Fast Robust Fuzzy Clustering Algorithm for Grayscale Image Segmentation

Hanane Barrah
Laboratory of Innovative Technologies
National School of Applied Sciences
Tangier, Morocco
Hananbarah@gmail.com

Abdeljabbar Cherkaoui
Laboratory of Innovative Technologies
National School of Applied Sciences
Tangier, Morocco
Cherkaoui.lti@gmail.com

***Abstract*—Image segmentation is a primordial step in the chain of image processing. A wide range of image segmentation methods are based on the FCM algorithm. However, the lack of any spatial information in this latter algorithm makes it very sensitive in the presence of noise. To overcome this problem and faster the segmentation process, we propose two versions of a fast and robust extension of the FCM algorithm.**

To prove the strength of the proposed algorithms we compare them, in a qualitative and quantitative standpoints, and four other fuzzy clustering algorithms. To this end, we use synthetic and real images as testing data.

***Index Terms*—Fuzzy clustering, FCM algorithm, c-means, image segmentation, spatial information.**

I. Introduction

Segmentation is one of the basic tasks in image processing, it is the process of extracting the most relevant information such as points, shapes or regions. In this context, several segmentation methods have been developed [1], [2], they can be divided into three main categories: Contour based, region based and hybrid methods. Each method has its own advantages and drawbacks according to the application domain.

Because of the fuzzy nature of some kind of images, the most useful image segmentation methods are based on the fuzzy clustering approach. FCM, or fuzzy *c-means*, is the well-known and the best-used fuzzy clustering algorithm [3], [4]. The major drawback of this algorithm lies on lack of any spatial information or constraints, which makes it sensitive to noise. To overcome this problem, many researchers have tried to include spatial constraints in many ways. Indeed, Dzung L. Pham [5] proposed an extension of the fuzzy c-means algorithm based on a generalized objective function that includes a spatial penalty term. Although its robustness to noise, this algorithm faces two major problems, the first one is the difficulty to select a parameter that controls the trade-off between minimizing the standard c-means objective function and obtaining smooth membership functions. The second problem is its time requirement. In order to handle noisy images and reduce the effect of intensity inhomogeneity, Mohamed N. Ahmed et al. [6] proposed a modified fuzzy c-means algorithm that uses the neighborhood information to influence the belongingness of each pixel. The authors demonstrated the effectiveness of their algorithm against noise on synthetic and real images. Obviously, including neighborhood information enforces the algorithm to be much more consuming in time. To overcome this latter problem, L. Szilágyi et al. [7] proposed a modification of the previous algorithm [6] by introducing a new factor ($0.5 \leq \gamma \leq 1.2$) which considerably reduces the required calculations. In addition to showing the robustness of this algorithm to noise, the authors demonstrated its quickness against its ancestors.

Our attention is this work is to provide a fuzzy clustering algorithm that is fast and robust to noisy images. To this end, we try to get benefits from two derivatives of the c-means algorithm. We get the fastness from the algorithm proposed by Jiu-Lun Fan et al. [8], where the authors introduced a parameter that controls the trade-off between the fastness of the hard clustering and the good quality of fuzzy clustering. Whereas the robustness to noise is gotten from the algorithm proposed by Songcan Chen and Daoqiang Zhang [9]. In fact, this latter algorithm is a direct extension of [6] and their authors proved that it performs faster and better than its ancestors.

The remainder of this paper is organized as follows: In section II, we present the FCM algorithm with spatial information [9]. The proposed algorithm is described in section III. Some experimental results and comparisons are presented in section IV. Section V is dedicated for some concluding remarks.

II. FCM with Spatial Information

FCM or c-means is an algorithm that consists of grouping data into the most homogeneous groups as much as possible [3], [10] by minimizing iteratively the following objective function:

$$J(D,U,C) = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m} . \left\| x_j - c_i \right\|^2 \qquad (1)$$

Where D is the input data, C is the set of clusters centers, $\| . \|$ is the Euclidean distance, m is the fuzziness exponent and $U = \left[u_{ij}\right]$ is the fuzzy partition matrix that satisfies the following condition:

Xème Conférence Internationale : Conception et Production Intégrées, CPI 2015, 2-4 Décembre 2015, Tanger - Maroc.

Xth International Conference on Integrated Design and Production, CPI 2015, December 2-4, 2015, Tangier - Morocco.

$$\left\{ u_{ij} \in [0,1] \middle| \sum_{i=1}^{c} u_{ij} = 1, \forall j \text{ and } 0 < \sum_{j=1}^{N} u_{ij} < N, \forall i \right\}.$$

The objective function does not include any spatial information, which makes the original FCM algorithm very sensitive in the presence of noise. To overcome this problem Songcan Chen and Daoqiang Zhang [9] proposed a direct modification of Eq. 1. The modified objective function is defined as follows:

$$J(D,U,C) = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m}.\left\| x_j - c_i \right\|^2 + \alpha \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m}.\left\| \bar{x}_j - c_i \right\|^2 \qquad (2)$$

Where α is a parameter that controls the effect of the penalty term. As the standard FCM, the clustering is carried out by minimization of the objective function presented in Eq. 2 under the same condition of FCM. The membership values and the clusters centers are updated as follows:

$$u_{ij} = \frac{\left(\left\| x_j - c_i \right\|^2 + \alpha \left\| \bar{x}_j - c_i \right\|^2 \right)^{-\frac{1}{(m-1)}}}{\sum_{k=1}^{C} \left(\left\| x_j - c_k \right\|^2 + \alpha \left\| \bar{x}_j - c_k \right\|^2 \right)^{-\frac{1}{(m-1)}}} \qquad (3)$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} \left(x_j + \alpha \bar{x}_j \right)}{(1+\alpha) \sum_{j=1}^{N} u_{ij}^{m}} \qquad (4)$$

$\bar{x}_j$ can be chosen as the mean or the median of the neighbors within a specified window around x_j. When α is set to zero, the algorithm is equivalent to the original FCM, while it approaches infinite, the algorithm acquires the same effect as the original FCM on the mean or median filtered image, respectively. Thus, this new objective function (Eq. 2) leads to two algorithms that will be noted in this manuscript by FCM_S1, uses the mean values, and FCM_S2, uses the median values.

Algorithm1

- Step 0. Fix the clustering parameters (the converging error ε , the fuzziness exponent m and the number of clusters C) and initialize the clusters centers.
- Step 1. For FCM_S1 (FCM_S2) compute the mean (median resp.) filtered image.
- Step 2. Update the partition matrix using (Eq. 3).
- Step 3. Update the clusters centers using (Eq. 4).

Repeat steps 2-3 until the following criterion is satisfied:

$$\left\| C_{new} - C_{old} \right\| < \varepsilon$$

III. Fast and Robust FCM Algorithm

The modifications in the updating equations (Eq. 3 and Eq. 4) alter the algorithm speed in a negative way. To deal with this problem, we have adopted the idea of Jiu-Lun Fan et al. [8]. Actually, we modified the previous algorithms, FCM_S1 and FCM_S1, by introducing a parameter γ that controls the trade-off between the fastness of the hard clustering and the robustness of FCM_S1/S2 to noise. the idea behind Jiu-Lun Fan's algorithm is to prize the biggest membership and suppress the others.

Let x_j be a pixel and u_{bj} be its degree of belongingness to the b[th] cluster. If u_{bj} is the biggest value of all the clusters, then the membership degrees of x_j will be modified as follows:

$$u_{bj} = 1 - \gamma \sum_{i \neq b} u_{ij} = 1 - \gamma + \gamma u_{bj} \qquad (5)$$

$$u_{ij} = \gamma u_{ij} \text{ , } i \neq b \qquad (6)$$

Where $\gamma \in [0,1]$.

When γ gets closer to 0, the algorithm becomes more hard and when it approaches 1 the algorithm tends to the fuzzy version.

This modification has to be done immediately after updating the fuzzy partition matrix. Thus, we come up with two algorithms that are slightly different from the previous ones, FRFCM1 (FRFCM2) that uses the mean (median resp.) filtered image.

Algorithm2

- Step 0. Fix the clustering parameters (the converging error ε , the fuzziness exponent m and the number of clusters C) and initialize the fuzzy partition matrix and the new parameters α and γ .
- Step 1. For FRFCM1 (FRFCM2) compute the mean (median resp.) filtered image.
- Step 2. Update the clusters centers using (Eq. 4).
- Step 3. Update the partition matrix using (Eq. 3).
- Step 4. Modify the partition matrix using Eq. 5 and Eq. 6.

Repeat steps 2-4 until the following criterion is satisfied:

$$\left\| U_{new} - U_{old} \right\| < \varepsilon$$

IV. Experimental Results

To demonstrate the strength of our resulting algorithms, in both running times and accuracy standpoints, we compare them and four other fuzzy clustering algorithms: The standard FCM, the algorithm proposed in [8] (S_FCM), FCM_S1 and FCM_S2. To do so, we test the six algorithms upon synthetic and real data.

A. Synthetic data

The synthetic data is an image of 250X250 pixels divided into three clusters with three intensity values taken as 0, 200 and 255. We test the algorithms when this synthetic image is corrupted by Gaussian and 'Salt and pepper' noises respectively. The results are shown in Fig. 1 and Table I. The clustering parameters are initialized as follows:

$c=3, \varepsilon=10^{-8}, m=2, \alpha=5$ and $\gamma=0.5$

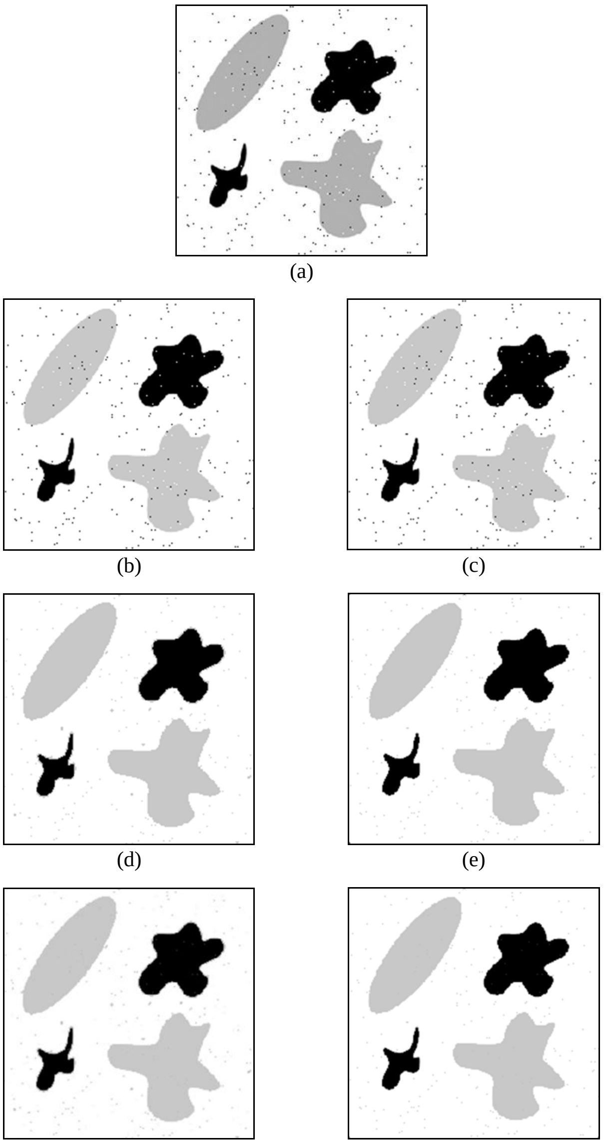


Fig. 1.Segmentation results on the synthetic image. (a) The original image corrupted by salt and pepper noise. (b) FCM result. (c) S_FCM result. (d) FCM_S_1 result. (e) FCM_S_2 result. (f) $FRFCM_1$ result. (g) $FRFCM_2$ result.

TABLE I. SEGMENTATION ERRORS AND RUNNING TIMES OF SIX ALGORITHM

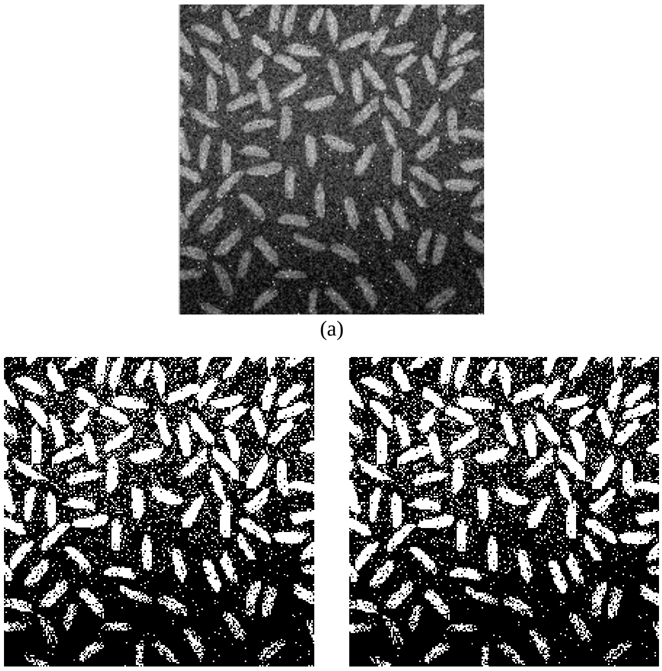
	Salt and pepper noise		Gaussian noise	
	Segmentation Error	***Running Time***	***Segmentation Error***	***Running Time***
FCM	5.844	1.221	7.503	3.603
S_FCM	5.844	0.907	7.393	2.281
FCM_S1	5.096	0.63	4.919	0.653
FCM_S2	**4.822**	0.478	**4.72**	0.716
FRFCM1	5.095	0.650	4.917	0.688
FRFCM2	**4.822**	**0.472**	**4.72**	**0.609**

From Fig. 1 and Table I, we note that the segmentation results of FRFCM1 (FRFCM2) and FCM_S1 (FCM_S2 resp.) are similar and superior to those of the FCM and S_FCM. It is remarkable that FRFCM2 and FCM_S2 are more robust to "salt and pepper" noise than FRFCM1 and FCM_S_1, this is due to the usage of the median filtered image.

In a running time point of view, and from the values depicted in Table I, we find out that our algorithms, FRFCM1 and FRFCM2, are the fastest.

B. Real data

In this sub-section we test the previous algorithms upon a real image [11] corrupted at the same time by Gaussian and "salt and pepper" with the same clustering parameters as fixed before. The segmentation results and the running times are presented in Fig. 2 and Fig. 3 respectively.



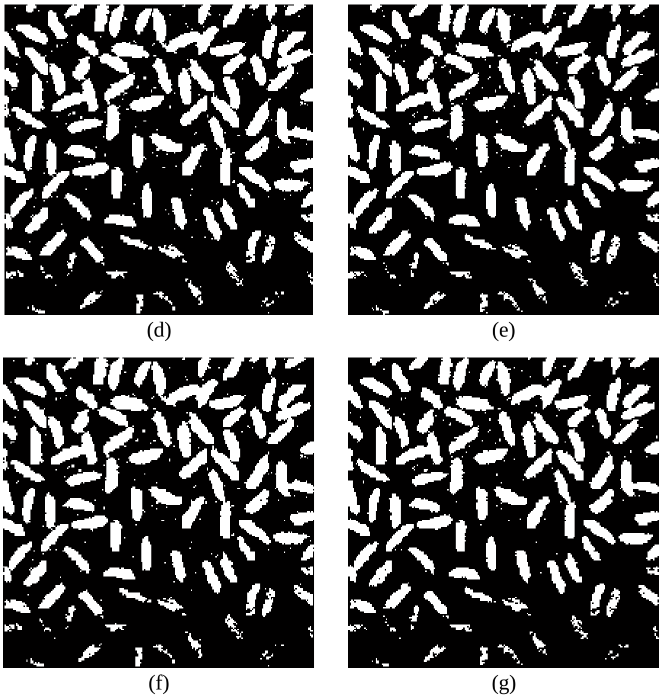


Fig. 2. Segmentation results on real image. (a) The original image corrupted by Gaussian and "salt and pepper" noise. (b) FCM result. (c) S_FCM result. (d) FCM_S_1 result. (e) FCM_S_2 result. (f) $FRFCM_1$ result. (g) $FRFCM_2$ result.

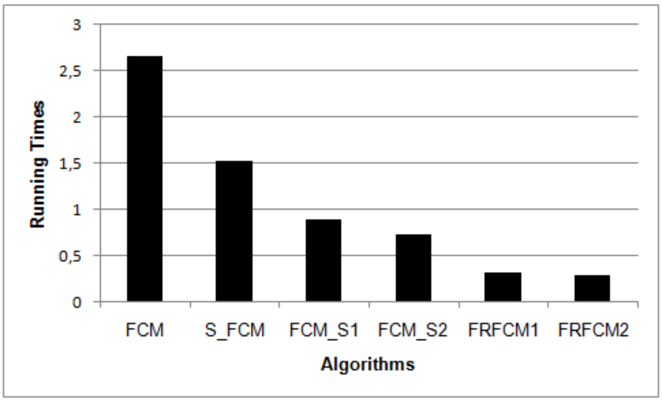


Fig. 3. Running times of the all algorithms to segment the real image.

From Fig. 2, we notice that the algorithms FCM and S_FCM are very sensitive to noise. However, the four other algorithms FCM_S1, FCM_S2, FRFCM1 and FRFCM2 have succeeded to handle noise and their segmentation results are very close.

The results depicted in Fig. 3 show that our algorithms FRFCM1 and FRFCM2 are the fastest. Which prove their effectiveness over the other algorithms.

V. CONCLUSION

To provide a fast and robust fuzzy clustering algorithm for image segmentation, we merged the advantages of two fuzzy clustering algorithms, FCM_S1/2 and S_FCM. By testing the resulting algorithms upon synthetic and real data, we concluded that they outperform their ancestors.

It is worth mentioning that the selection of α and γ plays an important role in the segmentation accuracy. In fact, a reasonable choice of (α, γ) leads to an accurate segmentation in a reasonable time. Thus, the selection of the couple (α, γ) is left as perspective

REFERENCES

[1] S. Kumar and R. Srinivas, "A Study on Image Segmentation and its Methods," Int. J., vol. 3, no. 9, 2013.

[2] R. P. NIKHL and K. P. SANKAR, "A Review on Image Segmentation Techniques." 1993.

[3] L. Rokach and O. Maimon, "Clustering methods," in Data mining and knowledge discovery handbook, Springer, 2005, pp. 321–352.

[4] S. Ghosh and S. K. Dubey, "Comparative Analysis of K-Means and Fuzzy C-Means Algorithms," IJACSA, vol. 4, pp. 35–38, 2013.

[5] D. Pham, "Spatial Models for Fuzzy Clustering," Comput. Vis. Image Underst., vol. 84, no. 2, pp. 285–297, Nov. 2001.

[6] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data," Med. Imaging IEEE Trans. On, vol. 21, no. 3, pp. 193–199, 2002.

[7] L. Szilagyi, Z. Benyo, S. M. Szilágyi, and H. S. Adam, "MR brain image segmentation using an enhanced fuzzy c-means algorithm," in Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, 2003, vol. 1, pp. 724–726.

[8] J.-L. Fan, W.-Z. Zhen, and W.-X. Xie, "Suppressed fuzzy c-means clustering algorithm," Pattern Recognit. Lett., vol. 24, no. 9–10, pp. 1607–1612, Jun. 2003.

[9] S. Chen and D. Zhang, "Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel-Induced Distance Measure," IEEE Trans. Syst. Man Cybern. Part B Cybern., vol. 34, no. 4, pp. 1907–1916, Aug. 2004.

[10] D. Lam and D. C. Wunsch, "Chapter 20 - Clustering," in Academic Press Library in Signal Processing, vol. Volume 1, J. A. K. S., Rama Chellappa and Sergios Theodoridis Paulo S.R. Diniz, Ed. Elsevier, 2014, pp. 1115–1149.

[11] Mathworks, Natick, MA. Image Processing Toolbox. [Online] Available: htttp://www.mathworks.com