

# Revived Fuzzy K-Means Clustering Technique for Image Segmentation

Jadhav Swapnil N.  
Embedded Systems, BAMU University,  
MIT College of Engineering,  
Beed Bypass road, Aurangabad, India

Prof. Sarita V. Verma  
Dept. of Electronics & Communication, BAMU  
University, MIT College of Engineering,  
Beed Bypass road, Aurangabad, India

**Abstract**— In this paper we have presented an approach in regards with image segmentation by applying K-means clustering algorithm. K-means algorithm is most widely used algorithm in the many fields. The paper proposes some examples of simple images taken by normal mobile phone and then that image segmentation is done using K-means clustering algorithm. K-means algorithm brings forward exact segmentation results as compared to other clustering techniques. Initially, the pixels are grouped in a particular cluster based on its colour and distance between other pixels, where the clustering process is achieved. Afterwards the clusters are blended to specific regions. Therefore this algorithm is providing new aspect of image segmentation which will be widely used for image retrieval. The algorithm shows intended results of our tentativeness for improving clustering with respect to computational time.

**Keywords**— Clustering, K-means Algorithm, Segmentation, Fuzzy clustering.

## I. INTRODUCTION

Normally images are way of conveying emotions, feelings or certain information in a easier way. We can easily get the information from the image by just looking at it, but the same information should be extracted by processing unit for performing certain task related to it. So the image is segmented into different clusters for getting related information from the image. Thus due to image clustering it becomes easy for passing certain information as well as it also helps in understanding in a proper way.

Clustering is a technique in which the image is divided into number of parts depending upon the number of clusters or in other words the value of K specified earlier. The clusters or parts of image are formed taking into consideration certain parameters. In a cluster the number of pixels present has a particular similarity depending upon the constraint applied. But if we use fuzzy clustering technique, then a particular pixel might be present in more one cluster. As stated earlier the clustering process can be used to retrieve information to create clusters by assigning certain pixels to a specific cluster having similar parameters. The cluster formed by the collection of certain number of pixels having similar parameters such as for example brightness, colour, phosphorescent etc. By creating clusters of image using any of the segmentation technique, image analysis and processing becomes easier, simple and faster. Because of more efficiency

and simplicity, revived K-means clustering technique used for segmentation of mobile images [1]. In a cluster all the pixels are interlinked with each other. For creating such clusters the process used is called as Fuzzy K-means [2] (Lloyd's algorithm). Image segmentation is used in many areas such as compression, retrieval, detection and enhancement of image.

## II. LITERATURE SURVEY

There are different techniques of segmentation, as in segmentation means in simple words its grouping of similar parameter pixels grouped together for forming a segment. Different user may think of different parameter for forming cluster, depending upon the requirement or area of interest. But in fuzzy cluster formation a particular pixel may be present in more than one cluster. The parameters which are normally taken into consideration for forming cluster or segments of a particular image can be colour, gray level (if the image is black and white), texture, brightness etc [3][4]. After the segmentation process all the clusters collectively completes the complete image. A effective K means algorithm performs good clustering with exact number of segments by reducing suggested cost function [5]. The algorithm is based on influenced that automatically get the value of K and selecting the good initial points [6]. K-means algorithm is reported to be implemented in spatial clustering [7], in Support Vector Machines [8], using simple partitioning [9] and for clustering categorical data [10] etc. K means algorithm is taken as an account which shows that modified K means algorithm can give better results [10].

In earlier days K means clustering algorithm was only used in number of applications related to the genetical studies. The changes in K means clustering algorithm for genetical studies, genetic K-means clustering algorithm called GKMCA is a combination of a genetic algorithm (GA) and the iterative optimal K-means algorithm (IOKMA) for clustering in genetical datasets. The GKMCA is better as compared to IOKMA [11]. Fast Genetic K-means Algorithm (FGKA) inspired by the Genetic K-means Algorithm (GKA) is considered to execute much faster than GKA [12].

As there are number of different clustering algorithms, for example, hierarchical clustering, self-organizing maps (SOM), modified k-means clustering etc is effective tool of analysis for data and images. Hierarchical clustering

algorithm is used for gene expression patterns. This algorithm creates hierarchy from top to bottom principle for self-organizing tree. As compared to traditional hierarchical agglomerative clustering (HAC) algorithm it was considered more effective [13]. Fuzzy adaptive resonance theory (Fuzzy ART) or Fuzzy K means is used to gene clustering of DNA microarray data and the clustering result was better than other methods of clustering including hierarchical clustering, SOM, and k-means clustering [13]. Lloyd's K-means Clustering and the Progressive Greedy K-means are used for recognizing pixels with same parameters, such as genes with similar expression patterns. The analysis was performed using both a gene expression level sample and on randomly-generated datasets in three-dimensional space. Of the two types, Lloyd's K-means Clustering algorithm was observed to be more efficient and gives better results, so we have implemented K-means clustering algorithm.

### III. FLOW OF SYSTEM

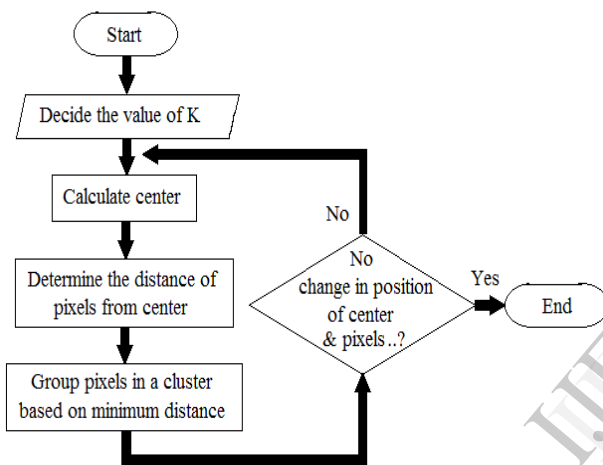


Fig. 1. Flowchart of K-means clustering algorithm.

In K-means clustering algorithm that is used to segment an image into K clusters. The above flowchart works as follows:

Here we initially decide the value of K, which can give relatively better result for analysis. Then we have to calculate K cluster centers (randomly). Then we allot every pixel in the image to the cluster that will reduce the distance between the pixel and the cluster center. After this we have to again calculate the cluster centers by averaging all of the pixels in the cluster. Then we have to repeat the process again and again unless the algorithm converges (i.e. neither the pixels nor the centers changes their position).

K-means algorithm is guaranteed to converge, but it may not give optimal results. The quality of the output may rely on the initial set of clusters and the number of clusters.

### IV. THE FUZZY K-MEANS CLUSTERING ALGORITHM

This is a new type of clustering algorithm Fuzzy K-means clustering algorithm is presented. For implementation of FKM, consider a image with  $R \times S$  pixels (i.e.,  $R$  denotes columns and  $S$  denotes rows of image) to be clustered. Let  $p(x,y)$  be the considered pixel and  $c_j$  as the  $j$ -th centre, where  $x = 1, 2, \dots, R, y = 1, 2, \dots, S$  and  $j = 1, 2, \dots, n_c$ . Based on the

initially mentioned consideration, for the conventional K-means, all pixels will be assigned to the nearest centre based on Euclidean distance. The new position for each centre is calculated using [14]:

$$c_j = \frac{1}{n_{c_j}} \sum_{x \in c_j} \sum_{y \in c_j} p(x, y) \quad (1)$$

Just like conventional FCM, the process of assigning every pixel to more than one cluster is based on the following function [14]:

$$M_{jp(x,y)}^m = \frac{1}{\sum_{k=1}^{n_c} \left( \frac{d_{jp(x,y)}}{d_{kp(x,y)}} \right)^{2/(m-1)}}; \text{ if } d_{kp(x,y)} > 0, \forall j, p(x,y) \quad (2)$$

$$\left. \begin{aligned} M_{kp(x,y)}^m &= 1 \\ M_{jp(x,y)}^m &= 0; \text{ for } p(x,y) \neq k \end{aligned} \right\} \text{ if } d_{kp(x,y)} = 0; \quad (3)$$

where  $d_{jp(x,y)}$  is distance from point  $(x,y)$  to the current cluster centre  $j$ ,  $d_{kp(x,y)}$  is distance from point  $(x,y)$  to other cluster centres  $k$ ,  $n_c$  is number of centres and  $m$  is an integer,  $m > 1$  which determine the degree of fuzziness. The normal FCM, the pixels are allocated to a particular cluster based on the distance between two points. While, in K-means each cluster must have meaningful number of pixels and the differences among the clusters should be minimized to ensure a good clustered output [16]. Initially centers are assigned to a certain value. Fuzzy cluster formation is applied here as it makes a particular pixel to be present in more than one cluster. The membership function  $M_{jp(x,y)}^m$  is calculated by using (2) and (3). In FKM, the membership function is performed within the pixels. Normally, the nearer the pixel, maximum is the effect on the clustering output. Thus, we get a very good effect on the degree of membership, which is afterwards used in finding out new location of center. In effective clustering process, changing of center position concept is put forward in the FKM algorithm which tells, that every cluster must have a certain value of belongingness which shows the bounding between the centre and its members. So, after indicating the membership for each pixel, the degree of belongingness,  $B_j$  for every cluster is determined. This relationship indicates a very good effect on the clustering output through the degree of belongingness. Maximum degree of belongingness gives a good relationship between the centre and its pixels in the same cluster, thus we get a good data clustering results. The degree of belongingness,  $B_j$  is determined by:

$$B_j = \frac{c_j}{M_{jp(x,y)}^m} \quad (4)$$

For improving the clustering technique, it is necessary for the degree of membership is used effectively based on amount of belongingness. While putting forward the concept of belongingness, Fuzzy K means states that each cluster should have members with a strong relationship formed between them and the difference of belongingness among the clusters

should be reduced [16]. The value of  $M^{m_{jp}(x,y)}$ , is updated in the iteration according to:

$$\left(M^{m_{jp}(x,y)}\right) = M^{m_{jp}(x,y)} + \Delta M^{m_{jp}(x,y)} \quad (5)$$

Where  $M^{m_{jp}(x,y)}$  is the new membership.  $M^{m_{jp}(x,y)}$ , is given as:

$$\Delta M^{m_{jp}(x,y)} = \alpha(c_j)(e_j) \quad (6)$$

where  $\alpha$  is a considered having constant value between 0 and 1 and it is generally set to 0.1. Simultaneously, calculate the value of  $e_j$  considering :

$$e_j = B_j - \hat{B}_j \quad (7)$$

where  $B_j$  is the normalized value for belongingness. At last new centre positions of the entire existing clusters are calculated based on the new membership function according to:

$$c_j = \frac{\sum_{x \in c_j} \sum_{y \in c_j} (M^{m_{jp}(x,y)}) p(x,y)}{\sum_{x \in c_j} \sum_{y \in c_j} (M^{m_{jp}(x,y)})} \quad (8)$$

(Note: Amount of fuzziness,  $m$  is set to 2) All the steps are repeated until the centre position does not change.

## V. OVERVIEW OF SYSTEM

Let  $X = \{x_1, x_2, x_3, \dots, x_n\}$  be the set of data points and  $V = \{v_1, v_2, \dots, v_c\}$  be the set of centers.

- 1) Without a particular aim select 'c' cluster centers.
  - 2) Find the distance between each data point and cluster centers.
  - 3) Assign the data point to the cluster center whose distance from the cluster center is minimum of all the cluster centers.
  - 4) Again calculate the new cluster center using:
 
$$v_i = (1/c_i) \sum_{j=1}^{c_i} x_j$$
- Where ' $c_i$ ' represents the number of data points in  $i^{th}$  cluster.
- 5) Recalculate the distance between each data point and new obtained cluster centers.
  - 6) If no data point was reassigned then stop, otherwise repeat from step 3.

## VI. STIMULATION RESULTS

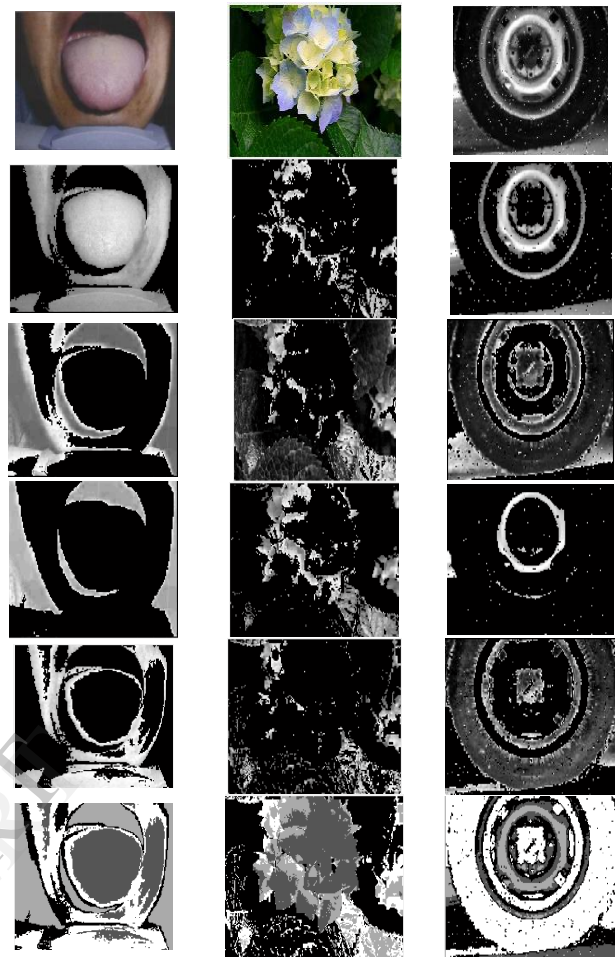


Fig. 2. First row shows original images (a) Mouth (b) Flowers (c) Tyre , Second, third, fourth and fifth rows contain K=1,2,3,4, Final row shows addition of all earlier rows output.

In this paper we have applied Fuzzy K-means clustering algorithm for certain images which are captured using normal camera for example mouth, flowers and tyre. The second row shows the results, where we have applied K=1, for third row we have applied K=2 and so on.. But the last row shows the addition of all rows i.e. from row second to fifth.

## VII. ADVANTAGES OF K-MEANS CLUSTERING.

- Can be easily implemented.
- Will defiantly converge in few initial iteration.
- It is linear complexity.

## VIII. DISADVANTAGES OF K-MEANS CLUSTERING.

- Need to specify value of K in advance.
- Sensitive to outliers.
- May yield empty clusters.

## IX. APPLICATIONS OF K-MEANS CLUSTERING.

- Web page classification.
- Classification analysis.

- Artificial intelligence.
- Image processing.
- Machine vision.
- Email filtering.
- Pattern recognition.
- Unsupervised learning of neural networks.

#### X. FUTURE SCOPE

- This clustering process can be completely automated on data, is done in automated fashion.
- The speed of making clusters of particular image can be increased by using nearest neighbor graph.
- We can increase the system productivity with minimum efforts.

#### XI. CONCLUSION

We have implemented K-means clustering algorithm for smaller values of K. This algorithm gave us better results and it converged at much faster rate. If we consider larger values of k, then the results might have been very coarse and many clusters would have appeared in the images at distinct areas. This is due to as Euclidean distance is not a very good parameter for cluster formation. Different initial clustering may result in different final clusters. So it is necessary to run the code several number of times for same and different values of K in-order to compare the quality of clusters obtained. We make clusters of image into different regions with the help of "Fuzzy K-means algorithm" and then we check the consistency of regions according to predicate.

Aim of developing an exact and more consistent image which can be used in finding cracks, face recognition, finger print recognition and in locating an object clearly from a satellite image etc [15]. It works nicely when segments are not well separated from each other. We revived an idea of clustering of images based on the different parameters of the image. It minimizes cluster to cluster variance, but it does not make sure that the final result has a global minimum of variance.

#### REFERENCES

- [1] S. P. Lloyd, "Least squares quantization in PCM," IEEE Trans. Inf.Theory, vol. IT-28, no. 2, pp. 129-136, Mar.1982.
- [2] J. Shi and J. Malik, Normalized cuts and image segmentation, IEEE Trans. Pattern Anal.Mach. Intell., vol. 22, no. 8, pp. 888-905, Aug.2000.
- [3] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Sonar image segmentation using a hierarchical MRFmodel," IEEE Trans. Image Process., vol. 9, no. 7, pp.1216-1231, Jul. 2000.
- [4] F. Destrempe, J.-F. Angers, and M. Mignotte, "Fusion of hidden Markov random field models and its Bayesian estimation," IEEE Trans. Image Process., vol. 15, no. 10,pp. 2920-2935, Oct. 2006.
- [5] Mingwei Leng, Haitao Tang, Xiaoyun Chen, "An Efficient K-means Clustering Algorithm Based on Influence," Proceedings of the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD 2007) - Volume 02 815-820.
- [6] Zhe Zhang Junxi Zhang Huifeng Xue, "Improved K-Means Clustering Algorithm," Proceedings of the 2008 Congress on Image and Signal Processing, Vol. 5 - Volume 05 Pages 169-172, 2008.
- [7] Tian, Da-Dong; Deng, Wei Jisuanji Gongcheng yu Yingyong, "Application of the improved K-means clustering algorithm in support vector machine.," Computer Engineering and Applications, Vol. 42, no. 32, pp. 161-163. 11 Nov. 2007.
- [8] Hung, Ming-Chuan; Wu, Jungpin; Chang, Jin-Hua; Yang, Don-Lin, "An efficient k-means clustering algorithm using simple partitioning," Journal of Information Science and Engineering. Vol. 21, no. 6, pp. 1157-1177. Nov. 2005.
- [9] Ming Lei, Pilian He, Zhichao Li, "An Improved K-means Algorithm for Clustering Categorical Data," Journal of Communication and Computer, Aug. 2006, Volume 3, No.8 (Serial No.21).
- [10] Fang-Xiang Wu, "A genetic weighted k-means algorithm for clustering gene expression data," IMSCCS 2007. Second International Multi-Symposiums on Volume , Issue , 13-15 Aug. 2007 Page(s):68-75.
- [11] Yi Lu, Shiyong Lu, Farshad Fotouhi, Youping Deng, Susan J. Brown, "FGKA: A Fast Genetic K-means Clustering Algorithm," Symposium on Applied Computing, Proceedings of the 2004 ACM symposium on Applied computing, Pages: 622-623, 2004.
- [12] Adil M. Bagirov and Karim Mardaneh, "Modified global k-means algorithm for clustering in gene expression data sets," ACM International Conference Proceeding Series; Vol. 246 Pages: 23 - 28, 2006.
- [13] Pablo Tamayo, Donna Slonim, Jill Mesirov, Qing Zhu, Sutsik Kitareewan, Ethan Dmitrovsky, Eric S. Lander and Todd R. Golub, "Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation," PNASMarch 16, 1999 vol. 96 no. 6 2907-2912.
- [14] M. Y. Mashor, "Hybrid Training Algorithm for RBF Network," International Journal of Computer, Internet and Management, vol. 8, no. 2, pp. 50-65, 2000.
- [15] S.Mary Praveena, Dr.IlaVennila, "Optimization Fusion Approach for Image Segmentation Using K-Means Algorithm," International Journal of Computer Applications (0975 - 8887)Volume 2 - No.7, June 2010.
- [16] Siti Noraini Sulaiman and Nor Ashidi Mat Isa, "Adaptive Fuzzy-K-means Clustering Algorithm for Image Segmentation," IEEE Transactions on Consumer Electronics, Vol. 56, No. 4, November 2010.