66 (1979)
5. Finlayson, G.D., Schiele, B., Crowley, J.L.: Comprehensive colour image normalization. In: Burkhardt, H.-J., Neumann, B. (eds.) *ECCV 1998*. LNCS, vol. 1406, p. 475. Springer, Heidelberg (1998)
6. Tek, F., Dempster, A., Kale, I.: A colour normalization method for giemsa-stained blood cell images. In: *IEEE 14th Signal Processing and Communications Applications* (2006)
7. Wolpert, D.H., Macready, W.G.: No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* 1, 67–82 (1997)
8. Platt, J.: Machines using sequential minimal optimization. In: Schoelkopf, B., Burges, C., Smola, A. (eds.) *Advances in Kernel Methods - Support Vector Learning*. MIT Press, Cambridge (1998)
9. Soille, P.: Morphological image analysis applied to crop field mapping. *Image and Vision computing*, 1025–1032 (2000)
10. Fu, H., Chi, Z.: A two-stage approach for leaf vein extraction. In: *IEEE International conference on neural networks and signal processing* (2003)
11. Li, Y., Chi, Z., Feng, D.D.: Leaf vein extraction using independent component analysis. In: *IEEE Conference on Systems, Man, and Cybernetics* (2006)
12. Nam, Y., Hwang, E., Kim, D.: A similarity-based leaf image retrieval scheme: Joining shape and venation features. *Computer Vision and Image Understanding* 110, 245–259 (2008)
13. Boese, B.L., Clinton, P.J., Dennis, D., Golden, R.C., Kim, B.: Digital image analysis of *zostera marina* leaf injury. *Aquatic Botany* 88, 87–90 (2008)
14. Clarke, J., Barman, S., Remagnino, P., Bailey, K., Kirkup, D., Mayo, S., Wilkin, P.: Venation pattern analysis of leaf images. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Remagnino, P., Nefian, A., Meenakshisundaram, G., Pascucci, V., Zara, J., Molineros, J., Theisel, H., Malzbender, T. (eds.) *ISVC 2006*. LNCS, vol. 4292, pp. 427–436. Springer, Heidelberg (2006)
15. Li, Y., Zhu, Q., Cao, Y., Wang, C.: A leaf vein extraction method based on snakes technique. In: *International Conference on Neural Networks and Brain* (2005)