# Unsupervised Segmentation of Images Based on Shuffled Frog-Leaping Algorithm

Amel Tehami\* and Hadria Fizazi\*

#### Abstract

The image segmentation is the most important operation in an image processing system. It is located at the joint between the processing and analysis of the images. Unsupervised segmentation aims to automatically separate the image into natural clusters. However, because of its complexity several methods have been proposed, specifically methods of optimization. In our work we are interested to the technique SFLA (Shuffled Frog-Leaping Algorithm). It's a memetic meta-heuristic algorithm that is based on frog populations in nature searching for food. This paper proposes a new approach of unsupervised image segmentation based on SFLA method. It is implemented and applied to different types of images. To validate the performances of our approach, we performed experiments which were compared to the method of K-means.

#### Keywords

Image, K-means, Meta-Heuristic, Optimization, SFLA, Unsupervised Segmentation

### **1. Introduction**

In an image analysis system, the segmentation appears as the most important step because all subsequent tasks such as extraction of primitives, detecting a position of an object or the object recognition, strongly depend on the quality of the segmentation. The aim is to facilitate the extraction of its components [1].

The segmentation is generally defined as a process of partitioning an image into homogeneous regions, as each region is homogeneous and the union of two adjacent regions is not homogeneous. There are several ways to categorize the image segmentation methods. Shanker [2] classified them into four classes: contour approach, pixels approach, regions approach, and hybrid approach. Ameur and Ameur [3] into two main classes—contour approach and regions approach, similarly for Guo et al. [4] into two classes (color and texture).

Many methods have been devised to solve the problem of unsupervised image segmentation. However, they have drawbacks: great sensitivity to the initial configuration or premature convergence to a local optimum. Consequently researches have adapted the segmentation problem to an optimization problem. This allowed to apply meta-heuristics, inspired biological and physical phenomena of nature, to the field of images segmentation. Among the most known meta-heuristics used in image segmentation

Manuscript received December 21, 2015; accepted March 20, 2016. Corresponding Author: Amel Tehami (tehami.amel@univ-usto.dz)

<sup>\*</sup> This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dept. of Computer Science, Faculty of Mathematics and Computer Science, University of Science and Technology of Oran-Mohamed

we find genetic algorithm [5], ant colony optimization [6], simulated annealing [7], and particle swarm optimization [8].

This article presents a new method of unsupervised segmentation of images using SFLA. The advantage of this approach is to segment the image with the best split according to an objective function.

This paper is organized as follows: Section 2 formulates the segmentation problem, Section 3 presents SFLA method. Section 4 describes the application of Shuffled Frog-Leaping Algorithm (SFLA) for unsupervised image segmentation. Section 5 defines the K-means algorithm. Section 6 shows the experiments, interpretation of the approach and comparison with K-means. In Section 7, we highlight the influence of one of algorithm parameters on image segmentation. Section 8 concludes this paper.

### 2. Formalizing the Segmentation Problem

Generally, segmentation seeks to partition an image *I* into disjoint subsets and related, called regions  $R_i$ . Each region is homogeneous and that the union of two adjacent regions is not. Therefore, segmentation is a partition of the image into regions, respecting this definition for P(.) such as *P* is a given predicate (often related to a criterion of homogeneity). From a mathematical point of view, Zucker [9] defines segmentation *S* of all pixels in an image *I*, as an ensemble of regions  $R_i$  (*i* from 1 to n);  $S = \{R_1, R_2, ..., R_n\}$  such as:

$$\bigcup_{i=1}^{n} R_{i} = I \tag{1}$$

with  $R_i \cap R_j = \{\}$  for  $i \neq j$  and  $P(R_i) = \text{True}$ ,  $P(R_i \cup R_j) = \text{False}(R_i \text{ adjacent to } R_j)$ . A segmentation algorithm tries to find a partition  $R_i$  such as the similarity between pixels of the same region is maximal and between the pixels of different groups is minimal. So from the original image, multiple partitions can be proposed, hence the need to define an objective function which must evaluate a region based on similarity and dissimilarity measures of pixels.

## 3. Shuffled Frog-Leaping Algorithm

This algorithm known by its name shuffled frog-leaping algorithm is an optimization meta-heuristic that mimics the evolution of frogs group looking for place which has a maximum amount of food. This latter is spread randomly on stones in a pond [10].

Each frog is a solution to the problem. Total population is divided into groups of frogs called memeplexes. They evolve independently to browse the solution space in different directions. The information between different memeplexes flows through a jump processes.

Assume that the initial population of frogs F is randomly defines in space  $(X_n, n=1, 2, ..., F)$ ,  $f_n$  represents the fitness value of the  $n^{th}$  frog. All frogs are sorted in descending order and are divided into q memeplexes, each containing p frogs  $(F=p^*q)$ . In this process, the first frog goes to the first

memeplex, the second frog goes to the second memeplex, frog m goes to the q memeplex and frog m+1 goes to the first memeplex [11], as shown in Fig. 1.



Fig. 1. Dividing *F* frogs into the *q* memeplexes.

In each memeplex, the frogs with the best and the worst fitness are considered as  $X_b$  and  $X_w$ , respectively. Also, the frog with the best fitness in all population is considered as  $X_g$  [12].

During the evolution of memeplex, in other words, for local exploration, the location of the frog with the worst fitness is regulated as follow:

$$S = H \times (X_b - X_w) \tag{2}$$

$$X_{new} = X_w + S \tag{3}$$

where *S* represents the value of change in the position, such as  $(-S_{max} < S < S_{max})$ , with  $S_{max}$  the maximum allowed change in a frog's position, *H* random number between 0 and 1. If this process produces better solution so it replaces the worst. Else, the same rule is applied by replacing the  $X_b$  by global solution  $X_g$ :

$$S = H \times \left( X_g - X_w \right). \tag{4}$$

After have obtained *S*, we recalculate  $X_{new}$  according to Eq. (3). If this new solution is worse than the worst frog, we randomly generate better solution than  $X_w$  and we replace  $X_w$  by  $X_{new}$  [12].

After a number of iterations, the different groups combined and share their ideas with themselves through a shuffling process. The local search and the shuffling processes continue until defined convergence criteria are satisfied.

# 4. Unsupervised Image Segmentation Using SFLA

In the context of an unsupervised segmentation of images each frog is composed of a representative  $\mu_i$  of each region or class, called gravity center.

Consequently, a frog jump gives several possible segmentations of the image, representing candidate solutions. So the need to assess just to keep only one, judged as best, according to a predefined objective function [13].

$$f = \frac{1}{E} \tag{5}$$

*E* represents the quadratic error, whose minimum is an index of good segmentation. It is expressed by Eq. (6):

$$E = \sum_{i=1}^{K} \sum_{j=1}^{Q_i} d(x_j^{(i)}, \mu_i)^2$$
(6)

where *K* is the number of classes (regions) desired,  $Q_i$  the number of pixels in the class *i*, *d* represents the distance between the pixel  $x_i^{(i)}$  belonging to class *i* and the gravity center  $\mu_i$  of this class.

For better segmentation, it is necessary to maximize *f*. The maximum value of the fitness corresponds to segmentation with minimum distance between the pixels belonging to the same region.

The main steps of the SFLA algorithm for unsupervised segmentation of images can be summarized as follows:

- Step 1 (Set the initial parameters). Initialize the size of the population *F*, the number of memeplexes *q*, the number of frogs in each memeplex *p*, the *H* parameter, the number of iterations *N1* for local search of each memeplex and the number of iterations *N2* for program execution.
- Step 2 (Generate population of frogs). To implement the SFLA algorithm, an initial population of *F* frog is generated. Each frog X<sub>n</sub> corresponds to a vector *B* of dimension *D*\**K* such *D* is the dimension of the search space and *K* the number of regions presents on the image. Indeed, each frog is composed of one representative μ<sub>i</sub> of each region. The generation of these gravity centers is done in random manner.
- Step 3 (Evaluate fitness of each frog). After generation of initial population, each pixel is assigned to the class whose center is the nearest. All frogs are then evaluated using the fitness function shown by Eq. (5). For each frog is associated a fitness value *f*.
- Step 4 (Sorting the population). The population of frogs is sorted in descending order according to the fitness value in order and determine the best frog  $X_g$  in this population.
- Step 5 (Partition of population in q memeplexes). After Have sorted the population of frogs, each memeplex contains *p* frogs.
- Step 6 (The local search). For each memeplex: Determine the best frog  $X_b$  and the worst frog  $X_w$ . The worst frog changes position, his new place  $X_{new}$  is calculated and evaluating a corresponding fitness function *f*. If  $f(X_{new}) > f(X_w)$  then, this solution replaces the worst. Else,  $X_w$  makes another jump determined from Eq. (4). So we recalculate the new position and its fitness *f*. The new position of the frog will replace the worst, if it produces better solution. Else generate randomly  $X_{new}$  better than  $X_w$ .
- Step 7. The different groups of memeplexes are combined to form again the population of frogs.
- Step 8. Go to the step 4 if the number of maximal iteration *N2* is not reached.
- Step 9. Show the best solution by pixels labeling to their nearest centers.

Fig. 2 shows the solution of an optimization problem using SFLA approach for unsupervised image segmentation.



Fig. 2. Flowchart of Shuffled Frog-Leaping Algorithm (SFLA) for unsupervised image segmentation.

### 5. The K-Means Algorithm

K-means is an unsupervised segmentation algorithm the most used, saw its simplicity. It partitions the data of image into K clusters. Unlike other so-called hierarchical methods, that create a structure in "tree of clusters" to describe groups. The algorithm returns a data partition, wherein the objects inside each cluster are as near as possible to each other and as far as possible from the objects of the other

clusters. Each cluster of the partition is defined by its objects and its centroid [14,15].

K-means is an iterative algorithm that minimizes the sum of distances between each object and the centroid of its cluster. The main steps of the K-means algorithm are:

- Step 1. Random choice from the initial position of the K clusters.
- Step 2. Affect the objects to cluster following the distances minimization criterion (generally according to a measure of Euclidean distance).
- Step 3. Once all objects placed, recalculate K centroid.
- Step 4. Repeat steps 2 and 3 until no more affectation is made.

## 6. Experiments and Results

In this section, segmentation results of the proposed method are evaluated on some test images. It has been tested on more than 50 images taken from public image. The regions included in the images are visible to the naked eye that which will assess the accuracy of obtained results.

To appreciate the efficiency and performances of the SFLA approach used in our work, we have compared with segmentation results of K-means. For all experiments, the parameters values of the proposed method determined after several tests and ensuring good convergence are recorded in Table 1.

Parameter	Description	Value
F	Number of frogs	20
q	Number of memeplexes	4
p	Number of frogs in each memeplex	5
N1	Number of iterations (local search)	50
N2	Number of iterations	100
Н	The parameter	0.05

Table 1. Initial parameters of SFLA

The experiments conducted have been applied on three synthetics images selected to evaluate the efficiency of the proposed algorithm (Fig. 3). The size of each tested image is 256×256 pixels.







Fig. 3. Images used (image 1, 2 and 3 from left to right).

Figs. 4–6 show the results obtained for unsupervised segmentation of the synthetics images by varying K (the number of desired region in the image), respectively.



**Fig. 4.** Unsupervised segmentation of image 1. (a) SFLA and K=2, (b) K-means and K=2, (c) SFLA and K=4, (d) K-means and K=4.



**Fig. 5.** Unsupervised segmentation of image 2. (a) SFLA and K=2, (b) K-means and K=2, (c) SFLA and K=4, (d) K-means and K=4.

From the results of Fig. 4, the image is very well segmented with the proposed approach for the different values of *K* unlike at that segmented with the K-means. For better test the performances of our algorithm SFLA, we used a complex satellite image. The result is presented in Fig. 5.

In Fig. 6, we show the comparison of unsupervised segmentation of image 3 used SFLA and K-means for *K*=3, *K*=4, *K*=5, and *K*=6.



**Fig. 6.** Unsupervised segmentation of image 3. (a) SFLA and K=3, (b) K-means and K=3, (c) SFLA and K=4, (d) K-means and K=4, (e) SFLA and K=5, (f) K-means and K=5, (g) SFLA and K=6, (h) K-means and K=6.

Clearly show that, the proposed approach can generate correct results of segmentation, the contours of regions are clearly visible. The image segmented by K-means is less good because we see that, for K=2; the blue color appears with a large portion in the red region. For K=4, the yellow color is found in the red part, and blue color represents the noise class.

To better distinguish the regions that make up the image, we have affected the colors to regions and we calculated the percentage of pixels in each class (Table 2).

Classes		SFLA (%)	K-means (%)	
<i>K</i> =2	Red	56.671143	53.897095	
	Blue	43.334960	46.109010	
<i>K</i> =4	Red	19.833374	24.473572	
	Blue	19.036865	6.0073853	
	Green	18.135070	30.799866	
	Yellow	43.006897	38.731384	

Table 2. Pixels percentages in each class of image 2

Table 3 presents the percentage of pixels in each class found in image 3 with variation of K (K=3, 4, 5, and 6).

Classes		SFLA (%)	K-means (%)
K=3	Red	29.183960	5.5496216
	Blue	24.093628	75.418090
	Green	46.731567	19.041443
K=4	Red	7.2036743	2.0828247
	Blue	24.502563	59.190370
	Green	33.970642	18.162537
	Yellow	34.335327	20.576477
<i>K</i> =5	Red	10.797119	20.24231
	Blue	16.096497	0.0030517578
	Green	27.339172	57.322693
	Yellow	26.745605	22.444153
	Pink	19.036865	0.0030517578
<i>K</i> =6	Red	21.339417	20.228577
	Blue	24.588013	18.432617
	Green	11.924744	5.0231934
	Yellow	23.452759	42.279053
	Pink	11.679077	9.7808840
	Magenta	7.0343018	4.2739870

Table 3. Pixels percentages in each class of image 3

The synthetic image 3 is very complex because it has several regions. Note that, more we increase the parameter *K*, more the objects of image are identified and particularly frog object for the two methods SFLA and K-means. However unsupervised segmentation of the image 3 with SFLA approach is better than that obtained by K-means.

We can conclude that the results (obtained in Figs. 4-6 and Tables 2 and 3) show that the different synthetics images are well segmented with SFLA algorithm whatever the variation of the parameter *K*, by against, those obtained by K-means are less satisfactory.

### 7. Influence of Parameter H on the Segmentation

The H parameter used in Eqs. (2) and (4) calculates the value of the jump made by the frog. This last allow recalculating again its new position.

According to the articles [11,12,16-18], *H* parameter is a random number that varies between 0 and 1. To improve the performance of the proposed algorithm and to highlight the influence of this parameter on the results obtained, we varied *H* as shown in the following Table 4.

Table 4.	Variation	of parameter	Η
----------	-----------	--------------	---

Н	0.05	0.1	1	2	3	4

In the following experiments, we initialize parameters as is present in Table 5.

Parameter	Description	Value
F	Number of frogs	15
9	Number of memeplexes	3
Р	Number of frogs in each memeplex	5
N1	Number of iterations (local search)	50
N2	Number of iterations	100

Table 5. Set parameters of proposed approach

#### 7.1 Experiment 1

In this experiment, we use an image of size 256×256 pixels, presents in Fig. 7.



Fig. 7. Image 4 used for this experiment.

Figs. 8 and 9 summary the different results obtained for unsupervised image 4 segmentation using SFLA with the variation of H parameter.



H=2 H=3**Fig. 8.** The SFLA unsupervised segmentation of image 4 for K=2.



Fig. 9. The SFLA unsupervised segmentation of image 4 for K=4.

We note that, for H=0.05, 0.1, 1 and 2 the results of unsupervised segmentation for the image 4 using the proposed method, are satisfactory.

To better distinguish the regions that make up the image 4, for K=2 and K=4, we have affected the colors to classes and we calculated the percentage of pixels in each class. Table 6 records the results obtained.

It is noted that, for H=3 and 4, the proposed approach does not show the various regions in the image 4. This is demonstrated in Table 6, where the percentage shows ignorance of some classes by the algorithm, for H=3 and K=4 (class 2, 4) and for H=4 and K=4 (class 3).

According to the results of the unsupervised segmentation for this experiment with the variation of H parameter, we can deduce that: for H varies of 0 to 2, the results of the unsupervised image

segmentation are convincing, by against, if *H* exceeds the value of 2, the resulting image does not reflect the original image.

Classes	<i>H</i> =0.05	<i>H</i> =0.1	<i>H</i> =1	<i>H</i> =2	<i>H</i> =3	<i>H</i> =4
<i>K</i> =2						
Red	42.272950	52.516174	52.516174	50.169373	84.471130	84.471130
Blue	57.733154	47.489930	47.489930	49.836730	15.534973	15.534973
<i>K</i> =4						
Red	29.115295	5.1635740	27.709960	10.444641	39.278107	78.610230
Blue	32.061768	30.047607	23.081970	26.220703	7.39645E-4	12.083435
Green	17.317200	34.651184	35.636900	37.057495	60.723373	0.00305175
Yellow	21.517944	30.149841	13.583374	26.289368	7.39645E-4	9.3154910

Table 6. Pixels percentages in each class of image 4 (unit: %)

It is noted that, for H=3 and 4, the proposed approach does not show the various regions in the image 4. This is demonstrated in Table 6, where the percentage shows ignorance of some classes by the algorithm, for H=3 and K=4 (class 2, 4) and for H=4 and K=4 (class 3).

According to the results of the unsupervised segmentation for this experiment with the variation of H parameter, we can deduce that: for H varies of 0 to 2, the results of the unsupervised image segmentation are convincing, by against, if H exceeds the value of 2, the resulting image does not reflect the original image.

#### 7.2 Experiment 2

In this experiment, we increase the size of image 4 (presented in Fig. 7) to  $520 \times 520$  pixels and use the same initialization parameters shown in Table 5.

For *H*=3 and *H*=4 the various results are shown in Figs. 10 and 11.







Fig. 11. The SFLA unsupervised segmentation of image 4 (520×520) for K=4.

We find that, the results are satisfactory because it can differentiate the regions of image 4. For K=2, the two images of unsupervised segmentation using SFLA are similar.

The results obtained for the percentage of pixels classified in each region of image 4 ( $520 \times 520$  pixels) are summaries in Table 7.

(	Classes	Н=3	<i>H</i> =4
<i>K</i> =2	Red	72.3162	72.3162
	Blue	27.685282	27.685282
<i>K</i> =4	Red	24.866863	39.41679
	Blue	15.7655325	15.6198225
	Green	29.58358	26.716347
	Yellow	29.786982	18.25

Table 7. Pixels percentages in each class of image 4 (520×520 pixels)

So from the experiment showing in Figs. 10 and 11, the proposed algorithm has a good performance in unsupervised segmentation of images (good detection of classes).

We deduce that, the value assigned to the parameter H of the SFLA method depends on the image size to be segmented.

### 8. Conclusions and Future Work

In this paper we have proposed a new unsupervised segmentation approach of images based on SFLA. This last has given satisfactory results compared to those obtained by K-means. The various tests carried have shown that the choice of parameters have a significant impact on the results.

Also, the initialization of H parameter depends on the image size to be segmented and it strongly influences on the segmentation quality. Therefore, it is important to study more detailed manner the choice of parameters.

In perspectives, we propose interesting improvements that can be made to our approach, like for example using multiple objective functions to compare the different results obtained, hybrid with other meta-heuristic can be beneficial to the algorithm and to apply the unsupervised segmentation algorithm based SFLA on other types of images.

## References

- P. Taler and K. Sabo, "Color image segmentation based on intensity and hue clustering," *Croatian Operational Research Review*, vol. 3, no. 5, pp. 375-385, 2014.
- [2] B. U. Shankar, "Novel classification and segmentation techniques with application to remotely sensed images," in *Transactions on Rough Sets VII*. Heidelberg: Springer, 2007, pp. 295-380.

- [3] S. Ameur and Z. Ameur, "Revue des approches de segmentation d'images texturees: exemple des images meteorologique," in *Proceedings of the 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunication (SETIT)*, Tunisia, Algeria, 2005, pp. 1-14.
- [4] D. Guo, V. Atluri, and N. Adam, "Texture-based remote sensing image segmentation," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, Amsterdam, The Nederlands, 2005, pp. 1472-1475.
- [5] A. Halder, S. Pramanik, and A. Kar, "Dynamic image segmentation using fuzzy c-means based genetic algorithm," *International Journal of Computer Applications*, vol. 28, no. 6, pp. 15-20, 2011.
- [6] S. Bansal and D. Aggarwal, "Color image segmentation using CIELab color space using ant colony optimization," *International Journal of Computer Applications*, vol. 29, no. 9, pp. 28-34, 2011.
- [7] S. Bandyopadhyay, S. Saha, U. Maulik, and K. Deb, "A simulated annealing-based multiobjective optimization algorithm: AMOSA," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 3, pp. 269-283, 2008.
- [8] F. Mohsen, M. Hadhoud, K. Mostafa, and K. Amin, "A new image segmentation method based on particle swarm optimization," *International Arab Journal of Information Technology*, vol. 9, no. 5, pp. 487-493, 2012.
- S. Zucker, "Region growing: childhood and adolescence," *Computer Graphics and Image Processing*, vol. 5, no. 3, pp. 382-399, 1976.
- [10] M. Eusuff, K. Lansey, and F. Pasha, "Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization," *Engineering Optimization*, vol. 32, no. 2, pp. 129-154, 2006.
- [11] M. R. Narimani, "A new modified shuffle frog leaping algorithm for non-smooth economic dispatch," *World Applied Sciences Journal*, vol. 12, no. 6, pp. 803-814, 2011.
- [12] D. Tang, Y. Cai, and J. Zhao, "An improved shuffled frog leaping algorithm with single step search strategy and interactive learning rule for continuous optimization," *Journal of Computers*, vol. 9, no. 6, pp. 1300-1308, 2014.
- [13] E. Dominguez and J. Munoz, "Applying bio-inspired techniques to the p-median problem," in *Computational Intelligence Bioinspired Systems*. Heidelberg: Springer, 2005, pp. 67-74.
- [14] A. Agrawal and H. Gupta, "Global K-means (GKM) clustering algorithm: a survey," International Journal of Computer Applications, vol. 79, no. 2, pp. 20-24, 2013.
- [15] J. Xie, S. Jiang, W. Xie, and X. Gao, "An efficient global K-means clustering algorithm," *Journal of Computers*, vol. 6, no. 2, pp. 271-279, 2011.
- [16] M. Farahani, S. B. Movahhed, and S. F. Ghaderi, "A hybrid meta-heuristic optimization algorithm based on SFLA," in *Proceedings of the 2nd International Conference on Engineering Optimization*, Lisbon, Portugal, 2010, pp. 1-8.
- [17] Y. Li and C. Zhang, "A hybrid intelligent optimization algorithm of fast convergence," *International Journal of Hybrid Information Technology*, vol. 8, no. 1, pp. 295-304, 2015.
- [18] J. Zhao and L. Lv, "Two-phases learning shuffled frog leaping algorithm," *International Journal of Hybrid Information Technology*, vol. 8, no. 5, pp. 195-206, 2015.



#### Amel Tehami

She received Engineer degree in Computer Science from the University of Mostaganem and Magister degree from the University of Science and Technology of Oran-Mohamed Boudiaf in 2008 and 2011, respectively. Since September 2011, she is with the Faculty of Computer Science from University of Mohamed Boudiaf as a PhD candidate. Her research interests include image processing, segmentation algorithms development for analysis of spatial data.



#### Hadria Fizazi

She is currently a Professor in the Faculty of Computer Science in the University of Science and Technology of Oran-Mohamed Boudiaf. Engineer degree in Electrical Engineering from the University of Mohamed Boudiaf in 1981, the "Doctor Engineer" degree in Industrial Computer Science and Automation from the University of Lille 1 France in 1987 and the Doctorate degree in Computer Science from the University of Mohamed Boudiaf in 2005. Her research interests cover image processing and the application of pattern recognition to remotely sensed data.