TOWARD A GENERIC PERFORMANCE EVALUATION OF IMAGE SEGMENTATION USING FUZZY C-MEANS, K-MEANS AND K-MEANS USING OTSU THRESHOLDING

T.Sam Pradeepraj[1], V.Anusuya Devi[2]

Associate Professor[1,2], Department of Computer Science and Engineering, School of Computing, Kalasalingam Academy of Research and Education, Krishnankoil, Tamil Nadu, India

sampradeepraj@gmail.com,samanusuya23508@gmail.com

Abstract: The goal of image segmentation is the splitting of an image into a set of disjoint regions with uniform and homogeneous attributes such as intensity, color, tone or texture etc. In many real locales, for images, issues such as limited spatial resolution, indigent contrast, overlapping intensities, noise and intensity inhomogeneities introduce fuzziness in the object boundaries in the image. Due to this the fuzzy set theory was proposed. This work consists of three segmentation algorithm such as fuzzy c-means, k-means and k-means using Otsu thresholding. The proposed work's aim is to evaluate the performance of above three image segmentation algorithms on synthetic images. Experiments show that the K means using Otsu thresholding algorithm effectively segments the given image other than two algorithms

Keywords: fuzzy c-means, k-means and Otsu thresholding.

1. INTRODUCTION

Computer algorithms are applied to digital images to perform image processing. Image segmentation refers to the process of partitioning a digital image into multiple regions (set of pixels). Segmentation is used to change the image into a meaningful thing for further work[5]. Image segmentation is used in various applications: 1. Medical Imaging are: Locate tumours, Measure tissue volumes, Computer guided surgery, Diagnosis, Treatment Planning, Study of anatomical structures, 2. Locate objects in satellite images (roads, forests, etc.), 3. Face recognition, 4. Fingerprint recognition, 5. Automatic traffic controlling systems, 6. Machine vision.

For all applications, a single method cannot produce the desired result because images have different properties and some other factors like noise, brightness, colour, shape and size etc. Most of the image segmentation algorithms are based on two characters of pixel grey level:

(i) Discontinuity - The approach is based on a portion an image based on an abrupt change in intensity. For example, Gray level discontinuities – points, lines, edges. (ii) Similarity - The approach is based on portioning an image into regions that are similar according to a set of predefined criteria.

There are five main methods of image segmentation: Cluster-based method, Thresholding based method, Edge detection method, Region-based method, Graph based method. But this paper focuses on Cluster-based method and Thresholding-based method.

i) Cluster-based method- Clustering refers to the process of grouping samples so that the samples are similar within each group. The groups are called clusters [2]. Clustering is a technique used in statistical data analysis, data mining, pattern recognition, image analysis etc. The clustering based method is further divided into three classes: Hierarchical Clustering, Overlapping(fuzzy c-means) clustering, and Exclusive(k-means) clustering. ii) Thresholding-based method- Thresholding is the operation of converting a multilevel image into a binary image, i.e. Based on the comparison of a threshold value T, assigns the value of 0(background) or 1(objects) to each pixel of an image. When T is constant, the approach is called global Thresholding. Threshold selection is typically done iteratively. However, it is possible to derive an automatic threshold selection algorithm

1.1 MOTIVATION AND JUSTIFICATION

Image segmentation is used to classify the pixels of an image in decision-oriented applications. The clustering method, the unsupervised classification of patterns into groups, is one of the most important tasks in exploratory data analysis. There is a number of clustering algorithms that could be interested in finding representatives of homogeneous groups (data reduction). This work investigates and evaluates the performance analysis of three image segmentation.

Algorithms such as Fuzzy c-means clustering, K- means clustering and K-means Otsu thresholding, by comparison, using the performance metrics such as PSNR, MSE, segmentation accuracy and elapsed time. The primary goals of each of the above three clustering algorithms include gaining insight into data (detecting anomalies, identifying salient features, etc.), grouping data, and compressing data. The motivation of this work is to find which one of the above three algorithms is really the best and most efficient clustering algorithm for image segmentation. Hoel Le Capitaine and Carl Frélicot presented a novel, A fast fuzzy c-means algorithm for colour image segmentation [6]. This algorithm introduces a fuzzy iterative algorithm allowing fast segmentation of images. Anil Z Chitade and Dr S.K. Katiyar presented the Colour Based Image Segmentation Using K-Means Clustering [3]. This procedure follows a simple and easy way to classify a given data set through a certain number of clusters. Vijay Jump et al. presented Color Image Segmentation Using K-Means Clustering and Otsu's Adaptive Thresholding [17]. This algorithm segments the image by automatically selecting the threshold value based on the largest inter-class variance between the target and background.

Most of the researchers have analyzed the performance of the existing clustering techniques [9],[15],[16],[18]. Till now, no comparative analysis has been done for clustering methods over thresholding techniques. This paper concentrates on a comparative analysis of clustering methods over thresholding techniques..

1.2 OUTLINE OF THE PROPOSED WORK

The outline of this paper is depicted in fig 1.1. This section describes the overall methodology of image segmentation using clustering algorithms such as fuzzy c-means, k-means and k- means using Otsu thresholding techniques using performance metrics (Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR)) [13], Segmentation Accuracy and Elapsed Time to evaluate above three algorithms. Before segmenting the given input image, some pre-processing method is used to remove noise from the input image. The median filter is a nonlinear digital filtering technique often used to detach noise. Noise reduction is a pre-processing step to improve the results of later image processing.

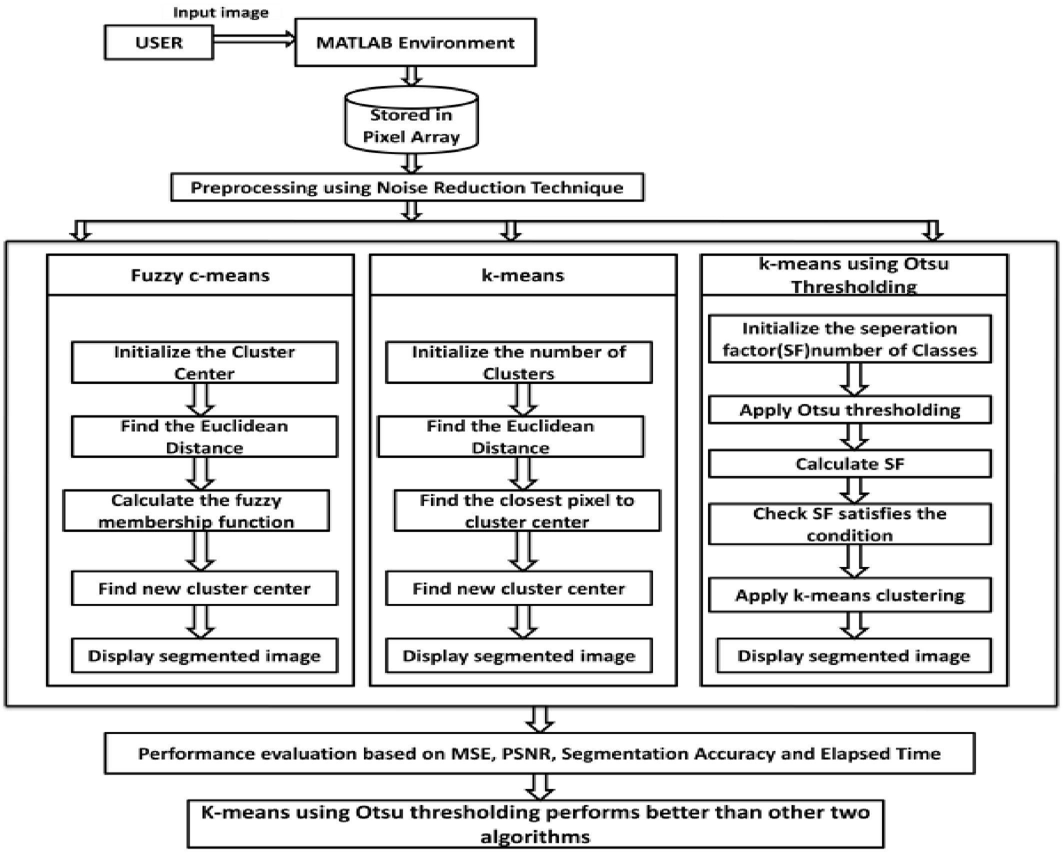

Figure 1.1 Outline of the proposed work

1.3 ORGANIZATION OF THE CHAPTER

This paper is organized as follows. Section 2 describes the fuzzy c-means clustering, k-means clustering and k-means using Otsu thresholding algorithms. Section 3 explains about performance evaluation of algorithms as mentioned above. Finally, the conclusion and perspectives are drawn in section 4.

2. ALGORITHMS

2.1Fuzzy C-Means Clustering Algorithm

Fuzzy C-means algorithm was developed by Dunn in 1973 and improved by Bezdek in 1981. A single piece of data is clustered into two or more clusters using Fuzzy c-means (FCM) [1],[6]. FCM is a famous method in pattern recognition. This algorithm works by assigning membership to each data point corresponding to each cluster centre on the basis of the distance between the cluster and the data point [7],[8]. More data is near the cluster centre more is its membership towards the particular cluster centre. An aggregate of the membership of each data point should be equal to one. Different membership coefficients of two or more clusters are framed based on the fuzzy c-means algorithm.

Fuzzy C-Means clustering is an iterative process [10]. First, the initial fuzzy partial matrix is generated, and the initial fuzzy cluster centres are calculated. At each step of the repetition, the best location of the cluster finding process is minimized, and cluster centres and membership grade points are updated. This process ends when the maximum number of repetitions is reached or when the objective function improvement between two consecutive repetitions is less than the minimum amount of progress specified.

Moreover, the update in the iteration is done using the membership degree and the cluster's centre, the two-parameter change. Steps are repeated once the threshold is reached. Once the maximum number of repetitions is reached, the process stops. From this, the improvement of the objective function between two consecutive iterations is less than the minimum improvement specified. In addition, a fuzziness coefficient 'm' is chosen, which may be any actual number greater than 1 [11],[12].

The algorithm (fig 2.1) is comprised of the following steps:

Step 1: Read the input image in the Matlab environment.

Step 2: Either randomly or fixedly initializes the centre or centroid of the cluster.

Step 3: Generating distance matrix from centroids to each data point in an image using euclidean distance.

Step 4: Generate membership matrix using inverse distance weighting.

Step 5: The group matrix can be obtained based on the membership matrix.

Step 6: Knowing the members of each group new centroid of each group can be obtained.

Step 7: This process continues until there is no change in centroid points.

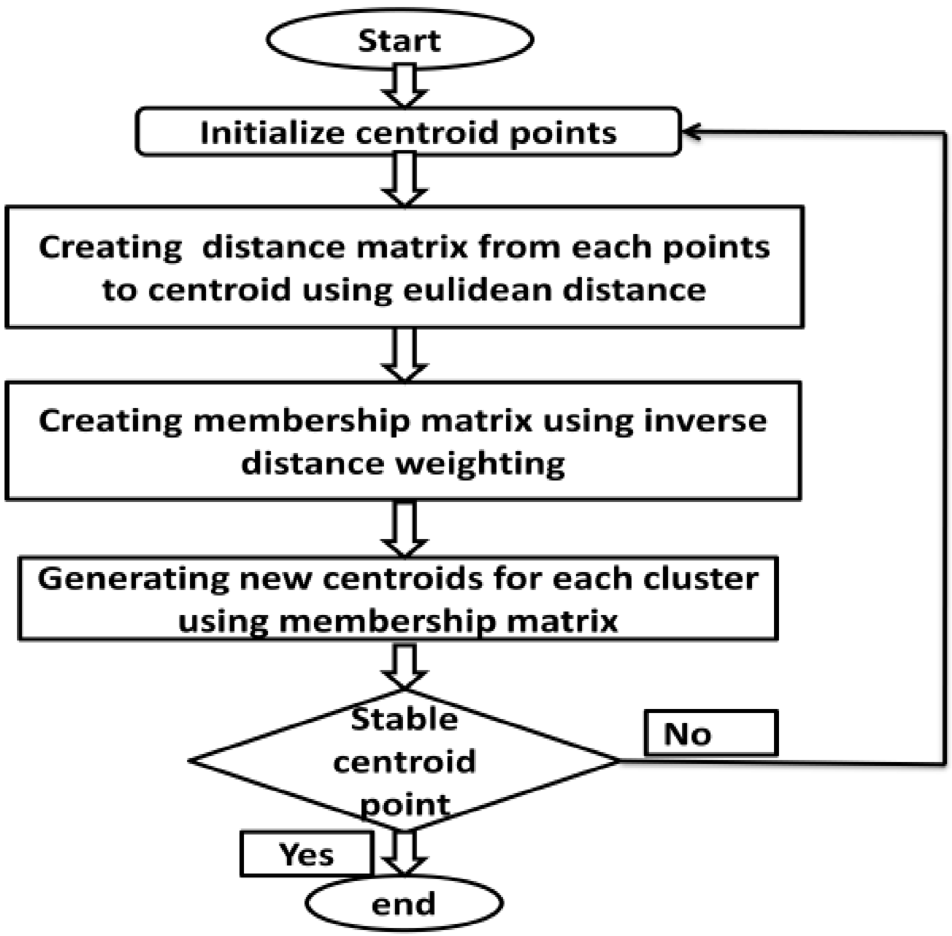

Figure 2.1 Flowchart of Fuzzy c means clustering algorithm

2.2 K-Means Clustering

Well, known clustering problem is solved with the help of K-means unsupervised learning algorithms(Macqueen, 1967) [3]. It can follow the simple and easy way to classify a given data set through a certain number of clusters fixed a priori. Disjoint clusters are generated with the help of k centroids. The low computational complexity algorithm is a K-means clustering algorithm compared to FCM[16]. A cluster's closeness does not always involve the 'centre' of clusters because every cluster member is closer to its cluster.

The algorithm (fig 2.2) is comprised of the following steps:

Step 1: Read the input image in the MATLAB environment.

Step 2: Randomly select the number of clusters and initialize the cluster centres.

Step 3: Compute the distance matrix using Euclidean distance.

Step 4: Generate a group matrix based on the distance matrix.

Step 5: Obtain new centroids by grouping the members of each cluster.

Step 6: This process continues until it reaches stable centroids.

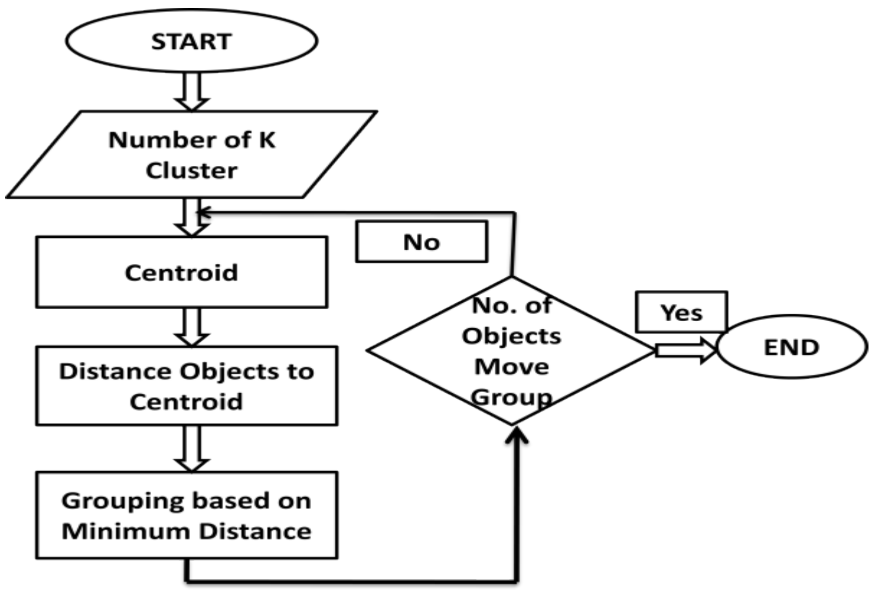

Figure 2.2 Flowchart of K-means clustering algorithm

2.3 K-Means Using Otsu Thresholding

Otsu thresholding (Christopher Mei, Anthony Joshua, Tony Collins) based segmentation is one of the oldest and most powerful techniques since the threshold value divides the pixels. Pixel's intensity value less than the threshold belongs to one class, while pixels whose intensity value is more significant than the threshold belongs to another class [17].

Region-based methods divide an image into different regions that are similar according to a set of some predefined conditions. Without operator intervention, An unsupervised segmentation method automatically partitions the images without [4]. An unsupervised segmentation can partition the images without operator intervention [14].

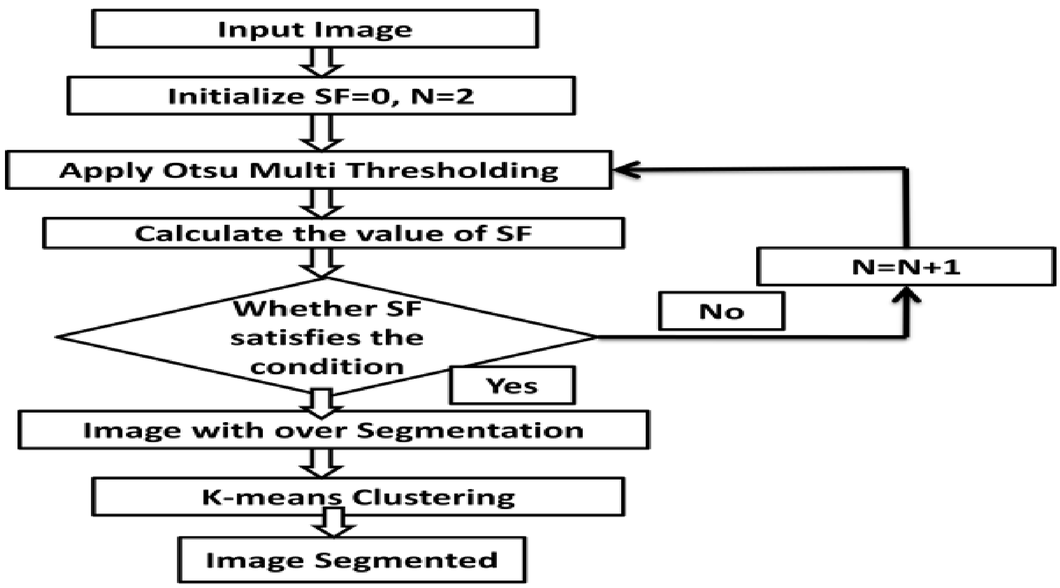

Figure 2.3 Flowchart of K-means using Otsu thresholding algorithm

The algorithm (fig 2.3) is comprised of the following steps:

Step 1: Consider the grayscale image as an input for this system.

Step 2: Initialize the Separation Factor (SF=0) and N=2, where N implies a number of classes. Otsu's thresholding is used to segment an image using the Separation Factor. The value of SF remains between 0 and 1. The higher value of SF implies that the image has been segmented absolutely.

Step 3: Otsu proposed the concept of the maximum classes' variance method. Otsu's method is greatly used to segment the image for a reason being simple calculation, less time-consuming and effective. This method segments the image by automatically selecting the threshold value based on the largest inter-class variance between the target and background. Otsu method selects the optimal threshold t by maximizing the between-class variance, as the total variance, which is the sum of the within-class variance and the between-class variance, remains constant for a given image.

Step 4: Otsu's method determines the value of SF. If the value of SF tends to 1, then that image is segmented; otherwise, the number of classes is Increased by 1, i.e. N=N+1, and again Otsu's thresholding is applied to this new value of N.

Step 5: The output of Otsu's thresholding may lead to over-segmentation. Hence this method needs some technique to merge the over-segmented regions. K- means clustering. This is a partitioning method for grouping objects so that within-group variance is trivialized. This method works as follows:

Initialize two class centres randomly; these centres represent initial group centroids.

Calculate the value of histogram bin value distance between each image pixel and class centroids; assign each image pixel to its nearest class centroid.

Recalculate the new positions of the centroids by calculating the mean histogram bin value of the same group.

Repeat steps b and c till the value of centroids changes.

3. PERFORMANCE EVALUATION

In this section, the Fuzzy c means, K-means and K-means using Otsu thresholding are tested with synthetic images.

Performance metrics:

The performance evaluation of the three algorithms can be measured using three metrics such as Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), Segmentation Accuracy(SA).

Peak Signal to Noise Ratio (PSNR):

Reconstruction of lossy and lossless image compression quality is measured with the help of PSNR. The signal, in this case, is the actual data, and the noise is the error introduced by compression [13]. When comparing compression codecs, PSNR approximates human interpretation of reconstruction quality. Although a higher PSNR generally shows that the reconstruction is of higher quality, in some cases, it may not. PSNR is most simply defined via the mean squared error.

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (1)$$

An image's maximum possible pixel value (MAXI) is defined in terms of pixels that are expressed by 8 bits per sample; this is 255.

Mean Square Error (MSE):

The mean square error (MSE) is one way to evaluate the difference between an estimator and the estimated quantity [13]. MSE measures the average square of the "error" with the error. It is the amount by which the estimator differs from the quantity to be estimated.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2$$

Segmentation Accuracy (SA):

The segmentation accuracy (SA) is defined as follows:

SA=(Number of correctly classified pixels/Total number of pixels) × 100.

3.1 Performance analysis

3.1.1 Analysis for MSE

Table 3.1 shows the experimental analysis for MSE of Fuzzy c-means, K-means, and K-means using Otsu thresholding algorithms. From the table, it is observed that K-means using Otsu thresholding produces better results as compared to the other two algorithms.

Means, K-means using Otsu thresholding algorithms. From the table, it is observed that K-means using Otsu thresholding produces better results as compared to the other two algorithms.

.

Table 3.1. Analysis for MSE

Images	Before Segmentation	After Segmentation		
		FCM	K-means	K-means Otsu
Boat	253.69	213.74	3.35	0.81
Lena	255.00	176.08	5.25	0.73
Cameraman	237.85	203.71	9.41	0.75
Cell	170.24	114.47	14.29	0.46
Barbara	255.00	160.51	15.86	0.71

3.1.2 Analysis for PSNR

Table 3.2 shows the experimental analysis for PSNR of Fuzzy c-means, K-means, and K-means using Otsu thresholding algorithms. From the table, it is observed that K-means using Otsu thresholding produces better results as compared to the other two algorithms.

Table 3.2. Analysis for PSNR

Images	Before Segmentation	After Segmentation		
		FCM	K-means	K-means Otsu
Boat	24.12	24.87	42.92	49.08
Lena	24.10	25.71	40.965	49.54
Cameraman	24.40	25.07	38.429	49.51
Cell	25.85	27.58	36.613	51.50
Barbara	24.10	26.10	36.16	49.66

Table 3.3 Analysis for segmentation accuracy

Images	FCM	K-means	K-means Otsu
Boat	99.94	97.22	99.98

Lena	99.95	97.83	99.98
Cameraman	99.75	94.19	99.91
Cell	99.35	98.45	99.93
Barbara	99.96	98.75	99.98

3.2.4 Analysis for elapsed time

Figure 3.1 Analysis for elapsed time

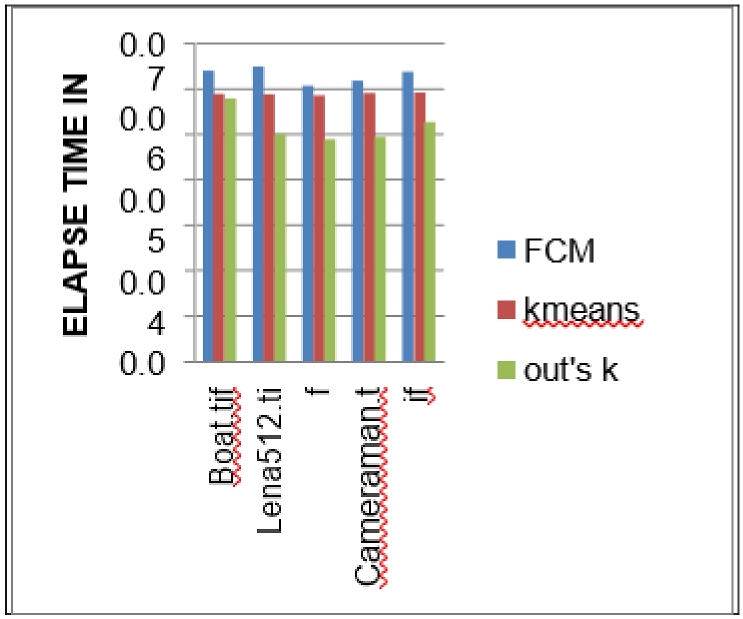

Figure 3.1 shows the experimental analysis chart for an elapsed time of Fuzzy c-means, K-means, and K-means using Otsu thresholding algorithms. From the figure, it is observed that K-means using Otsu thresholding produces better results as compared to the other two algorithms.

3.2 Experimental results:

Figure 3.2 : Experimental input and output images for FCM,K-means and K-means using Otsu thresholding

4. CONCLUSION

This paper analyzed the performance of three image segmentation algorithms, such as Fuzzy c-means, K-means and K-means, using Otsu thresholding. K-means using the Otsu thresholding method is greatly used to segment the image for a reason being simple calculation, less time consuming and effective. This method segments the image by automatically selecting the threshold value based on the largest inter-class variance between the target and background.K-means using the Otsu thresholding algorithm provides greater segmentation accuracy, low elapsed time and low error rate when compared with the fuzzy c-means clustering and K-means clustering algorithms. This paper concludes that K-means using Otsu thresholding performs better than the other two algorithms.

REFERENCES

1. Ahmed M. N., Yamany S.M et al. "A modified fuzzy c-means algorithm for bias field estimation
   and segmentation of MRI data", IEEE transactions on medical imaging, Volume 21, Issue 3, March ,2002, pp. 193 – 199
2. Ambroise, C.—Govaert, G.: "Convergence of an EM-Type Algorithm for Spatial Clustering. Pattern Recognition Letters", Vol. 19, 1998, pp. 919–927.
3. Anil Z Chitade, Dr. S.K. Katiyar .: "Colour Based Image Segmentation Using K-Means Clustering". International Journal of Engineering Science and Technology Vol. 2(10), 2010,pp. 5319-5325.
4. Cheriet M., Said J.N. and Suen C.Y. "A recursive thresholding technique for image segmentation", IEEE transactions on image processing, vol. 7, no.6, June 1998, pp. 918-921.
   Fu, S.K.—Mui, J.K.: "A Survey on Image Segmentation. Pattern Recognition", Vol. 13, 1981, pp. 3–16.
5. Hoel Le Capitaine and Carl Frélicot .: "A fast fuzzy c-means algorithm for color image Segmentation". EUSFLAT-LFA 2011, pp. 1074 – 1081.
6. Kang Jiayin, Min Lequan et al. "Novel modified fuzzy c-means algorithm with applications." Digital Signal Processing. March (2009),vol 19, no. 2, pp. 309-319.
7. Karmakar, G., Dooley, L., Murshed, M. "Fuzzy rule for image segmentation incorporating texture features", Proceedings of international conference on image processing, Vol.1, 2002, pp. 797-800. Khan.A.M, Ravi. S.: "Image Segmentation Methods: A Comparative Study." International Journal of Soft Computing and Engineering (IJSCE)

ISSN: 2231-2307, Volume-3, Issue-4, September 2013, pp.85-92.

8. Ma Li and Stuanton R. C. “A modified fuzzy c- means image segmentation algorithm for use with uneven illumination patterns”, Pattern recognition, vol. 40, 2007, pp. 3005-3011.
9. Mohamed N. A., Ahmed M. N., “Modified fuzzy c-means in medical image segmentation”, Proceedings of IEEE international conference on Acoustics, Speech, and Signal Processing, vol. 6, 1999, pp. 3429-3432.
10. Pham, D. L.—Prince, J. L.: “Adaptive Fuzzy Segmentation of Magnetic Resonance Images.” IEEE Trans. Medical Imaging, Vol. 18, 1999, pp. 737–752.
11. Pooja Kaushik , Yuvraj Sharma.: “Comparison Of Different Image Enhancement Techniques Based Upon Psnr & Mse.” International Journal of Applied Engineering Research, ISSN 0973-4562 Vol.7 No.11 (2012).
12. Sahoo P. K., Soltani S., and Wong A. K. C. “A survey of thresholding techniques”, Comput. Vision Graph. Image Processing, vol. 41, pp. 233-260, 1988.
13. ShahNilima,PatelDhanesh,JivaniAnjali.:”Review on Image Segmentation, Clustering and Boundary Encoding”. International Journal of Innovative Research in Science, Engineering and Technology, ISSN: 2319-8753, Vol. 2, Issue 11, November 2013, pp. 6309 – 6314.
14. Soumi Ghosh, Sanjay Kumar Dubey.: “Comparative Analysis of K-Means and Fuzzy C- Means Algorithms.” ((IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 4, No.4, 2013, pp. 35 – 39.
15. Vijay Jumb, Mandar Sohani, Avinash Shrivas.: “Color Image Segmentation Using K-Means Clustering and Otsu’s Adaptive Thresholding.” International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-3, Issue-9, February 2014, pp. 72 – 76.
16. Velmurugan.T, Santhanam.T.: “A Comparative Analysis Between K-Medoids And Fuzzy C-Means Clustering Algorithms For Statistically Distributed Data Points”. Journal of theoretical and Applied Information Technology,May 2011,Vol.27, pp. 19-30.