International Journal of Advances in Electronics Engineering
Vol:1 Issue:1 ISSN 2278 - 215X

Comparison of K-means and Adaptive K-means using MATLAB Simulation

Prof. Sachin M. Jambhulkar
Department of Electronics Engg.
R.T.M. Nagpur University, India.
Email: Sach_jambhulkar@yahoo.co.in

Dr. Pankaj Agrawal.
Department of Electronics Engg.
R.T.M. Nagpur University, India
Email:pankaj_rknec@rediffmail.com

***Abstract*— Image Segmentation based on K-means algorithm is presented in GUI (Graphical User Interface). The K-means algorithm and adaptive k-means clustering is used to obtain high performance and efficiency in image segmentation. In addition, it has a resolving capability of one image into different planes by selecting the number of clusters using datasets of image. And also the advance of K-means is adaptive K-means which will give the frame size and the absolute value between the means of an image.**

The iteration time on image segmentation is determined by using Adaptive k-means clustering. Adaptive k-means is used for better image segmentation that has been shown in MATLAB Simulation.

Keywords- Image Segmentation, K-means, Adaptive k-means

I. Introduction

Image Segmentation can be grouped under a general framework of image engineering. It consists of three layers as follows:-

(1)Image Processing (Low Layer)
(2)Image Analysis (Middle Layer)
(3)Image Understanding (High Layer)

Image Segmentation is most critical tasks in image analysis. It is used in image either to distinguish objects or to partition an image onto the related regions. It responds more quickly to human eye and produces result accurately so as the related scene is concerned.

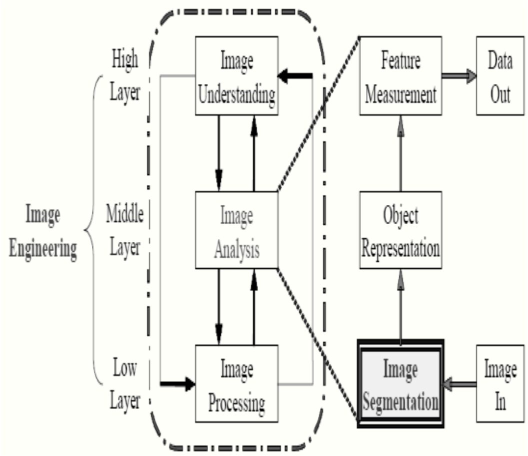

Figure 1: Basic Image Data Transfer Analysis

Colour is helpful in making many objects in "stand out". It is also considered to be one of the most popular problems in a computer vision. There are different techniques to solve image segmentation problems such as threshold approach, contour based approaches, region based approaches, cluster based approaches and other optimizations approaches using image framework. The clustering approach can be divided into two general groups such as a Partition and Hierarchical Clustering algorithm. Partition algorithms such as K-means and Adaptive Clustering are widely used in many applications such as Data Mining, Compression, Image Segmentation and Machine Learning. The advantages of Clustering algorithm are that the classification is simple and easy to implementation for clusters. A cluster is a collection of data objects that are similar to one another within the same cluster and dissimilar to the object in other cluster. Cluster's analysis has been widely used in a numerous application including Pattern Recognition, Data Analysis, Image Processing and Market Research.

The k-means algorithm takes the input parameter K and partitions a set of N objects into K cluster so that the resulting intra cluster similarity is high but the intercluster similarity is low. Cluster similarity is measured in regard to the mean value of objects in a cluster. The basic steps of K-means clustering are simple. In the beginning we determine number of cluster and we assume the centre of these clusters. We can take any random objects as the initial centroids. Each region is characterised by a slowly varying intensity that is distribution of function. Therefore to avoid this function of K-means clustering, Adaptive K-means clustering is used to decrease the number of iterations. The segmented image consists of very few levels .By seg-

Vol:1 Issue:1 ISSN 2278 - 215X

mented recognition system, the segmented image retains a crude representation of the image. The technique we develop can be regarded as a generalization of the K-means clustering algorithm in two respects as it is adaptive and includes spatial constraints.

The intensity function are constant in each region and equal to the K-means cluster centres as the algorithm progresses, the intensities are updated by averaging over a sliding window whose size progressively decreases in terms of image. The performance of the technique is clearly superior to K-means for the class of image we are considering. By using adaptive K-means clustering, we can overcome associate regions with the detected boundaries within a very less time.

II. RELATED WORKS

A. K-means-

K-means depends on following four main steps:

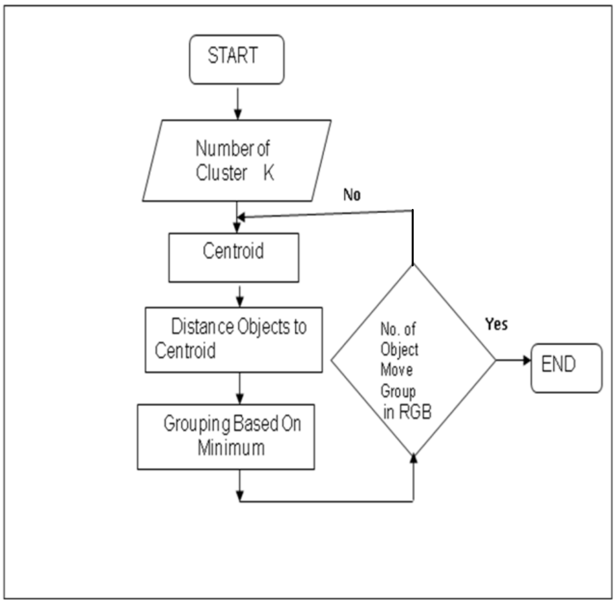

Figure 2: Standard K-means Analysis

(1)Initialisation
(2)Classification
(3)Computation
(4)Convergence Condition.

At first, it randomly selects K objects, each initially represents a cluster mean or centre. For each object a cluster is assigned to which it is most similar based on the distances between the object and the cluster mean. It then computes the new mean for each cluster. It repeats till the K-means algorithm convergences.The K-means algorithm for portioning based on the mean value of the object in the cluster. Input is the number of cluster K and a database containing N objects. Output is set of K clusters that minimize the square error criterion. Every data in the subclass shares a common trait.

The algorithm is iterative in nature i.e. $x\ldots x_n$ are data points or vectors or observations. Each observation will be assigned to one and only one cluster. C(i) denotes cluster number for the i^{th} observations. K-means minimizes the distance within cluster.

$$W(C) = \frac{1}{2}\sum_{k=1}^{K}\sum_{C(i)=k}\sum_{C(j)=k}\left\|x_i - x_j\right\|^2 = \sum_{k=1}^{K} N_k \sum_{C(i)=k}\left\|x_i - m_k\right\|^2$$

Where,

m_k = is the mean vector of the K^{th} cluster.

N_k= is the number of observations in K^{th} cluster.

Then cluster mean can be computed as follows:

$$m_k = \frac{\sum_{i:C(i)=k} x_i}{N_k},\ k = 1,\ldots,K.$$

The membership for each data point belongs to nearest centre, depending on the minimum distances. This membership distance is as follows:

$$C(i) = \arg\min_{1\le k\le K}\left\|x_i - m_k\right\|^2,\ i = 1,\ldots,N$$

For initialization step, each data point x_i computes its minimum distance with each centre x_j. For each centre x_j, it computes the new centre from all data point Xi belong to this centre.

B. Adaptive K-means-

The grey-scale image is a collection of regions of uniform or slowly varying intensity. This typically takes values between 0 to 255. Segmentation of the image into regions will be denoted by X. Where Xs=I means that the pixel at s belongs to region I. The different regions types or classes are K.

$$p\left(\frac{x}{y}\right) \alpha\ p\left(\frac{x}{y}\right) p(x)$$

Where,
p(x) = priori density of the region in process.
p(x/y) = conditional density of the observed image.

We model the regions process by a Markov random field. That is if Ns as a neighbourhood of the pixel at S, then

$$p\left(\frac{Xs}{Xq} all\ q \neq s\right) = P\left(\frac{Xs}{Xq}, q \in Ns\right)$$

We consider images defined on the Cartesian grid and a neighbourhood consisting of the 8 nearest pixels.

$$p(x) = \frac{1}{z} \exp\left(\sum_{c} Vc(x)\right)$$

Where, Z is normalizing constant and the summation is over all cliques C.

We consider images defined on the Cartesian grid and neighbourhood consisting of the 8 nearest pixels. The clique potential Vc depend only on the pixels that belong to clique C.

The two-point clique potential is defined as follows:

$$Vc(x) = \begin{cases} -\beta, if\ Xs = Xq\ and\ s, q\ \varepsilon\ C \\ +\beta, if\ Xs \neq Xq\ and\ s, q \in C \end{cases}$$

The parameter β is positive, so that two neighbouring pixels are more likely to belong to same class than to different ones.

III. INITIALIZATION METHOD

A. K-means-

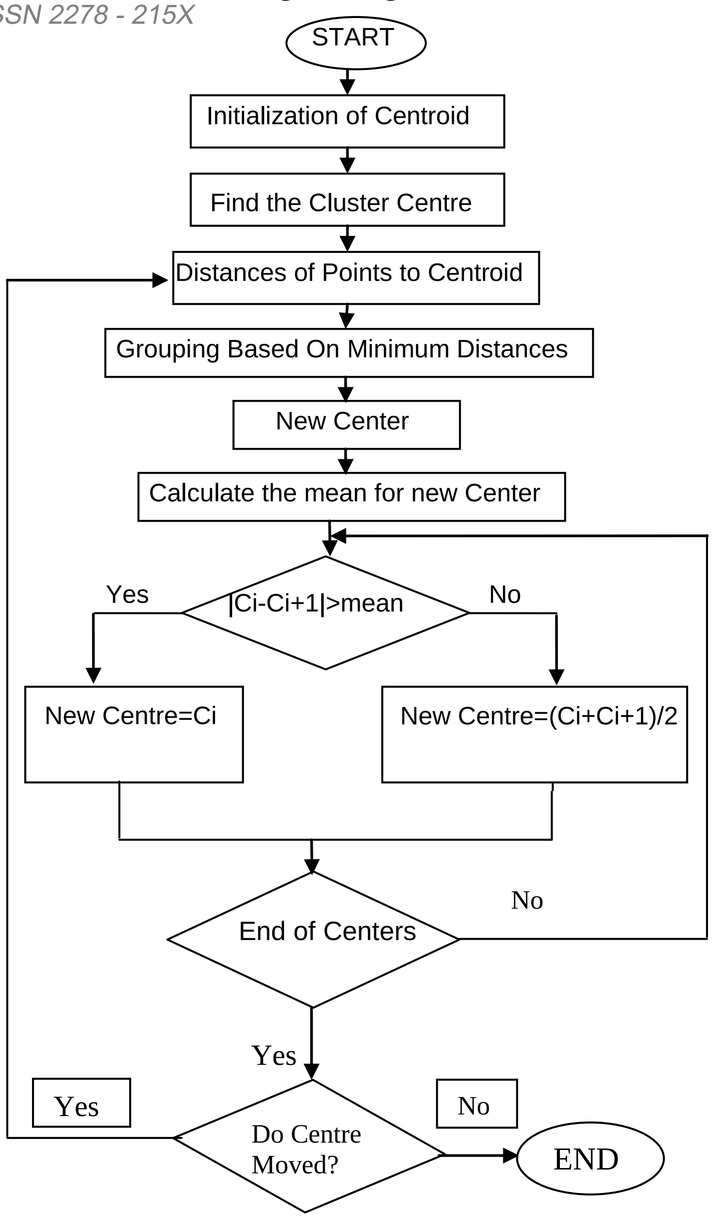

Figure 3: K-means cluter

It starts by reading the data as 2D matrix and then calculates the mean of the first frame size as F1=300×300,F2=150×150,F3=100×100,F4=50×50,F5=30×30, F6=10×10 or F7=7×7.

This number is the number of clusters and their values are the centroid values as indicated in the steps below:
(1)Read the data set as a matrix.
(2)Calculate the means of each frame depending on the frame size and putting them in array.
(3)Sort the means array in an ascending order.
(4)Compare between the current and the next element in the means array. If they are equal then keep the current element and remove the next otherwise, keep both.
(5)Repeat step 4 till the end of the means array. These are

Vol:1 Issue:1 ISSN 2278 - 215X

equal to the number of clusters and their values.

B. Adaptive K-means-

For estimating the distribution of region X and the intensity function µs. µs are defined on the same grid as the original gray-scale image Y and the region distribution X. When the number of pixels of the type i inside the window, cantered at S is too small, then the estimate µs is not very reliable. A reasonable choice for the minimum of pixels necessary for estimating µs is Tmin=W, the window width. The maximization is done at every point in the image and the cycle is repeated until convergence occurs. This is iterated conditional mode approached to propose adaptive K-means clustering. Now we consider overall adaptive clustering algorithm .We obtain an initial estimate of X by the K-means algorithm it alternates between estimating of X and Intensity function µs. We define an iteration consist of one update of X and, one update of function of µs. The window size for the intensity function estimation is kept constant until this procedure convergence occurs; usually in less than ten iterations it gets completed. Our stopping criterion is that the last iteration convergences in one cycle. The whole procedure iteration carried out with a new window size. The adaptation is achieved by varying the window size W. Initially the window for estimating the intensity functions is the whole image and the intensity function of each region are constant. Segmentation becomes better and smaller window size gives more reliable and accurate estimates. Thus, the algorithm starting from global estimates, slowly adapts to the local characteristics of each region.

The algorithm stops when the minimum window size reached. Typically we keep reducing the window size by a factor of two, until a minimum window size W=7 pixels is not obtained. If window size kept constant and equal to whole image, then the result is same as in the approach. Adaptive clustering algorithm is only an intermediate result of clustering algorithm. It results in segmentation similar to the K-means, only difference is that the edge is smoother. The strength of the approach lies in the variation of the window size which allows the intensity function to adjust to the local characteristics of the image.. Also there is no need to split out regions into a set of connected pixels. The noise variance is known easily or can be estimated independently by the algorithm. Thus, image segmentation is the most important to choose the variance. The noise variance controls the amount of detail by algorithm. In the following section we discuss the choice of K more extensively. The amount of computation for the estimation of X depends on the image dimensions, the number of different window used, the number of classes and course & image content. The amount of computation for each local intensity calculation does not depend on the image size.

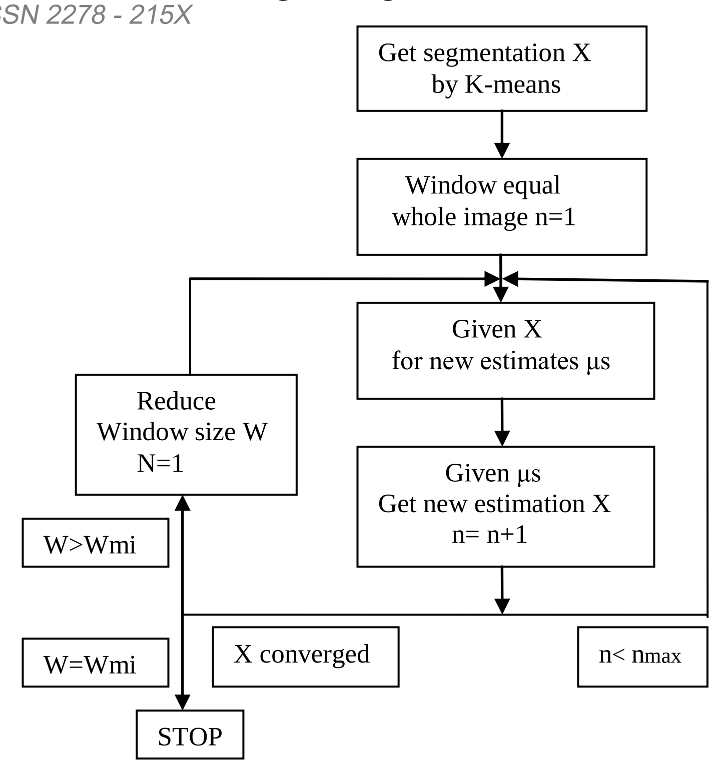

Figure 4: Adaptive Cluster

IV.SIMULATION RESULTS

Figure 5: Original image

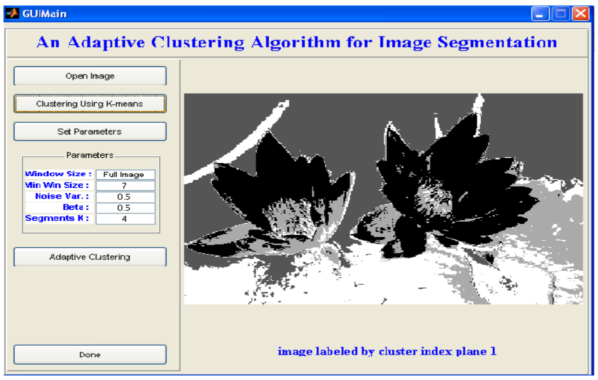

Figure 6- Image using K-means of Plane1

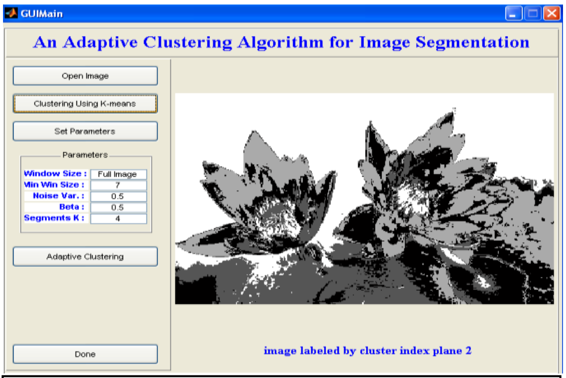

Figure 7: Image using K-means of plane2

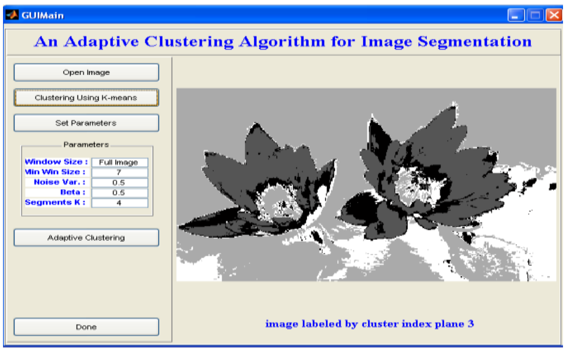

Figure 8: Image using K-means of plane3

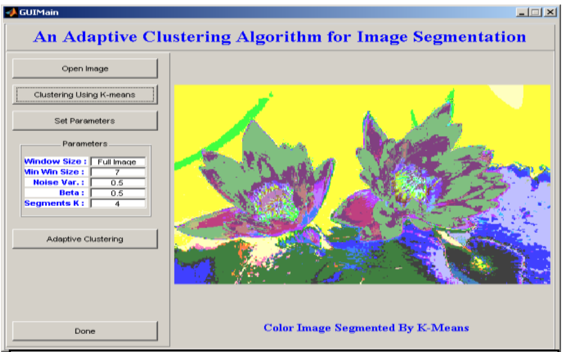

Figure 9: Image using K-means

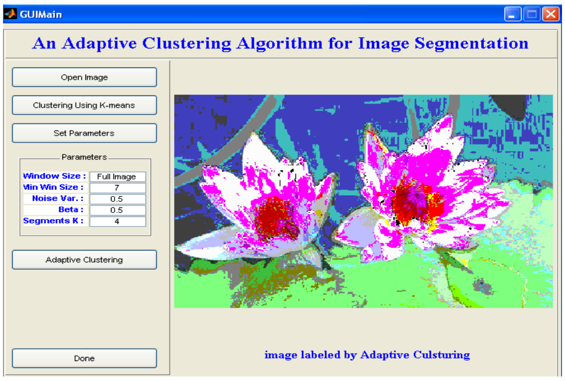

Figure 10: Image using Adaptive K-means

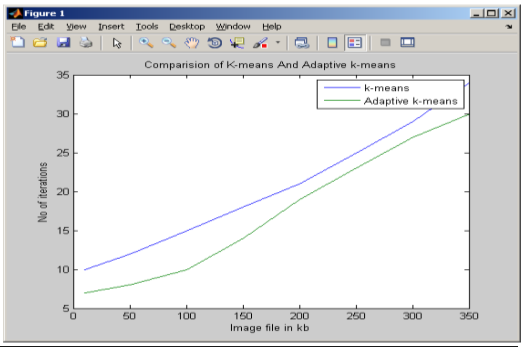

Fig 10: Comparison of k-means & Adaptive K-means

Original Image Cluster =242 pixels
Image Cluster 1 =121 pixels
Image Cluster 2 =60 pixels
Image Cluster 3 =30 pixels
Image Cluster 4 =15 pixels
Image Cluster 5 =7 pixels

TABLE 1: Data sets of Adaptive K-means

V. CONCLUSION

The number of iteration affects different methods for its computational cost performance for reaching convergence. Therefore adaptive K-means algorithm is proposed to improve the standard K-means algorithm using modification and additional steps for convergence condition. It can be concluded that the proposed scheme was successfully carried out in estimating the number of cluster by decreasing the number of iteration in K-means clustering, due to which the speed of execution has been increased.

Thus it can be conclude that if large absolute values are taken then small number of clusters can be obtained. Adaptive K-means clustering offers segmentation of objects with smooth surface. The intensity of each region is modeled as slowly varying function with white Gaussian Noise. The adaptive K-means clustering technique we developed can be regarded as a generalization of the K-means clustering algorithm to visualize Industrial objects, Building, Aerial Photography, Faces and Optical characters. The segmented images are useful for image display.

REFERENCES

[1] Rafael C. Gonzalez, Richard E. Woods "Digital Image Processing" Second Edition Prentice-Hall Inc, 2002.

[2] M.G.H. Omran, A. Salman and A.P. Engelbrecht "Dynamic clustering using particle swarm optimization with application in image segmentation" Pattern and Applica, 2005.

[3] B. Jeon, Y. Yung and K. Hong "Image segmentation by unsupervised sparse clustering" pattern recombination letters27science direct, 2006.

[4] A.K. Jain , M.N. Murty ,P.J.Flynn "Data Clustering: a review" ACM Compute, Surveys 31,1996.

[5] J. Bhaskar, "VHDL primer", second edition BS Publication, 2005.

[6] Shi Zhong and Joydeep Ghosh, "A Unified Framework for Model-based Clustering", Journal of Machine Learning Research 4(USA), 1001-1037, 2003.

[7] J.B.MacQueen "Some methods for classification and analysis of multivariate observation", In: Le Cam, L.M. Neyman, J.(Eds.), University of California,1967.

[8] C. Bouman and B. Liu, "Multiple resolution Segmentation of textured images", IEEE Trans. Patt. Anal. Machine Intell. Vol.13 no.3 pp.99-113, Feb.1991.

[9] Chalmond, "Image restoration using an estimated Markov Model" using Processing. vol.15 no.2 pp.115-129, 1998.