A new Expert System for greenness identification in agricultural images

J. Romeo [a], G. Pajares [a,*], M. Montalvo [b], J.M. Guerrero [a], M. Guijarro [a], J.M. de la Cruz [b]

[a] Dpt. Software Engineering and Artificial Intelligence, Faculty of Computer Science, Complutense University, 28040 Madrid, Spain
[b] Dpt. Computer Architecture and Automatic Control, Faculty of Computer Science, Complutense University, 28040 Madrid, Spain

ARTICLE INFO

ABSTRACT

It is well-known that one important issue emerging strongly in agriculture is related with the automation of tasks, where camera-based sensors play an important role. They provide images that must be conveniently processed. The most relevant image processing procedures require the identification of green plants, in our experiments they comes from barley and maize fields including weeds, so that some type of action can be carried out, including site-specific treatments with chemical products or mechanical manipulations.

The images come from outdoor environments, which are affected for a high variability of illumination conditions because of sunny or cloudy days or both with high rate of changes.

Several indices have been proposed in the literature for greenness identification, but under adverse environmental conditions most of them fail or do not work properly. This is true even for camera devices with auto-image white balance.

This paper proposes a new automatic and robust Expert System for greenness identification. It consists of two main modules: (1) decision making, based on image histogram analysis and (2) greenness identification, where two different strategies are proposed, the first based on classical greenness identification methods and the second inspired on the Fuzzy Clustering approach. The Expert System design as a whole makes a contribution, but the Fuzzy Clustering strategy makes the main finding of this paper. The system is tested for different images captured with several camera devices.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Problem statement

Camera based devices is an excellent sensor for several applications. One of them is in agriculture where autonomous vehicles equipped with cameras are demanding solutions to distinguish plants (crops and weeds) with the aim of applying treatments over site-specific areas in a larger field (Davies, Casady, & Massey, 1998). Focusing on maize or barley fields, one of the most important treatments is weeds killing. Different methods and strategies for plant identification have been applied in different works (Burgos-Artizzu, Ribeiro, Tellaeche, Pajares, & Fernández-Quintanilla, 2009; Guerrero, Pajares, Montalvo, Romeo, & Guijarro, 2012; Guijarro et al. 2011; Montalvo et al., 2012; Onyango & Marchant, 2003; Tellaeche, Burgos-Artizzu, Pajares, & Ribeiro, 2008; Tellaeche, Burgos-Artizzu, Pajares, Ribeiro, & Fernández-Quintanilla, 2008). Lopez-Granados (2011) makes a revision of methods where plant identification is a key step in the process. Most existing strategies address the problem of green identification under the assumption that plants display a clear high degree of greenness. The images come from outdoor environments, which are affected for a high variability of illumination conditions: sunny or cloudy days or both with high rate of changes. In sunny days the position of the sun with respect the camera makes the illumination component impact differently and plants have different shades and shadows in the image. Furthermore, the greater the solar illumination the more intense infrared and ultraviolet radiations are. Also, high components of illumination could cause sensor saturation. These situations are absolutely normal in agricultural environments. Although modern camera devices have the ability to make auto image corrections through auto-image white balance or enhancement, most times this is insufficient, particularly in industrial cameras, where most camera settings must be dynamically adjusted (exposure time, auto-image white balance) or fixed in advance by the user (focal length, iris aperture).

This paper proposes a new automatic and robust Expert System for greenness identification which is able to cope with the identification of green plants even though adverse environmental conditions. It consists of two main modules: (1) decision making, based on histogram analysis and (2) greenness identification, where two different strategies are proposed, the first based on classical methods and the second inspired on the Fuzzy Clustering approach. The

---

* Corresponding author. Tel.: +34 1 3 94 75 46; fax: +34 1 3 94 75 47.
  E-mail address: pajares@fdi.ucm.es (G. Pajares).

Expert System design as a whole and the Fuzzy Clustering strategy make the contributions of this paper. The performance of the method allows to verify its viability for automatic tasks in agriculture, involving the identification of green plants.

1.2. Revision of methods for greenness identification

Some years ago Tian and Slaughter (1998) considered the images captured under different conditions such as sunny or cloudy days, affecting illumination variability and assuming that they are typical situations in agricultural images coming from outdoor environments. Later several strategies have been proposed for segmenting crop canopy images, specifically oriented towards green segmentation:

(1) Visible spectral-index based, including the excess green index (ExG, Ribeiro, Fernández-Quintanilla, Barroso, & García-Alegre, 2005; Woebbecke, Meyer, von Bargen, & Mortensen, 1995), the excess red index (ExR, Meyer, Hindman, & Lakshmi, 1998), the color index of vegetation extraction (CIVE, Kataoka, Kaneko, Okamoto, & Hata, 2003), the excess green minus excess red index (ExGR, Neto, 2004) and the vegetative index (VEG) described in Hague, Tillet, and Wheeler (2006), which is designed to cope with the variability of natural daylight illumination. ExG, ExGR, CIVE and VEG have been applied under a combined form in Guijarro et al. (2011) gaining in performance with respect to their individual application. All these approaches need to fix a threshold for final segmentation, i.e. to discriminate between plants and other parts (soil, sky).
(2) Specific threshold-based approaches, including dynamic thresholding. Generally, these techniques assume a two-class problem where plants and soil are to be identified. Reid and Searcy (1987) estimate a decision function under the assumption that the classes follow Gaussian distributions. The Otsu's method (Otsu, 1979) is also applied considering a bi-class problem (Ling & Ruzhitsky, 1996; Shrestha, Steward, & Birrell, 2004). These algorithms are applied to gray images. Gebhardt, Schellberg, Lock, and Kaühbauch (2006) apply also thresholding for segmentation, and transform the images from RGB to gray scale intensity. This approach was later improved using local homogeneity and morphological operations in Gebhardt and Kaühbauch (2007). Kirk, Andersen, Thomsen, and Jørgensen (2009) apply a combination of greenness and intensity derived from the red and green spectral bands and compute an automatic threshold for a two-class problem assuming two Gaussian probability density functions associated to soil and vegetation respectively; this procedure requires the previous estimation of an angle to rotate the hypothetical greenness axis. Meyer and Camargo-Neto (2008) have applied the automatic Otsu's thresholding method for binarizing ExG and the normalized difference index (NDI), where a comparison is established against the segmentation obtained from ExGR determining that in this last case, a value of zero suffices for the threshold, therefore the Otsu's method is not required. Guijarro et al. (2011) and Burgos-Artizzu, Ribeiro, Guijarro, and Pajares (2011) have applied the statistical mean value of the transformed image obtained with the vegetation indices instead of automatic thresholding such as Otsu. They justify its choice because Otsu's method gives a threshold value higher than the mean and produces infra-segmentation, i.e. some plants are not conveniently identified.
(3) Learning-based, Meyer, Camargo-Neto, Jones, and Hindman (2004) have applied unsupervised approaches, including Fuzzy Clustering, for segmenting regions of interest from

ExR and ExG. Tian and Slaughter (1998) proposed the environmentally adaptive segmentation algorithm (EASA) for detecting plants through a supervised learning process. Ruiz-Ruiz, Gómez-Gil, and Navas-Gracia (2009) applied the EASA under the HSI (hue-saturation-intensity) color space to deal with the illumination variability. Zheng, Zhang, and Wang (2009) and Zheng, Shi, and Zhang (2010) use a supervised mean-shift algorithm under the assumption that the segmentation of green vegetation from a background can be treated as a two-class segmentation problem; the class separability is validated through a neural network and the Fisher linear discriminant respectively, the color spaces used were RGB, LUV and HSI. Guerrero et al. (2012) apply Support vector machines as the learning.

1.3. Motivational research of the proposed strategy

The above methods are intended for plant identification through their greenness, based on the accentuation of the green color (Meyer & Camargo-Neto, 2008), but their effectiveness drops when, in the plants analyzed, the green spectral component loss its relevance because of adverse environmental conditions where such component takes similar values and sometimes lower than the red one.

Moreover agricultural images contain not only green plants but other structures, (soil, sky, debris, crop residues or shades), that must be conveniently addressed.

The main direct effect caused by the illumination is reflected in the image histogram, where the contrast is a decisive factor for greenness identification. Classical methods (ExG, ExGR, CIVE or VEG) work appropriately for well-contrasted images but fail miserably when images are insufficiently contrasted although they are later enhanced. This situation occurs most often in images captured with industrial cameras connected to a computer for image processing. Another undesired situation appears when important parts in the images become highly saturated.

Thus, our system is designed with a first decision making module based on image histogram analysis, which determines if the incoming image contains sufficient quality to apply classical greenness identification methods. Otherwise, the image is to be processed by a new greenness strategy, which is a method specifically developed with such purpose.

We focus this specific method as a bi-classification approach where green plants are considered as belonging to a class and the remainder elements in the image are assigned to the other class. The Fuzzy Clustering (FC) approach is conveniently adapted for this purpose. FC consists of two phases, namely: learning and classification. The learning phase is exploited to determine a dynamic threshold for each image and the classification is reduced to a simple decision making process. Previous to this process we apply image down sampling to achieve an image resolution with the aim of save processing time so that the image fulfil real time requirements if any.

Thus, the idea is to apply an automatic strategy for image segmentation based on the potential ability of the FC approach, where its learning phase is exploited to obtain a specific threshold valid for each image. No learning is required in the general sense; this avoids the need of prior training.

1.4. Paper organization

This paper is organized as follows. In Section 2 we explain the design of the proposed automatic Expert System with its stages and the corresponding image preprocessing procedures. In Section 3 the performance of the proposed strategy is evaluated. Finally in Section 4, the most relevant conclusions are extracted.

2. Expert System design

The Expert System (ES) architecture is outlined in Fig. 1, where two main modules are identified: Decision making and Greenness identification. The incoming image is transferred to the decision making module, where based on histogram analysis a decision is made according to the contrast of the histogram. If sufficient contrast, we apply the combined method described in Guijarro et al. (2011) where at the end a binary image is obtained with the green plants identified; otherwise when the image displays what we consider an insufficient contrast, it is firstly image processed by applying a down-sampling together with smoothing. After this preprocessing, a procedure inspired on the FC approach allows to determine a threshold, during a phase equivalent to what is considered as learning in the classical FC approach. Once the threshold is available, through a simple decision rule, a binary image is obtained, also with the green plants identified. In what follows we explain all process involved in the ES.

2.1. Decision making

This module is designed with the aim of analyze the quality of the incoming image based on its histogram. After the analysis, a decision is made and the image is processed according to one of two available processes in the greenness identification module. The histograms for the three spectral channels from the RGB image are processed. Each histogram is defined by its corresponding discrete function $h(a) = n_a$, where a represents the gray level and n_a is the number of pixels with gray level a. In our 8-bit RGB channel representation a ranges in $[0, 255]$. Dividing $h(a)$ by the total number of pixels in the image, the estimation of the probability is obtained, i.e. $p(a) = h(a)/M \times N$, where M and N are both sizes of the image. The histogram provides useful image statistics that are exploited for determining the image quality. Let a be a random variable denoting gray levels, the nth moment of a about the mean is defined as (Gonzalez & Woods, 2008):

$$\mu_n(a) = \sum_a (a - m)^n p(a) \tag{1}$$

where m is the mean value of a, i.e. the average gray level

Fig. 1. Expert System architecture.

$$m = \sum_a a \cdot p(a) \tag{2}$$

The most important statistical parameters from the point of view of image quality are (Holub & Ferreira, 2006): mean, m; variance, $v = \mu_2(a)$; skewness, $\gamma = \frac{\mu_3(a)}{\mu_2^{\frac{3}{2}}(a)}$ and kurtosis $\kappa = \frac{\mu_4(a)}{\mu_2^2(a)}$. The mean determines the average level of brightness, where low, high and medium values indicate the degree of light which has impacted the device. The moments provide information about the distribution of values around the mean. Variance is a measure of gray-level contrast, where high values indicate dispersion of values around the mean and low values are indicative of a high concentration of values around the mean. The skewness measures the asymmetry in the distribution. A right skewness is presented when the histogram displays a large tail oriented towards high brightness values and high concentration in the part of low brightness values (positive skewness). In the opposite case the skewness is negative. The kurtosis provides information about the peakedness in the distribution; low kurtosis indicates flat top parts in the histogram around the mean but high values are indicative of peaks around the mean with high slopes and large tails. Skewness and kurtosis are both zero for Gaussian distributions.

Once we have defined the above parameters, the question is: how can we determine the image quality and what means quality from the point of view of our agricultural images? The answer comes from the experience in this area. Images highly contrasted are considered as images with sufficient quality and vice versa. An image with sufficient contrast should be identified by mean values in the central part of histogram, high variance, low skewness (positive or negative) and high kurtosis. On the contrary, an image with insufficient contrast is identified by mean values either low or high, high skewness (positive or negative) and low kurtosis. The next step is to determine the ranges of variability for the above parameters and accordingly the derivation of the corresponding decision rule. With such purpose we have analyzed a set of images described in Section 3, such images are classified as images with sufficient and insufficient contrast from which we obtain their mean, variance, skewness and kurtosis values. A set of three rules, each with four premises suffices for making decisions

If $m_{ai} < m_i < m_{bi}$ and $v_i > v_{ai}$ and $|\gamma_i| < \gamma_{ai}$ and $\kappa_i > \kappa_{ai}$
$\Rightarrow$ sufficient contrast (3)
otherwise $\Rightarrow$ insufficient contrast

where i = R, G, B; i.e. three values of each of the three parameters are obtained according to the three RGB spectral channels. Parameters with subindices a and b are upper and lower limits to be fixed in Section 3.

2.2. Greenness identification

According to the scheme displayed in Fig. 1, if the incoming image contains sufficient contrast a combination of well-tested methods is applied based on the experiments carried out in Guijarro et al. (2011). Otherwise, a new strategy inspired on the FC approach is proposed for image thresholding.

2.2.1. Combination of methods

Given an original input image in the RGB color space, we apply the following normalization scheme, which is usually applied in agronomic image segmentation (Gée, Bossu, Jones, & Truchetet, 2008),

$$r = \frac{R_n}{R_n + G_n + B_n}, \quad g = \frac{G_n}{R_n + G_n + B_n}, \quad b = \frac{B_n}{R_n + G_n + B_n} \tag{4}$$

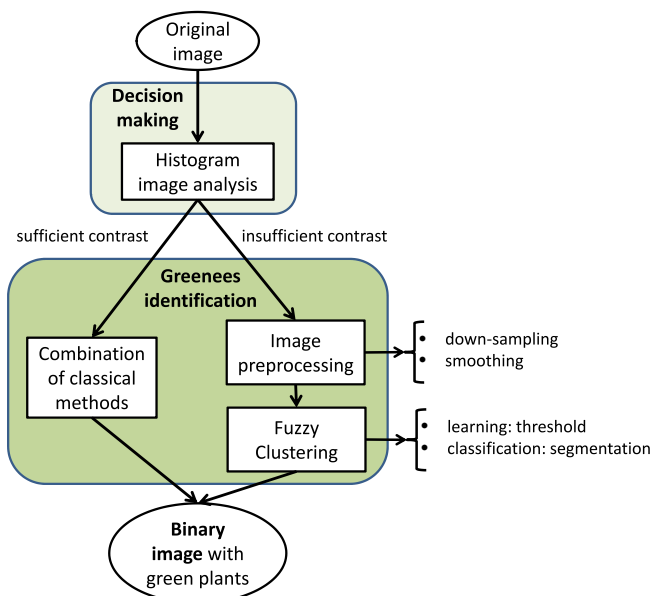where R, G and B are the normalized RGB coordinates ranging from 0 to 1 and are obtained as follows:

$$R_n = \frac{R}{R_{max}}, \quad G_n = \frac{G}{G_{max}}, \quad B_n = \frac{B}{B_{max}} \tag{5}$$

where $R_{max} = G_{max} = B_{max} = 255$ for our 24-bit color images. Vegetation indices to be combined are computed as follows (see references above in Section 1.2),

Excess green : $\quad ExG = 2g - r - b \tag{6}$

Excess green minus excess red : $\quad ExGR = ExG - 1.4r - g \tag{7}$

Color index of vegetation extraction

$$CIVE = 0.441r - 0.811g + 0.385b + 18.78745 \tag{8}$$

Vegetative $\quad VEG = \frac{g}{r^a b^{1-a}},$

with a set to 0.667 as in Hague etal., 2006 $\tag{9}$

Based on Guijarro et al. (2011) the above three indices are combined to obtain the resulting value COM as follows,

Combined :
$$COM = w_{ExG}ExG + w_{ExGR}ExGR + w_{CIVE}CIVE + w_{VEG}VEG \tag{10}$$

where w_{ExG}, w_{CIVE} and w_{VEG} are weights for each index, representing the relative importance of the index. Guijarro et al. (2011) provide the values for the four weights participating in the combination, which are the following: $w_{ExG} = 0.25$, $w_{ExGR} = 0.30 w_{CIVE} = 0.33$ and $w_{VEG} = 0.12$.

The resulting combined image COM, is linearly mapped to range in [0,1], after which, it is thresholded by applying the Otsu's (1979) method, obtaining a binary image, where white pixels identify plants in the original image, with clear spectral RGB components associated to unmasked plants. On the contrary, black pixels identify those pixels in the original image belonging to masked plants, soil and other materials present in the field.

2.2.2. Preprocessing and Fuzzy Clustering

Two consecutive strategies are proposed when the images are insufficiently contrasted, i.e. image preprocessing and FC. The former is proposed mainly for image reduction, although it could be avoided if restrictive real-time problems exist. Indeed, because FC involves a first phase associated to a learning process (iterative) the number of samples, which are pixels coming from the image, determine the computational cost. The lower the number of pixels to process lower the computation time. The second (FC) is the kernel of this process, where after the learning phase, a threshold is obtained which allows the image binarization.

(a) *Image preprocessing* The original image is down sampled by applying successive Discrete Wavelets Transform (DWT) (Pajares & de la Cruz, 2004). Only the DWT is of interest applied over the approximation coefficients at each resolution level and for each spectral R, G, B channel. In this way, as the same time the signal is decomposed, a smoothing process is applied to minimize possible noise or spurious pixels caused by the adverse environmental outdoor conditions.
(b) *Image thresholding by Fuzzy Clustering*
(b.1) *Learning phase and thresholding*: given the down sampled RGB image, it contains n pixels which are stored in X i.e., $X = \{x_1, x_2, \ldots, x_n\} \in \Re^d$, where d is the data dimensionality. Therefore, each sample vector $\boldsymbol{x_i}$ represents an image pixel, where its components are the three RGB spectral components of that pixel at the original image location (x,y). This means that in our experiments the data dimensionality is $d = 3$. Each sample is to be assigned to a given cluster w_j, where the number of possible clusters is c, i.e., $j = 1,2,\ldots,c$. In the proposed approach c is set to 2 because

we were only interested on two types of textures, i.e. green plants (crop/weeds) and the remainder (soil, debris, stones).

The samples in X are to be classified based on the well-known Fuzzy Clustering approach that receives the input training samples $\boldsymbol{x_i}$ and establishes a partition, assuming the number of clusters c is known. The fuzzy partitioning problem for $b > 1$ is minimize the following variance criterion,

$$\min J_b(\mu_i^j, \boldsymbol{v}, t) = \sum_{i=1}^n \sum_{j=1}^c (\mu_i^j(t))^b \|\boldsymbol{x}_i - \boldsymbol{v}_j\|^2 \tag{11}$$

The process computes for each $\boldsymbol{x}_i$ at the iteration t, its degree of membership in the cluster $w_j(\mu_i^j)$ and updates the cluster centers $\boldsymbol{v}_j$ as follows (Zimmermann, 1991):

$$\mu_i^j(t+1) = \frac{1}{\sum_{r=1}^c (d_{ij}(t)/d_{ir}(t))^{2/(b-1)}};$$
$$\boldsymbol{v}_j(t+1) = \frac{\sum_{i=1}^n [\mu_i^j(t)]^b \boldsymbol{x}_i}{\sum_{i=1}^n [\mu_i^j(t)]^b} \tag{12}$$

$d_{ij}^2 \equiv d^2(\boldsymbol{x}_i, \boldsymbol{v}_j)$ is the squared Euclidean distance. The number b is called the exponential weight (Bezdek, 1981; Duda, Hart, & Stork, 2000), $b > 1$. The stopping criterion of the iteration process is achieved when $\|\mu_i^j(t+1) - \mu_i^j(t)\| < \varepsilon \; \forall ij$ or a number t_{max} of iterations is reached.

The method requires the initialization of the cluster centers, so that the Eq. (11) can be applied at the iteration $t = 1$. Because our interest consists in the identification of green plants among other parts one ideal center should be the one with the pure green color, without loss of generality this center is identified as $\boldsymbol{v}_1 = (0, 255, 0)$; the second center only requires that the green spectral component is not dominant over the red and blue, the following values suffice $\boldsymbol{v}_2 = (255, 0, 128)$.

Once the learning process is finished we obtain two cluster centers $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ associated to clusters w_1 and w_2 with components $\boldsymbol{v}_1 \equiv \{\boldsymbol{v}_{1R}, \boldsymbol{v}_{1G}, \boldsymbol{v}_{1B}\}$ and $\boldsymbol{v}_2 \equiv \{\boldsymbol{v}_{2R}, \boldsymbol{v}_{2G}, \boldsymbol{v}_{2B}\}$ respectively; where v_{1R}, v_{1G} and v_{1B} or v_{2R}, v_{2G} and v_{2B} represent the averaged values for the corresponding RGB spectral components at each cluster obtained according to Eq. (12). Thus, because $\boldsymbol{v}_1$ is associated to the cluster where green plants belongs to, we compute the threshold value $T = v_{1G}/(v_{1R} + v_{1G} + v_{1B})$, which represents the proportion of greenness with respect the three spectral components. It is expected green plants tend to have their green spectral components in the RGB color model the greatest. Thus we chose T as the threshold for the posterior decision making process.

(b.2) *Classification phase and segmentation*: given the original RGB image, we obtain the three r, g and b normalized values through Eq. (4). Considering the threshold T computed as above, the following rule allows to obtain a binary image, W, where white pixels are identified as green pixels in the original image:

$$W = \begin{cases} 1 & if \; \frac{g}{r+g+b} > T \\ 0 & otherwise \end{cases} \tag{13}$$

3. Results

The images used for this study were acquired with the following digital color camera devices and dates: (a) HPR817 with image resolution of 600×800 in March/April/May 2007 (330 images) for barley and maize crops; (b) Canon EOS 400D with resolution of 1944×2592 during April/May 2011 (350 images); (c) Basler

17FC 1400 connected to a laptop through FireWire IEEE 1394 and resolution of 1392 × 1038 in April/May 2011 (320 images); (d) Kodak EasyShare M380 with resolution 1280 × 960 in March 2011 for barley crop and May 2012 for maize crop (210 images); (e) SVS-VISTEK with GigaEthernet connection to a laptop and resolution of 2336 × 1752 equipped with UV-IR cut Schneider filter 486 in March/April/May 2012 (340 images) for barley and maize crops. This makes a total of 1550 images.

All acquisitions for the same crop were spaced by different days varying from five/six/ten until thirty days, under a high variety of illumination conditions, including: cloudy and sunny days with sun incident around all angles with respect the camera; during the morning, afternoon, midday and late afternoon.

These digital images were captured under perspective projection in barley and maize fields containing soil with different elements (stones, debris or old crop residues), plants (crop and weeds), sky (with and without clouds) and other elements such as buildings. They were stored as 24-bit color images and saved in RGB (Red, Green and Blue) color space in the BMP and JPG formats. The Expert System was implemented in Matlab R2010b (The Mathworks, 2012) and the images were processed with its Image Processing Toolbox.

3.1. Decision making module: identification of parameters

From the set of available images, we have randomly selected the 5% of images obtained with each camera device, including images of barley and maize when available. According to the Greenness identification module in Fig. 1, for each image we obtain two binary images, one by applying the combination of classical methods according to Eq. (10) and the second by applying image preprocessing and Fuzzy Clustering, Eqs. (11) and (12). The binary images are visually analyzed by an expert in order to determine the best result, considering how the green plants, present in the original images, have been identified. Figs. 2–4(a) display three representative images where the best binarization is achieved by applying combination of classical methods. On the contrary, images Figs. 5–7(a) display three representative images in which the best binarized results have been obtained with the processing involving FC, i.e. they are considered images with insufficient contrast. Figs. 2–7 labeled with (b) display the corresponding histograms for the three RGB spectral channels.

Figs. 2–4 display images with histograms with distributions covering the whole range of values and with maximum values located at the central part. On the contrary, histograms associated to Figs. 4–7 are heeled left or right. This represents an important finding, which is decisive for making a decision. The only question is to quantify this observation in some way.
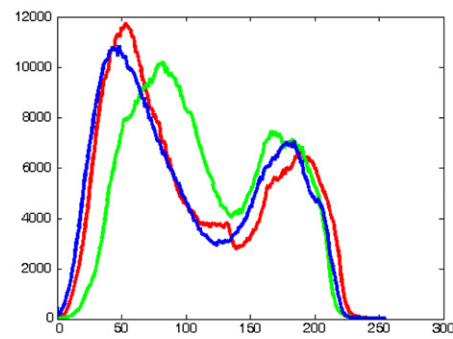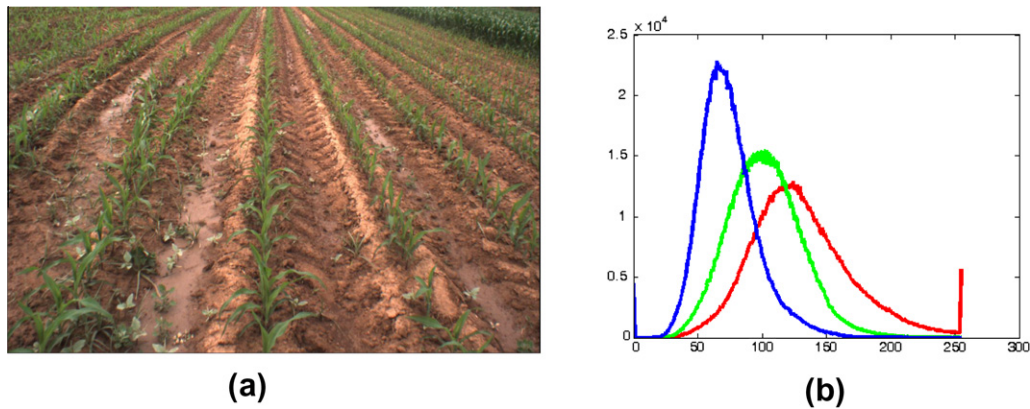
Table 1 displays the average parameter values derived from the corresponding histograms on the set of images analyzed according to the best binary image obtained. Minimum and maximum values for each R, G, B spectral channel are identified for the mean (m), variance (v), skewness ($|\gamma|$) and kurtosis (κ). They are distinguished as images with sufficient contrast and images with insufficient contrast and inside these last ones we still can distinguish between images with the histogram heeled left or right, this is because the mean histogram values ranges differently.

From results in Table 1, two important conclusions can be extracted. There are two parameters which are really discriminating (m and $|\gamma|$) and two that are not (v and κ). Indeed, considering the Red spectral channel, we can see that the mean for images with sufficient contrast ranges in [106,130] and for images with insufficient contrast there are two intervals, one for left heeled [60,75] and one for right heeled [174,217]. These intervals clearly appear as no overlapped with wide margins between them. This reasoning can be extended to the Green and Blue spectral channels, where this same behavior appears, i.e. [102,136] and [65,81]–[168,213] for Green and [78,112] and [49,60]–[144,167] for Blue. With respect skewness and the Red spectral channel we can see that images with sufficient contrast the range is [0.22,0.45] and with insufficient contrast [0.85,0.89] and [0.79,0.98], i.e. with overlapping between these intervals. Similar behavior is applicable for the Green spectral channel with [0.10,0.21] and [0.89,0.93]–[0.78,0.87] and for the Blue channel is [0.33,0.41] and [0.82,0.90]–[0.68,0.78]. Fig. 8 clarifies this situation, where we can verify indeed that not overlapping exists on each for the averaged mean and skewness values. Each channel is identified with the subindices R, G and B accordingly.

The behavior displayed by mean and skewness does not occur for variance and kurtosis where overlapping occurs and therefore, these last ones are not discriminant. Hence only setting of mean and skewness parameters involved in Eq. (3) for each spectral channel (R,G,B) is possible as follows: m_{aR} = 106; m_{aG} = 102; m_{aB} = 78; m_{bR} = 130; m_{bG} = 136; m_{bB} = 112, we can assume a certain margin of tolerance because there are sufficient distances for no overlapping; for skewness it is clear that appropriate values for the three spectral channels should be between 0.50 and 0.60, i.e. $\gamma_{aR} = \gamma_{aG} = \gamma_{aB}$ = 0.55. With respect variance and kurtosis they are no longer considered and they can be removed from Eq. (3).

### 3.2. Greenness identification module

Once the above parameters have been established, we have randomly selected the 15% of available images, i.e. 232, about half with enough contrast and the other half with insufficient contrast, which are all different from the ones used for parameter's identification. The performance of this combined strategy and also the one

Fig. 2. (a) Original image (maize) captured with the Kodak EasyShare M380 camera and (b) Histogram for the three RGB spectral channels.

Fig. 3. (a) Original image (maize) captured with the 17FC 1400 Basler camera and (b) Histogram for the three RGB spectral channels.
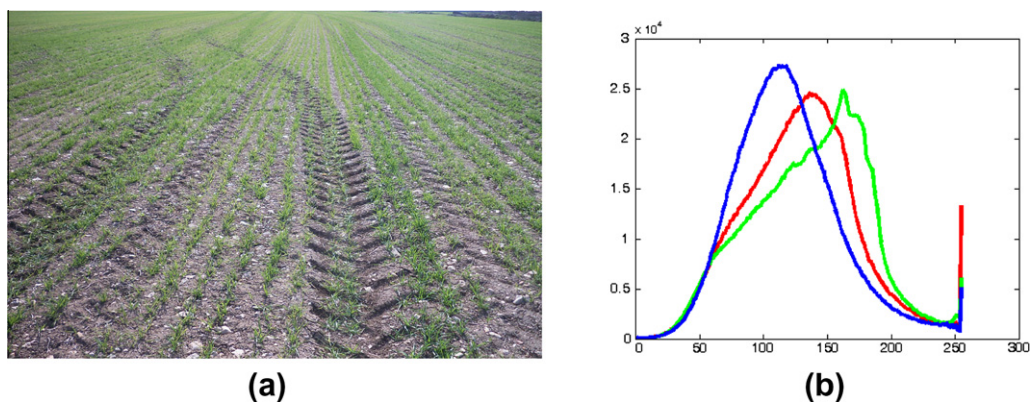
Fig. 4. (a) Original image (barley) captured with the HPR817 camera and (b) Histogram for the three RGB spectral channels.
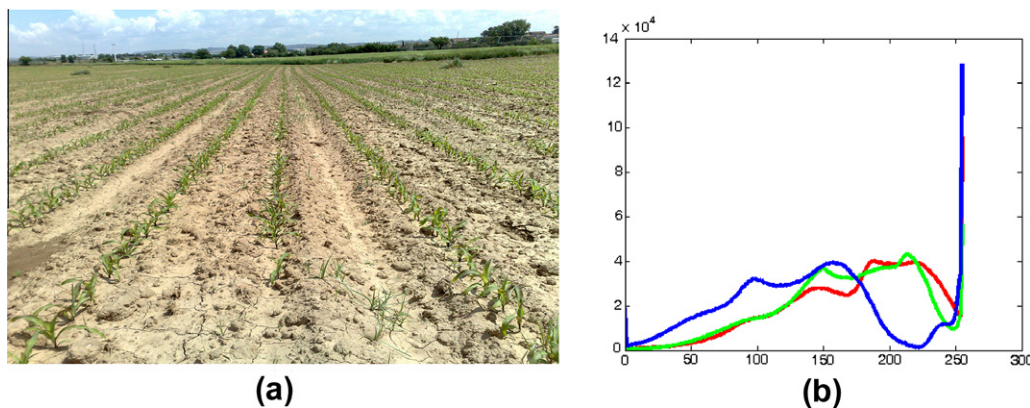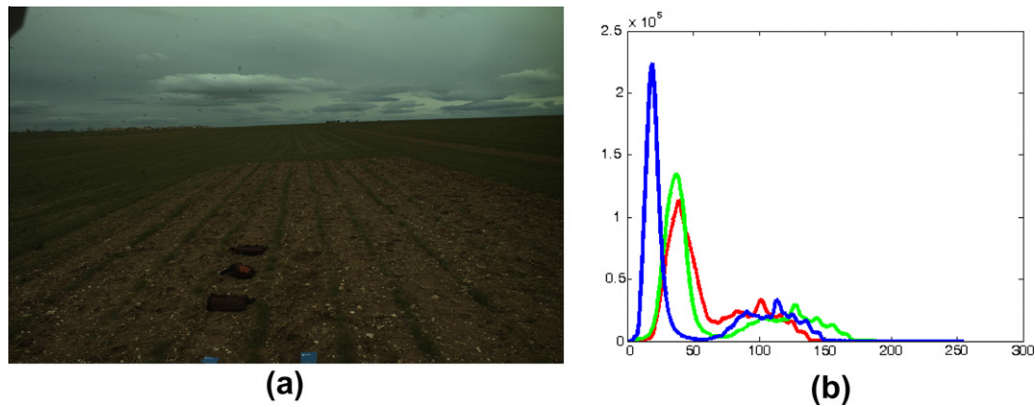

Fig. 5. (a) Original image (maize) captured with Canon EOS 400D and (b) Histogram for the three RGB spectral channels.

based on FC is tested against the corresponding ground-truth images which are built following the procedure described in Guijarro et al. (2011), which is summarized here for facility as follows: (1) given the original image in the RGB color space, we extracted all pixels where the Green component is greater than the Red and Blue ones simultaneously at least in a 10%. This allows us to obtain an important part of green plants, which are labeled as white; (2) also from the original image we extracted those parts where the Red and Blue components are dominant, i.e. they overpass at least in a value of 10% the other two. This allows extracting the Red and Blue parts with a certain degree of confidence, being

labeled as black; (3) the pixels that still were not labeled were manually touched up and labeled according to the human expert criterion by comparing them with the original image. Nevertheless, when the images are highly complex, such as the one in Fig. 6(a) or extremely complex as the one displayed in Fig. 10(a), we select parts containing green plants and other elements which are easily distinguishable.

3.2.1. Combination of classical methods

We have applied exactly the parameters proposed in Guijarro et al. (2011) obtaining similar results.

Fig. 6. (a) Original image (barley with wide row crops simulating maize) captured with SVS-VISTEK camera and (b) Histogram for the three RGB spectral channels.
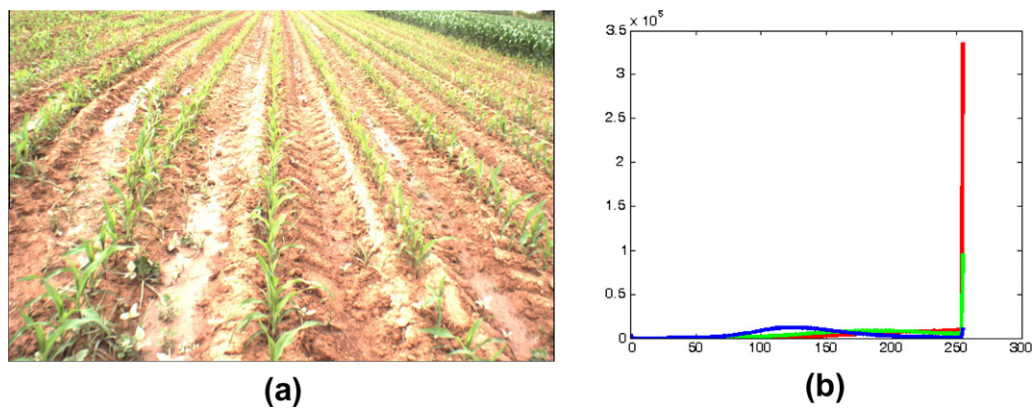
Fig. 7. (a) Original image (maize) with the 17FC 1400 Basler and (b) Histogram for the three RGB spectral channels.

Table 1
Histogram parameters (mean, variance, skewness, kurtosis) identifying minimum and maximum values for images with sufficient/insufficient contrast in the three R, G, B spectral channels.

	Sufficient contrast		Insufficient contrast			
	Minimum (R, G, B)	Maximum (R, G, B)	Left heeled		Right heeled	
			Minimum (R, G, B)	Maximum (R, G, B)	Minimum (R, G, B)	Maximum (R, G, B)
m	$(106, 102, 78)$	$(130, 136, 112)$	$(60, 65, 49)$	$(75, 81, 60)$	$(174, 168, 144)$	$(217, 213, 167)$
$v \times 10^3$	$(1.51, 0.83, 0.52)$	$(3.73, 2.64, 3.65)$	$(0.91, 1.83, 1.85)$	$(2.91, 3.55, 3.89)$	$(1.83, 2.20, 1.72)$	$(2.72, 2.53, 3.33)$
$|\gamma|$	$(0.22, 0.10, 0.33)$	$(0.45, 0.21, 0.41)$	$(0.85, 0.89, 0.82)$	$(0.89, 0.93, 0.90)$	$(0.79, 0.78, 0.68)$	$(0.98, 0.87, 0.78)$
κ	$(1.75, 1.84, 1.72)$	$(3.69, 3.73, 4.50)$	$(2.46, 2.26, 2.10)$	$(4.45, 3.98, 4.60)$	$(2.77, 2.74, 2.69)$	$(5.42, 3.08, 3.66)$

3.2.2. Image preprocessing

We have applied successive DWT decomposition. As mentioned before, this procedure involves both down-sampling and filtering processes. Images with maximum resolution used in this paper are those provided by Canon EOS 400D camera, i.e. 1944×2592. We have verified that reductions of $\frac{1}{16}$ suffices for these images to achieve appropriate performances in terms of effectiveness, as compared to the corresponding ground-truth, while reducing the computational cost during this process and also during the subsequent FC process because the number of samples to be processed. This leads to resolutions of 121×162 i.e. the number of samples for our experiments is finally fixed to 19,602. For original images with resolutions other than the above, the DWT decomposition level should be the appropriate to achieve a similar number of samples. Regarding the family of wavelets used, we have verified that this aspect is not relevant, thus we have used the one that involves less computational cost. The Haar transform fulfill this requirement because its associated low pass filter contains the less number of coefficients and the number of computational operations required for filtering is less.

3.2.3. Fuzzy Clustering

As mentioned above, this process starts with the cluster centers already initialized, then applying successive iterations we estimate the final cluster centers, Eq. (12), until convergence through ε or a maximum number of iterations is reached t_{max}.

In order to fix these parameters we have selected 10 images identified as images with insufficient contrast and its corresponding ground-truth. Initially we fix $t_{max} = 100$, which is a number relatively high. We vary ε from 10^{-2} to 10^{-5} in steps of 5×10^{-5}, once we achieve the best results according to the ground-truth for each image, we annotate the number of iterations and the ε value for each image. Table 2 displays these values for the ten images tested.
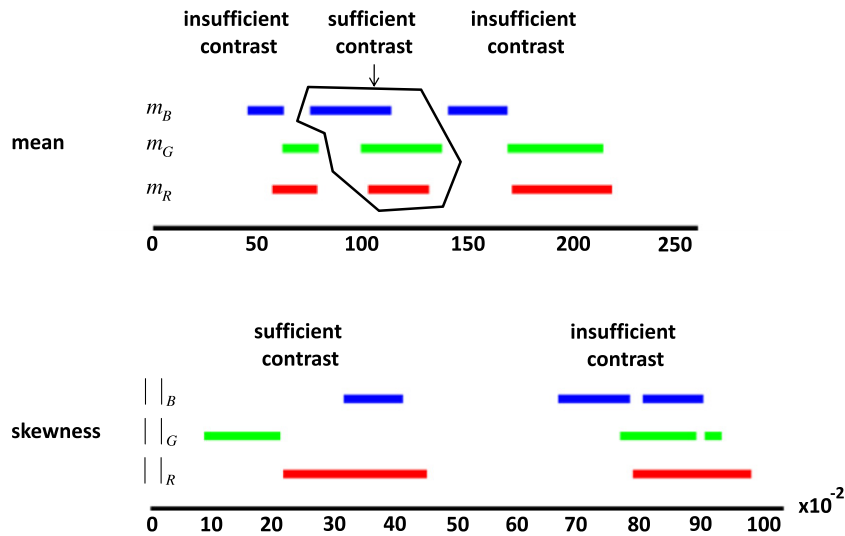
Fig. 8. Representation of averaged values for mean and skewness for the three RGB spectral channels.

Table 2
Iterations and ε values for ten images analyzed.

# Image	1	2	3	4	5	6	7	8	9	10
Iterations (t)	3	8	12	2	3	1	6	7	3	2
$\varepsilon \times 10^{-5}$	6	11	14	77	73	24	18	10	12	15

Table 3
Averaged CPU computational times in seconds for the process involved in the ES.

	Histogram analysis	COM	DWT	FC (per iteration)
CPU time (s)	0.19	0.72	1.01	0.03

With values in Table 2, we finally fix t_{max} = 14 and ε = 10^{-5}, which guarantee a certain degree of confidence, because they are the most restrictive. These values are the ones used for posterior analysis.

3.2.4. Global performance: quantitative analysis
The proposed ES is quantitatively analyzed with 222 images with their corresponding ground-truth, where the ones used above for fixing t_{max} and ε are now discarded.

Because these images have been previously identified as images with sufficient (108) and insufficient contrast (114) by the expert, the performance of the *decision making* module is evaluated based on the following magnitudes:

- SC: number of images identified correctly with sufficient contrast.
- IC: number of images identified correctly with insufficient contrast.
- SCE: number of images identified with sufficient contrast being of insufficient contrast (error).
- ICE: number of images identified with insufficient contrast being of sufficient contrast (error).

From these quantities we compute the percentage of correct classification as follows,

$$PCC = \frac{SC + IC}{SC + SCE + IC + ICE} \qquad (14)$$

Finally, according to the decision rule in (3) excluding variance and kurtosis as explained, we obtained SC = 101, IC = 109, SCE = 7, ICE = 5, i.e. PCC = 0.95, which is an acceptable result.

The combination of classical methods is tested with the SC and ICE images, from which we have achieved the 91% as percentage of success, averaged over the number of images tested and also based on the corresponding ground truth, which is a similar result to the one reported in Guijarro et al. (2011).

The performance of our proposed FC method is tested with the IC and SCE images obtaining a percentage of 89.9% of success.

Table 3 displays the averaged CPU computational times for each process involved in the ES for the set of images processed. The value in FC is averaged over the number of iterations, i.e. it represents the average time per iteration.

Based on results in Table 3, when an image is identified with sufficient contrast the total CPU time is 0.91s; when the image is identified as with insufficient contrast and in the worst case, t_{max} = 14, the total time for FC is 0.6s. This means that the total time for the full process in this case is 1.62 s. Common agricultural vehicles working in agricultural tasks and requiring green identification, navigate at speeds between 4 and 6 km/h (1.1–1.7 m/s). Thus, assuming that vision systems, on board these vehicles, have the ability to cover areas of 4 meters long without any difficult we have a window between 3.60 and 2.35 s to process the image, which is sufficient according to times above.

3.2.5. Global performance: qualitative analysis
Figs. 9–11(a) display representative original images from the set of available images of maize fields; in (b) appears their histograms. Results and conclusions extracted from this analysis are valid for the full set of images tested. In Fig. 10(a) we can see a limited field area with different weeds densities in the inter-crop rows spaces; alternatively appear low/high densities, which were manually prepared by farmers for testing purposes. Table 4 displays the relevant parameters (mean, m_i, skewness, γ_i, cluster center $v_1 \equiv \{v_{1R}, v_{1G}, v_{1B}\}$ and threshold T) to make a decision about the process to apply, i.e. COM or FC; for each image we identify the camera device with which it has been captured.

From results in Table 4 we can see that according to the decision rule in (3), image in Fig. 9 is identified as belonging to the set of images with sufficient contrast. It should be processed by applying the COM procedure according to Eq. (10). Images in Figs. 10 and 11
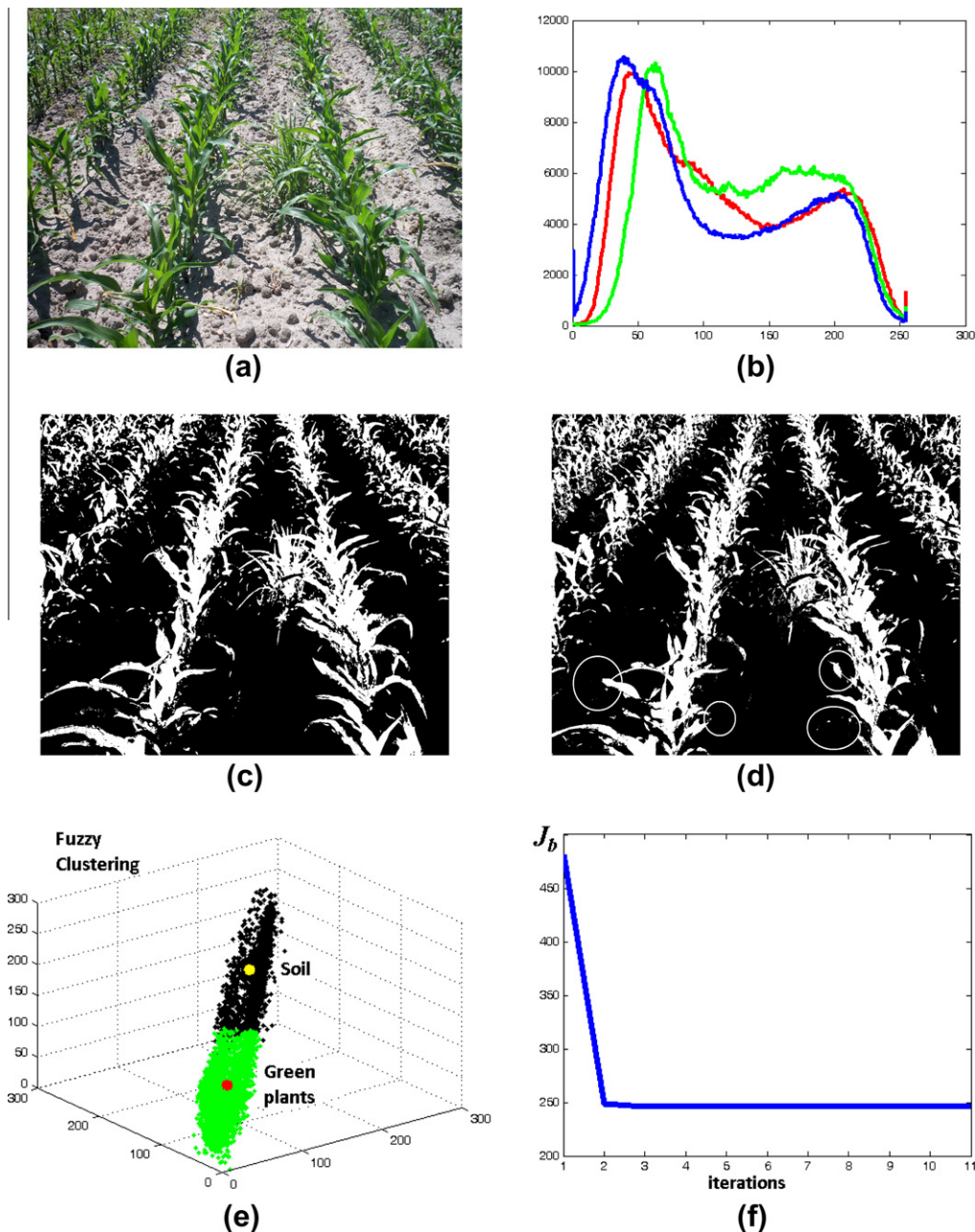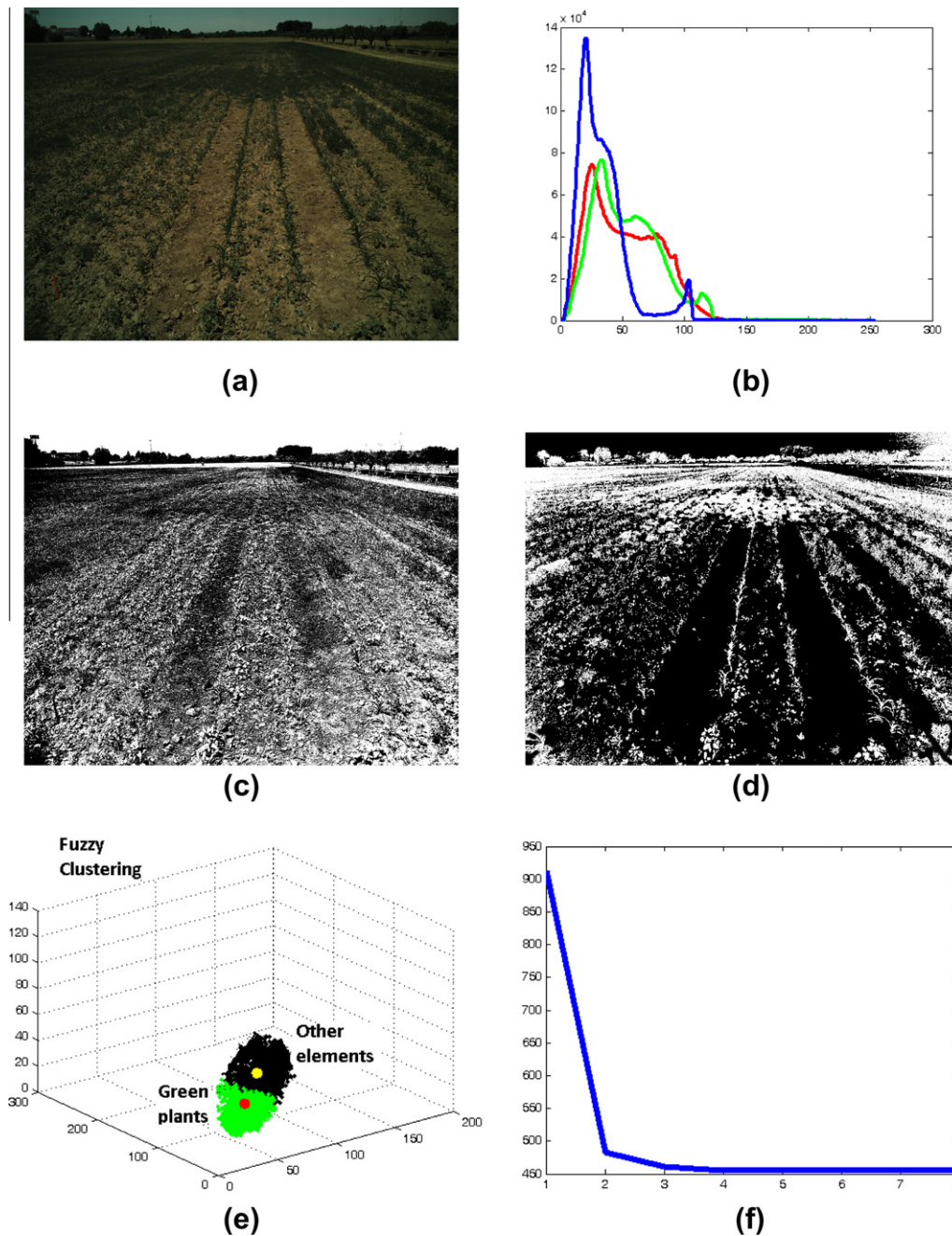
Fig. 9. (a) Original image representative of the set of images with sufficient contrast; (b) its histogram for the three RGB spectral channels; (c) binary image obtained with COM; (d) binary image obtained with FC; (e) distribution of clusters and centers obtained through FC and (f) variation of the criterion function against the number of iterations.

are identified as belonging to the set of images with insufficient contrast; they should be processed with the FC-based procedure. According to the expert criterion the three decisions made for these images are correct and the histograms distributions confirm these decisions. An important aspect to remark is the variability of the threshold T which in images with insufficient contrast varies from 0.34 to 0.52 depending on the type of image. Thus, fixing a value as a universal threshold becomes unfeasible. This makes an important contribution of this paper, because we provide a procedure for adjusting a dynamic threshold adapted to the image quality under processing, measured by its contrast. Additionally, some relevant information can be extracted from the histogram. This is the case for the image in Fig. 11(a) and its histogram in (b), where one can easily infer that this image presents oversaturation. Although the proposed FC based approach provides acceptable

results which allow to work under critical adverse conditions, perhaps the best alternative should be to control the input of illumination either applying iris modification or decreasing the exposure time in the sensor device. This makes an additional contribution.

Only for comparison purpose we have processed the three images with both COM and FC, obtaining respectively the corresponding binary images displayed in Figs. 9–11 labeled with (c) and (d). Also, for the same purpose, we display in Figs. 9–11 labeled with (e) the clustering of the samples with the cluster centers identifying the clusters with green plants and the remainder with other elements (soil, sky). Finally, in figures with labels (f) it is displayed the variation of the criterion function defined in Eq. (10) against the number of iterations. It is important to remark once again that during the ES normal operation images with sufficient contrast are
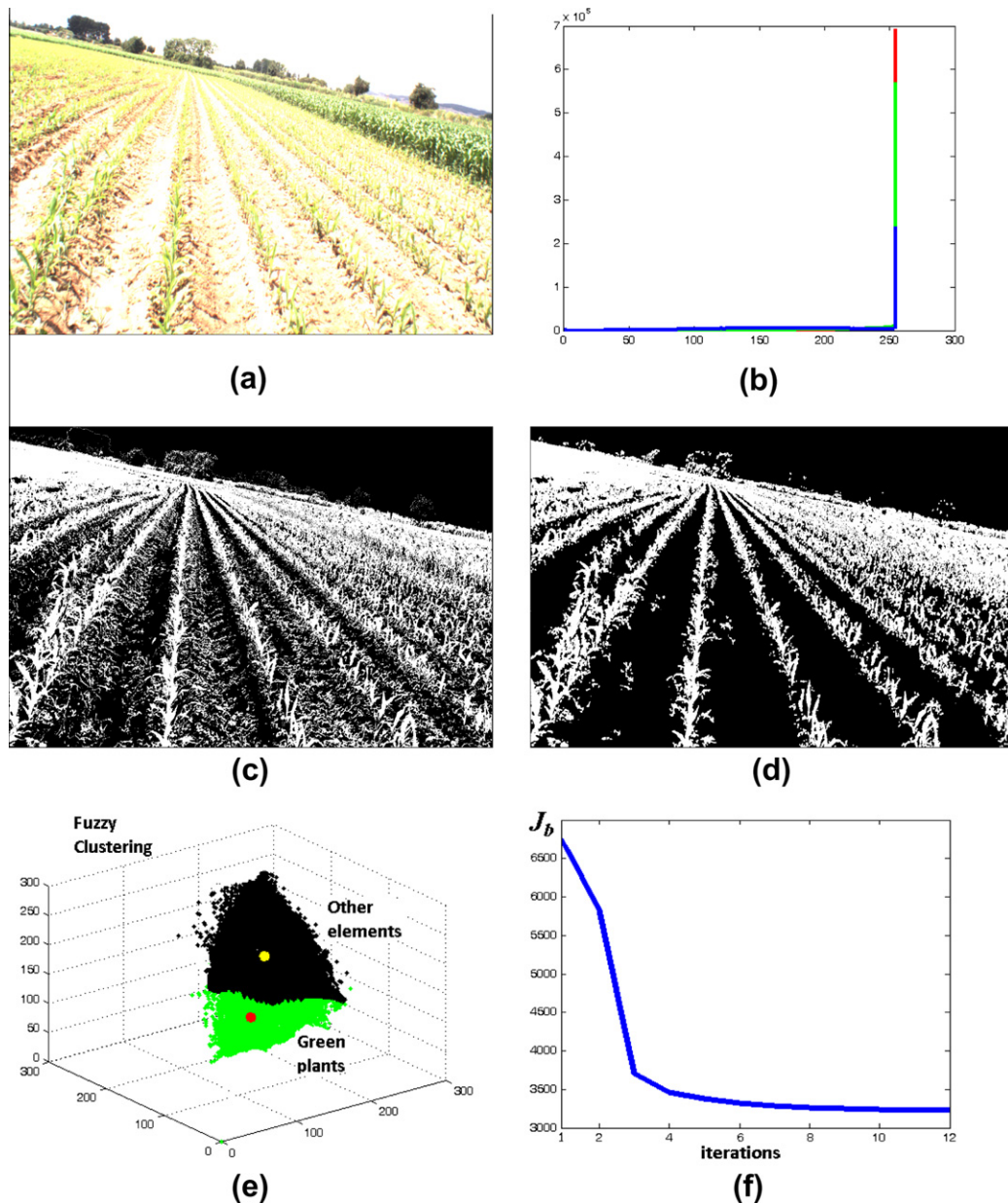
Fig. 10. (a) Original image representative of the set of images with insufficient contrast; (b) its histogram for the three RGB spectral channels; (c) binary image obtained with COM; (d) binary image obtained with FC; (e) distribution of clusters and centers obtained through FC and (f) variation of the criterion function against the number of iterations.

not processed under the FC approach and with insufficient contrast are not processed with COM. *Analysis of binary images*: by comparing the binary images obtained with COM and FC, we can easily verify the outperformance of the appropriate procedure according to the decision made. Indeed, the quality of the binary image in Fig. 9(c) is better than the one in Fig. 9(d) where important deficiencies are identified with circles, i.e. real plants which have not been identified. This means that COM outperforms FC as expected. On the contrary, the quality of binary images in Figs. 10 and 11(d) is clearly superior to the one obtained in binary images in Figs. 10 and 11(c). Now, this implies that FC for images with insufficient quality is the appropriate procedure but not COM. These results confirm in a qualitative way that the proposed ES is appropriate

for identifying greenness in agricultural images, captured even under adverse environmental conditions.

Analysis of FC (clustering and convergence): Figs. 9–11(e) display the clustering of green plants and the remainder elements in the image, basically soil, sky, trees and some building. As we can see in the three figures, samples belonging to each cluster appear grouped together. Perhaps a greater compactness can be seen in images with insufficient contrast, Figs. 10 and 11(e), which is a direct consequence of the histograms and their concentration either left or right. A different measure allows us to verify that in images with insufficient contrast, Figs. 10 and 11(e), the Euclidean distance between cluster centers is less than in images with sufficient contrast, Fig. 9(e), this could serve as an additional measurement to

Fig. 11. (a) Original image representative of the set of images with insufficient contrast; (b) its histogram for the three RGB spectral channels; (c) binary image obtained with COM; (d) binary image obtained with FC; (e) distribution of clusters and centers obtained through FC and (f) variation of the criterion function against the number of iterations.

Table 4
Histogram parameter values (mean, skewness, cluster center v_1 and threshold T) for the images displayed in Figs. 9–11.

	Fig. 9 (Kodak)	Fig. 10 (SVS-VISTEK)	Fig. 11 (Basler)
(m_R, m_G, m_B)	$(116, 124, 105)$	$(52, 53, 33)$	$(236, 230, 187)$
$(\gamma_R, \gamma_G, \gamma_B)$	$(0.32, 0.17, 0.38)$	$(0.68, 0.89, 1.65)$	$(-2.74, -2.13, -1.43)$
$v_1 \equiv \{v_{1R}, v_{1G}, v_{1B}\}$	$(71.2, 91.3, 57.1)$	$(32.0, 34.6, 19.9)$	$(176.8, 155.0, 107.2)$
T	0.42	0.40	0.35

verify that the image has been correctly processed if this distance is sufficiently small. With respect the convergence criterion one can see that the convergence is relatively fast, J_b drops drastically at the first iterations, iteration two at Figs. 9 and 10(f) and three at Fig. 11(f). Although this convergence process is only applied for images classified as with insufficient contrast, we can infer again that the number of iterations is relatively low. This confirms what we had already advanced with regard the maximum number of iterations set to $t_{max} = 14$, which validates the setting of this parameter.

4. Conclusions

We propose a new automatic ES for greenness image identification in agricultural images in maize and barley fields. Its open architecture allows its application to any type of crops in growth states where green plants appear. The proposed system is also applicable to other environments where the main goal is the greenness detection, by example urban areas where trees or lawn densities are of interest.

The proposed ES is based on two main modules, where the first one decides about the image quality by histogram analysis. Depending on this decision, each is image is conveniently processed either by applying a combination of classical greenness strategies or through a module where the FC approach supports the adjusting of a threshold which is dynamic for each image. The proposed strategy has been tested with a broad set of images coming from different camera devices, all oriented toward automatic applications in agricultural images for site specific treatments and captured under very different environmental conditions.

The ES has been designed with an open architecture, so that in the future be possible to replace or add new modules, being of particular interest to study new automatic thresholding methods (Avci & Avci, 2009) or improve the decision rule defined in Eq. (3) by applying a decision-tree inductive learning strategy such as ID3, CART or C4.5 among others (Duda et al., 2000). In this paper we have applied the DWT to the three RGB spectral channels; the down-sampling process could be only applied to the intensity image after the image transformation from the RGB color space to HSI (H-hue, S-saturation, I-intensity) among others.

In addition, once green plants are identified, the remainder parts, mainly the soil, could be analyzed to identify relevant ecologically categories (Luscier, Thompson, Wilson, Gorham, & Dragut, 2006), thus the proposed ES could be extended to deal with the analysis of soil materials with the feedback of samples belonging to this category.

Acknowledgments

References

Avci, E., & Avci, D. (2009). An expert system based on fuzzy entropy for automatic threshold selection in image process. *Expert Systems with Applications, 36*(2), 3077–3085.

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. New York, NY, USA: Plenum Press.

Burgos-Artizzu, X. P., Ribeiro, A., Guijarro, M., & Pajares, G. (2011). Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture, 75*, 337–346.

Burgos-Artizzu, X. P., Ribeiro, A., Tellaeche, A., Pajares, G., & Fernández-Quintanilla, C. (2009). Improving weed pressure assessment using digital images from an experience-based reasoning approach. *Computers and Electronics in Agriculture, 65*, 176–185.

Davies, G., Casady, W., & Massey, R. (1998). Precision agriculture: An introduction. Water Quality Focus Guide (WQ450, Available from: http://extension.missouri.edu/explorepdf/envqual/wq0450.pdf).

Duda, R. O., Hart, P. E., & Stork, D. S. (2000). *Pattern classification*. Wiley.

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

Gebhardt, S., Schellberg, J., Lock, R., & Kaühbauch, W. A. (2006). Identification of broad-leaved dock (*Rumex obtusifolius* L.) on grassland by means of digital image processing. *Precision Agriculture, 7*(3), 165–178.

Gée, Ch., Bossu, J., Jones, G., & Truchetet, F. (2008). Crop/weed discrimination in perspective agronomic images. *Computers and Electronics in Agriculture, 60*, 49–59.

Gonzalez, R. C., & Woods, R. E. (2008). *Digital image processing*. Upper Saddle River, NJ: Pearson Prentice Hall.

Guerrero, J. M., Pajares, G., Montalvo, M., Romeo, J., & Guijarro, M. (2012). Support vector machines for crop/weeds identification in maize fields. *Expert Systems with Applications, 39*, 11149–11155.

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P. J., Burgos-Artizzu, X. P., & Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture, 75*, 75–83.

Hague, T., Tillet, N., & Wheeler, H. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precision Agriculture, 1*(1), 95–113.

Holub, O., & Ferreira, S. T. (2006). Quantitative histogram analysis of images. *Computer Physics Communications, 175*, 620–623.

Kataoka, T., Kaneko, T., Okamoto, H., & Hata, S. (2003). Crop growth estimation system using machine vision. In *The 2003 IEEE/ASME international conference on advanced intelligent mechatronics*.

Kirk, K., Andersen, H. J., Thomsen, A. G., & Jørgensen, J. R. (2009). Estimation of leaf area index in cereal crops using red–green images. *Biosystems Engineering, 104*, 308–317.

Ling, P. P., & Ruzhitsky, V. N. (1996). Machine vision techniques for measuring the canopy of tomato seedling. *Journal Agricultural Engineering Research, 65*(2), 85–95.

Lopez-Granados, F. (2011). Weed detection for site-specific weed management: Mapping and real-time approaches. *Weed Research, 51*, 1–11.

Luscier, J. D., Thompson, W. L., Wilson, J. M., Gorham, B. E., & Dragut, L. D. (2006). Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots. *Frontiers in Ecology and the Environment, 4*(8), 408–413.

Meyer, G. E., & Camargo-Neto, J. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture, 63*, 282–293.

Meyer, G. E., Camargo-Neto, J., Jones, D. D., & Hindman, T. W. (2004). Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Computers and Electronics in Agriculture, 42*, 161–180.

Meyer, G. E., Hindman, T. W., & Lakshmi, K. (1998). *Machine vision detection parameters for plant species identification*. Bellingham, WA: SPIE.

Montalvo, M., Pajares, G., Guerrero, J. M., Romeo, J., Guijarro, M., Ribeiro, A., et al. (2012). Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Systems with Applications, 39*, 11889–11897.

Neto, J. C. (2004). *A combined statistical – soft computing approach for classification and mapping weed species in minimum tillage systems*. Lincoln, NE: University of Nebraska.

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

Otsu, N. (1979). A threshold selection method from gray-level histogram. *IEEE Transactions on System Man and Cybernetics, 9*, 62–66.

Pajares, G., & de la Cruz, J. M. (2004). A wavelet-based image fusion tutorial. *Pattern Recognition, 37*, 1855–1872.

Reid, J. F., & Searcy, S. W. (1987). Vision-based guidance of an agricultural tractor. *IEEE Control Systems, 7*(12), 39–43.

Ribeiro, A., Fernández-Quintanilla, C., Barroso, J., & García-Alegre, M.C. (2005). Development of an image analysis system for estimation of weed. In *Proceedings of the fifth European conference on, precision agriculture (5ECPA)* (pp. 169–174).

Ruiz-Ruiz, G., Gómez-Gil, J., & Navas-Gracia, L. M. (2009). Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *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. *Biosystems Engineering, 89*(2), 119–129.

Tellaeche, A., Burgos-Artizzu, X. P., Pajares, G., & Ribeiro, A. (2008). A vision-based method for weeds identification through the Bayesian decision theory. *Pattern Recognition, 41*, 521–530.

Tellaeche, A., Burgos-Artizzu, X., Pajares, G., Ribeiro, A., & Fernández-Quintanilla, C. (2008). A new vision-based approach to differential spraying in precision agriculture. *Computers and Electronics in Agriculture, 60*(2), 144–155.

The Mathworks. (2012). Available from: http://www.mathworks.com/.

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

Woebbecke, D. M., Meyer, G. E., von Bargen, K., & Mortensen, D. A. (1995). Shape features for identifying young weeds using image analysis. *Transactions of the American Society of Agricultural Engineers, 38*(1), 271–281.

Zheng, L., Shi, D., & Zhang, J. (2010). Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminate. *Pattern Recognition Letters, 31*(9), 920–925.

Zheng, L., Zhang, J., & Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture, 65*, 93–98.

Zimmermann, H. J. (1991). *Fuzzy set theory and its applications*. Norwell, MA, USA: Kluwer Academic.