

# Improving satellite image segmentation using evolutionary computation

Mohamad M. Awad

National Council for Scientific Research, Beirut, Lebanon

**Email address:**

mawad@cnrs.edu.lb (M. M. Awad)

**To cite this article:**

Mohamad M. Awad. Improving Satellite Image Segmentation Using Evolutionary Computation, *American Journal of Remote Sensing*. Vol. 1, No. 2, 2013, pp. 13-20. doi: 10.11648/j.ajrs.20130102.11

---

**Abstract:** Segmentation is the process of dividing an image into disjoint regions. It is the most important task in image processing where the success of the object recognition depends strongly on the efficiency of the segmentation process. The most popular and important segmentation methods are clustering such as Fuzzy c-Means (FCM), Iterative Self-Organizing Data (ISODATA) and K-means. Clustering methods depend strongly on the selection of the initial spectral signatures which represent initial cluster centers. Normally, this is done either manually or randomly based on statistical operations. In either case the outcome is unpredictable and sometime inaccurate. In this paper an unsupervised method based on Multi-Objective Genetic Algorithm (MO-GA) for the selection of spectral signature from satellite images is implemented. The new method works by maximizing the number of the selected pixels and by maximizing homogeneity through the minimizing of the difference between the pixels and their spectral signature. The objective is to create best cluster centers as an initial population for any segmentation technique. Experimental results are conducted using high resolution SPOT V satellite image and the verification of the segmentation results is based on a very high resolution satellite image of type Quickbird. The spectral signatures provided to K-means and Fuzzy c-means by MO-GA process increased the speed of the clustering algorithm to approximately 4 times the speed of the random based selection of signatures. At the same time MO-GA improved the accuracy of the results of clustering algorithm to more than 10% compared to the random statistical cluster centers selection methods.

**Keywords:** Optimization, Multi-Objective Genetic Algorithm, Spectral Signature, Clustering, Segmentation, Satellite Image, Software Development

---

## 1. Introduction

Image segmentation is the process of division of the image into regions with similar attributes [1]. It is an important step toward robust image analysis. Segmentation is one of the most difficult tasks in image processing because it determines the quality of the analysis of the extracted information [2]. There are many methods of image segmentation such as edge detection [3-7], Artificial Neural Networks (ANNs) [8, 9], region growing [10, 11], etc. But, the most popular ones are the clustering methods because of their unsupervised random selection nature based on statistics.

Formally, clustering divides an image  $I$  into  $k$  non-overlapping subsets  $C = \{C_1, C_2, \dots, C_k\}$  such that  $C_1 \cup C_2 \cup \dots \cup C_k = I$  and  $C_i \cap C_j = \emptyset, i \neq j$ . In general, each pixel in an image  $I$  is represented by an  $N$ -dimensional vector of attributes or features  $A = \{a_1, \dots, a_N\}$ . The main goal of clustering is to maximize both the homogeneity within each cluster and the heterogeneity among different clusters

[12]. Sometimes clustering can be *fuzzy* in that each object belongs to one or more clusters to different degrees [13]. In general, because of the image complex structure, clustering can be formally seen as an instance of *NP*-hard grouping problems [14].

There are many clustering methods that are used in the segmentation of different type of images in different disciplines. Iterative Self-Organizing Data (ISODATA), K-means, Fuzzy c-means (FCM) are three well-known clustering algorithms in image segmentation [15-17]. However, they require *a priori* knowledge of the clusters number and some statistical information which must be adjusted after verification. The solution is to combine a clustering algorithm with another algorithm [18] to form a hybrid method or approach similar to a split and merge method. Another way is to use non-clustering algorithm as a supplier of *a priori* data for a clustering algorithm. The last approach is the subject of this research.

Normally, the selection of spectral signatures is done manually using either geometric forms such as squares,

circles, rectangles, or any other uniform and non-uniform shapes. Sometime a seed (a pixel) is selected and the selection expands to include neighboring pixels according to pre-selected thresholds of spectral distance with minimum and maximum number of pixels. The problem with these methods is the dependency on manual selection of spectral signatures without *a priori* knowledge about the structure of the data in the images. This manual way of selection reduces the efficiency and accuracy of the clustering method. So there is a need to create an unsupervised method which can optimize the selection of the spectral signatures. The optimization process includes maximizing the homogeneity between the pixels of and their spectral signature in a cluster and maximizing the number of pixels in order to reduce over-segmentation.

Genetic Algorithms have already proved their usefulness when combined with image clustering methods [19]. However, the capacity of a GA to explore and exploit given population of complicated types of images such as multi-component images is still in progress. Multi-component image is characterized by having multi-spectral, multi-spatial, multi-radiometric and multi-temporal components. The increased volume and use of these images has created a strong need for fast and efficient image processing algorithms capable of producing reliable results.

In general, it has been proven that complex image segmentation problems have multiple objectives: minimize the distances between pixels and a cluster center to maximize homogeneity (intra-cluster spread of data), and maximize distances between different clusters (inter-cluster connectivity). Consideration of these objectives in combination is a difficult problem, and a multi-objective optimization approach is an appropriate method to solve this problem [20]. The use of a multi-objective approach in satellite image segmentation is still limited, but has great potential [21-23].

In this paper a new unsupervised method is described and evaluated for selecting the signatures from a satellite image using Multi-Objective Genetic Algorithm (MO-GA). The solutions provided by the evolutionary algorithm are used by many segmentation methods such as Artificial Neural Network (ANN) methods, supervised statistical methods including Maximum Likelihood [24] and Minimum Distance [25].

The remainder of this paper is organized as follow. After the introduction, Section II explains the concept of multi-objective optimization and particularly the Non-dominated Sorting Genetic Algorithm II (NSGA-II). Section III covers in details the Non-dominated Solutions Genetic Algorithm II (NSGA-II) application and tuning of parameters. Section IV covers experimental results and discussion. Finally, Section V completes the research with conclusion and future perspectives.

## 2. Multi-Objective Optimization

A multi-objective optimization problem can formally be

stated as [26]: Find the vector  $v = [v_1, v_2, \dots, v_n]^T$  of decision variables which will satisfy a number of equality and inequality constraints and optimizes the following vector function:

$$f(v) = [f_1(v), f_2(v), \dots, f_k(v)] \quad (1)$$

The constraints define the feasible region  $F$  that contains all the admissible solutions. Any solution outside this region is inadmissible since it violates one or more constraints. The vector  $v$  denotes an optimal solution in  $F$ . The concept of *Pareto optimality* is useful in the domain of multi-objective optimization. A formal definition of Pareto optimality from the viewpoint of the minimization problem may be given as follows: A decision vector  $v$  is called Pareto optimal if and only if there is no  $v''$  that dominates  $v$ , i.e. There is no  $v''$  such that

$$\forall i \in \{1, 2, \dots, k\}, f_i(v'') \leq f_i(v) \text{ and } \exists j \in \{1, 2, \dots, k\}, f_j(v'') < f_j(v).$$

In other words,  $v$  is Pareto optimal if there exists no feasible vector  $v''$  which causes a reduction on some criterion without a simultaneous increase in at least one other. In general, Pareto optimality usually admits a set of solutions called *non-dominated* solutions.

There are a number of multi-objective optimization techniques available. Among them, the GA based techniques such as Non-Dominated Solutions Genetic Algorithm Version II (NSGA-II) developed by Deb [27] and the Strength Pareto Evolutionary Algorithm 2 (SPEA2) [28]. In this article, NSGA-II is used as a framework for the new method because it is more efficient than any other evolutionary multi-objective optimization method.

Non-Dominated Solutions Genetic Algorithm Version II (NSGA-II) uses elitism and a crowded comparison operator that keeps diversity without specifying any additional parameters. Table 1 shows the pseudo code of the NSGA-II. The code shows that the process starts by creating an initial population of size  $M$  then their fitness values are computed and according to these values the individuals in the population are assigned a rank. Then recombination, reproduction and mutation operators are applied which results in a larger population two times the initial population. The intermediate population is sorted again based on non-domination criteria before the replacement operator is performed on the intermediate population. Once the intermediate population is sorted then the best solution is selected based on crowding distance. This helps NSGA-II in creating diverse fronts by making sure each member stays a crowding distance apart. Each front is filled in ascending order until the addition of population size is reached. This process continues except the initial creation of population until it converges with a global optimal solution of size  $M$ .

NSGA-II requires less computation time  $O(MN^2)$  compared to the older version  $O(MN^3)$  where  $M$  is the population and  $N$  is the number of objective functions. The old version of the NSGA lacks the elitism that is necessary to speed up the performance of the GA. The new version NSGA-II is able to maintain a better spread of solutions and converge

better in the obtained non-dominated front compared to other elitist Multi-Objective Evolutionary Algorithms.

Table 1. Pseudo code of NSGA-II

Initialize population
Generate random parent population of size $M$
Evaluate objective values
Assign rank (level) based on Pareto dominance – then sort
For $i=1$ to number of generations
Generate child population of size $M$ using binary tournament selection, recombination and mutation.
With combined parent and child populations, assign rank level based on Pareto dominance and sort into sets of non-dominated fronts.
Select the $M$ survivors (next generation parents):
Starting from the first front until $M$ individuals found:
Determine crowding distance between points on each front.
Select points (elitist) on the lower front (with lower rank) and are outside a crowding distance.
End inner loop
End outer loop
Return $M$ individuals in final population as the estimation of the Pareto front.

### 3. Spectral Signature Selection Optimization Using NSGA-II

The goal of this research is to automate and optimize the selection process of spectral signatures from satellite images. Initially the maximum number of spectral signatures was held constant, but the number of pixels in each spectral signature is variable. This means that the internal NSGA-II representation is a fixed-length “chromosome” of  $m$  “genes”. Each gene defines a spectral signature selected from the image. To keep things simple, the shapes of the signatures are limited to square and circle. The shapes are described by simply specifying the upper left and lower right corners of the square or the radius and the center for the circle as it is shown in Figure 1a and 1b. The basic format can easily be extended to become a complex geometric shape such as the one shown in Figure 1c.

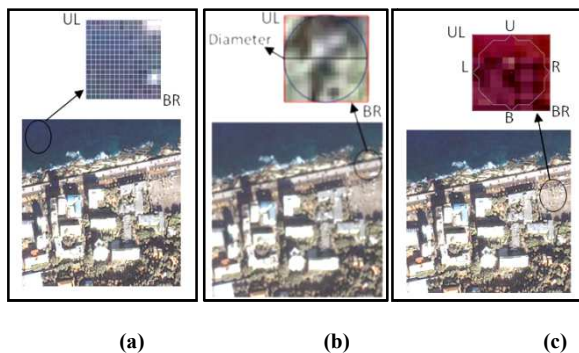


Figure 1. Spectral signature selection using (a) square (b) circle (c) complex shape

The process of signature selection starts by reading a satellite image which consists of multiple bands. The columns and rows of data in the image are used to create the initial population of chromosomes (individuals) with the same

length ( $m$  spectral signatures) but a variable number of pixels within each spectral signature.

In this research the multi-objective “fitness” of an individual involves two functions, each measuring a different characteristic of the satellite image. The first one represents the distance (Euclidean) between the mid pixel which represents the selected spectral signature  $Clust$  and the other pixels ( $Pixel$  (Equation 2)). The variables  $r_h, c_h$  are the number of rows and columns in a matrix (spectral signature), and  $n$  is the number of bands in the image. The goal is to minimize the sum of differences between a set of  $m$  spectral signatures and their neighboring pixels.

$$\min(\sum_{h=1}^m \sum_{i=1}^{r_h} \sum_{j=1}^{c_h} \sqrt{(\sum_{k=1}^n (\text{Clust}(h, \text{round}(\frac{r_h}{2}), \text{round}(\frac{c_h}{2}), k) - \text{Pixel}(i, j, k))^2)} \quad (2)$$

The second function measures the number of pixels for each spectral signature, the goal of which is to maximize the number of pixels in a set of spectral signatures (Equation 3). The degree of homogeneity in each spectral signature increases the efficiency of the clustering algorithm.

$$\max(\sum_{h=1}^m r_h \times c_h) \quad (3)$$

The spectral signatures are subject to the following constraints:

1-The minimum matrix size which represents a spectral signature is  $3 \times 3$ ,

2- The maximum is  $r_i \times c_i$  on condition that  $r_i = c_i$  ( $r_i$  and  $c_i$  are the number of rows and columns in an image).

The constraints on the minimum and maximum number of spectral signatures ( $m$ ) are  $m \geq 1$  and  $m \leq \frac{r \times c}{w \times w}$  where

$w \times w$  is the initial size of the spectral signature. In case  $r_i \neq c_i$  than columns and rows are padded with zeros until their sizes are equal.

Finally, two signatures are combined if the difference of the average spatial value of both neighbor signatures is less than  $\mathcal{E}_1$  and the difference between the two spectral values is less than  $\mathcal{E}_2$ . Such that  $\mathcal{E}_1$  and  $\mathcal{E}_2$  are calculated according to the following formula:

$$\mathcal{E}_1 = \frac{w^2}{\text{size}(im)} \text{ and } \mathcal{E}_2 = \frac{\min(im)}{\max(im)} * \text{mean}(im) \quad (4)$$

Where  $mean$ ,  $max$ ,  $min$ , and  $size$  are the average, maximum and minimum intensity values. The parameter  $size$  is the dimension of image  $im$ . Once a parent population is in place, a child population is created by selecting parents based on their non-dominated sorting fitness using binary tournament selection [29]. In this research two parents are selected to produce two offspring via uniform crossover [30-31]. The uniform crossover uses a fixed mixing ratio between two parents. Unlike, one and two-point crossover, the Uniform Crossover enables the parent chromosomes to

contribute the gene level rather than the segment level. If the mixing ratio is 0.5, the offspring has approximately half of the genes from first parent and the other half from second parent. The newly created children are then subject to mutation operator that modifies the size of a signature by incrementing/decrementing the Upper Left (UL) and/or Lower Bottom (BL) coordinates by 2 within the constraints specified earlier. Preliminary experiments indicated that a mutation rate between 0.05 and 0.1 can produce best results. The following pseudo code (Table 2) shows clearly how mutation operation works.

To provide additional diversity beyond the spectral signatures making up the initial population, a second mutation operator is implemented which replaces an existing spectral signature with a randomly generated one. This form of mutation is more explorative and may slow the process of convergence producing better global solutions.

**Table 2.** Pseudo code of the mutation operator

Select a gene that has a position as either 1 or modulo 4 If the position is greater than number of rows or columns then subtract 2 Else add 2
--

The replacement function replaces the individuals based on rank and crowding distance. After completing the process of producing an offspring population, the next generation of  $M$  parents is chosen from the current parent and offspring populations based on non-dominated ranking and crowding. The new parent population represents the current estimate of the entire Pareto front. The quality of that estimate is calculated using the standard Hypervolume indicator introduced by Coello *et al.* [32].

This process is repeated until the limit on the number of generation is reached or there is no further improvement in the Hypervolume indicator. Table 3 gives the pseudo-code of the overall process of applying MO-GA based on NSGA-II in the creation of the new automatic signature selection tool.

Notice that the notion of Pareto optimality implies a set of non-dominated solutions based on the objectives given. It is then the job of the Decision Maker (DM) to select which of the non-dominated solutions (if any) are preferable based on other aspects not specified as part of the multi-objective optimization problem [33].

**Table 3.** Signature selection method pseudo code

Read multi-component image. Initialize the parameters (including the number of spectral signatures). Choose the population size. Determine the number of generations. Initialize the population with random spectral signatures. Run NSGA-II until the stopping criteria is satisfied Return as a result the final population which represents the best estimate of the Pareto front.
---

## 4. Experimental Results

In order to evaluate the effectiveness of the new signature selection method, images from two different satellites are used. SPOT V and Quickbird images. SPOT V offers a resolution of 2.5 to 5 meters in panchromatic mode and 10 meters in multi-spectral mode. Quickbird satellite images have four multi-spectral bands with 2.4 meter resolution and one panchromatic band with 0.61 meter resolution. The results of the new signature selection algorithm are passed to K-means algorithm [34-36] and Fuzzy c-means [37-38] to complete the clustering process. K-means follows a simple procedure for classifying a given data set using a predefined number of  $k$  clusters. The number of clusters is decided by the supplied spectral signatures which are obtained by the multi-objective evolutionary algorithm. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method is frequently used in pattern recognition. It is based on minimization of a defined objective function.

### 4.1. Selecting the Best Parameters for MO-GA

The first step is to decide on the population size and the number of generations. Normally, the selection of these parameters depends strongly on the size of the image. A rule of thumb, if the image sizes are organized into four categories (a very large image is an image having a size greater than 2048×2048 pixels, large image is an image between 512×512 and 2048×2048 pixels, a medium is between 128×128 and 512×512 pixels and finally small one is less than 128×128 pixels) then the population size and the chromosome length increases 30 individuals and 4 genes for each increment of 128×128 pixels. It is preferable to start with a population of size 20 and individual length of 8 genes (each represents 4 coordinates) for small image.

The above rule is used to decide on the initial population size and individual length. However, experiments are still needed to find the optimal combination of the population size and individual (chromosome) length. In this research the population size is varied between 40 and 100 and the individual length is varied between 64 and 96 representing either 16 or 24 matrices (spectral signatures) with 4 coordinates. Finally, the number of generations varied between 200 and 600. The ratio of both objective functions ( $F1/F2$ ) is calculated such that the closer the ratio to zero the better, so one can conclude after several experiments that the population size 100 and the chromosome length 96 are the best combination (Figure 2).

In order to find the best results, the new selection method is run with different mutation and reproduction rates. The mutation rate is varied between 5% and 10% while crossover rate is varied between 60% and 80%. The population size is set to 100 and individual length is 96.

The outcomes of running the method with different parameters are compared with respect to many criteria such as: 1-The stable progress toward the final solutions, and 2-the fast or slow convergence toward global or local optimal solutions. From the results of the experiments, mutation and crossover operators with 5% and 60% rates are se-

lected. The smooth increase in the number of pixels accompanied by the increase in homogeneity in every selected area is the reason for selecting these rates. The other parameters are not selected because the number of pixels and homogeneity increase simultaneously. But suddenly the homogeneity of the selected pixels and spectral signatures decreases again. This means that the run of the algorithm should be aborted earlier and we have a premature convergence toward a local optimal solution.

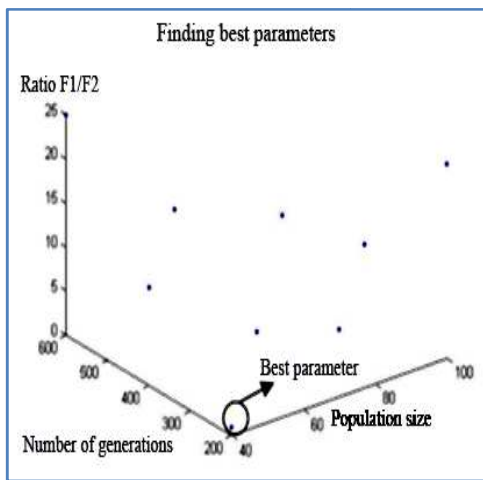


Figure 2. Different parameters with chromosome length of 96

After the convergence of MO-GA the population represents the final solution. Diversity in the solution is preferred (different number of classes or spectral signatures) because it helps the decision maker to select the best solution.

**4.2. Optimizing the Selection of the Spectral Signatures Using MO-GA**

In the experiment, a sharpened SPOT VXS image of size 304X304 pixels for different locations is used. The image is of 5 meters resolution and it represents a large area of urban and agricultural site (Figure 3a). The QuickBird image is used to verify the segmented SPOT image. It is selected from the same area during the same period of the year (Figure 3b). The size of the QuickBird image is 1900X1900 pixels with resolution of 0.61 meter.

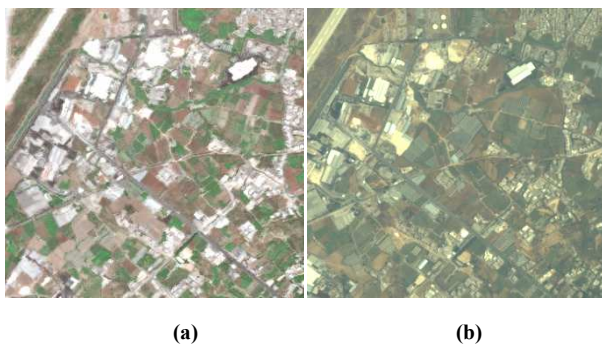


Figure 3. Satellite images (a) SPOT V image (b) QuickBird image.

As it is decided before the population consists of 100 individuals, the number of generations is 600, and the length of the chromosome “individual” is 96. The final solutions of the MO-GA are investigated and the matrices with 4 classes are selected. The reason for the selection of 4 classes is based on the preferences to reduce (generalize) the classes in the image. Normally, the verification of generalized classes using a very high resolution image is much simpler than detailed classes which require field verification. To find the best matrix in the group of matrices of 4 classes the one with the highest number of pixels and best homogeneity is selected. Then the spectral signatures are provided to K-means and Fuzzy c-means the initial cluster centers to segment the SPOT V image.

After the convergence of MO-GA one of 100 individuals (final solution) is selected. However, some spectral signatures have similar values which mean that the final optimal multi-objective solution may consist of different individuals with different numbers of spectral signatures. Figure 4a shows the satellite image with different solutions (10 spectral signatures) such that each solution has a different signature or class represented by a square. Investigating these squares one can see that each has a different size and different position. Sometime these signatures may have the same value or their difference is less than a certain threshold. Knowing the relationship between these different signatures will help in removing redundancies by using one spectral signature value for many similar signatures. In order to increase the efficiency of the selection method a procedure which combines similar spectral signatures (classes) and eliminates redundancy is implemented. Figure 4b shows the final individual (4 spectral signatures) after using the procedure which removes the redundancies in the original 10 spectral signatures.

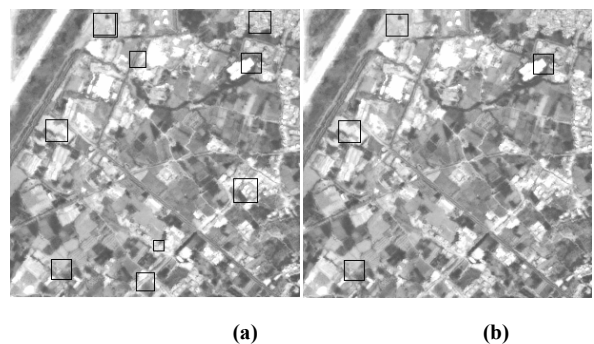
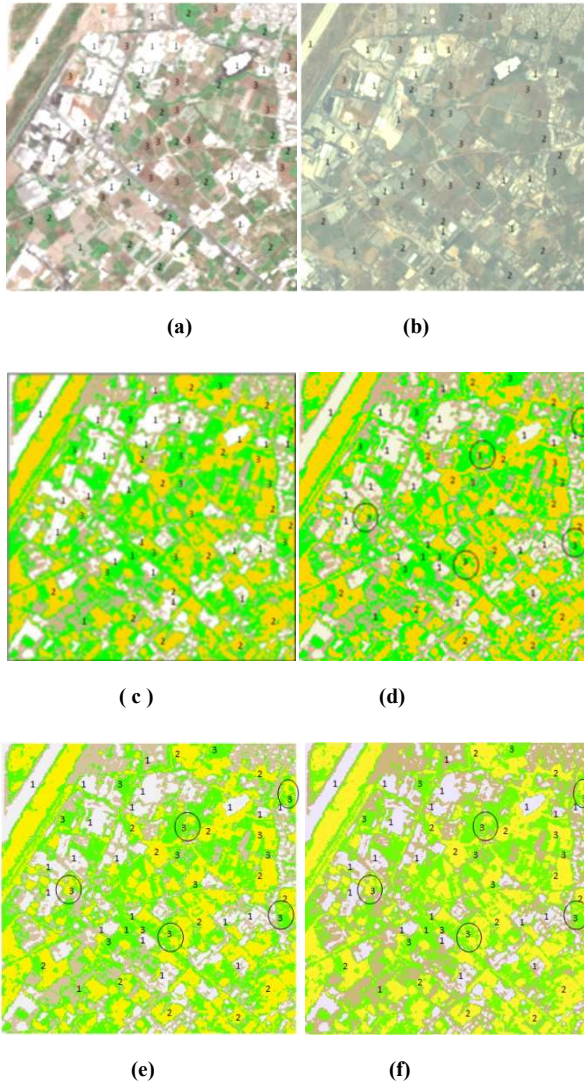


Figure 4. The selected signatures (a) original by MO-GA (b) after the redundancy elimination procedure.

Several areas are selected as samples for verification using the QuickBird high resolution images. The samples represent vegetation, urban and soil classes. The images with 4 classes generalize the object and combine different classes while the other segmented images with a higher number of classes give more details. Sometime, more spectral signatures indicate an over-segmentation. But, if the correct number of spectral signatures is selected then ho-

mogeneity will increase and over-segmentation will decrease. Out of 4 classes 3 are selected for verification because they represent different important and general features in the SPOT V satellite image. This makes the job of verification easier and eliminates the need for field survey. The classes represent urban and other features such as roads with green houses, vegetation, and soil. A number of objects are selected randomly from each class and then they are compared with the same objects which exist in the high resolution image (Figure 5a and 5b).



**Figure 5.** Selected samples (a) SPOT V image (b) QuickBird image (c) segmented image by MO-GA + K-means (d) segmented image by random K-means (e) Segmented image by MO-GA + FCM (f) segmented image by normal FCM.

The samples are distributed uniformly and represents the three classes 1- Urban and others (White), 2- Vegetation (Yellow), and 3- Soil (green). In the SPOT image the number of the first class is 24 objects, 15 objects for each other classes. Table 4 provides a way to compute the accuracy of the selected samples in a way similar to the confusion matrix of Kohavi and Provost [39]. The only difference is the dependability of the predicted and actual data on field verifi-

cation.

The normal K-means and Fuzzy c-means are run with the same initial number of iterations as that of MO-GA based method and the results of clustering are shown in Figures 5c to 5f which include some investigated spots (circles). In our case, MO-GA with K-means and Fuzzy c-means converged after less than 40 iterations. On the other hand, K-means and FCM with random selection process (normal K-means and FCM) converged after more than 100 iterations. In addition, it took between 10 and 17 seconds to complete the process for both algorithms compared to less than 4 seconds for MO-GA based process.

The percentage of the correct segmented features indicates the accuracy of the new selection method. According to the results in Table 4, 5, 6 and 7 the number of correctly identified objects is indicated in the diagonal of the table. The remaining cells in the table indicate the number of conflicts with other selected objects e.g. urban objects with soil or vegetation objects.

The total number of the samples in the diagonal divided by the total number of all samples indicates the percentage of accuracy. The segmented image with 4 classes is used for the verification process of MO-GA based signature selection method. The calculated accuracy of the new unsupervised selection method is more than 95% (K-means and Fuzzy c-means based on MO-GA). While the accuracy for the normal K-means and Fuzzy c-means with random process is less or equal to 85%.

**Table 4.** Verification of the clustered image by MO-GA and K-means.

	Urban and others	Vegetation	Soil	Total
Urban and others	23	0	2	25
Vegetation	0	15	0	15
Soil	0	1	14	15
Total	23	16	16	55

**Table 5.** Verification of the clustered image by random K-means.

	Urban and others	Vegetation	Soil	Total
Urban and others	23	0	2	25
Vegetation	0	13	2	15
Soil	0	5	10	15
Total	23	18	14	55

**Table 6.** Verification of the clustered image by MO-GA and Fuzzy c-means.

	Urban and others	Vegetation	Soil	Total
Urban and others	24	0	1	25
Vegetation	0	15	0	15
Soil	1	1	13	15
Total	25	16	14	55

**Table 7.** Verification of the clustered image by random Fuzzy c-means.

	Urban and others	Vegetation	Soil	Total
Urban and others	24	0	1	25
Vegetation	0	13	2	15
Soil	1	4	10	15
Total	25	17	13	55

## 5. Conclusions

In this research multi-objective Genetic Algorithm (MO-GA) is used to create a robust and efficient spectral signature selection method. The role of the Non-Dominated Sorting GA version II (NSGA-II) in the new method is to find a global optimal solution of spectral signatures. The solution consists of many individuals each is a possible group of spectral signatures which can be used with any segmentation/classification algorithm such as K-means and Fuzzy c-means. The obtained individuals represent the global optimal solutions with different characteristics such as the total number of pixels and the degree of homogeneity. Several experiments are conducted to find the best combinations of parameters which represent the size of the population, the size of the individuals and the reproduction rates. Other parameters such as the maximum and minimum size of the matrix which represents a gene in a chromosome (individual) are affected by the size of the image. These parameters must be computed each time the satellite image type and size are changed. In our case, the image is considered to be medium size image and the parameters are selected according to the criteria set in the experiments. The reproduction operator rates of the MO-GA are selected after several try and test experiments where the objective is to minimize the heterogeneity and maximize the number of pixels per spectral signature.

The application of this selection method to a Spot V multi-spectral satellite image and the verification of the results using another high resolution QuickBird satellite image have proved the reliability of this method. The verification method is based on the collection of sample objects representing different classes such as vegetation and urban. The results indicate that the accuracy of the new selection tool can reach more than 95 %. Although, the verified image consists of only 4 classes, other images can be verified too depending on the required details and the available verification data. The new selection method can be used as an unsupervised spectral signature selection tools. The speed of the new selection tool is more than 4 times faster compared to the traditional method of random selection of cluster centers by K-means and Fuzzy c-means with the same initial number of iterations as the MO-GA.

In the future there are many issues left to be solved such as how accurate this new method is compared to the supervised selection methods. In addition, there are unsupervised methods which depend on random selection of seeds and which can be compared in combination with the new selection

method.

The new selection method uses different geometric shape ( $N \times N$  matrix) which may represent a square or a circle. The next step is to improve the selection method and to modify it in order to use other different geometric shapes such as hexagon, octagon, non-uniform shapes.

## References

- [1] R. Demirci, Rule-based automatic segmentation of color images. *International Journal of Electronics and Communication* 60, 435 – 442(2006).
- [2] W. Pratt, *Digital Image Processing*. 2nd edition, (Wiley & Sons Inc, New York., 1991).
- [3] I. Sobel, G. Feldman, A 3x3 Isotropic Gradient Operator for Image Processing, In: Hart, *Pattern Classification and Scene Analysis*, (Wiley and Sons, New York, 1973), ed. by R. Duda, and P. Hart, pp. 271-2.
- [4] J. Canny, Computational approach to edge detection, *IEEE Trans. On Pattern Analysis and Machine Intelligence* 8( 6), 679-698 (1986).
- [5] J. Shen and S. Castan, An optimal linear operator for edge detection, In *Abstract of the Proceeding of IEEE Int. Conf. On Computer Vision and Pattern Recognition, USA, 1986*.
- [6] M. Kass, A. Witkin and D. Terzopoulos, Snakes: Active Contour models, *International Journal of Computer Vision* 1, 259-268 (1987).
- [7] R. Deriche, Using Canny's criteria to derive a recursively implemented optimal edge detector, *International Journal of Computer Vision* 1(2), 167-187(1987).
- [8] M. Awad, K. Chehdi, and A. Nasri, Multi-component Image Segmentation Using Genetic Algorithms and Artificial Neural Network, *IEEE Geosciences and Remote Sensing Letters* 4(4), 571-575(2007).
- [9] M. Awad, An unsupervised Artificial Neural Network method for satellite image segmentation, *International Arabic Journal of Information Technology (IAJIT)* 7( 2), 199-205 (2010).
- [10] F. Shih, and S. Cheng, Automatic seeded region growing for color image segmentation, *Image and Vision Computing* 23 (10), 877-886 (2005).
- [11] L. Garcia, E. Saber, S. Vantaram, V. Amuso, M. Shaw, R. Bhaskar, "Automatic Image Segmentation by Dynamic Region Growth and Multiresolution Merging," *IEEE Transaction on Image Processing* 18(10), 2275-2288 (2009).
- [12] L. J. Arabie, G. Hubert, and P. DeSoete, *Clustering and Classification*. (World Scientific, Singapore, 1999).
- [13] F. H'oppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*. (Wiley and Sons, 1999).
- [14] E. Falkenauer, *Genetic Algorithms and Grouping Problems*. Wiley and Sons, 1998).
- [15] J. Tou, and R. C. Gonzalez, *Pattern Recognition Principles*. (Addison-Wesley Publishing, Massachusetts, 1974).

- [16] H. Ng, S. Ong, K. Foong, P. Goh, W. Nowinski, Medical Image Segmentation Using K-Means Clustering and Improved Watershed Algorithm, In the abstracts of IEEE Southwest Symposium on Image Analysis and Interpretation, Colorado, USA, 2006.
- [17] J. Noordam, W. Broek, and L. Buydens, Geometrically Guided Fuzzy C-means Clustering for Multivariate Image Segmentation, In the Abstracts of 15th International Conference on Pattern Recognition (ICPR'00), Spain, 2000.
- [18] M. Awad, K. Chehdi, and A. Nasri, Multi-component image segmentation using Hybrid Dynamic Genetic Algorithm and Fuzzy C-Means, IET image processing 3 ( 2), 52-62 (2009).
- [19] M. Awad, K. Chehdi., Satellite image segmentation using variable hybrid Genetic Algorithm, International Journal of Imaging Systems and Technology (19), 199-207 ( 2009).
- [20] V. Guliashki, H. Toshev, and C. Korsemov, Survey of Evolutionary Algorithms used in multi-objective optimization, Problems of Engineering Cybernetics and Robotics 60, Bulgarian Academy of Sciences, 2008.
- [21] S. Bandyopadhyay, A. Srivastava and S. Pal, 'Multi-objective Variable String Genetic Classifier: Application to Remote Sensing Imagery, ed. A. Ghosh and S. Pal, Lecture Notes in Soft Computing, Image Processing and Pattern Recognition, (World Scientific, 2002), 65-94.
- [22] A. Mukhopadhyay, S. Bandyopadhyay, and U. Maulik, Clustering using multi-objective genetic algorithm and its application to image segmentation, In Abstracts of IEEE International Conference on Systems, Man and Cybernetics 3, Taipei, Taiwan, 2007.
- [23] N. Ghoggali, F. Melgani, and Y. Bazi, A Multi-objective Genetic SVM Approach for Classification Problems With Limited Training Samples, IEEE Trans. on Geosciences and Remote Sensing 47, 1707-1718 ( 2009).
- [24] J. Aldrich, R. A. Fisher and the making of maximum likelihood 1912–1922, Statistical Science 12( 3), 162–176 (1997).
- [25] P. Swain, and S. Davis. Remote Sensing: The Quantitative Approach. (McGraw Hill Book Company, New York, 1978).
- [26] K. Deb, Multi-objective Optimization Using Evolutionary Algorithms. (John Wiley and Sons, England, 2001).
- [27] K. Deb, S. Agrawal, A. Pratab, and T. Meyarivan, A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II, Springer Lecture Notes in Computer Science No. 1917, 849–858 ( 2000).
- [28] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm," Gloriosastrasse 35, CH-8092 Zurich, Switzerland, Tech. Rep. 103, 2001.
- [29] D. Goldberg, K. Deb, A comparison of selection schemes used in genetic algorithms, ed. by G.J.E. Rawlins, (Morgan Kaufmann Publishers, San Mateo, CA, 1991), p. 69-93.
- [30] G. Syswerda, Uniform Crossover in Genetic Algorithms, ed. D. J. Schaffer, in proceedings of the Third International Conference on Genetic Algorithms ICGA, (Morgan Kaufmann, Los Altos, Ca, 1989), p. 2-9.
- [31] W. Spears and K. De Jong, On the virtues of parameterized uniform crossover, ed. R. K. Belew and L. B. Booker, in proceedings of International Conference on Genetic Algorithms, Morgan Kaufmann, Los Altos, Ca, 1991), p. 230–236.
- [32] C. A. CoelloCoello, D. A. Van Veldhuizen, and G. B. Lamont, Evolutionary algorithms for solving Multi-objective problems. (Kluwer Academic Publishers, New York, 2002).
- [33] E. Zitzler, Evolutionary algorithms for multiobjective optimization: Methods and applications, PhD Dissertation, Swiss Federal Institute of Technology Zurich, 1999.
- [34] J. MacQueen, Some Methods for classification and Analysis of Multivariate Observations, In Abstracts of 5-th Berkeley Symposium on Mathematical Statistics and Probability 1, Berkeley, University of California, USA , 281-297 (1967).
- [35] R. Duda, P. Hart, and D. Stork D, Pattern Classification, Second Edition. (JohnWiley& Sons, New Jersey, 2000).
- [36] J. Tou and R. Gonzalez, Pattern Recognition Principles. ( Addison-Wesley, MA, 1974).
- [37] J. Dunn, A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters, Journal of Cybernetics 3, pp. 32-57, 1973.
- [38] J. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, (Plenum Press, New York, 1981).
- [39] R. Kohavi and F. Provost, Glossary of Terms, In Special Issue on Applications of Machine Learning and the Knowledge Discovery Process 30 (2/3), 1998.