Performance evaluation of Clustering Algorithms for Image segmentation in view of Image Processing

[1]Hemantkumar R. Turkar, [2]Nileshsingh V. Thakur

[1]Rajiv Gandhi College of Engineering & Research, Nagpur, India
[2]Principal Nagpur Institute of Technology Nagpur, India
[1]hemantturkar@rediffmail.com, [2]thakurnisvis@rediffmail.com

Abstract— ***Image segmentation is mostly used as a research area by various researchers in computer vision. Clustering is an unsupervised study and it has many applications in various fields of science and engineering. In literature, it is found that various clustering algorithms have been proposed for image segmentation. This paper presents various clustering algorithms for image segmentation with comparative analysis of k- means, Moving k-Means and enhanced moving k-means on natural images with YCbCr color space and Grey images. The purpose of clustering algorithms is mostly used for effective automated image segmentation. This study shows that the performance of Enhanced Moving K-means algorithm is the most suitable technique for image segmentation in qualitative and quantitative parameters.***

Index Terms— Image segmentation, Clustering algorithm, K-means, YCbCr, Grey image.

I. INTRODUCTION

The process of grouped samples [1] is called as clustering technique. Clustering is a process of creating groups of clusters of same objects whereas different objects are separated in different clusters. The approach in [2] deals with the new clustering technique called as enhanced moving K- means algorithm which is less sensitive to cluster variance. This algorithm has various applications in customer electronic products for image segmentation. Clustering technique widely used in data mining, image analysis, face recognition, finger print recognition, object recognition, optical character recognition and statistical data analysis etc. [3-4]. Fine edge preserving technique uses various gradient operators for the detection of accuracy of edges [5]. In an image processing domain, segmentation is the basic step for image classification and description [6]. Recently, the clustering algorithms are widely used in the medical image processing domain for image segmentation [7-8]. Clustering is divided into two main types namely, hard and soft clustering. Hard clustering belongs to single cluster while soft clustering belongs to multiple clusters [12-13]. In our approach we have studied various clustering techniques for color image segmentation and Grey images. The other content of the paper is described as follows. A short review of various clustering algorithms are presented in section 2, experimental results and relevant discussions based on YCbCr and grey images is presented in section 3 and the section 4 concludes the work carried out with future scope.

II. CLUSTERING BASED IMAGE SEGMENTATION TECHNIQUES

In recent years color image processing is having more applications due to the reason that human eyes are more responsive to brightness, so at a particular time human eye can identify more colors than gray scale [14-16]. Clustering algorithms have had many applications over the years, for example, numerical taxonomy, vector or color image quantization, image segmentation & image retrieval. As of now, numerous clustering algorithms have been proposed with the aim to produce enhanced segmentation as shown in figure 1.

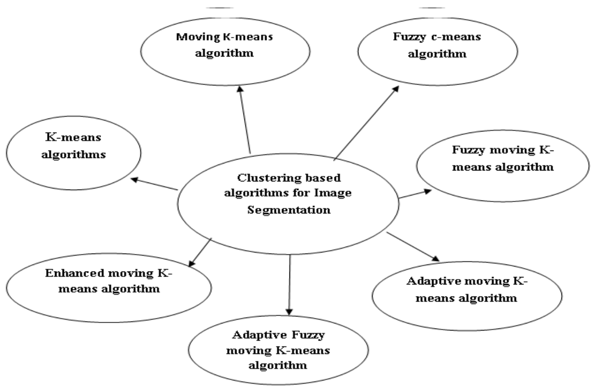


Figure1. Clustering based approaches for Image Segmentation.

A. *K-Means Algorithm:*

Due to its simplicity K-Means clustering algorithm is commonly used algorithm. K-Means (KM) is an exclusive clustering method [9] in which data belongs to a specific cluster and cannot be included in other cluster. The k-means algorithm is an iterative process [7,8]. The groups of patterns are divided into homogeneous and disjoint clusters where every cluster is characterized by center point. Though the k-means algorithm is known for it's a simple approach but, it has some disadvantages such as skewed distributions, sensitive to outliers and it depends on the initial declaration. [10, 11] .

B. Moving K-Means (MKM):

To overcome the drawbacks of K-Means several clustering algorithms were proposed. Mashor proposed a new approach for the clustering of images in image segmentation called as Moving K-Means (MKM) clustering [12]. The algorithm works on the basis of checking the center with fitness value of each center during the clustering process. The center is moved to the dataset with the most active center if the center does not fulfill the certain criteria. MKM algorithm works on the basis that every cluster must have specific number of members or final fitness value and so that the difference between the clusters is decrease. Thus the moving k-means overcomes the drawbacks of conventional k-means algorithm namely trapped centers, dead centers and the redundancy of center. Still the MKM algorithm has some drawbacks such as noise sensitivity and the centers are not located in middle of dataset. [13]

C. Fuzzy c-means algorithm:

FCM is an unsupervised clustering algorithm used in various fields such as, medical diagnosis, agricultural engineering, image analysis and image recognition. It works on the basis of image classification by grouping same data points into clusters. The pixels of images are closely related so that pixels in neighboring clusters have same data features. In FCM algorithm pixels are assigned to a category by using fuzzy membership. The FCM algorithm is most popular clustering methods based on a generalized least-squared errors function with minimization. Its advantages are it is an unsupervised clustering method and does not require prior knowledge of tested data, can be used with any number of classes and any number of features and distribution of membership values in done in normalized fashion. But it has some drawbacks such as noise sensitivity, computation time and sensitivity to initialization of cluster center.

D. Fuzzy Moving K-means Clustering algorithm:

The concept of fuzzy logic is used in fuzzy moving K-means clustering algorithm (FMKM). It works on the basis of assigning each data member to more than one class at the same time by different degree of membership. In medical science visual interpretation of medical images by radiologist is time consuming and depends upon experience of radiologist. This limitation can be overcome by using computer aided system depending upon artificial intelligence method like digital image processing combining with fuzzy logic in clustering algorithm of image segmentation.

E. Adaptive Moving K-means Clustering Algorithm:

The disadvantages of moving k-means algorithm is miss-clustering of data to unwanted noise clusters due to obligating the members of the center with the possible largest fitness value to become a member of the center with the smallest fitness value. This weakness of moving k-means algorithm is overcome by Adaptive Moving K-means Clustering Algorithm (AMKM). The AMKM offers the solution to the above problem by assigning of the members of center with the largest fitness value. This process of re-assigning varies from the conventional moving k-means algorithm as members of center with largest fitness value are assigned to center with the smallest fitness value.

F. Adaptive Fuzzy Moving K-means Clustering algorithm:

It is the combination of concepts of FMKM and AMKM. In Adaptive Fuzzy Moving K-means Clustering Algorithm the fuzzy logic of FMKM and fitness function of AMKM are used simultaneously. In Adaptive Fuzzy Moving K-Means, fitness function is presented where the members of the largest fitness value is allocated to the nearest cluster dependent on the Euclidean distance with smallest value. The adaptive fuzzy clustering algorithm is able to utilize local contextual information, thus impressing local spatial continuity and noise suppression to, to determine categorization uncertainty. The AFMKM has numerous advantages such as less noise sensitivity, less the initialization value sensitivity and avoid the problems related to center redundancy, dead centers and trapped center at local minima.

G. Enhanced Moving K-means Algorithm (EMKM):

Enhanced Moving K-means Algorithm (EMKM) [17-18] works on the basis of hard clustering membership function in which the data is clustered into non-overlapping regions. In conventional Moving K-Means Algorithm, the poor segmentation in image processing occurs due to increasing cluster variance as it forces the cluster center with smallest fitness value. In EMKM the moving concept of conventional moving k-means algorithm is enhanced and the defined range of members is re-defined as compare to the MKM algorithm in order to decrease segmentation variance difference among the clusters. It has several advantages over the above discuss various clustering algorithms such as it has less cluster variance and less sensitivity to initialization of center values. Thus EMKM has capability to generate better quality image segmentation [12].

III. EXPERIMENTAL RESULTS

Various experimentations are carried out on mobile devices with different values of *k* for YCbCR and Grey images. YCbCr represents colors in terms of one luminance component/luma (Y) and two chrominance components/chroma (Cb and Cr). YCbCr color space makes use of fact to achieve more efficient representation of images. Qualitative analysis is widely used in the probabilistic statement to carry out weakness and effectiveness of the algorithm on visual perception. Three images, namely, apple logo; Beach; and Red flower shown in Figure 3.1 are used for comparative analysis. Comparative analysis of KM, MKM, EMKM algorithms is carried out using Mean Square Error (MSE).

(a) (b) (c)

Figure 3.1. Original images: (a) Apple logo, (b) Beach and (c) Red flower

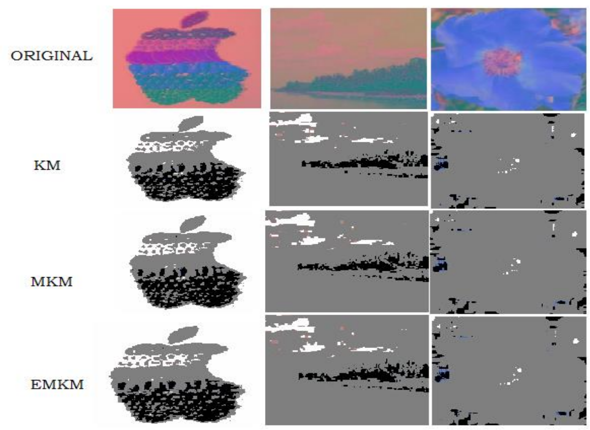


Figure 3.2. Segmentation of YCbCr images by selecting cluster value k=3

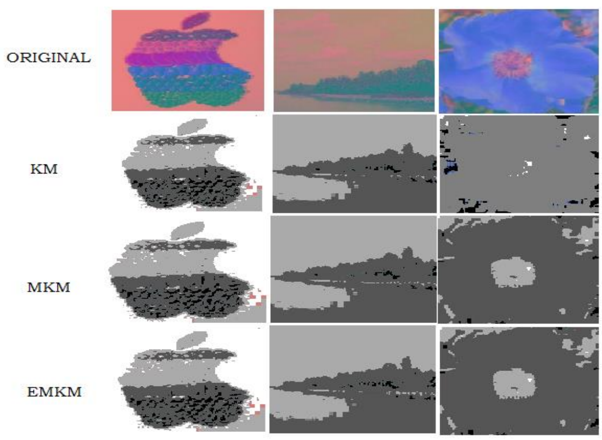


Figure 3.3. Segmentation of YCbCr images by selecting cluster value k=4

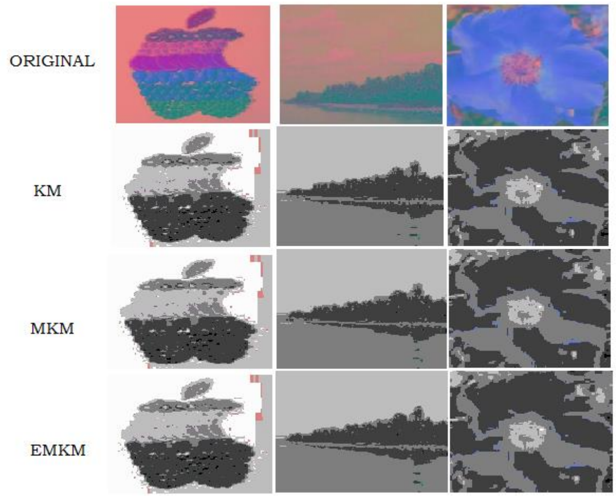


Figure 3.4. Segmentation of YCbCr images by selecting cluster value k=5

Quantitative Analysis is evaluated in terms of Mean Square Error (MSE) evaluation parameter to measure the performance of the algorithms. MSE is calculated by using an equation (1) as given below.

$$MSE = \frac{1}{N} \sum_{j=1}^{k} \sum_{i \in c_j} \left\| v_i - c_j \right\|^2 \quad \text{...equation (1)}$$

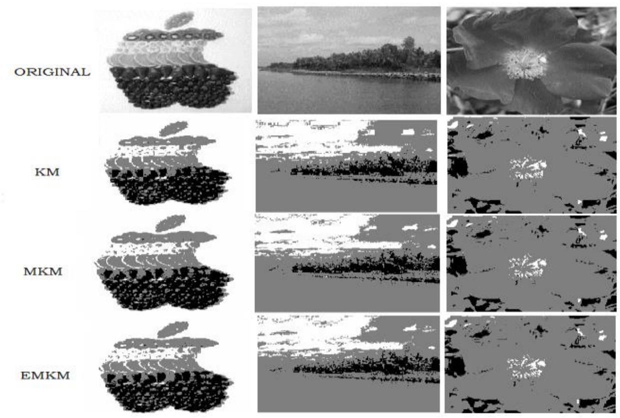


Figure 3.5. Segmentation of Grey images by selecting cluster value k=3

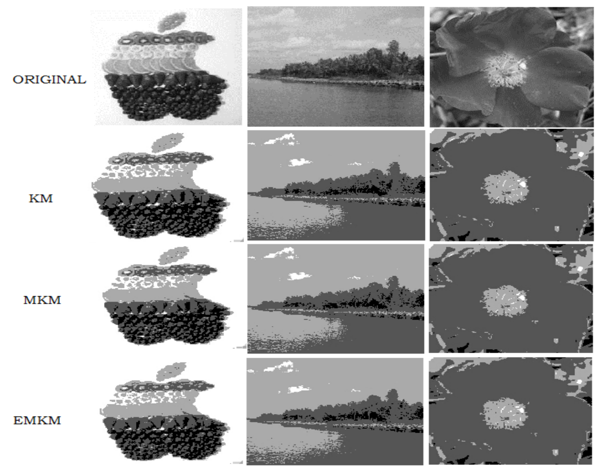


Figure 3.6. Segmentation of Grey images by selecting cluster value k=4

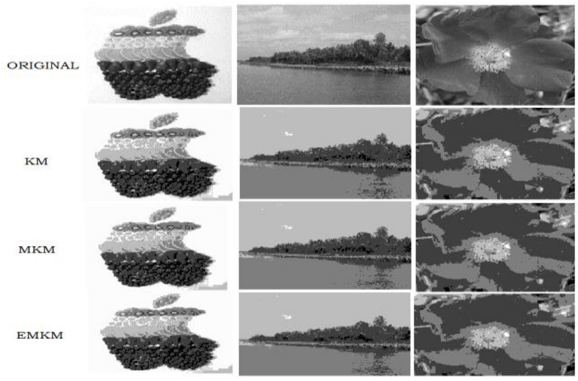


Figure 3.7. Segmentation of Grey images by selecting cluster value k=5

For calculation of MSE value V_i is the pixel which is part of j^{th} cluster and N is an image.

The lesser difference between the output and an original image conveys that all the data in the considered region is located near its center. The comparison of implemented algorithms is carried out in terms of Mean Square Error (MSE). MSE is calculated for KM, MKM and EMKM for YCbCr and Grey images with the cluster count of *k*, where *k* = 3, 4 and 5 respectively. Table 3.1 and Table 3.2 show the summary of MSE evaluation for the YCbCr color space and Grey images respectively.

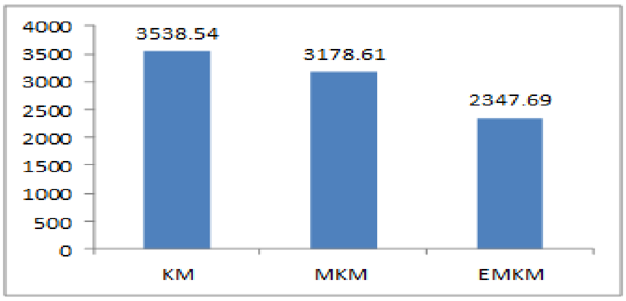
Table 3.1. Quantitative Evaluations: MSE on YCbCr Color Space

YCbCr color Space		Mean Square Error (MSE)		
No. of Cluster		*k*=3	*k*=4	*k* =5
KM	Apple logo	3538.54	26600.00	20921.99
	Beach	14392.13	9019.59	16372.97
	Red flower	2802.36	69489.47	30721.74
MKM	Apple logo	3178.61	25658.16	20259.48
	Beach	14341.77	8886.29	15515.95
	Red flower	2117.17	68930.89	30268.78
EMKM	Apple logo	2347.69	25479.16	19678.58
	Beach	13236.26	7958.60	14949.39
	Red flower	1458.48	68156.30	29593.94

Table 3.2. Quantitative Evaluations: MSE on Grey Images

Grey Image		Mean Square Error (MSE)		
No. of Cluster		*k*=3	*k*=4	*k* =5
KM	Apple logo	25275.74	14114.29	25973.59
	Beach	15644.95	16478.70	16282.02
	Red flower	16748.48	34641.41	827.91
MKM	Apple logo	24287.87	13380.03	25499.96
	Beach	15427.20	16242.11	15664.50
	Red flower	15993.37	34307.64	416.75
EMKM	Apple logo	24223.06	13081.45	24621.38
	Beach	14244.33	15231.85	14923.71
	Red flower	15669.25	33355.48	304.94

Graphical analysis of the obtained results of MSE evaluation of KM, MKM, and EMKM algorithms with k=3 for YCbCr color space and Grey images is shown in Figure 3.8 and Figure 3.9. Obtained results show that the EMKM algorithm has less MSE value for every image and for every type of cluster in YCbCr color space and Grey images. So, the EMKM algorithm performs better as compared to KM and MKM algorithms.



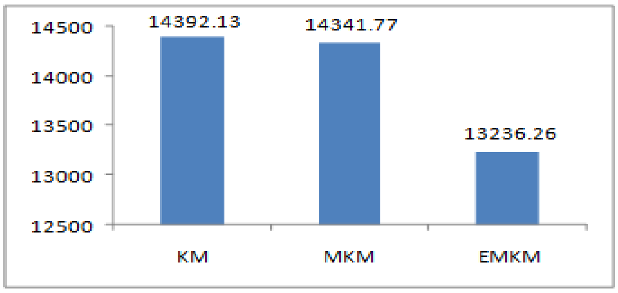


Fig.3.8. (a) and (b) MSE evaluation of KM, MKM, and EMKM algorithms with k=3 for YCbCr color space.

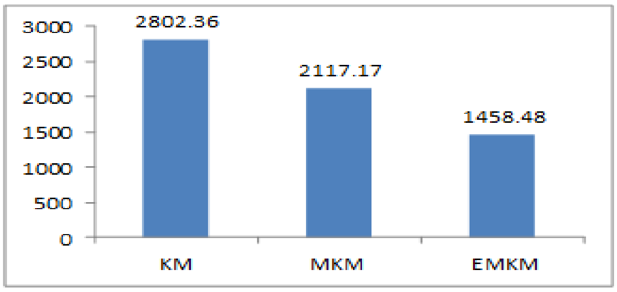


Fig.3.8. (c) MSE evaluation of KM, MKM, and EMKM algorithms with k=3 for YCbCr color space.

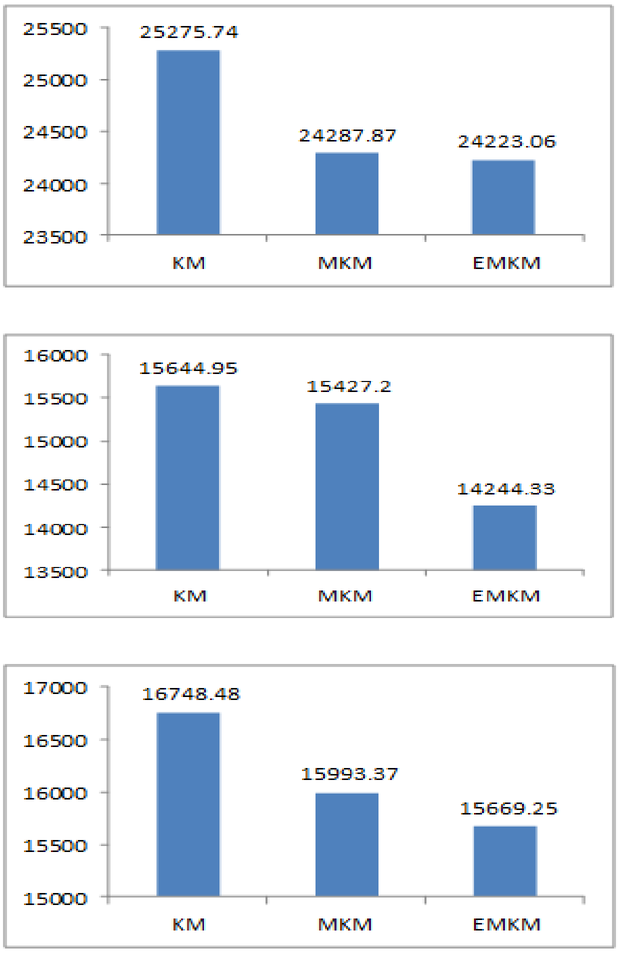


Fig.3.9. MSE evaluation of KM, MKM, and EMKM algorithms with k=3 for Grey images.

III. CONCLUSION AND FUTURE SCOPE

In this paper overview of different clustering algorithms has been presented. Although many clustering algorithms have been developed, each algorithm has some advantages and disadvantages. From the various studied clustering algorithms the enhanced moving k-means algorithm is having more promising concept with better clustering based image segmentation. Experimental results on qualitative and quantitative analysis on YCbCr color space and Grey images shows that EMKM outperforms with least MSE value. This paper provides a platform to develop new technique in clustering for image segmentation. Future work concerns an improvement and refinement with initialization of cluster centers and cluster variance.

REFERENCES

[1] Rafael C. Gonzalez, Richard E. Woods, “Digital Image Processing”, second edition, prentice hall.

[2] Nor Ashidi Mat **Isa**, Samy A. Salamah and Umi Kalthum Ngah, “Adaptive Fuzzy Moving K-means Clustering Algorithm for Image Segmentation”, IEEE Transactions on Consumer Electronics, Vol. 55, No. 4, NOVEMBER 2009.

[3] B. Bhanu, J. Peng, “Adaptive integrated image segmentation and object recognition” , IEEE Trans. On Systems, Man. and Cybernetics, vol. 30, pp.427 – 441, 2000.

[4] Y. Abe, and M. Hagiwara, “Hierarchical object recognition from a 2D image using a genetic algorithm”, IEEE International Conference On Systems, Man. and Cybernetics, vol. 3, pp. 2549 – 2554, 1997.

[5] S. Yang, Y. Kim, and J. Jeong, “Fine Edge-Preserving Technique for Display Devices," IEEE Trans. on Consumer Electronics, vol. 54, no. 4, pp. 1761-1769, 2008.

[6] D. Kim, I. Jeon, S. Yerl Lee, P. Rhee, and D. Chung, “Embeded Face Recognition based on Fast Genetic Agorithm for Intelligent Digital Photography”, IEEE Trans. on Consumer Electronics, vol. 52, no. 3, pp. 726-734, 2006.

[7] Nor Ashidi Mat Isa, Samy A. Salamah,, “Adaptive Fuzzy K-means Clustering Algorithm for Image Segmentation”, IEEE Transactions on Consumer Electronics, Vol. 56, 2010.

[8] P.A.Khaire, N. V. Thakur, “An Overview of Image Segmentation Algorithms”, International Journal of Image Processing and Vision Sciences 1(2) 62-68, 2012.

[9] Rose H Turi, “Clustering Base Colour Image Segmentation”, Monash University, Australia 2001.

[10] C. Ordonez, “Clustering binary data streams with K-means,” in Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery, San Diego, California 2003.

[11] P. K. Agarwal, and N. H. Mustafa, “k-means projective clustering,” in Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, Paris France 2004.

[12] Devarshi Naik , Pinal Shah, “A Review on Image Segmentation Clustering Algorithms”, International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014.

[13] N. A. Mat-Isa, M. Y. Mashor, and N. H. Othman, “Comparison of segmentation performance of clustering algorithms for Pap smear images”. International Conference on Robotics, Vision, Information and Signal processing (ROVISP2003). pp. 118-125, 2003.

[14] Dibya Jyoti Bora, Anil Kumar Gupta,” A Novel Approach Towards Clustering Based Image Segmentation” International Journal of Emerging Science and Engineering (IJESE) ISSN: 2319–6378, Volume-2 Issue-11, September 2014.

[15] Thodeti Srikanth, Pradeep Kumar, Ashwin Kumar, “Color Image Segmentation using Watershed Algorithm”, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (5) , 2011, pp. 2332-2334

[16] Peck, M.A Cell Image Segmentation of Gastric Cancer Based on Region-Growing and Watersheds.. 2002. Quebec City, Que.: Univelt Inc.

[17] Devarshi Naik , Pinal Shah, “A Review on Image Segmentation Clustering Algorithms”, International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014.

[18] Fasahat Ullah Siddiqui and Nor Ashidi Mat Isa, “Enhanced Moving K- Means (EMKM) Algorithm for Image Segmentation”, IEEE Transactions on Consumer Electronics, Vol. 57, No.2, 2011.

◈◈◈