Image Segmentation for Nature Images using K-Mean and Fuzzy C-Mean

Anita V. Gawand Prashant Lokhande Sulekha daware Umesh Kulkarni
MGMC.E.& T,Kamothe MGMCE&T ,Kamothe D.M.C.E. Airoli K.G.C.O.E.,Karjat

ABSTRACT
Clustering can be considered the most important unsupervised learning problem. It deals with finding a structure in a collection of unlabeled data. We defined cluster is process of organizing objects into group whose member are similar in some way. In my paper we taken natural image and we apply unsupervised learning algorithm k-mean and Fuzzy c-mean that solve the well known clustering problem.

Keyword: segmentation, k-mean, fuzzy c-mean

1. INTRODUCTION
In computer vision, segmentation refers to the process of partitioning a digital image into multiple segments (sets of pixels, also known as superpixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. These approaches can be roughly classified into three classes: (i) edge-based approaches, which rely on edges found in an image by edge detecting operators, such as Canny method [5], (ii) region-based approaches, which group pixels into homogeneous regions and segment the image to some major areas, such as normalized cut [1] and region growing [2], (iii) clustering-based methods, which segment the feature space of image to several clusters and get a sketch of the original image [4], such as mean shift [3] and fuzzy C-means (FCM) clustering.

Image clustering analysis is one of the core techniques for image indexing, classification, identification and segmentation for medical, natural, still images process. Here we apply k-mean and fuzzy c-mean to natural images.

Real world image segmentation problems actually do have multiple objectives, i.e., minimize overall deviation (intra-cluster spread of data), maximize connectivity (inter-cluster connectivity), minimize the number of features or minimize the error rate of the classifier etc.

Cluster is process which partition a given data set into homogeneous group base on given feature such that similar object are kept in a group where as dissimilar object are in different group. Clustering algorithm classified as

1) Exclusive clustering (hard clustering)-k-mean, k-medoids, X-mean, ISODATA
2) Overlapping clustering- FCM
3) Hierarchical clustering-agglomerative clustering, divisive cluster

2. SYSTEM WORK
We take image images as nature image. We first apply k-mean algorithm and get segmented image. On same image we apply Fuzzy c-mean standard algorithms and we get segmented image.

2.1). K-Means clustered algorithm
K-mean is unsupervised clustering algorithm its iterative technique that is used to partition an image into k-cluster. Some time k-mean algorithm is called hard cluster because data is divided into distinct cluster where each data element belongs to exactly one cluster[14]

2.1.1) Algorithmic steps for k-means clustering
Let \( X = \{x_1, x_2, x_3, \ldots, x_n\} \) be the set of data points and \( V = \{v_1, v_2, \ldots, v_c\} \) be the set of centers.

1) Randomly select ‘c’ cluster centers.

2) Calculate 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) Recalculate the new cluster center using:

\[ \mathbf{v}_j = \frac{1}{c_j} \sum_{i=1}^{c} \mathbf{x}_i \]

where, \( c_j \) represents the number of data points in the \( i^{th} \) cluster.

5) Recalculate the distance between each data point and the new obtained cluster centers.

6) If no data point was reassigned then stop, otherwise repeat from step 3.

2.3) Standard fuzzy C-mean

Fuzzy clustering defined by a probability of membership of each object in one cluster. Fuzzy clustering also called soft clustering because data element can belong to more than one cluster and associated with each element is a set of membership level. These indicate the strength of the association between that data element and a particular cluster. This algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point[15]. More the data is near to the cluster center more is its membership towards the particular cluster center. Clearly, summation of membership of each data point should be equal to one. After each iteration membership and cluster centers are updated according to the formula:

\[ \mu_{ij} = \frac{1}{\sum_{k=1}^{c} (d_{ij} / d_{ik})^{(2/m-1)}} \]

\[ \mathbf{v}_j = \left( \sum_{i=1}^{n} (\mu_{ij})^m \mathbf{x}_i \right) / \left( \sum_{i=1}^{n} (\mu_{ij})^m \right), \forall j = 1, 2, \ldots, c \]

where, \( n \) is the number of data point

\( \mathbf{v}_j \) represent \( j^{th} \) cluster center

\( m \) is fuzzindex where \( m \in [1, \infty] \)

\( c \) represents the number of cluster center

\( \mu_{ij} \) represents membership of \( i^{th} \) data to \( j^{th} \) cluster center

\( d_{ij} \) represents the Euclidean distance between \( i^{th} \) data and \( j^{th} \) cluster center.

Main objective of fuzzy c-means algorithm is to minimize:

\[ J(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (\mu_{ij})^m \| \mathbf{x}_i - \mathbf{v}_j \|^2 \]

Where, \( \| \mathbf{x}_i - \mathbf{v}_j \| \) is the Euclidean distance between \( i^{th} \) data and \( j^{th} \) cluster center.

2.1) Algorithmic steps for Fuzzy c-means clustering

Let \( X = \{x_1, x_2, x_3, \ldots, x_n\} \) be the set of data points and \( V = \{v_1, v_2, v_3, \ldots, v_c\} \) be the set of centers.

1) Randomly select \( c \) cluster centers.

2) calculate the fuzzy membership \( \mu_{ij} \) using:

\[ \mu_{ij} = \frac{1}{\sum_{k=1}^{c} (d_{ij} / d_{ik})^{(2/m-1)}} \]

3) Compute the centers \( \mathbf{v}_j \) using:

\[ \mathbf{v}_j = \left( \sum_{i=1}^{n} (\mu_{ij})^m \mathbf{x}_i \right) / \left( \sum_{i=1}^{n} (\mu_{ij})^m \right), \forall j = 1, 2, \ldots, c \]

4) Repeat step 2) and 3) until the minimum \( J \) value is achieved or \( \| U^{(k+1)} - U^{(k)} \| < \beta \).

where, \( k \) is the iteration step.

\( \beta \) is the termination criterion between \( [0, 1] \).

\( U = (\mu_{ij})_{n \times c} \) is the fuzzy membership matrix.

\( J \) is the objective function.

3. EXPERIMENT DATABASE (SOME PICTUER IN JPEG. FORMAT AND MOBILE IMAGE)
Here we take 37 natural images and we apply k-mean and fuzzy c-mean algorithm. Fuzzy c-mean can’t apply directly to color image. We must convert natural image into gray scale image then we apply standard fuzzy c-mean algorithm. Executed fuzzy algorithm required more time as compare k-mean.

4. EXPERIMENTAL RESULT

V. CONCLUSION
K-Means clustering algorithm is inherently iterative, and no guarantee can be made that it will converge to an optimum solution. The performance of the K-means algorithm depends on the initial positions of the cluster centers. Same thing for Fuzzy c-mean algorithm. Many approaches have been proposed for color image segmentation, but Fuzzy C-Means has been widely used, because it has not adequate for noisy images and it also take more time for execution as compare to method as K-Means.

5. FUTURE SCOPE
K-mean clustering, fuzzy c-mean algorithms we use various application... We can’t use standard fuzzy c-mean algorithm for natural images. Color image segmentation we need changes some part fuzzy c-mean equation and consider color property then we use fuzzy c-mean for color image.

6. ACKNOWLEDGMENTS
My thank to experts name is Azad Naik, Santle Camilus and www.weforanimal.com photo gallery who help me.

7. REFERENCES:


[8] The k-means algorithm - Notes by Tan, Steinbach, Kumar Ghosh.