

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Survey on Image Segmentation in Various Clustering Algorithms.

MD Naresh Karthik, I Kathirvelan, and K Srilatha*.

Department of Electronics and Communication Engineering, Sathyabama University, Chennai, Tamilnadu, India.

ABSTRACT

The initial goal of image segmentation is to cluster image pixels into salient image region. These clustering algorithms can be used for image segmentation. Clustering is process of identifying groups of related images by image segmentation. To achieve the super pixel formation many clustering algorithms can be classified. The reason for using clustering algorithms is to get proper result with high efficiency effective storage image. This deals with different clustering algorithms like Normalized cuts algorithm, K-means clustering algorithm, Fuzzy clustering algorithm and DBSCAN (Density Based Spatial Clustering of Application with Noise).

Keywords: survey, image, clustering algorithms.

**Corresponding author*



INTRODUCTION

An image is projection of three dimensional scene into a two dimensional projection plane. An image is two dimensional function that represents a measure of some characteristics such as brightness or color of viewed scene .Mathematical representation can be done by analog images for a continuous range of values representing position and intensity .Digital image is composed of picture elements called pixels Chun-Yan Yu et al [4]. Each pixel is a sample of an original image. More samples typically provide more accurate representations of the original image. The intensity of every pixel is variable. In color image systems, a color is typically represented by three or four component intensities such as red, green, and blue, or cyan, magenta, yellow, and black.

Compared to analog images digital image plays an important role in daily application such as Satellite Television, Computer Tomography etc., Digital image processing widely used in fields such as medicine, forensics, remote sensing, communication and automobile.

Segmentation

Image segmentation is a process of partitioning a digital image into multiple segments. The main aim of Image segmentation is domain independent partitioning of an image into a set of disjoint region that are visually different, homogeneous and meaningful with respect to some characteristics. Image segmentation is typically used to locate objects and boundaries (lines, curve) in images. The process of image segmentation is to divide the images into regions. Image segmentation is a major part in digital image processing Melissa, S., Srilatha, K et al [9]. Image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. There are two properties in image segmentation they are discontinuity and homogeneity. Methods based on discontinuities are called as boundary based methods and methods based on homogeneity are called as region based methods. The main objective of image segmentation is to cluster pixels and salient image regions.

There are two segmentation they are local segmentation and global segmentation. The process of segmenting sub-images which are small windows on a whole image is called as Local segmentation. The process of segmenting the whole image is called Global segmentation. The number of pixels used in local segmentation is very much lower than global segmentation Siva Kumar, B., Srilatha, K et al [6]. The main objective of image segmentation is to partition an image into a meaningful region with respect to particular application.

There are three approach in image segmentation they are region approach, boundary approach and edge approach. In region approach each pixel is assigned to a particular object or region. In boundary approach the boundaries exist between the regions. In edge approach edges are found first and they are linked to form required boundaries. Region growing is approach to image segmentation were neighboring pixels are examined and added to region class if no edge is detected. If adjacent regions are found then region merging algorithm are used in which weak edges are dissolved and strong edges are considered. Region growing offers several advantages over conventional segmentation techniques. The algorithms is also very stable with respect to noise. Region splitting and merging is a continuous approach after region growing. Region splitting and merging is image segmentation technique that takes spatial information into consideration. The segmentation is based on measurement taken from the image such as color, depth, gray level, texture. Image segmentation involves several process such as pattern representation, feature selection, pattern proximity.

Image segmentation is based on two properties of image (1) intensity values involving discontinuity that involves rapid or abrupt changes in intensity as edges. (2) Resemblance that refers to partitioning a digital image into region according to some pre-defined likeness criterion. Our summary is presented for K means clustering algorithm, Fuzzy c means clustering algorithm, Normalized cuts algorithm, Kernel Fuzzy C-means clustering algorithm and DBSCAN(Density Based Spatial Clustering of Application with Noise).

Properties of images can be grouped such as brightness, pixel value and color. By converting pixels into super pixels, it reduces the computational cost and complexity. Each super pixel can be treated as a node in graph and edge weights in group based algorithms between nodes the initial stage of clustering during each

iteration it extracts the cluster from the previous iteration to obtain the better segmentation in gradient ascent based algorithms. For simple images, the segmentation process is clear and effective due to small pixels variation. For complex images the utility for subsequent processing becomes very touch process.

Clustering

Clustering is process of grouping a set of objects in the same group are more similar to each other and dissimilar to those in the other group Jian Hou et al [1]. The requirement that should be satisfied by clustering algorithms are scalability, dealing with different type's quality, ability to deal with noise, interpretability and usability.

Clustering of data is a method in which large set of data's are grouped into clusters of smaller sets of similar data. A clustering algorithm endeavor to find natural group of components based on some resemblance. The centroid of group of data sets are found by the clustering algorithm. To evaluate cluster membership most algorithm evaluate the distance between cluster centroid and points. The clustering algorithm's output is basically a statistical description of cluster centroids with number of components in each cluster.

Some image may not have the appropriate keyword to describe them therefore the image search will be difficult. Relevance feedback is one of the solution to overcome this problem. This technique utilizes the user feedback and hence reduces the possible errors and redundancy. This technique uses a Bayesian classifier. Which also deals with positive and negative feedback.

Clustering algorithms

clustering algorithm helps us to access the relationship among pattern of data set by organizing the pattern into groups or clusters such the pattern within are more similar to each other patterns belong to different patterns Siva Kumar, B., Srilatha, K et al [6]. Clustering refers to the classification objects into group according to certain properties of these objects. There are two process categorized in clustering they are supervised clustering and unsupervised clustering. Supervised clustering works only if human interaction to decide clustering criteria and the unsupervised clustering decides the clustering criteria by itself. A variety of clustering algorithm have been introduced to make segmentation more effective. some of the clustering algorithm included in this paper are K means clustering algorithm, Fuzzy c means clustering algorithm, Normalized cuts algorithm, Kernel Fuzzy C-means clustering algorithm and DBSCAN(Density Based Spatial Clustering of Application with Noise).

Literature Survey

Normalized cuts algorithm

Image segmentation with low computational burden has been highly regarded as important goals for researchers. Using Normalized cuts algorithm is one of the popular image segmentation method. It is unfavorable for high resolution images to have its resolution reduced as high detail information is not fully made used when the critical objects with weak edges is coarsened undesirable after its resolution is reduced. Thus a graph based image segmentation method done in multistage manner is proposed in Normalized cuts algorithm.

K means clustering algorithm

It is one of the simplest method in supervised classification. It does not require training data. It is iterative procedure. K means clustering algorithm clusters the date by iteratively computing a mean intensity for each class and segmenting the image by classifying each pixel in the class with closest mean. Clustering based on the optimization of an overall measures is a fundamental approach explored since the early days of pattern recognition. The most popular method for pattern recognition is K means clustering a centroid vector is computed for every cluster Melissa.S, Srilatha.K et al [9]. The centroid must be chosen that it should minimize the total distance within the clusters.

Steps:

(i) Choose K initial clusters $z_1(s), z_2(s), \dots, z_k(s)$

(ii) At kth iterative step, distribute the samples x among K clusters using the relation $x \in C_i(k)$ if $||x - z_m(k)|| \leq ||x - z_n(k)||$

For $m=1, 2, \dots, K, m \neq n$, where $C_i(k)$ denotes the set of samples whose cluster centre is $z_n(k)$

(iii) Compute new cluster centers $z_n(k+1), n=1, 2, \dots, K$, such that the sum of squared distance from all points in $C_i(k)$ to new cluster is minimized. The measure which minimizes this is only the sample mean of $C_i(k)$. So the new cluster is given

by $Z_n(k+1) = 1/N_n \sum x, n = 1, 2, \dots, K$ Where N_n is the number of samples in $C_n(k)$

(iv) If $z_n(k+1), n = 1, 2, \dots, K$, the algorithm has converged and the procedure is terminated. If not go to step 2. shown in Fig.1

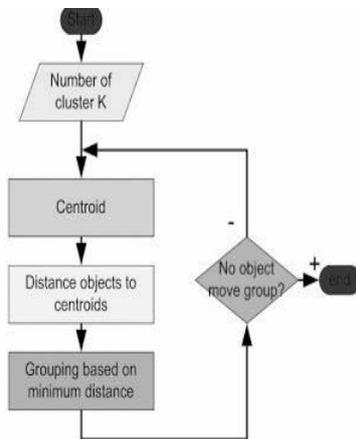


Fig.1: Flowchart of K means clustering algorithm

Fuzzy C means clustering algorithm

Clustering method can be c

lassified either hard or fuzzy depending on whether a pattern data belongs exclusively to simple cluster or several cluster with different values. In hard clustering membership values of zero or one is assigned to each pattern data Siva Kumar, B., Srilatha, K et al [6]. In Fuzzy clustering the values between zero and one is assigned to each pattern by membership function. Fuzzy clustering methods can be considered to the superior to those of the hard counterparts since they can represent the relationship between the input pattern data and cluster. It is used to minimize the heuristic global cost function. Fuzzy c means clustering algorithm iteratively updates the cluster centroid and estimates the class membership function by using gradient decent approach.

Steps

(i) Fix $c, m, A, ||k|| |A$. Choose an initial matrix $U^{\wedge}(0) \in M_{fc}$. Then at step $k, k = 0, 1 \dots LMAX$.

(ii) Compute c means $v^{\wedge}(k), l = 1, 2, \dots, c$.

(iii) Compute an updated membership matrix $U^{\wedge}(k+1) = [u_{ik}^{\wedge}(k+1)]$

(iv) Compare $U^{\wedge}(k+1)$ to $U^{\wedge}(k)$ in any convenient matrix form. If $|| U^{\wedge}(k+1) - U^{\wedge}(k) || < \epsilon$, stop. Otherwise, set $U^{\wedge}(k) = U^{\wedge}(k+1)$ and return to step(ii) shown in Fig.2

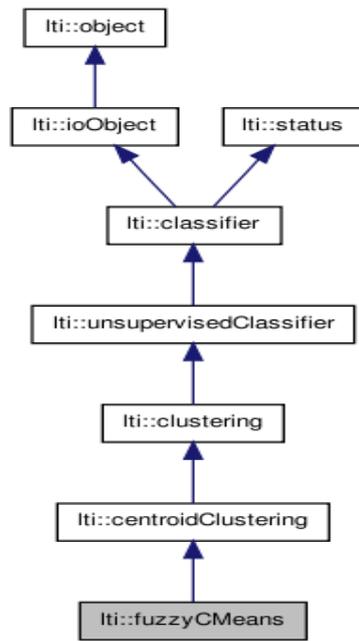


Fig.2: Fuzzy C means clustering algorithm

Fuzzy kernel C means clustering algorithm (FKCM)

The main idea of Fuzzy Kernel C means clustering algorithm is the combination of both k means clustering algorithm and Fuzzy c means clustering algorithm. Recently many traditional linear methods have been generalized to powerful corresponding nonlinear forms using Mercer kernel, including Principal Component Analysis (PCA). C means clustering and Fisher determination. Also many new methods will employ Mercer Kernel including SVM classification, SVM regression and Kernel PCA are proposed from Covers theorem, projecting into a high dimensional features space is assumed to making linearly non-separable problem more likely to be linearly separable Vinod Kumar Dehariya et al [3]. The Kernel based concept has been adopted for unsupervised learning. Basically Mercer Kernel, the Fuzzy Kernel c means clustering algorithm is derived from the Fuzzy c means clustering algorithm. The Fuzzy kernel c means clustering algorithm that provides image clustering can improve accuracy significantly compared to classical Fuzzy c means clustering algorithm

DBSCAN (Density Based Spatial Clustering of Application with Noise)

Clustering image pixel is a important image segmentation technique. While a large amount of clustering algorithm have been introduced and some of them generate impressive clustering results But their performance always depends on user specified parameter. To avoid is this problem a parameter free clustering algorithm has been produced it is known as DBSCAN (Density Based Spatial Clustering of Application with Noise) James C Bezdek et al [2]. It is density based algorithm giving a set of points in some space, it groups together points that are closely packed together, marking as outliers points that lie alone in low density based regions. DBSCAN requires to two parameters they are eps and minimum number of points required to form a dense region (min pts).

Finally, we present the procedures of the DSets-DBSCAN algorithm as follows.

- 1) Calculate the pair wise similarity matrix of data to be clustered.
- 2) Apply histogram equalization to transform the similarity matrix.
- 3) Extract a dominant set *D* with the transformed similarity matrix as input.
- 4) Set *MinPts* = 3 and determine *Eps* with Eq. (12).
- 5) Starting from an arbitrary point in *D*, retrieve all density reachable points, which form a

Cluster C.

- 6) Remove the rows and columns related to the data in C from the similarity matrix.
- 7) Go to Step 3, until the clustering is accomplished.

CONCLUSION

Segmentation is an important process in advance image analysis and computer vision and it is an ongoing research area although a dense literature is available. A survey of highlighting different clustering algorithm used for image segmentation has been presented. Image segmentation basics, Clustering basics and Clustering algorithm in image segmentation have been analyzed. Using clustering algorithms, image segmentation can be done in an effective way. Clustering algorithm are used to enhance the efficiency of the image retrieval process. Finally comparing the other algorithm mentioned in this survey paper, DBSCAN (Density Based Spatial Clustering of Application with Noise) is the leading clustering algorithm.

REFERENCES

- [1] Jian Hou, Huijun Gao, Xuelong Li, "Dsets – DBSCAN: A parameter free clustering algorithm", IEEE transaction on image processing, vol. 25, No. 7, July 2016.
- [2] James C Bezdek, Robert Ehrlich, William Full, "FCM : The Fuzzy C means clustering algorithm", IEEE transaction on image processing, vol.10, No.2-3, pp.191-201,1984.
- [3] Vinod Kumar Dehariya, Shailandra Kumar Srivastva, R.c.Jain, "Clustering of image data sets using K means and Fuzzy c means algorithm", IEEE transaction on 2010 International Conference on Computational Intelligence and Communication Networks.
- [4] Chun-Yan Yu, Ying Li, Ai-lian Liu, Jing-hong Liu, "A novel modified kernel fuzzy c means clustering algorithm on image segmentation", IEEE International Conference on the 14th IEEE International Conference on Computational Science and Engineering.
- [5] Mei Yeen Choong, Wei Leong Khong, Wei Yeang Kow, Lorita Angeline, Kenneth Tze Kin Teo, " Graph based image segmentation using k means clustering and normalized cuts algorithm", 2012 Fourth International Conference on Computational Intelligence, Communication Systems and Networks.
- [6] Siva Kumar, B., Srilatha, K, "A novel method to segment blood vessels and optic disc in the fundus retinal images", Research Journal of Pharmaceutical, Biological and Chemical Sciences, VOL 7/ISSUE 3/MAY 2016/PP 365-373.
- [7] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," Science, vol. 315, no. 5814, pp. 972–976, Feb. 2007.
- [8] Long Chen, Jing Zou and C. L. Philip Chen, "Kernel Non-local shadowed C means for image segmentation", 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) July 6-11, 2014.
- [9] Melissa, S.,Srilatha, K, " A novel approach for pigmented epidermis layer segmentation and classification" International Journal of Pharmacy and Technology, VOL 8/ISSUE 1/MAR 2016/PP 10449-10458..
- [10] Guifang Shao ,Shunxiang Wu, Tiejun Li, "cDNA microarray image segmentation with an improved moving k means clustering method", Proceedings of the 2015 IEEE 9th International conference on semantic computing.
- [11] X. Wang, Y. Tang, S. Masnou, and L. Chen, "A global/local affinity graph for image segmentation," IEEE Trans. Image Process., vol. 24,no. 4, pp. 1399–1411, Apr. 2015.
- [12] M. Pavan and M. Pelillo, "Dominant sets and pairwise clustering," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 1, pp. 167–172, Jan. 2007.
- [13] J. Hou, Q. Xia, and N.-M. Qi, "Experimental study on dominant sets clustering," IET Comput. Vis., vol. 9, no. 2, pp. 208–215, 2015.
- [14] J. Hou and M. Pelillo, "A simple feature combination method based on dominant sets," Pattern Recognition., vol. 46, no. 11, pp. 3129–3139,2013.