Computers and Electronics in Agriculture 78 (2011) 9–18

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Leaf classification in sunflower crops by computer vision and neural networks

Juan Ignacio Arribas [a,*], Gonzalo V. Sánchez-Ferrero [a], Gonzalo Ruiz-Ruiz [b], Jaime Gómez-Gil [a]

[a] Dept. of Teoría de la Señal y, Comunicaciones e Ingeniería Telemática, Univ. Valladolid, Valladolid 47011, Spain
[b] Dept. of Ingeniería Agrícola y Forestal, Univ. Valladolid, Valladolid 47011, Spain

ARTICLE INFO

Article history:
Received 11 December 2009
Received in revised form 19 April 2011
Accepted 10 May 2011

Keywords:
Classification
Computer vision
Learning machines
Model selection
Sunflower

ABSTRACT

In this article, we present an automatic leaves image classification system for sunflower crops using neural networks, which could be used in selective herbicide applications. The system is comprised of four main stages. First, a segmentation based on *rgb* color space is performed. Second, many different features are detected and then extracted from the segmented image. Third, the most discriminable set of features are selected. Finally, the Generalized Softmax Perceptron (GSP) neural network architecture is used in conjunction with the recently proposed Posterior Probability Model Selection (PPMS) algorithm for complexity selection in order to select the leaves in an image and then classify them either as sunflower or non-sunflower. The experimental results show that the proposed system achieves a high level of accuracy with only five selected discriminative features obtaining an average Correct Classification Rate of 85% and an area under the receiver operation curve over 90%, for the test set.

1. Introduction

Precision Agriculture (PA) refers to the application of new technologies in agricultural field work. PA mainly uses new communication technologies, as well as signal and image processing techniques. In addition, computer vision can be used for detecting the state of the components of the vehicle, its implements and the environment. Computer vision can also be used in PA to distinguish plants from weeds, so that a variable rate of herbicide can be applied depending on the size and number of plants detected. This work is focused on the application of computer vision for classification purposes between sunflower crops and weeds.

The automatic classification by computer vision of weeds and plants is receiving an increasing attention in the literature. For instance, some relevant machine vision algorithms can classify plants into either crop or weeds (Yang et al., 2000; Cho et al., 2002; Aitkenhead et al., 2003; Burks et al., 2005), so making it possible to perform mechanical weeding or to apply herbicide selectively.

The classification of crops and weeds was studied by Brivot and Marchant (1996), who used infrared images in low-light conditions to study transplanted crops. Lee et al. (1999) and Hemming and Rath (2001) conducted similar experiments with tomato crops and weeds, respectively, using controlled artificial lighting to identify morphologic features of plants. To solve illumination problems, Hague et al. (2006) more recently developed a method for cereal crop and weed segmentation using a vegetative index not affected by the lighting conditions. In the same way, Tian and Slaughter (1998) developed an adaptive algorithm for vegetation segmentation that works under various lighting conditions, although they did not use the algorithm to classify crops and weeds. In order to achieve this, they transformed color space, from *RGB* into *rgb*. Meanwhile, recently Ruiz-Ruiz et al. (2009) applied the same adaptive algorithm to both hue-saturation and hue color spaces.

Shrestha et al. (2004) also extracted morphologic features to distinguish between maize plants and weeds. They independently applied the Otsu method (Otsu, 1979) and the Maximum-Likelihood Estimation (MLE) to two leaf features, area and length, the latter of which is a non-invariant feature. Gebhardt et al. (2007) and Gebhardt and Kuhbauch (2007) added the features of shape, color and texture and used them in the MLE classification, still using controlled lighting. Finally, Tellaeche et al. (2008) used images under natural light and proposed a method for the selective spraying of crops using a complex combination of structural and area features. In this case, classification was not made at the plant level but at the coarse cell level, virtually dividing the ground into small regions. On the contrary, our proposed method defines a high number of features under natural conditions at the plant level.

Some of the work on plant row detection has been developed by Marchant and Brivot (1995), Marchant (1996), Southall et al. (1998), Sogaard and Olsen (2003) and Onyango and Marchant (2003). Most of them used color features for segmenting vegetation and Hough transform for detecting straight lines.

Neural networks appear in farming tasks too. For instance, Huang (2007) presented an application of neural networks and image processing for classifying seedling diseases. He successfully used color and texture features, 21 in all, in a classical Multi Layer

* Corresponding author. Tel.: +34 983423000; fax: +34 983423667.
E-mail address: jarribas@tel.uva.es (J.I. Arribas).

0168-1699/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.compag.2011.05.007

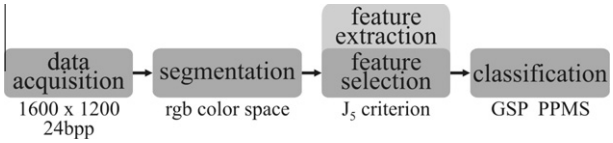

Fig. 1. A block diagram of the proposed pattern recognition system for sunflower crop automatic classification, with input image data with a resolution of 1600 × 1200 pixel and 24 bpp color depth, *rgb* color space segmentation, feature extraction and selection with the classical J_5 criterion and GSP–PPMS neural network classification subsystems.

Perceptron (MLP) with the back-propagation learning algorithm, though his process did not include feature selection. In our case, we have decided to use a more recently developed neural-genetic network architecture based on the Generalized Softmax Perceptron (GSP) and the Posterior Probability Model Selection (PPMS) algorithm (Arribas and Cid-Sueiro, 2005). GSP network has proven to obtain better classification results than the MLP. At the same time, PPMS algorithm provides a proper model selection of the network complexity.

Our main objective is to detect individual sunflower plants by distinguishing them from weeds. In continuation of the previous work done on crop segmentation, this paper presents a method for segmenting sunflower plants and weeds based on color space using the GSP neural network architecture trained with the help of the PPMS algorithm to classify input plants as either sunflowers or weeds. In addition, we have opted to implement a previous feature selection step in order to select the most distinct features. This helped to keep our neural-genetic classifier from producing overly-generalized results when an excessive number of input features are entered into the neural network. Finally, we believe this system might potentially be of help in the field of Precision Agriculture applications.

Fig. 1 shows the system block diagram, which illustrates the order and structure followed in this document. In the diagram one can see the different subsystems including: data image acquisition (c.f. Section 2), segmentation in the *rgb* color space (c.f. Section 2.1), extraction of 13 distinct features (c.f. Section 2.2), selection via the criterion (c.f. Section 2.3), and classification into either sunflower or weed output classes by means of the GSP neural network architecture and the PPMS complexity selection algorithm (c.f. Section 2.4). Finally, we will finish this document by showing the classification results in Section 3 and drawing the final conclusions in Section 4.

2. Materials and methods

Regarding the image data bank, we used a PANASONIC DMC-LX1 camera with objective Leica DC VARIO-ELMARIT 1:2.8-4.9/6.3-25.2 ASPH. The white balance was automatic, and the image format was JPG 1600 × 1200. We used the camera 'fine' mode to reduce image compression. Pictures included two rows of sunflower crop at a crop row distance of approx. 70 cm, thus comprising approx. 1 m in total width. Thus the approximate spatial resolution of the images was about 1600/1000 = 1.6 pixels/mm. For the development and evaluation of the process of plant classification, we used a data bank of 192 color images of sunflower crops. The images were stored in *RGB* format, with 24 bits per pixel, and their size was 1600 × 1200 pixels. All images were taken at an approximate height of one and a half meters, using the same focal distance. Fig. 2 shows four examples of crop images. All of the 192 images were taken in June of 2006 in the sunflower crop fields at Aguilar de Bureba, Burgos (Spain), with coordinates 42 degrees 35 min 27 s North and 3 degrees 19 min 42 s West. A total of 10,596 objects or shapes in the images were segmented (isolated)

Fig. 2. The image data bank: four examples of different and typical (ordinary) sunflower crop images under various capture conditions (day light, day time and background).

in the data bank, regardless of whether they were weeds or sunflowers. Of those, 717 were sunflowers. All methods were programmed in Matlab and C++.

It is worth noting that a new application has been developed in Matlab to accomplish the segmentation process described above, including a graphical user interface with options for an easy way to segment images, either one by one or several in folder at a time, including saving of the binary output image. In addition, the program has three main parameters used to control the segmentation process: the number of iterations in clustering, the size of

morphological erosion (closing) in the first segmentation and the size of morphological dilation (opening) in the final segmentation.

2.1. Segmentation

The objective of the first part of the process, the segmentation, is to obtain a binary image in which '1' represents vegetation and '0' represents ground. This process of segmentation is based on color information, which originally is represented by three color components: red, green, and blue (*RGB*). To get a binary image from an *RGB* image, an automatic algorithm (Fig. 3) has been developed following three main steps: 'seed' extraction from the *RGB* image, transformation of the *RGB* image into *rgb* format and clustering of 'seeds' in the *rgb* color space, which we explain in the next section.

2.1.1. 'Seed' extraction

The clustering step needs two initial 'seeds' in the *rgb* color space to start: one from the vegetation and one from the soil. The 'seed' extraction procedure is automatic and it starts with a segmentation of the *RGB* image, comparing the Green component (G_{ij}) with the Red (R_{ij}) and Blue (B_{ij}) in each pixel. A pixel is classified as vegetation if its Green component is greater than its Red and Blue components, whereas the pixel is classified as soil otherwise:

$$\begin{aligned} (G_{ij} > R_{ij}) \& (G_{ij} > B_{ij}) &\Rightarrow \text{pixel}_{ij} = \text{vegetation} \\ \text{rest} &\Rightarrow \text{pixel}_{ij} = \text{soil} \end{aligned}$$

Fig. 4 shows two examples of 'seed' extraction. For the first image, the result is quite good, but not for the second, in which there is segmentation noise due to complex lighting conditions; however, the result is good enough to properly extract the two desired 'seeds'. Finally, the 'seeds' from the vegetation and soil are calculated as the mean *RGB* values, or centroids, corresponding to the vegetation and soil segmented areas, respectively. These 'seeds' are processed in *RGB* format, but for clustering they need to be converted into *rgb* color space, as described in the next section.

In addition to the automation of 'seed' extraction, this first step makes the algorithm adaptive to each new image, because the 'seeds' are specifically calculated to reflect the lighting conditions, the greenness of plants and the type of soil.

2.1.2. Transformation into rgb format

Clustering pixels in the *RGB* color space for crop images does not produce good results in most cases, and it gets worse if there are both shaded and bright regions in the image. The pixels of the crop images are mainly distributed around the gray axis $R_{ij} = G_{ij} = B_{ij}$ (see Fig. 5a) , and there are no significant differences between vegetation and soil pixels with respect to R_{ij} ,G_{ij} and B_{ij} component values, so clusters in *RGB* are usually a mixture of vegetation and soil pixels. This could be due to the automatic white balancing of the camera, but since turning off the automatic white balance function would result in additional problems, for instance when

Fig. 4. Two examples of early-stage sunflower images from the first *RGB* sunflower leaf segmentation: original and segmented images, respectively.

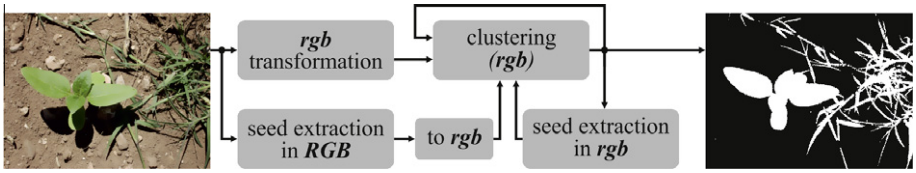

Fig. 3. The proposed algorithm for automatic segmentation: an overview with an example of the original and segmented output images, where iteration paths are depicted.

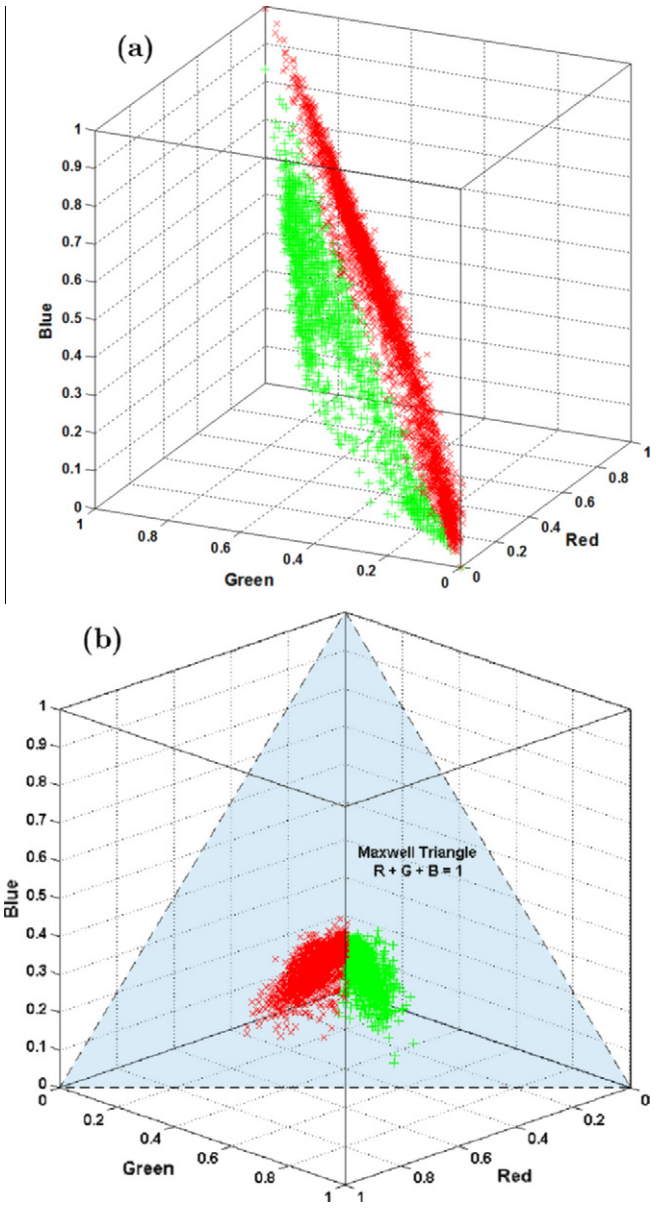

Fig. 5. A simple example of the distribution of image pixels from the same original sunflower crop image both in *RGB* (a) and *rgb* (b) color spaces: green leaves vegetation foreground (+) and soil background (x).

classifying both classes (weeds and sunflower) by a human expert, we preferred to leave the white balancing on. Thus, we believe the transformation from *RGB* to *rgb* space will resolve this problem, at least partially, and this step is commonly used in the literature. It has been shown in Tian and Slaughter (1998) that processing images in *rgb* color space yields the same results regardless of lighting. The transformation of the *RGB* format into the *rgb* format follows:

$$r_{ij} = \frac{R_{ij}}{R_{ij}+G_{ij}+B_{ij}}; g_{ij} = \frac{G_{ij}}{R_{ij}+G_{ij}+B_{ij}}; b_{ij} = \frac{B_{ij}}{R_{ij}+G_{ij}+B_{ij}}; \quad (1)$$

This transformation is a normalization of color space that projects image pixels into a triangular plane, perpendicular to the gray or lighting axis, as shown in Fig. 5b.

The new vegetation and soil distributions of pixels in *rgb* have a reduced intraclass distance because it is more compact than the corresponding distributions in *RGB*. For this reason, clustering in *rgb* using two 'seeds' is more efficient, and the resultant clusters complement the real vegetation and soil distributions.

2.1.3. Clustering in rgb color space

The last main step in the segmentation algorithm is the clustering of image pixels. This process assigns each pixel to a class (vegetation or soil) based on which 'seed', also called a centroid, is nearest to it, which can be determined by means of the Euclidean distance (d_E):

$$\begin{aligned} d_E[(r,g,b),(r_v,g_v,b_v)] &\leqslant d_E[(r,g,b),(r_s,g_s,b_s)] \Rightarrow \\ \Rightarrow \text{pixel} &= \text{vegetation} \\ d_E[(r,g,b),(r_v,g_v,b_v)] &> d_E[(r,g,b),(r_s,g_s,b_s)] \Rightarrow \\ \Rightarrow \text{pixel} &= \text{soil} \end{aligned} \quad (2)$$

(r,g,b) being the pixel under classification analysis (r_v,g_v,b_v), the vegetation centroid and (r_s,g_s,b_s) the corresponding soil centroid.

Regarding the illumination conditions, our system works under general (arbitrary) lighting conditions, obtaining proper image segmentation results in shady areas by means of the *rgb* color normalization clustering. In addition, given that the 'seeds' used in clustering are specifically calculated for each image, the results of the first clustering iteration are normally quite good. In most cases, it is only necessary to perform a few iterations of this process: typically three iterations were enough, calculating new 'seeds' or centroids and regrouping pixels around them.

Finally, it is possible to apply some morphological filtering using well-known dilation or erosion filters of size 3×3, 5×5, and 7×7, using typically the intermediate size 5×5 filter, in order to reduce the segmentation noise. The dimensions of the filters are variable and, as previously noted, are one of the three parameters that can be modified in the segmentation application that was designed for the present work. The easiest way to remove little groups of white pixels, misclassified as vegetation, is to apply an opening filter. Its size must be appropriate for the image resolution and plant size in order not to remove valuable information in the process of reducing noise.

2.2. Feature extraction

Image feature extraction is an important step for image pattern recognition. Feature extraction should consider which input features provide the most information for classification. Thus, a good feature extraction and selection must be performed in order to achieve the best classification results.

In formal mathematical terms, an input pattern may be represented by a set of d features as a d-dimensional feature vector extracted and computed from the segmented images. The d-dimensional feature vectors were used in the neural network training process. Invariant features (invariant to illumination and position) must be taken into consideration in order to identify objects with different lighting and position conditions. The aim of using invariant features is to identify objects independently of how and where they are observed (Smeulders et al., 2000).

Several kinds of features, such as color, shape and texture, are usually employed in pattern recognition. The selection of these features will determine the effectiveness of the classification. Object shapes have proven to be one of the most important ways of describing biological characteristics (Tian et al., 1997). Regarding the nature of the problem, we propose 13 morphological features:

- Perimeter (PRI). Number of boundary pixels.
- Centroid (CEN). The geometric center.
- Area (PXC). Number of pixels of the object.
- Major axis of the best fit ellipse (MJX).

- Minor axis of the best fit ellipse (MNX).
- Height (HET). Difference of the highest vertical coordinate and the smallest of the fitted ellipse.
- Width (WID). Difference of the highest horizontal coordinate and the smallest of the fitted ellipse.
- Area to length ratio (ATL). The ratio of area to length was defined as ALT $= \frac{PXC}{MJX}$.
- Compactness (CMP). Ratio of area to the perimeter squared, defined as CMP $= 16\frac{PXC}{PRI^2}$.
- Elongation (ELG). The differences between the lengths of the major and minor axes of the best fit ellipse, divided by the sum of the lengths: ELG $= \frac{MJX-MNX}{MJX+MNX}$.
- Logarithm height/width ratio (LHW). The logarithm of the height to the width ratio: LHW $= log_{10}(\frac{HET}{WID})$.
- Perimeter to broadness (PTB). Defined as: PTB $= \frac{PRI}{2(HET+WID)}$.
- Length to perimeter ratio (LTP). It is a measure of the 2-D distribution pattern of the boundary of the object: LTP $= \frac{MJX}{PRI}$.

The feature extraction method deals with binary images that were generated by the segmentation module. All contours in a binary image were extracted and analyzed. If these contours have holes inside their shape, we considered that this area belongs to the object, because sometimes the leaves have debris covering their surfaces. Next, we estimate the PRI, CEN and PXC. The best fit ellipse is computed in the least-squares sense to the set of 2D points of the contour. If there is no ellipse that matches the points, the following parameters are considered:

$$\text{Rotation} \quad \text{Angle} = 0^o$$
$$x_{ce} = \frac{x_{\max} + x_{\min}}{2}$$
$$y_{ce} = \frac{y_{\max} + y_{\min}}{2}$$
$$\text{MJX} = \frac{x_{\max} + x_{\min}}{2}$$
$$\text{MNX} = \frac{y_{\max} + y_{\min}}{2}$$

where x_{ce} and y_{ce} are the coordinates of the center, MJX is the major MNX axis, is the minor axis, $x_{\max}$ is the maximum x coordinate, $x_{\min}$ is the minimum x coordinate, $y_{\max}$ is the maximum y coordinate and $y_{\min}$ is the minimum y coordinate of the set of 2D points.

With previous data computations, we can now calculate all of the features described above. In addition, we can also expect that the number of features selected should be less than 7 due to the curse of dimensionality, given that it is a commonly accepted empirical rule, see (Haykin, 1999), that the ratio of input patterns to input features must remain above 100: here the number of patterns/number of features = 717/7 > 100 for each output class.

2.3. Feature selection

There are two main reasons to keep the value of the dimensionality of input patterns low (Jain et al., 2000):

1. *Measurement cost:* as this system could be used in real time, we have to compute features and classify patterns in a short span of time, so we need a fast system that does not require a lot of memory.
2. *Classification accuracy:* we must take into account that a reduction in the number of features may lead to a loss in discrimination capability. But an excessively high number of features, on the other hand, may prevent the neural network from properly generalizing to new input patterns in the test set, so the model selection (size) of the neural architecture is an important issue to consider (Watanabe, 1985; Arribas and Cid-Sueiro, 2005).

The classification capability of a feature vector in a class is based on the assumption that different classes have different positions, though overlapping, within the feature input space. Therefore, the ideal situation would be that vectors belonging to different classes are distant from one another in the feature space, whereas vectors of the same class are close. A proximity measurement is therefore necessary and a suitable choice is the statistical distance computed through the scatter matrix for the Between-class and Within-class distances. The criterion applied in this study is the following (Devijver and Kittler, 1982; Ghosal and Mehrotra, 1997):

$$J_5(\xi) = tr(S_W^{-1} S_b) \tag{3}$$

$$S_W = \sum_{i=1}^{C} \frac{P_i}{n_i} \sum_{k=1}^{n_i} (\xi_{ik} - m_i)(\xi_{ik} - m_i)^T \tag{4}$$

$$S_b = \sum_{i=1}^{C} P_i (m_i - m)(m_i - m)^T \tag{5}$$

In order to select an optimal feature set, multiple combinations of features must be taken into account so that better results can be obtained. Therefore, the problem is finding a transformation that maximizes the separability criterion. This transformation is given by the matrix W with size $d' \times d$, where d is the total number of features and d' is the desired number of features:

$$W_{\text{optima}} = \arg\max J_5(W\xi) \tag{6}$$

and where ξ is the input feature vector with size and is the output features vector with size d'. Finding the best features vector is equivalent to finding the transformation given by the matrix $W'(d \times d')$, so that $\xi' = W'\xi$. Analyzing all the possible sets of features then becomes completely unrealizable, so it is necessary to establish another suboptimal searching strategy:

1. The feature with the most significance is sought. It supposes d' separability calculus. Those operations were fast because the S_w and S_b matrices are scalars and the inversion of S_w is trivial.
2. After finding the most significance feature, the remainder $d-1$ features were analyzed in order to find the feature that adds more separability in conjunction with the previously selected feature. This requires $d-1$ separability computations. In this case, there were $d-1$ matrices of sizes 2×2.

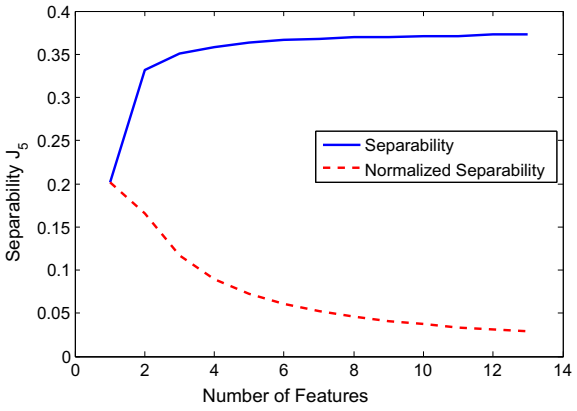

Fig. 6. Separability and normalized separability achieved in the feature selection phase. Note that as more features were considered, they add less information. Thus, the best point to select the number of features is the one closest to 0 features and to a value of 0 in the normalized separability, corresponding in this case to an approximate number of about five features.

3. These steps were repeated until there are no remaining features. In each step, the size of the matrices grows as the number of matrices decreases.

We know this process does not achieve a global maximum. However, the relatively low number of features, when compared to the great number of samples, will provide a good feature selection (Ghosal and Mehrotra, 1997). Fig. 6 shows both the separability and the normalized separability (by the number of features) added by each feature in the process explained before. It is clear that each feature adds less separability to the group. Fig. 6 illustrates how the normalized curve decreases more rapidly when it starts from feature number 5.

Thus, this was defined as the selected final number of features, best five, which are listed below:

1. PRI: perimeter.
2. PXC: area.
3. MJX: major ellipse axis.
4. MNX: minor ellipse axis.
5. LHW: logarithm of height/width ratio.

2.4. Classification

During classification, before training and testing, all features were normalized and the Karhumen–Loeve transformation was applied prior to the feature selection phase in order to eliminate all possible dependencies. Additionally, in order to study the effectiveness of the classifier, different sets of features were used for training and testing.

In this study, neural networks help us estimate the *a posteriori* probabilities of an input pattern to belong to each class. Specifically, we use the Generalized Softmax Perceptron (GSP) neural network with the Posterior Probability Model Selection (PPMS) algorithm, given that, in addition to an estimation of the posterior probabilities, they determine their optimal structure and complexity during training (Arribas and Cid-Sueiro, 2005). The structure or complexity model selection of the network, consists of pruning, splitting and merging the subclasses depending on the probability of belonging to each class.

Note that the PPMS algorithm defines three threshold parameters (Prune, Split and Merge). In order to avoid a dependency on the threshold selection, those thresholds were varied during the learning phase as follows: the Split Threshold (μ_{split}) increases by ($\delta\mu_{\text{split}}$) each time a split operation is accomplished, and conversely the Prune Threshold (μ_{prune}) decreases by ($\delta\mu_{\text{prune}}$) every time a pruning action takes place. The values of the initial thresholds has shown to be noncritical in the results. Nevertheless, they should be initialized properly in order to achieve a good estimate of the number of subclasses per class, thus determining the network model complexity, size or model selection.

As the initialization of thresholds is a problem that depends on the data, we decided to apply a genetic algorithm (GA). These algorithms have demonstrated good behavior and offer an attractive approach to multiple criteria optimization which cannot be handled by most of the other methods (Verma and Zhang, 2007). The genetic algorithm is a searching process based on the laws of natural selection and genetics. Usually, a simple GA consists of three operations applied to a population of possible solutions: selection (according to a fitness function), genetic operation (crossover, mutation), and generational-replacement. This cycle is repeated until a desired termination criterion is reached.

For this work, a GA algorithm was implemented with the following parameters: population was formed by 100 individuals from a subset of all possible thresholds. Each individual is represented by split and prune thresholds randomly chosen from 0 to 1.

In signal detection theory, the Receiver Operation Characteristic (ROC) is formally defined as a graphical plot of the sensitivity, or True Positive Fraction (TPF), versus the False Positive Fraction (FPF) (1 − specificity or 1 − True Negative Fraction (TNF)), for a binary classifier system as its discrimination (detection) threshold is varied.

At the same time, the Area Under the ROC Curve (AUC) is defined as a numerical measure for the precision in the decision that the binary classifier is taking, being AUC = 0.5 for random chance binary classifier (worst possible case) and AUC = 1 for the perfect error-free binary classifier, that is AUC $= \int_0^1 d(FPF)$.

The fitness function used in the evolutionary process was a function of AUC under the ROC curve (in short the 1 − False Negative (FN) versus the False Positive (FP) fraction) that describes the effectiveness of the classifier and the difference between the number of prune and split operations:

$$F = \text{AUC}e^{-k|Nm-Np|} \tag{7}$$

where is the Area Under the ROC Curve (AUC) as previously defined, is the number of neurons being merged in the neural network, N_p is the number of prune actions in the neural network architecture, and is a constant that reduces *AUC* by $\frac{1}{2}$ when $|Nm - Np|$ equals 10.

We believe that the Receiver Operation Characteristic (ROC) and the Area Under Curve (AUC) analysis together provides one of the most objective and commonly accepted ways to measure the performance in any binary classification problem, particularly in medical diagnostics (an ill versus healthy binary problem is commonly used by diagnosticians). Since this problem is binary (weeds vs. sunflower), we wanted to provide the AUC and ROC values, in addition to the typical classifier hard output Correct Classification Rate (CCR) or confusion matrices values, after the classifier decision threshold. So in summary, in addition to the hard classifier output after the (fixed) decision threshold (CCR, hits and/or errors) one can also provide what can be called the soft classifier output prior to the classifier decision threshold, simply by varying the classifier decision threshold, thus complementing the classification accuracy measured by the CCR. The meaning of the AUC values should also be clear, since the AUC = 0.5 is the worst possible classifier one can imagine (simply making decisions by chance) while the best is that in which AUC = 1, thus eliminating all false positive and false negative events.

In addition, the well-known roulette wheel strategy of selection (Fogel, 1994) was used in the process. The probability of crossing and mutation are $p_x = 0.5$ and $p_m = 0.01$, respectively. The number of generations is 300. The results of the GA applied to the initializing thresholds shows a correlation between the AUC and thresholds as shown in Table 1. The threshold values show that, although there is a noncritical relationship between the AUC and thresholds, the AUC is directly related to higher split thresholds and lower prune thresholds.

The results of the area under the ROC and the thresholds are depicted in Fig. 7. It is clear that, although there is an inherent random result, the population tends to accumulate around the high values of the split threshold (0.96) and the low values of the prune threshold (0.02).

In order to show the benefits of this estimation of thresholds and its influence, an experiment was implemented in which all the possible thresholds where considered in steps of 0.1 (20 independent neural networks were trained for each of the threshold

Table 1
Correlation between thresholds and AUC.

Correlation	μ_{split}	μ_{prune}	AUC
μ_{split}	1.00	−0.215	0.155
μ_{prune}	−0.215	1.00	−0.166
AUC	0.155	−0.166	1.00

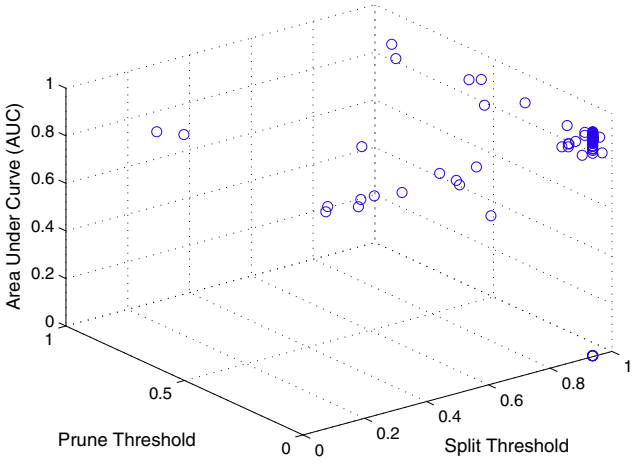

Fig. 7. Selected PPMS thresholds after 300 generations in a population of 100 individuals. Please note that many of the individuals from the population are placed under high values for the Split PPMS Threshold and low values for the prune PPMS threshold. This means that, despite the random nature of the network initialization, better results in the classification were achieved with these thresholds.

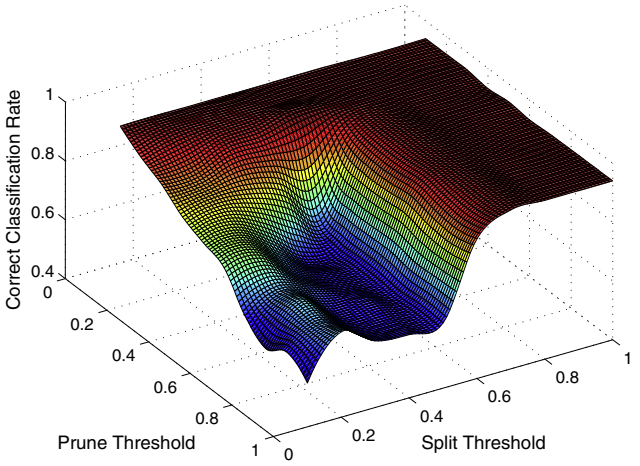

Fig. 8. Mean Correct Classification Rate (CCR) for 20 independent GSP neural networks. Please note that the higher values of the Correct Classification Rate are also in the higher values of the Split PPMS Threshold and in the lower values of the Prune PPMS Threshold. This fact once again validates the threshold selection provided by the genetic algorithm.

values). A surface graph of the Correct Classification Rate is depicted in Fig. 8, which shows that the selected thresholds were situated as far as possible from the region of the surface where the Correct Classification Rate is low. This result is consistent with those achieved with genetic algorithms.

3. Results

3.1. Segmentation algorithm

Fig. 9 shows the result of the whole process of segmentation. In this example, the shadowed vegetation has been correctly detected, since the normalized transformation from *RGB* into *rgb* makes the segmentation independent of lighting.

3.2. Receiver Operation Characteristic, area under curve and Correct Classification Rates

Since in the image bank only 717 were positive classifications, just 1434 samples out of the total bank were used in training:

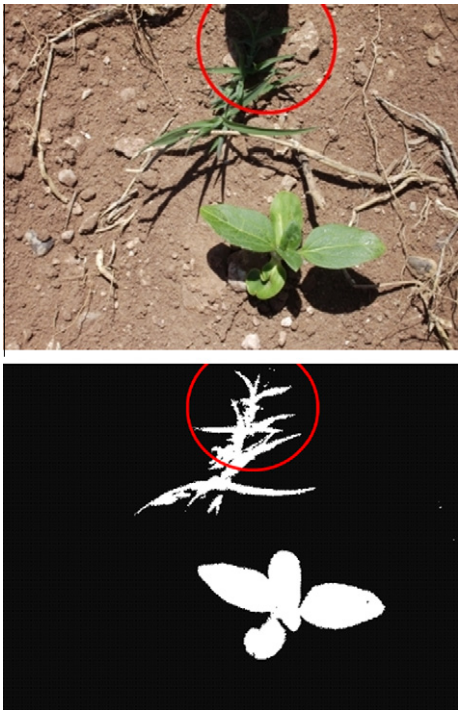

Fig. 9. An example of a sunflower crop original image and its final leaf segmentation image, even under severe low-high contrast and light-shadow constraints: please note the areas circled in red in both the original and segmented image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

50% positive (sunflower) and 50% negative (non sunflower or weeds). Please note that it is a common practice to balance un-balanced input patterns before the machine learning phase, especially when input patterns are highly unbalanced, like in the problem at hand, since otherwise it is almost impossible for the neural network to learn (supervised mode, known class labels) properly from the examples in the training set and thus will not be able to generalize correctly under the test set. Nevertheless, this is a major issue, regarding the number of input samples, the number of classes, the statistics behind those input sample sets, and the optimal size of the network while learning (estimated through the PPMS algorithm) since by providing that balanced dataset one is indeed not following the true prior probability distribution of input data samples. Thus, one can find plenty of controversy in the literature.

Since the neural network was trained in supervised mode (desired classification class labels were known in advance), we divided the input data patterns into two disjointed data sets: the training set to properly train the network under the supervised, known-label mode and the testing set to test the generalization capabilities of the network after the training step. Thus the images were classified manually by a human expert in order to provide supervised learning to the neural network classifier.

Of the previous samples (totaling 1434 input samples for both classes), 90% composed the training set, and the remaining 10% were used for testing purposes. The neural network training and testing was repeated 100 times to extract statistically significant results. Both the training and test sets were selected randomly, guaranteeing a balanced proportion between positive and negative outcomes. The probability threshold of classification was set to 0.5.

In order to analyze the performance of the feature selection stage, 100 independent neural networks were trained with an increasing number of features. It is expected that the AUC and

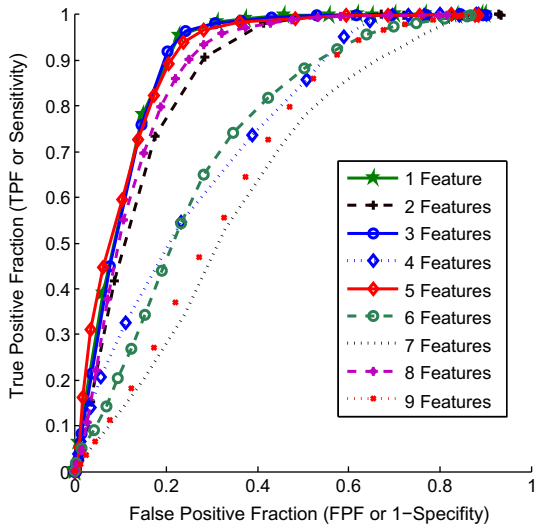

Fig. 10. Mean ROC test curves for 100 independently trained neural networks with an increasing number of features, starting with one selected feature and ending with nine features. The maximum value of the AUC is over 90% for the five features finally selected: PRI, PXC, MJX, MNX and LHW.

the Correct Classification Rate decrease due to the *curse of dimensionality* (Jain et al., 2000). Fig. 10 shows this effect, since as the number of features increases, more samples are needed to achieve a good classification. Fig. 10 shows in detail an AUC maximum value over 90% for the five finally selected features.

Finally, the Correct Classification Rate (CCR), the Square Error (SE), the Cross Entropy (CE) error and the ROC AUC mean and standard deviation values are shown in Table 2, computed for the 100 samples and the five selected input features, the results of which are promising.

4. Discussion and conclusion

Let us start with the discussion and later on present a conclusions list for this paper. To begin with, the proposed system could be employed into weed control with herbicide applications. The damage caused by the classification errors depends on the herbicide type: (i) weeds misclassified as sunflowers will be kept in the field whereas sunflowers misclassified as weeds will be eliminated if a total herbicide is deployed. (ii) employing a selective herbicide, weeds misclassified as sunflowers will be also kept in the field but sunflowers misclassified as weeds will not be eliminated, and thus only a little herbicide will be wasted. Once these misclassification damages are quantified, it is possible to choose the optimal operation point along the ROC curve.

After careful analysis of the numerical results from our system, and following the results for the ROC curves over the test set shown in Fig. 10 (diamonds solid-line plot ROC curve, five features, AUC = 0.9033), we can determine that a possible optimal decision threshold would result.

For instance, the approximately optimal threshold would be at the point in the ROC curve with FPF = 0.20 = 1−TNF and TPF = 0.95 = 1−FNF, meaning that in case of using a non-selective herbicide only is the probability of a sunflower plant to be misclassified as weeds and that TNF = 0.80 is the probability that weeds are classified as weed.

Thus, FPF = 0.20 being the probability that weeds are misclassified as sunflower plants. It must be clear that the most critical error in this system is the , since, in principle, the system could be potentially used in selective herbicide application, but depending upon the cost of each type of the two possible errors in the binary classification, one can choose an optimal operation point along the ROC curve by simply varying the output discrimination decision threshold of the neural-genetic classifier.

If we compare with previous numerical CCR classification results from the literature (we cannot compare in terms of AUC values since in the literature we could not find any system where the ROC analysis had been carried out), we see that our system has an acceptable CCR as compared with those claimed in other papers, although a direct comparison is not possible since each is using a different image data bank. Respect previous point, we must add that we on purpose used different and heterogeneous illumination conditions (day light, day time, etc.) in sunflower images, as should be clear after observing Figs. 2, 4 and 9, in order to obtain a system as general and automatic as possible for sunflower plants classification. Nevertheless, comparing these results to those obtained in the work of Yang et al. (2000), we can point out that a selection of morphological features allows obtaining better results than the color index. In that work a maximum rate of success of 80% was achieved whereas, in our work, the average CCR was 85% and the maximum rate of success was over 90% which evidences the better performance of morphological features.

The importance of these morphological features was suggested by Aitkenhead et al. (2003). In that work the relation between perimeter and area was used as a feature for classification purposes. CCR Results obtained with that feature varied between 52% and 74%. These results were not enough for practical applications but showed the importance of morphological features which could be improved: "the use of more sophisticated morphological measurements than the relatively simple one used here could improve the performance and applicability of the method", (Aitkenhead et al., 2003).

Our work confirms the importance of area and perimeter since the results obtained by the feature selection method selected these features as the most discriminant ones and combined them with other morphological features increased the results to 85 ± 5%.

A common problem in neural networks classifiers is the structure of the neural networks. In the work of Yang et al. (2000) this problem makes the classification results to vary from 60% to 80% in the best case. The same problem appears in the work of Cho et al. (2002) since the maximum classification rate variation is 30% depending on the number of hidden nodes. One the advantages of our work is that the GSP–PPMS algorithm determines the optimal structure and complexity during the training stage.

Some works such as Burks et al. (2005) study the influence of different neural network models obtaining the best results with backpropagation models, however this work does not study the influence of the structure of the classifier. Our work uses a Generalized Softmax Perceptron architecture proposed by Arribas and Cid-Sueiro (2005) which is related to perceptron models but with the advantage of selecting the optimal structure, so it is expected to obtain at least the same results of Burks et al. (2005) in their data set.

Table 2
Correct Classification Rate (CCR), Cross Entropy (CE) cost, Square Error (SE) cost, and Receiver Operation Characteristic (ROC) Area Under Curve (AUC) mean and standard deviation values.

	CCR	SE	CE	ROC AUC
Mean	0.85 ∈ [0.84–0.86]	0.14 ∈ [0.13–0.14]	0.63 ∈ [0.60–0.65]	0.90 ∈ [0.90–0.91]
Deviation	0.05 ∈ [0.04–0.06]	0.03 ∈ [0.02–0.03]	0.10 ∈ [0.09–0.11]	0.04 ∈ [0.03–0.04]

Despite we believe that the performance, accuracy and results obtained are at least promising and have potential interest in a real application, in order to try to further improve the numerical results here presented for automatic classification, in terms of CCR and AUC values, one could possibly work further improving any of the following subsystems of the proposed system, in no particular order: using it in combination with a plant row detection system; incorporating stereoscopic computer vision, instead of single lens vision we are using now; employing spectral image technology, that is, to acquire and process the images in bands in or beyond the visible range; improving the segmentation algorithm, always a critical part of any computer vision system; using multiple modality and/or sources of image types and information (multimodal or hybrid systems). For example, combining infrared cameras with conventional RGB cameras; and finally, trying to ameliorate the high illumination variability and heterogeneous nature of the input image data base, obtaining more homogeneous images where possible by either automatic image preprocessing or by improving and controlling the image acquisition process in more detail.

Besides, as briefly indicated before, it is worth emphasizing that weed discrimination could help to improve these results, because they are usually outliers in plant rows. Thus, it is reasonable to believe that our proposed method might be used in combination with the previously mentioned plant row detection methods to increase performance. To end up with the discussion, and previous to the proper conclusions, we believe this system might potentially be of help in the field of Precision Agriculture applications such as mechanical weeding and the selective application of herbicide, but this can only be proven once a real prototype system is built.

Results have shown that a high accuracy system can be achieved by selecting the five most distinct (discriminant) features when a high number of different type features are considered for classification. For sunflower crops, this intelligent system of only five leaf features include: the perimeter (PRI), the area (PXC), the major ellipse axis (MJX), the minor ellipse axis (MNX) and the logarithm of the height to the width ratio (LHW). Thus, given the simplicity of the approach, both in the small number of classes (two) and the small number of simple input pattern features selected above mentioned (five), it is expected that the system can be easily implemented in real time and with little memory requirements using common and low cost digital signal processors that are readily available. In precise terms, given that the system for automatic classification will need only to compute five morphological features after feature selection, it is expected that it could be used in real time practical applications due to its low memory usage and fast classification once the neural network has learned from the patterns and after it has gone through supervised learning from the training set. The exact time in milliseconds is very hard to know until the system has been implemented and programmed using a real time system, and thus we prefer not to speculate. However, since the technology already exists, we believe that meeting the time constraints for this system should be relatively straightforward.

To conclude, we have proposed a complete system for the binary automatic classification of plants into either crop or weeds based on neural networks and proposed a method for selecting the most important features of classification. In precise terms, the following conclusions hold:

1. Good classification under heterogeneous and highly variable lighting can be obtained on the basis of five features only. These features are: the perimeter (PRI), the area (PXC), the major ellipse axis (MJX), the minor ellipse axis (MNX) and the logarithm of the height to width ratio (LHW).
2. GSP NN structure together with the PPMS algorithm in conjunction with the classical J_5 criterion yield an accurate numerical classification.
3. In addition, experimental results point out that a highly accurate classification can be achieved using a completely automatic system. The average CCR of 85% and the AUC value over 90% under the ROC curve for the test set indicate that neural networks, in conjunction with genetic algorithms, can be used to accurately classify sunflower crops.

Acknowledgements

This work was supported by Grant No. TEC07-67073 from the Comisión Interministerial de Ciencia y Tecnología and VA026A07, VA064A08 and VA008B08 from Junta de Castilla y León, Spain. The authors want to thank J.M. Hernández for his help with the experiments and while preparing the manuscript.

References

Aitkenhead, M.J., Dalgetty, I.A., Mullins, C.E., McDonald, A.J., Strachan, N.J., 2003. Weed and crop discrimination using image analysis and artificial intelligence methods. Computers and Electronics in Agriculture 39, 157–171.

Arribas, J.I., Cid-Sueiro, J., 2005. A model selection algorithm for a posteriori probability estimation with neural networks. IEEE Transactions on Neural Networks 16 (4), 799–809.

Brivot, R., Marchant, J.A., 1996. Segmentation of plants and weeds for a precision crop protection robot using infrared images. In: IEEE Proceedings Vision, Image & Signal Processing, vol. 143, pp. 118–124.

Burks, T.F., Shearer, S.A., Heath, J.R., Donohue, K.D., 2005. Evaluation of neural-network classifiers for weed species discrimination. Biosystems Engineering 91, 293–304.

Cho, S.I., Lee, D.S., Jeong, J.Y., 2002. Weed-plant discrimination by machine vision and artificial neural network. Biosystems Engineering 83, 275–280.

Devijver, P.A., Kittler, J., 1982. Pattern Recognition: A Statistical Approach. Prentice Hall International, London.

Fogel, D.B., 1994. An introduction to simulated evolutionary optimization. IEEE Transactions on Neural Networks 5 (1), 3–14.

Gebhardt, S., Kuhbauch, W., 2007. A new algorithm for automatic *Rumex obtusifolius* detection in digital images using colour and texture features and the influence of image resolution. Precision Agriculture 8 (1), 1–13.

Gebhardt, S., Schellberg, J., Lock, R., Kuhbauch, W., 2007. Identification of broad-leaved dock (*Rumex obtusifolius* L.) on grassland by means of digital image processing. Precision Agriculture 7 (3), 165–178.

Ghosal, S., Mehrotra, R., 1997. A moment-based unified approach to image feature detection. IEEE Transactions on Image Processing 6 (6), 781–793.

Hague, T., Tillett, N., Wheeler, H., 2006. Automated crop and weed monitoring in widely spaced cereals. Precision Agriculture 7 (1), 21–32.

Haykin, S., 1999. Neural Networks. A Comprehensive Foundation. Prentice/Hall.

Hemming, J., Rath, T., 2001. Computer-Vision-based Weed Identification under Field Conditions using Controlled Lighting. Journal of Agricultural Engineering Research 78 (3), 233–243.

Huang, K.-Y., 2007. Application of artificial neural network for detecting phalaenopsis seedling diseases using color and texture features. Computers and Electronics in Agriculture 57 (1), 3–11.

Jain, A.K., Duin, R., Mao, J., 2000. Statistical Pattern Recognition: A Review. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (1), 4–37.

Lee, W.S., Slaughter, D.C., Giles, D.K., 1999. Robotic Weed Control System for Tomatoes. Precision Agriculture 1, 95–113.

Marchant, J.A., 1996. Tracking of row structure in three crops using image analysis. Computers and Electronics in Agriculture 15, 161–179.

Marchant, J.A., Brivot, R., 1995. Real-Time Tracking of Plant Rows Using a Hough Transform. Real-Time Imaging, 363–371.

Onyango, C.M., Marchant, J.A., 2003. Segmentation of row crop plants from weeds using colour and morphology. Computers and Electronics in Agriculture 39 (3), 141–155.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics 9, 62–66.

Ruiz-Ruiz, G., Gomez-Gil, J., Navas-Gracia, L.M., 2009. Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm. Computers and Electronics in Agriculture 68, 88–96.

Shrestha, D.S., Steward, B.L., Birrell, S.J., 2004. Video Processing for Early Stage Maize Plant Detection. Biosystem Engineering 89 (2), 119–129.

Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R., 2000. Content based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (12), 1349–1380.

Sogaard, H.T., Olsen, H.J., 2003. Determination of crop rows by image analysis without segmentation. Computers and Electronics in Agriculture 38, 141–158.

Southall, B., Marchant, J.A., Hague, T., Buxton, B.F., 1998. Model based tracking for navigation and segmentation. In: Proceedings Fifth European Conference on Computer Vision, vol. 1, pp. 797–811.

Tellaeche, A., Burgos-Artizzu, X.P., Pajares, G., Ribeiro, A., Fernandez-Quintanilla, C., 2008. A new vision-based approach to differential spraying in precision agriculture. Computers and Electronics in Agriculture 60, 144–155.

Tian, L.F., Slaughter, D.C., 1998. Environmentally adaptive segmentation algorithm for outdoor image segmentation. Computers and Electronics in Agriculture 21 (3), 153–168.

Tian, L.F., Slaughter, D.C., Norris, R.F., 1997. Machine vision identification of tomato seedlings for automated weed control. Transactions of the ASAE 40 (6), 1761–1768.

Verma, B., Zhang, P., 2007. A novel neural-genetic algorithm to find the most significant combination of features in digital mammograms. Applied Soft Computing 7 (2), 612–625.

Watanabe, S., 1985. Pattern Recognition: Human and Mechanical. Wiley, New York.

Yang, C.C., Prasher, S.O., Landry, J.A., Ramaswamy, H., Ditommaso, A., 2000. Application of artificial neural networks in image recognition and classification of crop and weeds. Canadian Agriculture Engineering 42, 147–152.