International Journal of Electrical and Computer Engineering (IJECE)
Vol. 7, No. 2, April 2017, pp. 810~817
ISSN: 2088-8708, DOI: 10.11591/ijece.v7i2.pp810-817

Parametric Comparison of K-means and Adaptive K-means Clustering Performance on Different Images

Madhusmita Sahu[1], K. Parvathi[2], M. Vamsi Krishna[3]
[1]Dept. of ECE, CUTM, Paralakhmundi, Odisha, India
[2]School of Electronics, KIIT University, Bhubaneswar, India
[3]Dept. of CSE, CUTM, Paralakhmundi, Odisha, India

Article Info

Article history:

Received Nov 18, 2016
Revised Feb 23, 2017
Accepted Mar 7, 2017

Keyword:

Adaptive K-means clustering
Image segmentation
K-means clustering
MRI color and gray color image
PSNR & RMSE
Satellite image

ABSTRACT

Image segmentation takes a major role for analyzing the area of interest in image processing. Many researchers have used different types of techniques for analyzing the image. One of the widely used techniques is K-means clustering. In this paper we use two algorithms K-means and the advance of K-means is called as adaptive K-means clustering. Both the algorithms are using in different types of image and got a successful result. By comparing the Time period, PSNR and RMSE value from the result of both algorithms we prove that the Adaptive K-means clustering algorithm gives a best result as compard to K-means clustering in image segmentation.

Corresponding Author:

Smt. Madhusmita Sahu,
Dept. of ECE,
Paralakhmundi, Odisha, India.
Email: madhusmitas79@gmail.com

1. INTRODUCTION

Now a days cancer is a common diseases which spreaded all over world. It is very much essential for improving the survival rates because it is one of the most demanding and deliberated topics which are being discussed among the research scientist. Multiplication of a set of cells in a particular area of our body and the growth of the cells are uncontrollable, this is one types of structural distrortion of cells that are mentitoned as tumor. It is very important to know the perfect spreading area of tumor cell for cure the diseases. [1]

In image processing some techniques are used to detect the tumor cell among that image segmentation takes a major role to analyzing the image. In segmentation by using different techniques we detect the tumor cell and extracted the area for analyzing. Out of different techniques clustering is the most efficient methods. There are different types of clustering K-means clustering, Fuzzy C-means clustering, hierarchal clustering and Adaptive K-means clustering.

Among all the clustering K-means and Adaptive K-means clustering method is widely used in image segmentation. These algorithms are simple and faster than other clustering. Both are used for work in a large number of variables. In case of K-means, initialize the number of K takes an important task. Whereas in Adaptive K-means it is not required to initialize the number of K, based upon some features and characteristics the image will automatically finalize the number of K.

Image segmentation takes an important role to analyze the extract region in medical images or MRI (Magnetic resonance image) in medical area. So medical images are segmented and for further analyze by using different techniques in image segmentation. For analyzing the medical images one of the fundamental problems is to identify the boundaries of an object such as organs, or abnormal regions in images. From the result of segmentation it is possible for shape analysis, detecting volume change and making accurate radiation therapy treatment plan. [2]

In this paper, by using both the algorithm we point out the tumor spreading area in brain MR gray and color images whereas in the Satellite image we find out the spreading area of lake. After the detection of the changes from the segmented images we have compared some parameters based on PSNR, RMSE and Time Period.

After the comparison we concluded that Adaptive K-means clustering gives the best result as compared to K-means clustering for the accurate detection of tumor cell in MR gray and color images and volume detection in satellite images. So by this way we can detect the accurate detection of tumor cell and cure the diseases through radiation therapy treatment.

2. K-MEANS CLUSTERING

K-means algorithm is used for initialization of parameters since it is simple and works well for large data sets when compared with hierarchical clustering. [3]

According to some similarity measure (e.g. Euclidean) the data are divides into a specific number of groups that group number is noted as 'K' variable where K=1, 2, 3…………n. it is a user defined number. The group of same data is called as clustering and the method is called as K-means clustering.

This algorithm is consisting of two separate phases. At first it calculates the K centroid and secondly it takes each point to the cluster which has nearest centroid from the adjacent data point. To find out the distance of nearest centroid it will used the Euclidean distance method. After the grouping of the data it finds out centroid of each cluster. The Euclidean distance is calculated among each center and each data point assigns in the cluster which have minimum Euclidean distance, after reassigning the points each time the centroid of the clusters are recalculated. This process is continued till the end of the rest points then finally it calculated the final centroid or mean of the clusters.

So K-means is an iterative algorithm in which it means the sum of distance from each object to its cluster centroid, over all clusters. [2, 4]

Let us consider an image m × n and the image has to be cluster into K number of cluster. Suppose a point f(x, y) which is an input pixels to be cluster and C_k be the cluster center. The algorithm of K-means clustering is shown as follows:

a. First, initialize the number of cluster K and calculate their center C_k by using the formula

$$\text{Ck} = \frac{\text{Number of cluster K} * \text{no. of pixels}}{K+1}$$

b. Calculate the Euclidean distance E_d between the center and each pixel of an image using the relation given below.

$$E_d = \|f(x,y) - C_k\| \tag{1}$$

c. Based upon the Euclidean distance E_d assigns the rest pixels to the nearest cluster.
d. After assigning all the pixels in the appropriate cluster the value of the cluster center should be changed and it is calculated by using the relation given below:

$$\text{Xk} = \frac{1}{K}\sum_{y\in Ck}\sum_{x\in Ck} f(x,y) \tag{2}$$

e. Reiterate the process tilll it satisfies the tolerance or error value.
f. Update the cluster mean i.e. calculate the mean of the object for each cluster.

3. ADAPTIVE K- MEANS CLUSTERING

In Adaptive K-means clustering algorithm the number of K is selected from data set and it is randomly selected based upon some characteristics and features. The properties of each element also form the properties of the cluster that is consisting by the element.

Like K-means clustering this algorithm is also compute the distance between a given element and a cluster. So the function of the algorithm is based upon to compute the distance between two elements. This function has an important consideration i.e. it should be able to count for the distance based on properties that have been normalized so that distance is not dominated by one property or some of the property is not ignored in the computation of distance. It uses Euclidean distance formula for compute the distance.

For example: when spectral data specified by n-dimensions, the distance between two data elements. $D_1=\{D_{11},D_{12},\ldots\ldots D_{1n}\}$ and $D_2=\{D_{21},D_{22},\ldots\ldots D_{2n}\}$

The Euclidean distance is given by

$$E.D=\sqrt{D11-D21)^2 + (D12-D22)^2 + \ldots + (D1n-D2n)^2} \quad (3)$$

Based upon the distance the adaptive K-means clustering algorithm is as follows:

At first find out the distance of each cluster from every other cluster and it is stored in a 2D array as a triangular matrix. The minimum distance D_{min} between any two clusters C1 and C2 was noted. Compute the E.D distance from each cluster to each un clustered element. In this case there are three possibilities for assignment of the rest element to the suitable cluster that are like below four forms.

a. When the distance is zero from a element to a cluster then that element is assigned in that cluster, then proceed for next element.
b. The next element is allocated to its closest cluster if the distance of a element from a cluster is less than D_{min}. After reassigned the element in the cluster the centroid of the cluster is changed. So according to the properties of the elements the centroid is recomputed. As well as we recompute the distance of the affected cluster from every other cluster, also the minimum distance between any two clusters and the two clusters they are closest to each other.
c. Finally, it is computed when the distance D_{min} is less than the distance of the element from the nearest cluster.
d. We take two closest clusters C1 and C2, and then merge C2 into C1. Now we get only one cluster C1 and we destroy the elements as well as the representation of C2, so C2 called as empty cluster. We get a new cluster successfully. Again the distances between all clusters are recomputed and two closest clusters recognized again.

Till all the elements has been clustered the above three steps are to be repeated. By this way we should cluster the elements without defining the number of k in adaptive K-means clustering. [5]

4. PROPOSED ALGORITHM

In the proposed method we applied both the techniques K-means clustering and Adaptive K-means clustering to track the tumor objects in MRI brain gray color image, color image and also these techniques are applied to find out the lake region in satellite image.

The block diagram of the proposed method for K-means and Adaptive K-means clustering in gray color and color image is:

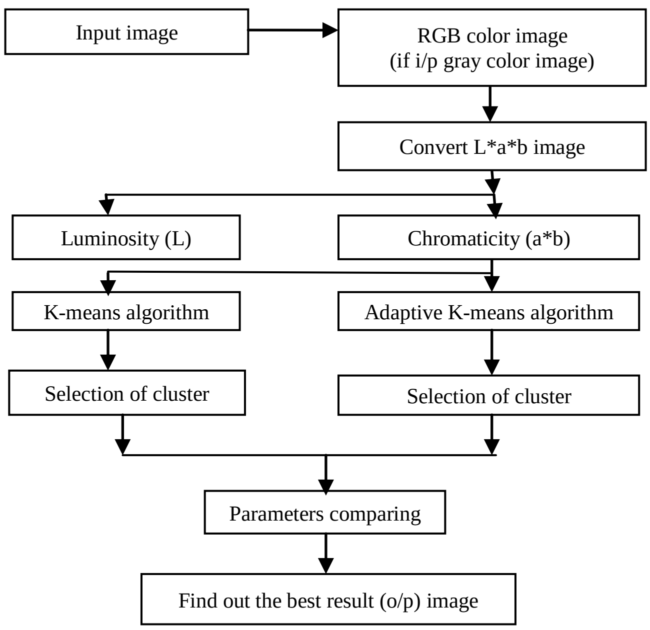


Figure 1. Block Diagram of proposed method

The algorithm of the proposed method is given below from step1- step14:

step 1 Read the images which are in JPEG format, that are collected from the medical brain images and satellite images in the goggle site.

step 2 Resize the image

step 3 Convert the image into RGB color space image, this step is only applicable for gray color images.

step 4 Then after convert the image into L*a*b color space image. When we ignoring the brightness in the images. We can easily visually distinguish the colors from the image. The L*a*b* color space (also well-known as CIELAB or CIE L*a*b*) enables us to quantify these visual differences. The L*a*b* color space based on luminocity layer and Chromaticity layer, where the luminosity layer is 'L*' where as chromaticity-layer 'a*' representing color falls along the red-green axis and chromaticity-layer 'b*' representing the color falls along the blue-yellow axis. All of the color related information is in the 'a*' and 'b*' layers. Using the Euclidean distance metric we can measure the difference between two colors. Convert the previous image in to L*a*b* color space. [6]

step 5 By Using K-Means Clustering classify the Colour in 'a*b*' Space.
A group of objects are separated by using clustering method. K-means clustering treats each element as having a location in space. The partitions where the objects inside each cluster were closer to each other as possible, and as distance from the objects in other clusters as possible were selected by it. In K-means clustering we have to define the number of clusters to be partitioned and a distance metric to compute how closer two objects each other. As the color information based in the 'a*b*' space, your objects are pixels by means of 'a*' and 'b*' values. By using K-means cluster the objects into K clusters using the Euclidean distance metric. [7, 8]

step 6 Label Every Pixel in the Image Using the Results from K-Means.
Without index input K-means returns to a corresponding cluster for every object. Label every pixel in the image with its cluster index.

step 7 According to the number of K which defined by the user based on that image we will get segmented images or clustered images.

step 8 Select one clustered image.

step 9 In the same way, Classify the Color in 'a*b*' Space Using Adaptive K-Means Clustering. Adaptive K-means clustering doesn't need specifying the number of clusters to be partitioned. It automatically finds out the number of cluster based upon some properties or features. Since the color information exists in the 'a*b*' space and the objects are pixels with 'a*' and 'b*' values. Use Adaptive K-means to cluster the objects into K clusters using the Euclidean distance metric.

step 10 In the Image label Every Pixel by using the Results from Adaptive K-Means clustering.
With our index input, Adaptive K-means clustered every object. Label every pixel in the image with its cluster index.

step 11 As per the number of objects in that image we will get the segmented images (As per the number of centroid).

step 12 Select one clustered image or segmented image.

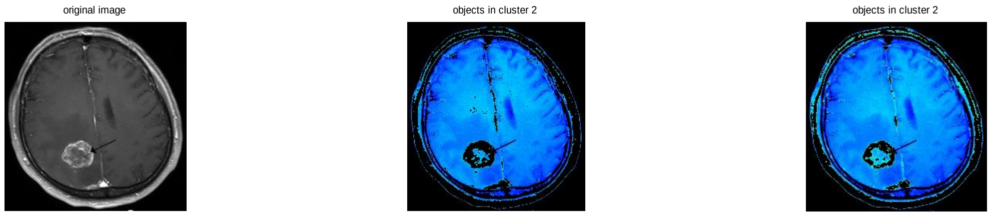
step 13 Compare the parameters (TIME, PSNR, and RMSE) of the original image with the segmented image, find out the best result.

step 14 Finally, from the segmented image we can detect the appropriate tumour spreading area by using Adaptive K-means clustering in case of MRI images and from the satellite image we find out the area covered by the lake.

5. RESULTS AND ANALYSIS

In the proposed method, we convert a gray-level MR brain image into an RGB color image but when we use color MRI and Satellite image we have not converted the input image into RGB color space, then the image converted into CIE Lab color model. Colors in both the *a** and *b** spaces are feature vectors for K-means clustering and Adaptive K-means clustering. To express the detection performance of the proposed method, we take a MR brain image of gray color and color image containing the pathological change area shown in Figure 2(a) & Figure 3(a) served as a original image, here also we take a satellite color image for detection of geographical area covered by the water shown in Figure 4(a). Generally, an MR brain image consists of regions that represent the bone, soft tissue, affected cancer cell area, fat and background [6]. Based upon some visual judgement and some features the K-means and Adaptive K-means clustering algorithm clustered the gray color and color brain image and the selected cluster images are shown in Figure 2(b), 2(c) & Figure 3(b), 3(c).

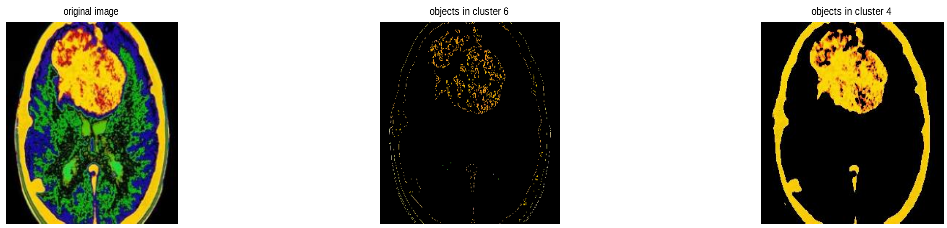
After select the cluster we calculate the time period, RMSE value and PSNR value and then we comparing the results of the entire image. Which results are shown in parameters comparing table.



(a) (b) (c)

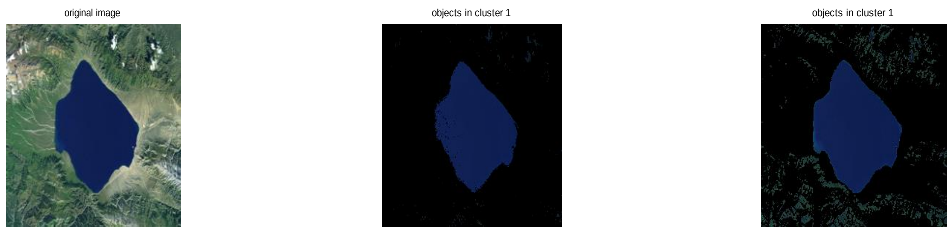
Figure 2. (a) Original image (b) K-means clustering (c) Adaptive K-means clustering

Here we discuss the MRI brain color image and Satellite image. When we applied both the algorithm in those images in color brain MRI image we track the tumor cell area clearly by using Adaptive k-means clustering is compared to K-means clustering. In Other case when both algorithms are implemented in Satellite image the segmentation of lake from the satellite image is clearly isolated in Adaptive K-means clustering algorithm which we can visualize in the cluster image 4(c).



(a) (b) (c)

Figure 3. (a) Original image (b) K-means clustering (c) Adaptive K-means clustering



(a) (b) (c)

Figure 4. (a) Original image (b) K-means clustering (c) Adaptive K-means clustering

From the Figure 2, Figure 3 and Figure 4 we observed that the output result using Adaptive K-means clustering algorithm has better segmentation result as compared to the k-means algorithm.

The quality of the segmented image is analyzed using the measurement amount of Peak to Signal Noise Ration (PSNR) and Root Mean Square Error (RMSE) [2].

5.1 RMSE (Root Mean Square Error)

It has been used for the standard performance measurement of the output image. It gives how much output image is deviated from the input image.

$$\text{RMSE}=\sqrt{\frac{1}{n_x n_y}\frac{\sum_0^{n_x-1}\sum_0^{n_y-1}[(r(x,y))]^2}{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y)-t(x,y)]^2}} \quad (4)$$

5.2 PSNR(Peak to Signal Noise Ratio)

The peak to signal noise ratio is the proportion between maximum attainable powers and the corrupting noise that effect likeness of image. It is used to measure the quality of the output image.

$$\text{PSNR}= 10.\log_1 0\left[\frac{max(r(x,y))^2}{\sqrt{\frac{1}{n_x n_y}\frac{\sum_0^{n_x-1}\sum_0^{n_y-1}[(r(x,y))]^2}{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y)-t(x,y)]^2}}}\right] \quad (5)$$

5.3 Comparison Analysis of the Parameters

In the Table 1 we have compared the parameters results of the algorithm used in different images.

Table 1. Parametric Comparison Table

Image	RMSE		PSNR		Elapsed time	
	K-means	Adaptive K-means	K-means	Adaptive K-means	K-means	Adaptive K-means
Gray color Brain (MRI)	14.5914	13.7564	9.6887	10.3540	0.235163	0.152034
Color Brain (MRI)	22.7310	20.0241	8.2613	12.4905	0.314162	0.158180
Satellite	25.0647	24.1429	8.5738	8.7060	0.283999	0.067358

5.4 Graphical representation of Parameters

Figure 5 graph represents the RMSE in K-means is more as compared to Adaptive K-means algorithm. It defines that the Adaptive K-means clustering image is less deviated as compared to K-means clustering image.

Where as in PSNR, which defines the quality of the image, is more in Adaptive K-means as compared to K-means clustering algorithm that shows in the Figure6 graph [10, 11]. Also in case of time period Adaptive K-means takes very less time as compared to K-means that will define from the figure 7.

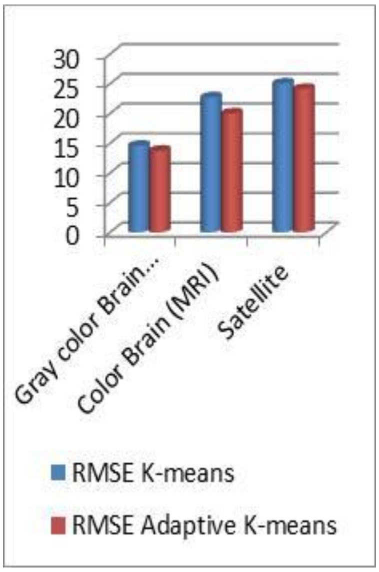


Figure 5. Graphical representation of RMSE

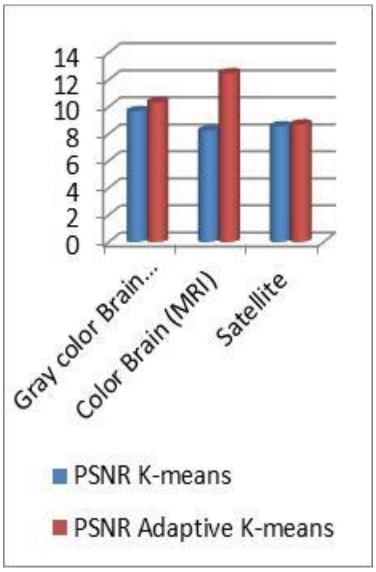


Figure 6. Graphical representation of PSNR

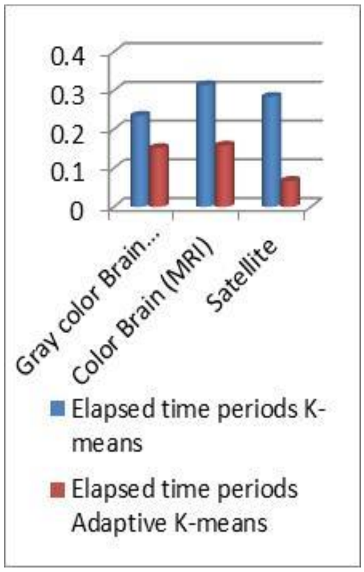


Figure 7. Graphical representation of Elapsed time periods

6. CONCLUSION

Segmentation being an important area of research, determining its performance is also important [9]. By using K-means and Adaptive K-means algorithm we have segmented the brain gray color & color MRI image also Satellite image of a lake. During such time these algorithm uses the color separation and conversion techniques i.e. RGB to L*a*b conversion based upon the a*b chromaticity the algorithms are used and segmented the image. We can successfully find out the tumor cell spreading area from the gray and color brain MRI clustered image after the segmentation along with identification of the lake cover area from the satellite image.

The RMSE and PSNR values of the images both in K-means & Adaptive K-means were found and compared in Table 1. From which it is observed that the RMSE value in Adaptive K-means is less then that of K-means, also the PSNR value is higher in Adaptive K-means then that of K-means. This is the condition of good image segmentation quality. From the comparison of RMSE, PSNR, time period & clustering parameters Adaptive K-means clustering gives best and better results as compared to K-means algorithm.

REFERENCES

[1] Sushma S J, S C Prasanna Kumar, et al., "Advancement in Research Techniques on Medical Imaging Processing for Breast Cancer Detection", published in *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 6, No. 2, April 2016, pp. 717~724.

[2] Nameirakpam Dhanachandra, et al., "Image Segmentation using K-means Clustering Algorithm and Subtractive Clustering Algorithm", Published in *Elsevier at (IMCIP-2015) Procedia Computer Science*, pp.764 – 771.

[3] T. Jyothirmayi1, K. Srinivasa Rao, P. Srinivasa Rao, Ch. Satyanarayana, *et al.*, "Image Segmentation Based on Doubly Truncated Generalized Laplace Mixture Model and K Means Clustering", published in *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 6, No. 5, October 2016, pp. 2188~2196

[4] Ming-Ni Wu, et al., "Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation", published in IEEE Xplore 2009.

[5] Sanjiv K. Bhatia, et al., "Adaptive K-Means Clustering", published in *American Association for Artificial Intelligence*, 2004.

[6] Anil Z chitade, et al., "Color based image segmentation using k-means clustering", published in *IJEST*,vol.2,2010

[7] M. Sahu & Dr. K.Parvathi, et al., "Classification and Detection of Tumor cell from color Brain MRI image by Implementing K-means clustering and Histogram Processing", published in *PISER journal 2014*, PISER 13, Vol.02, Issue: 03/06 May- June, Page(s) 019-023.

[8] M.sahu & Dr. K.Parvathi, et al., "Segmentation of Color Data Base Image by Implementing K-Means Clustering", published in *IJIET*, vol.2 issue 4 Aug 2013.

[9] M. Akhil, R. Aishwarya, Vandhana Lal and Shanthi Mahesh, et al. "Comparison and Evaluation of Segmentation Techniques for Brain MRI using Gold Standard", published in *IJST journal*, Vol-9, December 2016.

[10] S. Mahalakshmi and T. Velmurugan, et al. "Detection of Brain Tumor by Particle Swarm Optimization using Image Segmentation", published in *IJST journal*, Vol-8, Septmber 2015.

[11] A. Naveen and T. Velmurugan, et al. "Identification of Calcification in MRI Brain Images by k-Means Algorithm", published in *IJST journal*, Vol-8, November 2015.

BIOGRAPHIES OF AUTHORS

Smt. Madhusmita Sahu received the M.Sc. (Electronic sciences) from Berhampur university, Ganjam. M.tech (Communication system of enggینering) from VSS university of Technology, Burla, Sambalpur, Odisha

Dr. K. Parvathi received the B E (ece) from Osmania university, MTech & PhD from Andhra University, AP. Research interests are signal & Image Processing, Computational Intelligence.

Dr. M. Vamsi Krishna received the M Tech (CS) in Allahabad University, M.Tech (AI &R) degree in Andhra University, and Ph.D from Centurion University, Odisha. Currently he is working as Professor & HOD in Department of Computer Science and Engineering. He has 15 years of experience in teaching. His research interests include Artificial intelligence, computer networks, image processing and Data Analytics.