

Extended Fuzzy C-Means Clustering Algorithm in Segmentation of Noisy Images

Omprakash Dewangan¹, Prof. Asha Ambhaikar²

¹Rungta College of Engineering & Technology
Bhilai, Chhattisgarh, India
dewangan.omprakash@gmail.com

²Rungta College of Engineering & Technology
Bhilai, Chhattisgarh, India
asha31.a@rediffmail.com

Abstract: A major problem in noisy image processing is the effective segmentation of its components. In this work, we are proposing an Extended Fuzzy C means clustering algorithm for noisy image segmentation, which is able to segment all types of noisy images efficiently. As the presented clustering algorithm selects the centroids randomly hence it is less sensitive, to any type of noise as compare to other clustering algorithms. And we will try to prove that Extended Fuzzy C means clustering converges to approximate the optimal solution based on this criteria theoretically as well as experimentally. Here we will also compare the efficiency of available algorithm for segmentation of gray as well as noisy images.

Keywords: FCM, Clustering, Image Segmentation, Image noise, EFCM.

1. Introduction

1.1 Image Segmentation

Image segmentation [1], [2] as an important research area in Digital Image Processing is extremely large, often depending on a variety of techniques from a wide range of other mathematical fields.

Image segmentation is an essential step in series of processes designed at on the whole image understanding and used for computer vision and not for human being interpretation and the main goal of segmentation is to partition of an image. Image segmentation subdivides an image into its constituent regions or object with respect to a specific application. In segmentation, the computer scans the image in the form of various regions and image may be grey scale, color, texture, depth or motion. Image segmentation can be achieved in any of the two approaches: Segmentation based on discontinuous in intensity (Point Detection, Line Detection and Edge Detection) and Segmentation based on similarities in intensity (Region growing, Region splitting, Region merging and Split and merge). So image segmentation is the fundamental step in most regular pictographic pattern recognition and scene analysis applications. In image processing, two terms are usually seen very close to each other: clustering and segmentation. When analyzing the color information of the image, e.g. trying to separate regions or ranges of color components having same characteristics, the process is called color clustering. Mapping the clusters onto the spatial domain and physically separated regions in the image is called segmentation.

Segmentation implies the division of an image into different objects or connected regions that do not overlap. Thus, the union of all the regions is the image itself. A region often has a similar intensity or a discrete boundary.

1.2 FCM Clustering Algorithm

Fuzzy c-means (FCM) is developed by Dunn in 1973 and it is enhanced by J. C. Bezdek in 1981 and it is commonly used in pattern recognition [9]-[11]. It is based on fuzzy algorithm which is used for image segmentation and it allows data points to be assigned into more than one cluster and each one data point has a level of relation of belonging to each one cluster.

In fuzzy clustering [3], data elements can belong to more than one cluster, and connected with each element is a set of membership levels. These designate the strength of the involvement between that data element and a meticulous cluster. Fuzzy clustering is a method of passing on these membership levels, and then using them to allocate data elements to one or more clusters. It allows each attribute vector to belong to multiple clusters with various membership degrees (between 0 and 1) and boundaries of fuzzy between clusters. The numbers of fuzzy clustering algorithms are available in modern era and fuzzy clustering is one of best approach for solving the problem of color image segmentation.

2. Image Noise

Image noise is the casual dissimilarity of intensity or color information in images formed by the sensor and circuitry of a scanner or digital camera. It can be defined as any undesired information that contaminates an image and it appears in images from an array of sources. Image noise can also generate in film particle and in the necessary shot noise of an ideal photon detector [4]. Image noise is commonly regarded as an unwanted by-product of image capture. Noise is desirable electromagnetic energy that degrades the data and signal quality. Noise occurs in analog and digital system and it can be affect data, files and communications of all types of applications such as text, images, audio, video,

handwritten characters, fingerprints, and pictures. Images are affected by different types of noise:

2.1 Gaussian Noise

It occurs from electronic noise in image acquisition system. It is statistical noise that has a probability density function of the typical distribution. The values that the noise can capture on are Gaussian distributed. It is commonly used as additive white noise (stable power scale). It is also form natural noise process like electronic noise, occurs during image acquisition.

2.2 Poisson Noise

Poisson noise is also called shot noise. It has a probability density function of a poisson distribution.

2.3 Salt and Pepper Noise

It is also called impulse noise, e or spike noise. It represents itself as randomly occurring white and black pixels. It is an effective noise reduction method for this type of noise involves the behavior of a low and high pass filter.

2.4 Speckle Noise

Speckle noise is a rough noise that naturally exists in and degrades the image quality. It is a multiplicative noise. It is in direct proportion to the local grey level in any area. The signal and the noise are statistically independent of each other.

3. Clustering

Clustering is a kind of unsupervised learning. Clustering is a method of grouping data that share similar trend and patterns. Clustering of data is a method by which large sets of data are grouped into clusters of smaller sets of similar data. We can show this with a simple graphical example:

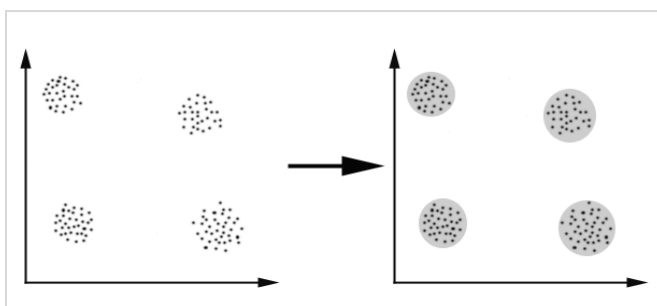


Figure 1. Graphical form of clustering

Clustering [9]-[11] is a mathematical tool that attempts to discover structures or certain patterns in a data set, where the objects inside each cluster show a certain degree of similarity.

The purposes of cluster are to:

- i. Group data points those are similar to each Other.
- ii. Identify such clusters in an unsupervised manner.
- iii. Unsupervised: no one information is provided to the algorithm on which data points belong to clusters.

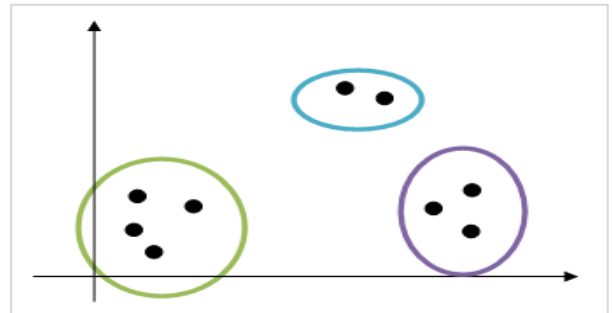


Figure 2. Graphical form of clustering

4. FCM

The Fuzzy c means (FCM) [5] is clustering algorithm which is mostly used in computer science and engineering such as digital image processing, real image, medical image, pattern recognition, bioinformatics science and data mining techniques.

- Minimize

$$E(U, V) = \sum_{i=1}^k \sum_{j=1}^n (u_{ij})^m \|x_j - v_i\|^2 \quad (1)$$

Subject to

$$\sum_{i=1}^k u_{ij} = 1 \quad \forall_j = 1, \dots, n \quad (2)$$

- Lagrangian multiplier is used to solve constrained optimization problem?

$$L_j(U, V) = \sum_{i=1}^k \sum_{j=1}^n (u_{ij})^m \|x_j - v_i\|^2 + \partial_j \left(\sum_{i=1}^k u_{ij} - 1 \right) \quad (3)$$

- Iterative optimization

o Fix V, optimize w.r.t. U

$$u_{ij} = \frac{1}{\sum_{i=1}^c \left(\frac{\|x_j - v_i\|}{\|x_j - v_i\|} \right)^{\frac{2}{m-1}}} \quad (4)$$

- Fix U, optimize w.r.t. V

$$v_i = \frac{\sum_{j=1}^n (u_{ij})^m x_j}{\sum_{j=1}^n (u_{ij})^m} \quad (5)$$

5. FCM clustering with density

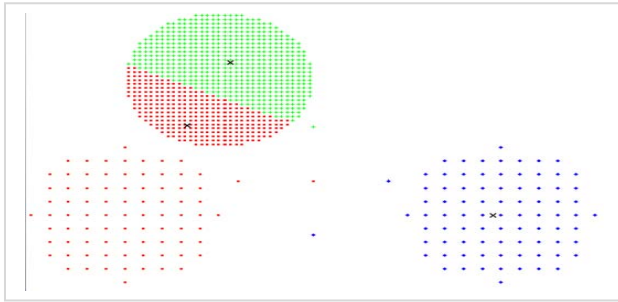


Figure 3. FCM clustering with varying density

The [9]-[11] higher density cluster attracts all other cluster prototypes so that the prototype of the right cluster is slightly drawn away from the original cluster center and the prototype of the left cluster migrates completely into the dense cluster.

6. The Proposed Technique, Extended Fuzzy C Means (EFCM)

FCM is a good segmentation technique if there is no noise in the image. Keeping this in view, we are proposing a very simple and effective method, which preprocess noisy pixels in the image and then apply conventional FCM algorithm for segmentation.

The pixels in an image are highly correlated i.e. the pixels in the immediate neighborhood possess nearly the same feature data. In the propose technique, we are trying to regenerate the value of noisy pixels by processing every pixel of an image based upon its immediate neighborhood pixels. We are considering an image in RGB format. Let size of an image is $R \times C \times K$, where R is the number of rows, C is the number of columns, and K is 3 (for color image). So every pixel will be processed using equation:

$$pixel(l, m, k) = \sum_{k=1}^3 \frac{1}{N} \left(\left(\sum_{i=l-1}^{i+1} \sum_{j=m-1}^{m+1} pixel(i, j, k) \right) - pixel(l, m, k) \right) \quad (10)$$

Where

$N = 8$, denoting 8 neighborhood points of a study pixel.

Pixel (i, j, k) = intensity value of pixel at i, j, k location.

Pixel (l, m, k) = intensity value of study pixel.

In the homogeneous region, this function simply fortifies the same characteristics and the clustering result remains unchanged. However, for a noisy pixel, this formula tries to regenerate the value of original pixel. As a result, misclassified pixels from noisy regions can easily be corrected. After processing the noisy pixels, FCM is used to do segmentation. Here, we are processing the image only once before segmentation, which does not increase the complexity of the program compare to Spatial FCM and other robust clustering algorithms for image segmentation. Other methods mostly changed the standard FCM equations and introduced new functions which are calculated for every pass of the FCM algorithm i.e. they are extremely

computationally intensive so the strength of the proposed technique is its least execution time, which is approximately equal to the conventional FCM method.

7. Result

Experiments are implemented and simulated using MATLAB Version 7.0 (Only for inserting the noise).



Figure 4. Original picture

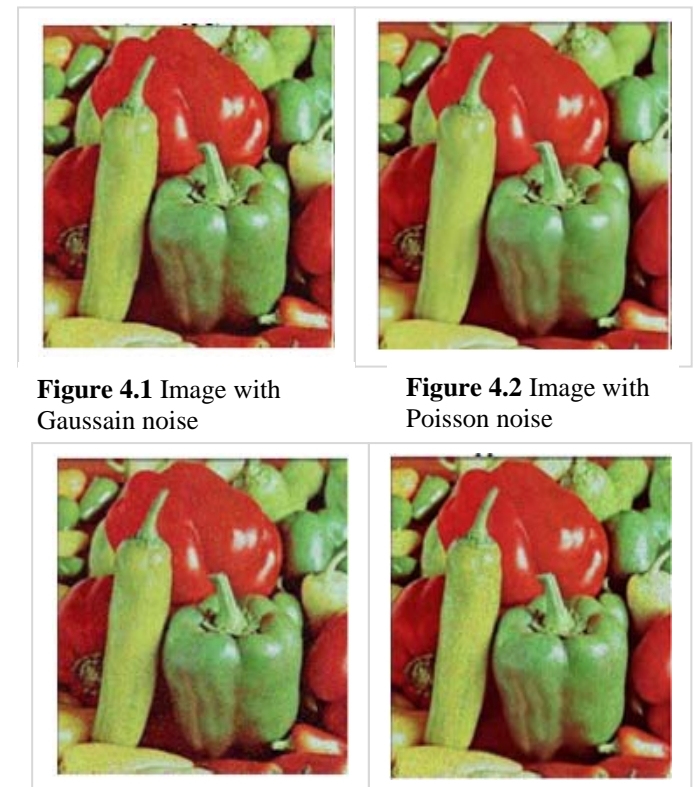


Figure 4.1 Image with Gaussain noise

Figure 4.2 Image with Poisson noise

Figure 4.3 Image with Salt and Pepper noise

Figure 4.4 Image with Speckle noise

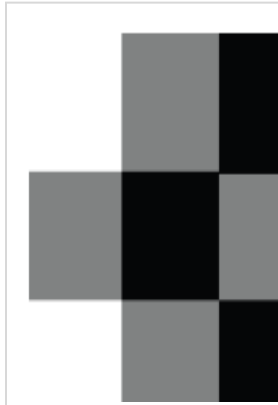


Figure 5.1 Original Image

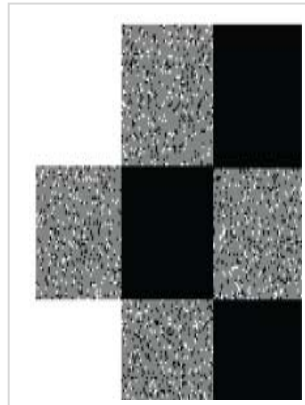


Figure 5.2 Image Noisy

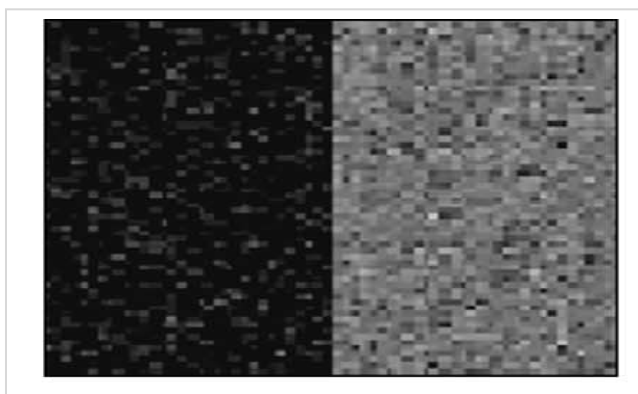


Figure 6. Original Image Homogenous intensity corrupted by 5% Gaussian noise)

8. Conclusion

FCM is a popular segmentation method for digital images. However, it is an intensity-based clustering algorithm which is not robust against noisy images. The pixels in an image are highly correlated and this spatial information is an important characteristic for color image segmentation which is not mentioned in previous work, we proposed an extended FCM (EFCM) algorithm which preprocesses the image before segmentation. Preprocessing of image is influenced by the direct eight neighborhood pixels of every pixel of an image under consideration. Its strength is its least execution time and fast convergence rate compared to other image segmentation techniques and clustering algorithm for noisy segmentation, which is able to segment all types of noisy images efficiently.

References

- [1] S.Pradeesh Hosea, S. Ranichandra, and T.K.P.Rajagopal, "Color Image Segmentation – An Approach," Color Image Segmentation – An Approach, vol. 2, no. 3, March 2011.
- [2] R. Rajesh N. sethikumar " A Note on Image Segmentation Techniques, " International J. of Recent trends in Engineering and Technology, vol. 3 no. 2 May 2010.

- [3] Nock, R. and Nielsen, F. (2006) "On Weighting Clustering", IEEE Trans. on Pattern Analysis and Machine Intelligence, 28 (8), 1–13
- [4] Manoj Gupta, Sumit Srivastava, Ashok Kumar Nagawat Pawan Patidar, "Image De-Noising By Various Filters For Different Noise," International Journal Of Computer Applications, Vol. 9, November 2010.
- [5] J.C. Bezdek (1981), "Pattern Recognition with Fuzzy Objective Function Algorithm", Plenum, NY.
- [6] F.C.H. Rhee, C. Hwang, A Type-2 fuzzy c means clustering algorithm, in: Proc. in Joint 9th IFSA World Congress and 20th NAFIPS International Conference 4, 2001, pp. 1926–1929.
- [7] Chuang Keh-Shih, Tzeng Hong-Long, Chen Sharon et al. (2006), "Fuzzy C-means clustering with Spatial Information for Image Segmentation," Computerized Medical Imaging and graphics, 30(2006) 9-15.
- [8] S.Pradeesh Hosea, S. Ranichandra, and T.K.P.Rajagopal,"Color Image Segmentation – An Approach," Color Image Segmentation – An Approach, vol. 2, no. 3, March 2011.
- [9] J. C. Dunn (1973): "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters", Journal of Cybernetics 3: 32-57.
- [10] J. C. Bezdek (1981): "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New.
- [11] A. K. Jain and M. N. Marty and P. J. Flynn, Data clustering: a review, ACM Computing Surveys, 31:3, pp. 264 - 323, 1999.
- [12] T. R. Golub et. al, Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring, Science, 286:5439, pp. 531 – 537, 1999.
- [13] Gasch, A.P. and Eisen, M.B. (2002) Exploring the conditional coregulation of yeast gene expression through fuzzy k-means clustering. Genome Biol., 3, 1–22.
- [14] M. Eisen et. Al, Cluster Analysis and Display of Genome-Wide Expression Patterns. Proc Natl Acad Sci U S A 95, 14863-8, 1998.