# A Comparison of Fuzzy Clustering Algorithms Applied to Feature Extraction on Vineyard

Christian Correa, Constantino Valero, Pilar Barreiro, Maria P. Diago, and Javier Tardáguila

E.T.S.I. Agrónomos. Departamento de Ingeniería Rural
Universidad Politécnica de Madrid. LPF-TAGRALIA.
Av. Complutense s/n Madrid, Spain
c.correaf@alumnos.upm.es,{constantino.valero,pilar.barreiro}@upm.es,
{maria-paz.diago,javier.tardaguila}@unirioja.es
http://www.etsia.upm.es

**Abstract.** Image segmentation is a process by which an image is partitioned into regions with similar features. Many approaches have been proposed for color image segmentation, but Fuzzy C-Means has been widely used, because it has a good performance in a large class of images. However, it is not adequate for noisy images and it also takes more time for execution as compared to other method as K-means. For this reason, several methods have been proposed to improve these weaknesses. Method like Possibilistic C-Means, Fuzzy Possibilistic C-Means, Robust Fuzzy Possibilistic C-Means and Fuzzy C-Means with Gustafson-Kessel algorithm. In this paper we perform a comparison of these clustering algorithms applied to feature extraction on vineyard images. Segmented images are evaluated using several quality parameters such as the rate of correctly classified area and runtime.

**Keywords:** FCM clustering, image segmentation

## 1 Introduction

Several segmentation techniques have been developed to object identification in industrial applications, (e.g. quality control in the size or volume of mechanical parts) and agricultural applications like, identification of defects and size of various horticultural products [2], [4]. For the identification of a product these techniques are based mainly on the identification by color and shape. However most of the applications are performed under controlled conditions (structured environment) of lighting, speed and distance to the product, making the algorithms valid only under those conditions. Thus, classical techniques are not applicable for unstructured environments [12], so recently, techniques from the area of artificial intelligence are being tested to increase the degree of generalization to identify objects [15].

Color classification techniques can be separated into supervised and unsupervised [14]. Supervised methods are those where the user specifies the number of classes and how the prototype of these classes will be. For unsupervised methods, the prototypes of the classes are not known a priori, but the idea behind the method is that the elements in each class will exhibit similar characteristics. In unstructured environments, the conditions are variable, so establishing a priori which features will correspond to the elements of a given class leads to a bias that limits the possible solutions, because the characteristics imposed are valid only for particular situations [9].

In the real world most of the time the distribution of classes is unknown, but if known, it is difficult to extract objects from each class to prepare the training set [7]. Therefore unsupervised classification techniques are of special interest in agricultural applications, since they make data groups (pixels in the case of images) without a pre-determined criteria by minimizing the distance between pixels within each group. So, after making groups, just remains to identify which data has been grouped in each class.

In this paper we propose the use fuzzy based classification techniques to images in the visible spectra for the identification of grape and leaves. The identification is performed in order to measure the area and location of grapes and leaves. These measures are generally used to predict yield, to asses the effect of early defoliation and for real-time spraying systems. We will determine which of these algorithms is best suited for real-time applications (e.g. spraying systems) in term of runtime.

## 2 Methodology

In order to evaluate the performance of the segmentation methods, from a set of 200 vineyard images a representative subset of 20 images (resolution of 640x480 and 320x240) was manually segmented.

As a first step, the subset was transformed to HSV, HSI, CMYK, L*a*b*, XYZ and Ohta color spaces [3]. This step is performed to determine which color space is most suitable for vineyard image segmentation.

As a second step, the following fuzzy clustering algorithms were applied to this subset: Fuzzy C-Means (FCM), Possibilistic C-Means (PCM), Fuzzy Possibilistic C-Means (FPCM), Robust Fuzzy Possibilistic C-Means (RFCM) and Fuzzy C-Means with Gustafson-Kessel algorithm (FCM-GK). The algorithms were selected due to their theoretical advantages as well as to asses the drawbacks when applied under field conditions. For this comparison 8 clusters was used as reference.

As a third step, these algorithms were applied over the subset modified with 5% of salt and pepper noise. This modification is performed in order to simulate images acquired in field conditions, like lower resolution camera, vibrations and dust.

As fourth step the algorithms was benchmarked by speed and accuracy against the manual segmentation.

### 2.1 Clustering Theoretical Background

**Fuzzy C-Means (FMC).** The FCM algorithm assigns membership values, which are inversely related to the relative distance of a point to the prototypes (cluster centers in the FCM model) [5]. In FCM, the closeness of each data, $x_k$, to the center a cluster, $v_i$, is defined as the membership $(u_{ki})$ of $x_k$ to the i cluster of X minimizing the following objective function:

$$J_m\left(U,V\right)=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}^{m}\|x_k-v_i\|^2 \tag{1}$$

where $\mathrm{X}_k$ a given set of unlabeled N data; $V_i$ are the cluster centers and $m = [1, \infty]$ is the weighting exponent which determines the fuzziness of the resulting clusters, U = $[\mu_{ik}]$ matrix $c \times n$, where $u_{ik}$ is membership of $x_k$ to the i cluster $\sum_{i=1}^{c}\mu_{ik}=1,\ \ \forall\ k=1,2,\ldots,n;$. The cluster centers and the memberships are computed as:

$$V_i=\sum_{k=1}^{n}\mu_{ik}^{m}X_k/\sum_{k=1}^{n}\mu_{ik}^{m} \tag{2}$$

$$u_{ik}=1/\sum_{j=1}^{c}\left(\frac{\|x_k-v_i\|}{\|x_k-v_j\|}\right)^{2/(m-1)} \tag{3}$$

**Possibilistic C-Means Algorithm (PCM).** The PCM algorithm considers the clustering problem from the viewpoint of possibility theory [10]. The approach adopted in PCM differs from the FCM algorithm because the resulting membership values can be interpreted as degrees of possibility (or compatibility) of the points belonging to the classes. The FPCM algorithm simultaneously produces both membership and typicality values. Outliers have low typicality values and automatically eliminated by the algorithm. The objective function for PCM is:

$$P_m\left(T,V;X,\gamma\right)=\sum_{i=1}^{n}\sum_{k=1}^{c}t_{ik}^{m}d_{ki}^{2}+\sum_{i=1}^{c}\gamma_i\sum_{k=1}^{n}(1-t_{ki})^m \tag{4}$$

where $t_{ki}$ is the typicality of $x_k$ to the cluster $i, v_i, T$ is the typicality matrix, defined as $T=[t_{ki}]_{NC}$, $d_{ki}$ is a distance measure between $x_k$ and $c_i$, and $\gamma_i$ denotes a user-defined constant: $\gamma_i > 0$, $1 < i < c$. By using an approximate optimization (AO) of $P_m$, PCM-AO algorithm, additional conditions are necessary for the solution of (4), $1 \le i \le c$, $1 \le k \le N$, as:

$$t_{ki}=1/\left(1+\frac{d_{ik}}{\gamma_i}\right)^{1/m-1},\forall i,k \tag{5}$$

$$v_i=\frac{\sum_{k=1}^{n}t_{ki}^{m}x_k}{\sum_{k=1}^{n}t_{ki}^{m}},\forall i \tag{6}$$

PCM algorithm solves (4) with (6) and adds the next condition on $\{\gamma_i\}$:

$$\gamma_i = K\frac{\sum_{k=1}^{n} u_{ki}^{m} d_{ki}^{2}}{\sum_{k=1}^{n} u_{ki}^{m}}, K > 0 \tag{7}$$

where $u_{ki}$ are membership values obtained in FCM and $K$=1 is mostly used.

**Fuzzy Possibilistic C-Means.** Using the same notation as in FCM, $\mu_{ik}$ is the membership value of the data point $x_k$ in cluster $i$, computed by the Ec. 3, while $t_{ik}$ is the typicality value of $x_k$ in cluster $i$. The objective of FPCM model is to find the partition of X into $c$ fuzzy subset by minimizing the Ec. (8) [11]:

$$J_{m,\eta}\left(U,T,V\right) = \sum_{i=1}^{c}\sum_{k=1}^{n}(\mu_{ik}^{m} + t_{ik}^{\eta})\|x_k - v_i\|^2 \tag{8}$$

subject to the constraints $m > 1$, $\eta > 1$, $0 \leq u_{ik}, t_{ik} \leq 1$, $\sum_{i=1}^{c} \mu_{ik} = 1, \forall k$ and $\sum_{k=1}^{n} t_{ik} = 1, \forall i$. Where $m$ and $\eta$ are both weighting exponents.

Under the constraints above and conditions established on c-means optimization problems, we will have the first order necessary conditions for extreme of $J_{m,\eta}(U,T,V)$ in terms of Lagrange multiplier theorem as follows.

$$t_{ik} = 1/\sum_{j=1}^{n}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/m-1}, \forall i, k \tag{9}$$

$$v_i = \frac{\sum_{k=1}^{n}\left(\mu_{ik}^{m} + t_{ik}^{\eta}\right)x_k}{\sum_{k=1}^{n}\left(\mu_{ik}^{m} + t_{ik}^{\eta}\right)}, \forall i \tag{10}$$

where $d_{ik}$ is the distance of the data point $x_k$ to the prototype $v_i$, computed as:

$$d_{ik} = \|x_k - v_i\| = (x_k - v_i)^T A(x_k - v_i) \tag{11}$$

where A is symmetric positive definite matrix. When A is identity matrix, $d_{ik}$ represents Euclidean distance which represents the similarity between data points and cluster center.

**Robust Fuzzy-Possibilistic C-Means.** FPCM algorithm, deals with the noise sensitivity of FCM algorithm and by keeping the constraint that the memberships of a data across classes sum to one, it could solve the coincident clusters problem of PCM. However FPCM algorithm still uses a norm-induced distance, so it does not have enough robustness. In order to improve the performance on the noisy data of the traditional FPCM algorithm, we now consider a kind of kernel-induce distance. And on the basis of it, [16] propose a new RFPCM algorithm.

Suppose that the data point $x_k \in R^s, k = 1, 2, \ldots, n$, is transformed from the original space to a feature space H by a nonlinear mapping $\Phi$, it becomes the following form $\Phi(x_1), \Phi(x_2), \ldots, \Phi(x_n)$. So the inner product in the original space could be expressed by the Mercer kernel [6] as

$$K(x_k, x_j) = (\Phi(x_k) \cdot \Phi(x_j)) \tag{12}$$

the Euclidean distance in the feature space could be denoted as follows:

$$d_H(x, y) = \sqrt{\|\Phi(x) - \Phi(y)\|^2} = \sqrt{\Phi(x) \cdot \Phi(x) - 2\Phi(x) \cdot \Phi(y) + \Phi(y) \cdot \Phi(y)} \tag{13}$$

So the objective function of the FPCM algorithm could be modified as [13]:

$$J_{m,\eta}(U, T, V) = 2 \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik}^{m} + t_{ik}^{\eta}) \left(\|\Phi(x_k) - \Phi(v_i)\|^2\right) \tag{14}$$

subject to $m > 1$, $\eta > 1$, $0 \leq u_{ik}, t_{ik} \leq 1$, $\sum_{i=1}^{c} \mu_{ik} = 1, \forall k$, $\sum_{k=1}^{n} t_{ik} = 1, \forall i$, where $m$ and $\eta$ are both weighting exponents.
Combining (11) and (12), one can obtain:

$$\|\Phi(x_k) - \Phi(v_i)\|^2 = K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i) \tag{15}$$

Using a Gaussian function $K(x, y) = e^{-\|x-y\|^2/\sigma^2}$, as kernel, $K(x_k, x_k) = 1$ and $K(v_k, v_k) = 1$. Therefore, Ec. (14) can be transformed into the following form through this kernelization.

$$J_{m,\eta}(U, T, V) = 2 \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik}^{m} + t_{ik}^{\eta})(1 - K(x_k, v_i)) \tag{16}$$

Under the same conditions of the FPCM algorithm, we will have the first order necessary conditions for extrema of $J_{m,\eta}(U, T, V)$ in terms of Lagrange multiplier theorem as follows.

$$u_{ik} = 1 / \sum_{j=1}^{c} \left( \frac{1 - K(x_k, v_i)}{1 - K(x_k, v_j)} \right)^{1/m-1}, \forall i, k \tag{17}$$

$$t_{ik} = 1 / \sum_{j=1}^{n} \left( \frac{1 - K(x_k, v_i)}{1 - K(x_j, v_i)} \right)^{1/m-1}, \forall i, k \tag{18}$$

$$v_i = \frac{\sum_{k=1}^{n} (\mu_{ik}^{m} + t_{ik}^{\eta})\, K(x_k, v_i) x_k}{\sum_{k=1}^{n} (\mu_{ik}^{m} + t_{ik}^{\eta})\, K(x_k, v_i)}, \forall i \tag{19}$$

For our analysis we use, $\sigma$ as the maximum $\sigma$ between $a^*$ and $b^*$ channels. Parameters $m = 2$ and $\eta = 2$ as [13] propose were used.

**Gustafson-Kessel Clustering** The main feature of the Gustafson-Kessel (GK) algorithm is the local adaptation of the distance metric to the shape of the cluster by estimating the cluster covariance matrix and adapting the distance-inducing matrix correspondingly [8]. The FCM-GK algorithm is based on iterative optimization of an objective functional of the c-means type:

$$J_m\left(U, V, \{A_i\}\right) = \sum_{i=1}^{c}\sum_{k=1}^{n} \mu_{ik}^{m} d_{ikA_i}^{2} \tag{20}$$

The distance norm $d_{ikA_i}$ as in the Ec. (11).

The metric of each cluster is defined by a local norm-inducing matrix $A_i$, which is used as an optimization variables in the functional. This allows the distance norm to adapt to the local topological structure of the data. The minimization of the GK objective functional is achieved by using the alternating optimization (AO) method according to the well-known algorithm [1].
For $l = 1, 2, \cdots s$ where $l$ represent every step in the iteration process and $s$ is the maximum number of iteration allowed (stop criteria).

$$v_i^{(l)} = \frac{\sum_{k=1}^{n} (u_{ki}^{(l-1)})^m x_k}{\sum_{k=1}^{n} (u_{ki}^{(l-1)})^m}, 1 \leq i \leq k \tag{21}$$

$$F_i = \frac{\sum_{k=1}^{n} (u_{ik}^{(l-1)})^m (x_k - v_i^{(l)})(x_k - v_i^{(l)})^T}{\sum_{k=1}^{n} (u_{ik}^{(l-1)})^m}, 1 \leq i \leq k \tag{22}$$

$$d_{ikA_i}^2 = (x_k - v_i^{(l)})^T \left[\rho_i det(F_i^{1/n} F_i^{-1})\right] (x_k - v_i^{(l)}), 1 \leq i \leq k \tag{23}$$

for $1 \leq k \leq N$

if $d_{ikA_i} > 0$ for $1 \leq i \leq K$,

$$u_{ik}^{(l)} = \sum_{j=1}^{k} \left(\frac{d_{ikA_i}}{d_{jkA_i}}\right)^{-2/(m-1)} \tag{24}$$

otherwise

$u_{ik}^{(l)} = 0$, if $d_{ikA_i} > 0$, and $u_{ik}^{(l)} \in [0, 1]$ with $\sum_{i=1}^{k} u_{ik}^{(l)} = 1$ otherwise.
Until $\|U^{(l)} - U^{(l-1)}\| < \epsilon$.

This algorithm was implemented using the improved covariance estimation proposed by [1]. In this analysis $m = 2$ and $\epsilon = 0.0001$ were used as [1] propose.

## 3 Results

Field experiments were conducted on a vineyard ( *V. vinifera* L. cv. Tempranillo) located in Ollauri, La Rioja (Spain) in October 2010. The canopy was photographed using a digital camera mounted on a tripod set normal to the canopy at 2 m from row axis and 1.05 m aboveground. A white screen was placed behind the canopy to avoid confounding effects from background vegetation. Images were captured at a resolution of 3504x2336.

In order to evaluate the performance of the segmentation techniques, from a set of 20 images, areas of leaves and grapes were manually segmented and corresponding area computed. These areas was compared with the areas provided by the five methods using the Ec. (25) and results are shows in Table 1.

$$Accuracy = \sum_{k=1}^{n} \frac{Correct\ segmented\ Area\ in\ Class\ k}{Total\ Area\ in\ Class\ k}, \forall\ k \tag{25}$$

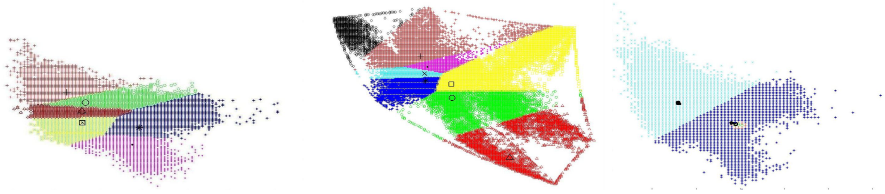
The color space L*a*b* has shown to be the best performing color space, more

**Table 1.** Accuracy of clustering techniques. Performance for $m = 2$.

| Method | FCM | PCM | FPCM | RFPCM | FCM-GK |
|---|---|---|---|---|---|
| Accuracy without noise | 90% | 2% | 88% | 5% | 91% |
| Accuracy with noise | 87% | 3% | 85% | 5% | 88% |

precisely the chroma channels a*b*. These channels allows the best clustering due to the elimination of the effects of the darkness or lighting in the images.

Regarding to the cluster number, Figure 1 shows some examples of cluster generated over the $a^*b^*$ space. Note that in the Fig. 1 only FCM and FCM-GK generates eight clusters, while PCM generates two and FPCM six, over the eight requested, as a consequence of coincident clusters.

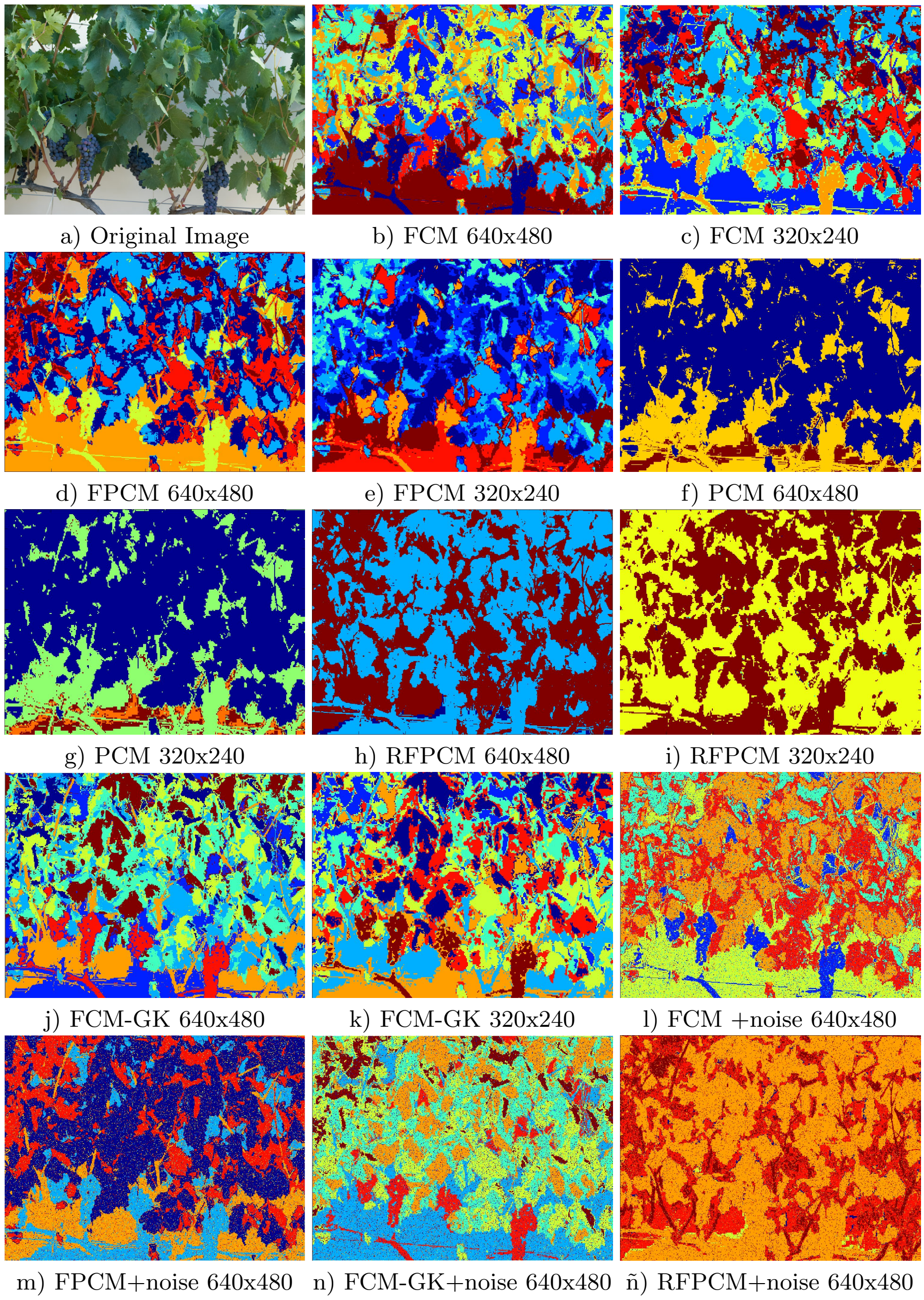


a) FPCM 640x480 b) FCM-GK+noise 640x480 c) PCM 320x240

**Fig. 1.** Example of some graphics outputs of the clustering segmentation using 8 classes over the plane a*b* and the cluster centers generated.

Referent to the noise effect, because the original images are not noisy the performance of the algorithms oriented to noise reduction do not show a better performance as compared with the normal FCM. Therefore we applied artificially a 5% of salt and pepper noise effect in order to simulate the field conditions. The effects over the cluster centers are depicted in Fig. 1 b) and the noise effect over the segmented images is shown in Fig. 2 l) to ñ).

Regarding to the resolution effect, in Fig. 2 b) to k) the different cluster segmentation applied over two resolutions 640x480 and 320x240 are depicted. The algorithms do not show a significant difference when the resolution is reduced. Finally the runtime was measure using Matlab 7.9 running on a processor Core 2 Duo at 2.6 GHz and the performance are shows in Table 2.



a) Original Image b) FCM 640x480 c) FCM 320x240
d) FPCM 640x480 e) FPCM 320x240 f) PCM 640x480
g) PCM 320x240 h) RFPCM 640x480 i) RFPCM 320x240
j) FCM-GK 640x480 k) FCM-GK 320x240 l) FCM +noise 640x480
m) FPCM+noise 640x480 n) FCM-GK+noise 640x480 ñ) RFPCM+noise 640x480

**Fig. 2.** Output of the clustering segmentation using 8 classes over the channels a*b* from the L*a*b* color space, in two resolutions 640x480 and 320x240.

**Table 2.** Runtime in seconds of the clustering techniques at different resolutions

| Method | FCM | PCM | FPCM | RFPCM | FCM-GK |
|---|---|---|---|---|---|
| Time at 640x480 | 50 | 56 | 100 | 5 | 600 |
| Time at 320x240 | 13 | 19 | 39 | 1 | 210 |

## 4 Discussion

In this paper, a comparative study of five clustering techniques was performed (FCM, PCM, FPCM, RFPCM and FCM-GK). A set of 20 vineyard images was selected to asses the performance of such techniques, using as parameters runtime and the percentage of area classified correctly. Based in the results we can draw the following conclusions.

- The reduction of resolution did not affect the performance of the clustering techniques, but improved significantly the runtime.
- Due to the drawback of generating coincident clusters in RFPCM and PCM they should not be considered as segmentation techniques for vineyard.
- FPCM and FCM solved the problem of the outlayers but these points didn't affect significantly the performance of the FCM. FPCM gave good results, but generated two coincident clusters, so it just provided six clusters for the eight.
- In relation to the runtime, the best performance was obtained for RFPCM, at expense of a poor classification. FCM remains the faster algorithm after RFPCM.
- Although, RFPCM considers small clusters as noise, that is, it is too robust against the noise, it should be take into account that RFPCM shows to be a fast classifier. Therefore given these characteristics, RFPCM could be implemented as a recursive binary classifier.
- Finally, the best compromise between speed and classification performance is FCM algorithm, and stands as a good candidate to be implemented in real-time processing. However, it is not enough fast to be implemented in real time applications, so our future work will be focus in to implement this algorithm on a Field Programmable Gate Array due his highest computational capabilities.

## 5 Acknowledgment

The authors want to thank to the Spanish Ministry of Science and Innovation who have provided support for this research work through projects PLAN NACIONAL AGL2011-23673.

## References


1. R. Babuka, P.J. van der Veen, and U. Kaymak. Improved covariance estimation for gustafson-kessel clustering. In *Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on*, volume 2, pages 1081 –1085, 2002.
2. Ron Berenstein, Ohad Shahar, Amir Shapiro, and Yael Edan. Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer. In *Intelligent Service Robotics*, pages 1–11, september 2010.
3. L. Busin, J. Shi, N. Vandenbroucke, and L. Macaire. Color space selection for color image segmentation by spectral clustering. In *Signal and Image Processing Applications, IEEE International Conference on*, pages 262 –267, 2009.
4. R. Chamelat, E. Rosso, A. Choksuriwong, C. Rosenberger, H. Laurent, and P. Bro. Use of zernike moments for grape detection with image processing. In *Conference of the IEEE Industrial Electronics Society (IECON)*, pages 3697 – 3702, 2006.
5. P. Ganesan and V. Rajini. A method to segment color images based on modified fuzzy-possibilistic-c-means clustering algorithm. In *Recent Advances in Space Technology Services and Climate Change (RSTSCC)*, pages 157 –163, nov. 2010.
6. Mark Girolami. Mercer kernel-based clustering in feature space. *IEEE Transactions on Neural Networks*, 13:780–784, 2002.
7. Damaris Pascual Gonzlez. *Algoritmos de agrupamiento basados en densidad y validacin de clusters*. PhD thesis, Universitat Jaume I, 2010.
8. Daniel Graves and Witold Pedrycz. Fuzzy c-means, gustafson-kessel fcm, and kernel-based fcm: A comparative study. *Advances in Soft Computing*, 41:140–149, 2007.
9. S B Kotsiantis. Supervised machine learning: A review of classification techniques. *Informatica*, 31(3):249–268, 2007.
10. R. Krishnapuram and J.M. Keller. A possibilistic approach to clustering. *Fuzzy Systems, IEEE Transactions on*, 1(2):98 –110, may 1993.
11. N.R. Pal, K. Pal, and J.C. Bezdek. A mixed c-means clustering model. In *Fuzzy Systems, 1997., Proceedings of the Sixth IEEE International Conference on*, volume 1, pages 11 –21 vol.1, jul 1997.
12. Lei F Tian and David C Slaughter. Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computers and Electronics in Agriculture*, 21(3):153–168, 1998.
13. Xiao-Hong Wu and Jian-Jiang Zhou. An improved possibilistic c-means algorithm based on kernel methods. In *SSPR/SPR*, pages 783–791, 2006.
14. Yixin Yan, Yongbin Shen, and Shengming Li. Unsupervised color-texture image segmentation based on a new clustering method. In *New Trends in Information and Service Science, 2009. NISS '09. International Conference on*, pages 784 –787, 302009-july2 2009.
15. Miin-Shen Yang and Hsu-Shen Tsai. A gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction. *Pattern Recognition Letters*, 29(12):1713 – 1725, 2008.
16. Zhou Yong, Li Yue'e, and Xia Shixiong. Robust fuzzy-possibilistic c-means algorithm. In *Intelligent Information Technology Application, 2008. IITA '08. Second International Symposium on*, volume 1, pages 669 –673, dec. 2008.