IMPROVEMENTS OF IMAGE PROCESSING FOR WHEAT EAR COUNTING

Frédéric Cointault [1], Ludovic Journaux[2], Johel Miteran[3], Marie-France Destain[2], Xavier Tizon[4]

[1] ENESAD – UP GAP – 26, Bd Dr Petitjean – BP 87999 – 21079 Dijon cedex, France f.cointault@enesad.fr

[2] Faculté Universitaire des Sciences Agronomiques de Gembloux – Passage des Déportés 2 - B-5030 Gembloux, Belgique journaux.l@fsagx.ac.be ; destain.mf@fsagx.ac.be

[3] Le2i – UFR Sciences et Techniques – Université de Bourgogne – B.P. 47870 – 21078 Dijon Cedex, France miteranj@u-bourgogne.fr

[4] Oncodesign SA- 20 rue Jean Mazen – B.P. 27627 – 21076 Dijon Cedex, France xtizon@oncodesign.com

Abstract

One of the most important activity of agricultural research insititutes concerns the agronomical experiments done under different conditions needing many land observations and valuations to quantify several variables. These observations, although generally accurate, are visually done by the agriculturist technicians and present numerous drawbacks: penibility, weak productivity, numerous labor force, limited sampling … Two feasibility studies lead in our laboratory recently have shown that some of the previous observations, and particularly the counting of the number of wheat ear per m², can be done by color and/or texture image processing for images taken directly in the field with a specific acquisition system. This paper describes the improvements of the previous studies concerning the image acquisition system, and especially the illumination control, and the justification of different hypothesis on the number of classes to detect in an image.

The use of a cluster validity index has allowed to prove that 3 classes to determine all the objects in a wheat ear image are not sufficient. A correlation with a study based on the size of the analysis window is currently under investigation to improve the ear detection, which is now of 6%, compared to manual counting done by agriculturist technicians.

Introduction

Since a lot of decades, the actors of agricultural domain have profited of the advantages of the new technologies. This development allowed improvement of the production management with environmental respect. This vision is reflected in the concept of Precision Agriculture which uses new information and communication technologies (GPS, micro-computing, embeddable electronics …) to obtain a maximum of information concerning the fields, the plants, the soils … The non-destructive methods put in place use remote sensors to characterize the intra-field variability.

Nevertheless, agronomical experiments, which are the most important activity of the agricultural organisms, can be done under different conditions needing many land observations and valuations to quantify several variables such as disease rate, yield components, weed rate … These observations, although generally accurate, are at present visually done by the technicians of research institutes and present numerous drawbacks: penibility, weak productivity of the in-field measurements, numerous labor force, subjectivity, limited sampling …

A lot of work has been done on the detection of weed in a field to propose precision spraying (Gée et al., 2007) or to evaluate a leaf area (Lu et al., 2004), from classical or satellital images. However few works have been done on the wheat ear counting (Germain et al., 1995), which constitutes one of the most important component of the yield. Yield prediction of cereal, especially the wheat, is a great waiting for research institutes, because its manual evaluation takes a lot of time. Its evaluation is a light stake of the current agriculture. Even if it exists yield sensors available directly on a combine harvester, the yield is always determined *a posteriori* and its prediction before the harvest could allow the French cooperatives to better manage their harvest.

Therefore, we proposed to use image processing methods to first count the number of wheat ears per m², allowing the agriculturist technicians to simplify their work.

State of the art

Two feasibility studies lead in our laboratory recently have shown that the number of wheat ears per m² could be determined by image processing for images took directly in the field with a specific acquisition system

Particularly, Cointault et al. (2008a) have shown that the combination of texture and color information allows to detect and count the number of wheat ears per m². During this study, each image took in RGB space is represented in an other specific hybrid space based on *a priori* knowledge on color and texture in the image. After segmentation by classical distance measurement methods, mathematical morphology tools are used to count the ears. First results obtained on few images appeared satisfying and provided ear counting closed to ground truth (about 80% of well recognition) but this technique remains supervised, needs a learning, and provides a non-recurrent hybrid space for image representation due to variations in lighting conditions and wheat growth stages.

Since the color information in the images was not sufficiently significant, we focused our research on textural information. The second feasibility study done (Cointault et al., 2008b) has then first improved the image acquisition system (figure 1) and developed image processing algorithms based on higher order statistical methods, especially the use of Run

Length, for which a mean error of 6% was generally obtained for the countings compared to manual countings. The errors observed are particularly due to the influence of image acquisition conditions, but also to a bad detection (stems and/or leaves seen as ears, number of classes to justify) or to a bad counting due to overlapping or dense clustering.
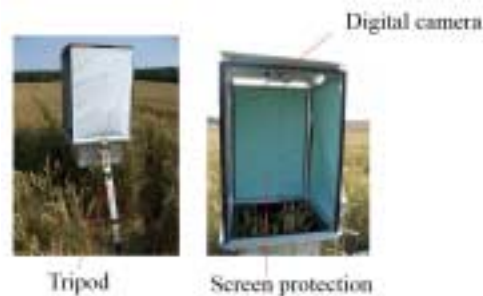

Figure 1. The first simple image acquisition system.

The following paragraphs will describe the modifications on the image acquisition system we done, especially on the illumination control, and will justify the choice of the number of classes we used for learning of the image processing algorithms.

Materials and methods

Image acquisition system

In order to acquire the images under controlled conditions, a specific image acquisition system has been developed based on the previous system. Our choice was to completely control the illumination conditions by firstly protecting the global system and propose a closed system to allow the taking of photographs for all wheat growth stages, and secondly to define the best illumination system.

On the figure 1, the system is not stable because of the tripod and it appeared that the light reflection on the soil implied to protect the whole system of the sunlight illumination from the bottom to the top of the box. To avoid these constraints, a new system has been conceived with two separated frames of 1,10m each, completely opaques. Inside this box, the previous image acquisition system can slide (figures 2 and 3).
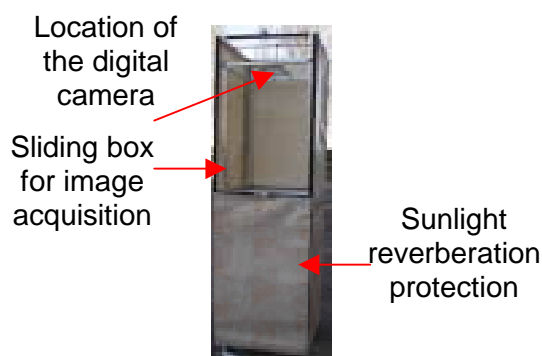

Figure 2. The final global system.   Figure 3. The inside of the system with the sliding box.

The different elements are in aluminium material and the dimensions of this second box have been calculated to be sure we can embed this system on the autoguided mobile platform of the figure 4.
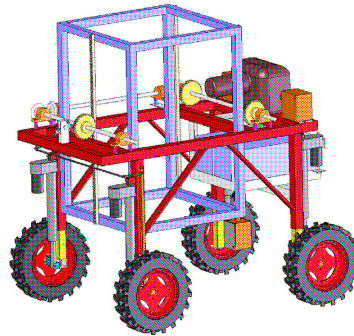

Figure 4. Mobile platform conceived under SolidWorks ®.

The system developed is very simple and will allow us to take photographs at every wheat growth stage for ear countings and disease rate evaluation.

Concerning, the illumination, different solutions have been envisaged: fluorescent lamp; leds; flashes. Illumination with fluorescent lamps is the cheaper system but however these lamps are difficult to handle because they do not support shocks and the lighting intensity cannot vary. The flashes are a bad solution because we risk to overexpose the centre of the image and therefore to induce errors on the different image processing. Finally, the use of Leds appear to be the most appropriate system because the lighting intensity can be easily controlled, and the use of power-Leds (from 3W to 20W) provides sufficient illumination (figure 5).

Figure 5. The Power-leds used (left); their location (middle); one image took with the two Leds (right).

However this system is the most expensive of the three and other tests have to be done to characterize the most appropriate location and number of Leds to be used for our application.

Image processing

The image processing can be decomposed in two steps: detection and counting of the wheat ears. To detect correctly the ears, we have shown previously that a textural study could be more efficient than a color one. In that way, higher order statistical methods based on Run

Length techniques, combined with mathematical morphology tools allowed to evaluate a number of wheat ears per m² closed to those obtained by manual countings done by technicians.

Nevertheless, even if the results are quite good, they are obtained on averages of wheat ear number and no current study has been done to evaluate the quality of our detection. Indeed, some objects detected with our statistical methods can be some stems and/or leaves for which the illumination and/or the location according to the digital camera let appear them as ears. The opposite is also true, so that the mean number of object detected as ears can be close to the reality, with however some errors for the detection.

Before to focus our research on the detection step, it is for us to better justify the hypothesis done on the number of classes contained in an image and on the size of the analysis window. Indeed, in a classical image, we considered, based on purely subjective information, only three classes: ears, stems and leaves, soil. Although this *a priori* observation allowed the unsupervised image processing methods developed easier to implement, to avoid the subjectivity, we proposed an image processing in three steps. We first determine the effective number of classes (corresponding to the number of clusters in multidimensional RGB color space) of several images, through the estimation of a cluster validity index as well proven by Journaux et al. (2006). The second step consists in the extraction of the textural features by the Generalized Fourier Descriptors (GFD). Finally these features are used as input in an unsupervised classification method.

Finding the number of clusters

Despite many attempts, the problem of finding a suitable number of clusters for a given dataset still remains tricky (Maulik and Bandyopadhyay, 2002). We have chosen, in this study, to restrict ourselves to a simple clustering algorithm (K-means), and the choice of K, the number of clusters, is crucial to ensure a satisfying clustering. Following Kim and Ramakrishna (2005), we use two simple measures of cluster validity, that reflects two desirable properties of a clustering result. First, it should produce compact clusters, well grouped around their centroid. One possible way to measure this notion of intra-cluster validity criterion is to compute the average point to centroid distance:

$$M_{\text{intra}} = \max_k \sum_{x_i \in C_k} \|x_i - z_k\| \qquad (1)$$

Where z_k is the centroid of cluster C_k. M_{intra} should be as small as possible.

Second, it should produce separable clusters, where data points belonging to different clusters are as far away as possible, as indicated by the inimum distance between inter-clusters data points:

$$M_{\text{inter}} = \min_{\substack{i,\,j \in \langle 1,K \rangle \\ i \neq j}} \|z_i - z_j\| \qquad (2)$$

5

M_{inter} should be large. The overall cluster validity index (CVI) is then estimated by $val = M_{int\,ra}/M_{int\,er}$, and should be minimal for better clusterings. To find the optimal number of clusters, we run the K-means algorithm for increasing values of K. Several runs of the K-means algorithm are executed for a fixed value of K, to avoid the tendency of K-means of getting stuck in local minima. We then plot the obtained validity index against K. Theoretically, this plot should exhibit a minimum around the optimal value of K (figure 6).
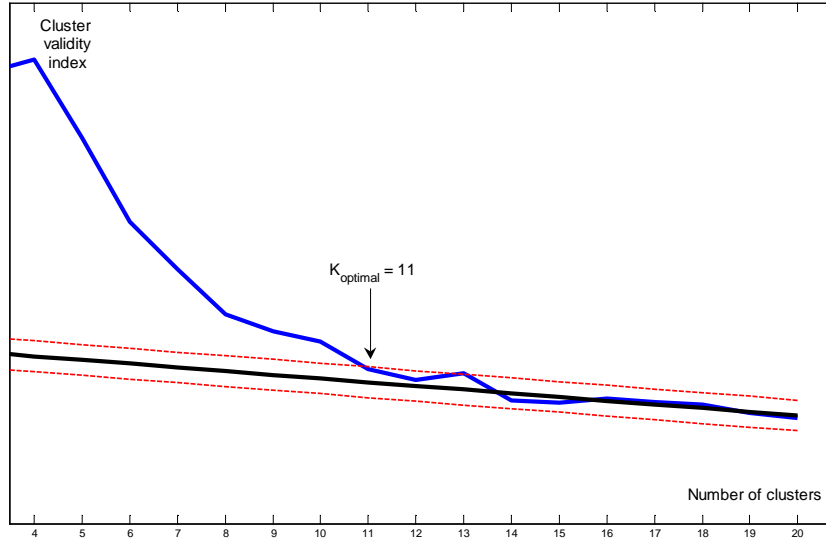

Figure 6. Schematic representation of the method used to find the knee of the CVI curve.

However, it only shows a "knee", that is a reduction of the decreasing tendency of the curve. To identify the value of K corresponding to the knee, we proceed as follows:

1. Robustly fit a line to the five last data points, to identify the second part of the L-curve.

2. Find the minimum value of K for which the previous fit is not satisfactory, that is the point where the curve falls outside a confidence interval around the fitted line. This interval is defined as plus or minus five times the mean difference between the curve and the fitted line.

Generalized Fourier Descriptors

From a convolution mask, we extract the textural parameters by using the Generalized Fourier Descriptors (GFD) (Smach et al., 2007) obtained for each R, G, B band. The length of the Fourier vector is dependant on the size of the analysis window. For each pixel a DF vector is obtained on the three bands (the size of a vector is equal to half of the analysis window size). The GFD are defined as follows. Let f be a square summable function on the plane, and $\hat{f}$ its Fourier transform:

$$\hat{f}(\xi) = \int_{\mathbb{R}^2} f(x) \exp\left(-j\langle x | \xi \rangle\right) dx. \qquad (3)$$

Where $\langle .|. \rangle$ is the scalar product in $\mathbb{R}^2$

If (λ, θ) are polar coordinates of the point ξ, we shall denote again $\hat{f}(\lambda, \theta)$ the Fourier transform of f at the point (λ, θ). Gauthier et *al.* (1991) defined the mapping D_f from $\mathbb{R}_+$ into $\mathbb{R}_+$ by:

$$D_f(\lambda) = \int_0^{2\pi} \left| \hat{f}(\lambda, \theta) \right|^2 d\theta. \qquad (4)$$

So, D_f is the feature vector which describes each texture image and will be used as an input of the unsupervised classification method. Motion descriptors, calculated according to equation (2), have several properties useful for object recognition: they are translation, rotation and reflexion-invariant.

K-means clustering

The K-means algorithm (Duda et al., 2001) is among the most popular and cost-effective clustering techniques. It finds the clustering result that minimizes the sum of squared Euclidean distances between data points and cluster centroids. We choose this unsupervised method for its efficiency and simplicity.

Results and discussion

To evaluate automatically the number of classes to extract textural parameters, twelve images took just before the harvest have been tested (figure 7).
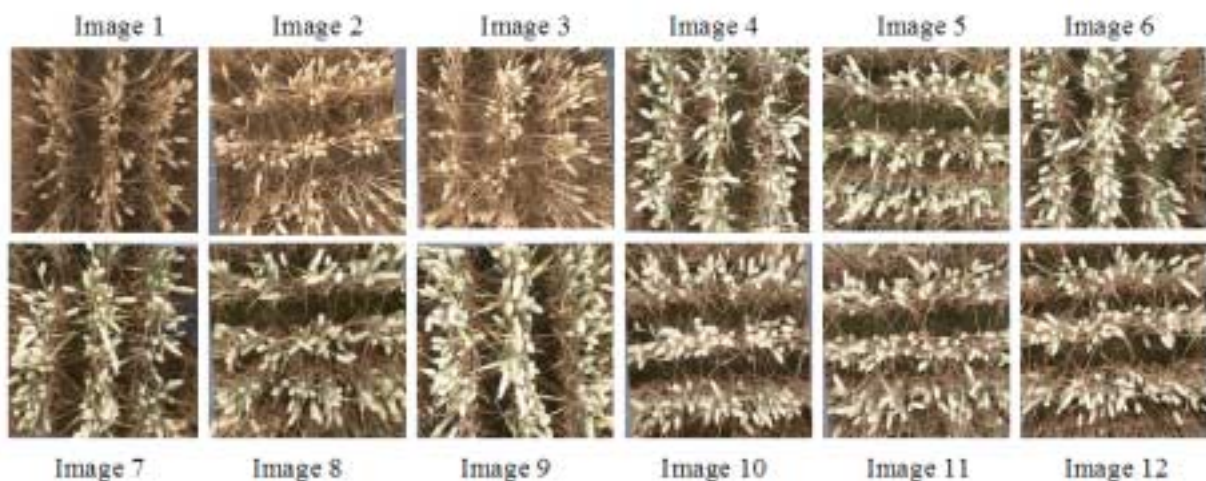

Figure 7. The 12 images tested in terms of number of classes.

The results on the number of classes determined with the previous clustering index are given in table 1. It must be noticed that the number of classes includes a class related to the side of the image.

Table 1. Results on the number of classes for each previous images.

Image	1	2	3	4	5	6	7	8	9	10	11	12
Number of classes	10	12	12	11	10	9	13	12	11	13	12	9

These first results show that the mean number of classes in such this kind of wheat image is of 11, which seem to prove that our choice of 3 classes does not reflect the reality and is therefore not well adapted.

The figure 8 gives also for the first previous image the results on image classification for different class numbers for a fixed analysis window size, arbitrarily took as 9x9 pixels.
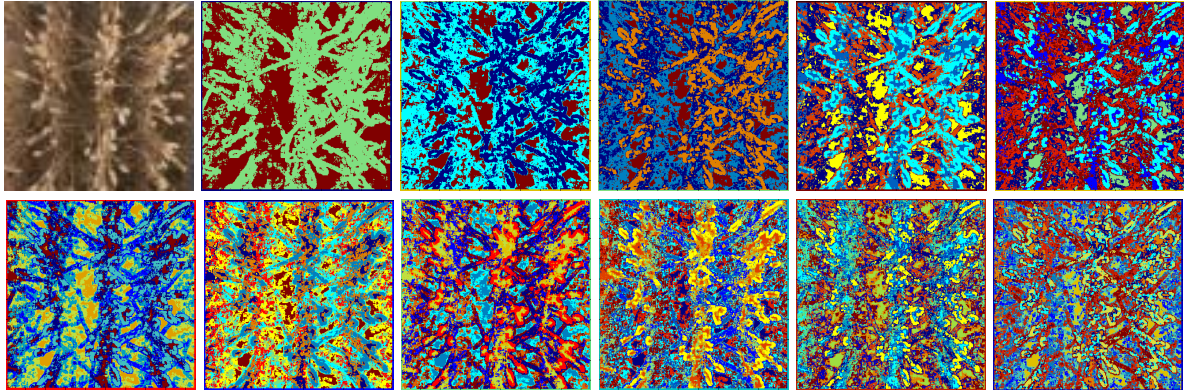

Figure 8. From left to right: part of an original image; results of detection with k=2 to k=12 with a 9x9 analysis window size.

Although several other tests must be conducted, especially by taking into account the illumination conditions and the wheat growth stage, some conclusions can be done. With a number of classes lower than 4, it seems to be quite difficult to detect correctly the different objects into the scene. If the number of classes is included between 4 and 9, the classification appears to give better results with not a lot of over-segmentation areas. These last problems are devoted to a number of classes greater than 9.

Our conclusions are that a number of classes equals to 3 is not well-adapted to our application because of illumination and wheat growth stage problems. Indeed, if the lighting conditions are not controlled, some stems and/or leaves can appear as ears and be detected as them. To be sure to avoid these problems of detection, a solution could be to use a first primary segmentation to eliminate the soil and the waste, before to refine the detection with a more precise classification.

To complete the first previous results, we are currently analysing each of the class previously extracted to gather together all the classes characterizing the ears. This study will be combined with information on acquisition conditions (illumination, variety …). This will also constitute a database of characteristics on the ears for each of the wheat growth stages.

In parallel, as part of the texture analysis, we are studying the influence of the size of the analysis window, the dimension of which needs to be adapted to the size of the image and/or to the representative resolution of the ears contained in the image. The main goal is then to define an optimal size of the window allowing a good detection. Some tests have been done

for a window size included between 7x7 and 19x19 pixels, and k parameter variations from 2 to 12, the class characterizing the side effect being not took into account. Some results are presented in figure 9, for which we varied the values of the k parameter (k-means) from 8 to 10, with analysis windows of sizes between 9x9 and 13x13 pixels.
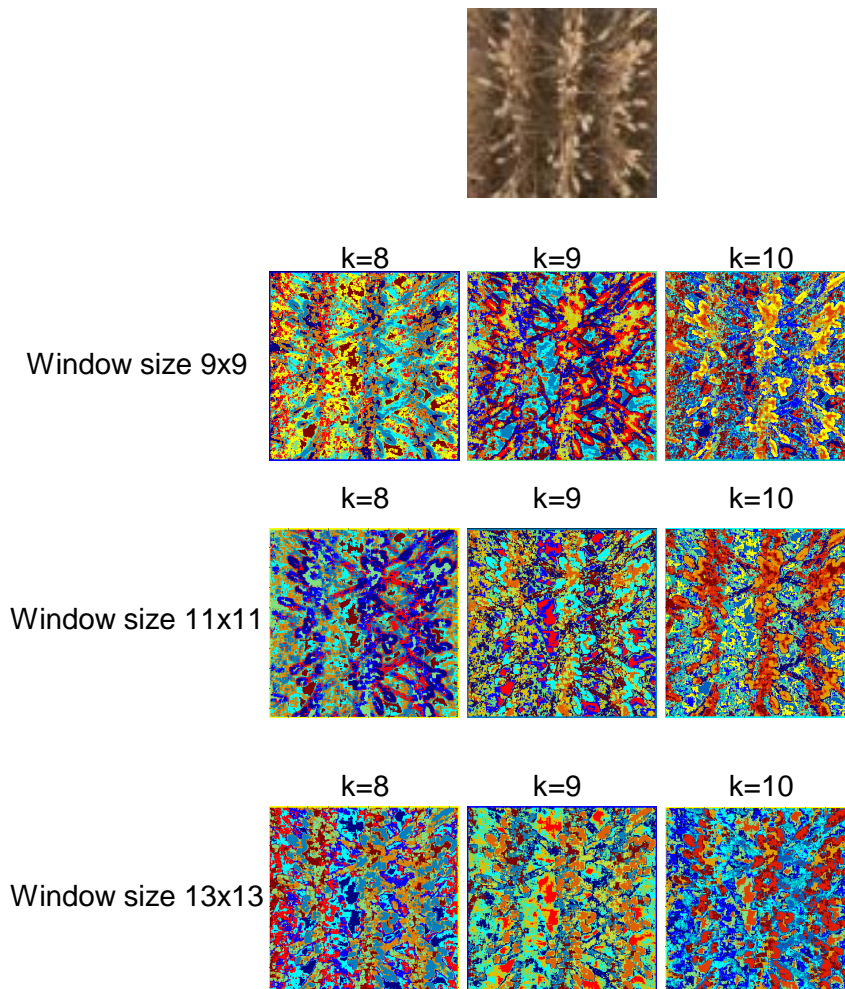

Figure 9. Examples of segmented images with 9x9, 11x11 and 13x13 pixel window sizes.

The results obtained from the segmented images show that an analysis window with a size between 9x9 and 13x13, and a mean k value would allow to provide an appropriate preliminary wheat ear detection for our application.

Conclusion

In this paper we proposed improvements of a feasibility study based on the use of textural image processing, and especially Run Length technique, to detect and count the number of wheat ear per m² to simplify manual countings. Results on the repercussion of these improvements on the results will be provided in the conference.

First of all, the image acquisition system and the control of the lighting conditions have been studied and modified, before to justify the different hypothesis done for the image processing methods, especially the number of classes for k-means algorithm and the size of the analysis window. The assumption of three classes, naturally correlated to the three different objects of

a wheat ear image (stems-leaves, soil, ear) is not so evident. The use of a cluster validity index has shown that the real number of classes is included between 4 and 9, due to illumination problems and wheat growth stages.

A parallel study on the determination of the best value for the analysis window size has also been done and proved that a window size between 9x9 and 13x13 pixels appears to give the best results, even if several tests must again been conducted. Continuation will be done to propose a more global solution.

Improvements of the whole study is now envisaged at two fondamental levels. The first one concerns the acquisition system which constitutes the essential point. A study is currently investigated concerning the use of stereoscopic processes and more efficient controlled illumination systems. Moreover, we envisage the use of multispectral images which will certainly offer more accurate information than the only use of color component.

The second one deals with the image processing decomposed into two stages: 1) the determination of efficient parameters to extract the objects, which is currently studying with approaches based on spatio-frequential analyses of textures (combination of Markov field and Gabor filters, Karhunen-Loeve); 2) the implementation of pattern recognition methods adapted to our application and the comparison of different supervised and non-supervised classification algorithms. Particularly, some works are at present focused on the use of the efficient and well-known SVM (Support Vector Machine) approach which appears to be very useful currently (Vapnik, 1998).

Aknowledgments

References

Cointault F., Guérin D., Guillemin J.P. and Chopinet B. 2008a. "In-Field Wheat ears Counting Using Color-Texture Image Analysis", *NZJCH (New Zealand Journal of Crop and Horticultural Sciences),* **36**(1) (*to be published*).

Cointault F., Journaux L., Destain M.F. and Gouton P. 2008b. "Wheat ear detection by textural analysis for improving the manual countings", *Proceedings of the 5th IASTED International Conference on Signal Processing, Pattern Recognition and Applications (SPPRA)*, 13-15 February 2008, Innsbrück, Austria.

Duda R.O., Hart P.E. and Stork D.G. 2001. *Pattern Classification (2nd Edition)*, ed. W. Interscience. 654p.

Gauthier J.P., G. Bornard and Silbermann M. 1991. "Harmonic analysis on motion groups and their homogeneous spaces*", IEEE Transactions on Systems, Man and Cybernetics*, **21**(1): pp. 159-172.

Gée Ch., Bossu J., Jones G. and Truchetet F. 2008. "Crop/Weed discrimination in perspective agronomic images", *Computers and Electronics in Agriculture*, Vol 60/1, p.49-59.

Germain C., Rousseau R. and Grenier G. 1995. "Non destructive counting of wheatears with picture analysis", *Image Processing and its application*, pp. 453-439.

Journaux L., Tizon X., Foucherot I. and Gouton P. 2006. "Dimensionality Reduction Techniques: An Operational Comparison On Multispectral Satellite Images Using Unsupervised Clustering", *Proceeding of the IEEE 7th Nordic Signal Processing Symposium (NORSIG)*, Reykjavik, Iceland, pp. 242-245.

Lu H.Y., Lu C.T., Wei M.L. and Chan L.F. 2004. "Comparison of different models for nondestructive leaf area estimation in taro", *Agronomy journal*, vol. 96, n°2, pp. 448-453.

Kim M. and Ramakrishna R.S. 2005. "New indices for cluster validity assessment", *Pattern Recognition Letters*, vol. 26, pp. 2353-2363.

Maulik U. and Bandyopadhyay S. 2002. "Performance evaluation of some clustering algorithms and validity indices", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 1650-1654.

Smach F., Lemaître C., Gauthier J.P., Miteran J. and Atri M. 2007. "Generalized Fourier Descriptors with Applications to Objects Recognition in SVM Context", *Journal of Mathematical Imaging and Vision (JMIV)*, Springer.

Vapnik V.N. 1998. "The statistical Learning Theory", Springer.