NOVEL APPROACH BASED AN IMAGE SEGMENTATION USING K-MEAN CLUSTERING

Mrs. PARAMESWARI. A[1], Dr. V.J. ARUL KARTHICK[2]

[1]Electronics and Communication Engineering Department
[2]JCT College of Engineering and Technology, Coimbatore.

ABSTRACT - Now a day's image processing is an important role in real life application. Image segmentation is the maneuver of partitioning an image into a set of connected sets of pixels. Partitioning covers the whole image. Each row of data represents one pixel. Each column of data represents one image. Have k-means decide which cluster each pixel belongs to. Here, we convert the pixel into images for all class index. It is supposed to be monochrome, get the dimensions of the image. If it is not a gray image then that notifies the colour image. Use subjective computation of all channels to generate a gray scale image. Convert it to gray scale by taking only the green channel, Convert it to gray scale by taking only the green channel, get the data for doing k-means. We will have one column for each colour channel. These papers describe an image segmentation method based on k-mean clustering algorithm to get a segmented image. Finally we get the segmented image.

1. Introduction

Image segmentation is used to separate an image into several "meaningful" parts. It is an old research topic, which started around 1970, but there is still no robust solution toward it. There are two main reasons, the first is that the content variety of images is too large, and the second one is that there is no benchmark standard to judge the performance. For example, we show an original image and two segmented images based on different kinds of image segmentation methods. The one separates the sky into several parts while that misses some detail in the original image. Every technique has its own advantages also disadvantages, so it's hard to tell which one is better. There are tons of previous works about image segmentation, great survey resources could be found. From these surveys, we could simply separate the image segmentation techniques into three different classes (1) feature-space based method, (2) image-domain based method, and (3) edge-based method. The feature-space based method is composed of two steps, feature extraction and clustering. Feature extraction is the process to find some characteristics of each pixel or of the region around each pixel, for example, pixel value, pixel color component, windowed average pixel value, windowed variance, Law's filter feature, Tamura feature, and Gabor wavelet feature, etc.. After we get some symbolic properties around each pixel, clustering process is executed to separate the image into several "meaningful" parts based on these properties. This is just like what we have tried from DIP homework 4,
where we used Law's feature combined with K-means clustering algorithm.

Image-domain based method goes through the image and finds the boundary
between segments by some rules. The main consideration to separate two pixels into different segments is the pixel value difference, so this kind of methods couldn't deal with textures very well. Split and merge, region growing, and watershed are the most popular methods in this class. The third class is edge-based image segmentation method, which consists of edge detection and edge linking.

Although there have been many kinds of existed methods, some common problem still can't be solved. For the accurate boundaries between segments are still hard to determine because features take properties around but not exactly on each pixel. Then only uses the pixel value information, which may result in over-segmentation on texture regions. Finally the edge detection process makes always suffer the over-segmentation problem.

2. K-means clustering

We have two goals to achieve: maintain large distance among data points in different clusters and small distance among data points in the same clusters. This is the famous tool for unsupervised classification problems.

2.1 K-Means Algorithm

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as barycentre of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop we may notice that the k centroids change their location step by step until no more changes are done.

2.2 Algorithm for K Means Algorithm

Step 1:Read the image.

Step 2:Get the number of clusters to be formed.

Step 3:Convert the color image into its corresponding gray image.

Step 4:Resize the two dimensional image into one dimensional array of length "r×c".

Step 5:Find the intensity range of the image.

Range = [(Maximum intensity value) - (Minimum intensity value)]

Step 6:Find the centroid value Centroid 1 = Range/Number of clusters

Centroid 2 = (2 × Centroid 1)

Step 7:Find the difference between the first intensity value and the various centroid values.

Step 8:Based on the minimum difference obtained, group the intensity values into the corresponding clusters.

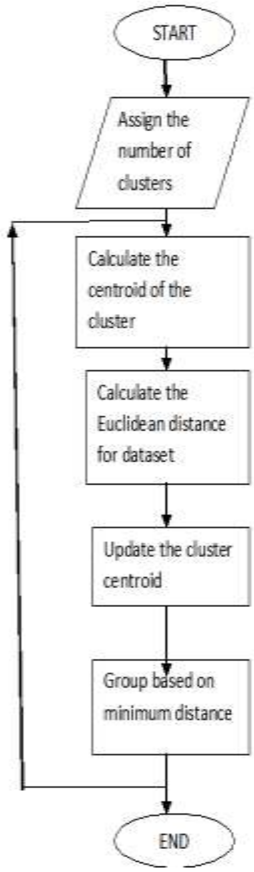
Step 9:Repeat step 1 & 2 for all the other intensity values of the image.

Centroid 3 = (3 × Centroid 1)
Centroid n=(n×Centroid1)

2.3 Flowchart for k-mean algorithm

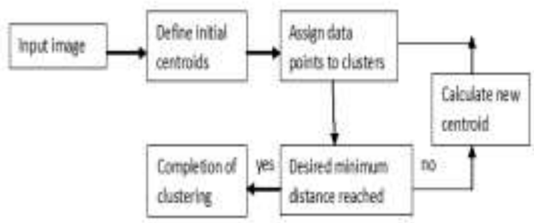
With this flowchart, clearly explain how we get minimum distance from the dataset of the given sample.

a) Initially we should assign the cluster for the given sampled dataset.
b) We should calculate the centroid for the cluster in the dataset.
c) Then we calculate the Euclidean distance for the dataset.
d) Update this Euclidean distance to the clusters.
e) Next find the minimum distance from the groups
f) Repeat the above all the steps up to we minimize the clusters into two groups.



2.4 Block diagram for K-Mean clustering

Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, 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.



K- means clustering is an algorithm to classify or to group the objects based on attributes/features into K groups. ... the position of the centroid is updated by the means of the data points assigned to that cluster. In other words, the centroid is moved toward the center of its assigned points.

K- means clustering is an algorithm to classify or to group the objects based on attributes/features into K groups. ... the position of the centroid is updated by the means of the data points assigned to that cluster. In other

words, the centroid is moved toward the center of its assigned points.

1. Pick 2 points Y and Z that are furthest apart in the measurement space and make them initial cluster means.

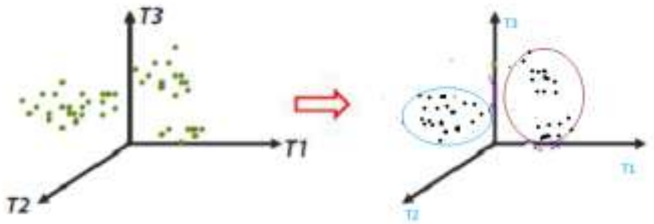
2. Assign all points to the cluster whose mean they are closest to and recompute means.

3. Let d be the max distance from each point to its cluster mean and let X be the point with this distance.

4. Let q be the average distance between each pair of means.
5. If d > q / 2, make X a new cluster mean.
6. If a new cluster was formed, repeat from step 2.

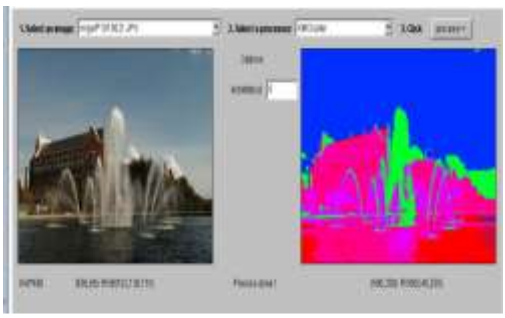
3. Texton and feature representation

The K centroids after the K-means clustering operation in the before subsection are called the K textons of the image. We compute the texton-index histogram in a sliding window (containing several pixels inside) around each pixel as the final feature representation of this pixel. So finally we have a histogram for each pixel in the image with each bin as the occurring probability of each texton inside the sliding window, and the number of bins of the histogram is equal to the number of groups after the K-means clustering operation. After all, we have histogram as the feature representation for each pixel. We can use this information to compute the similarity matrix and then segment this image.



4. Experimental result

We perform our image segmentation method on several kinds of images, including the cartoon images, landscape images, and the texture images, etc., and in this section we'll show the result.



5. Conclusion and Future works

In this section we'll show the result. There are still some things we can do for future works. At first, we will improve the stability of program. We want to modify the code of function to let programs more stable. Secondly, it's a chance to get a better adaptive method for image segmentation. The third one is to generate the post-processing mechanism for k-mean clustering. We can write a code about merging groups with the same texture into a single group. K-Means extensions will be assess for link with more type of images for successful justification investigational results.

6. References

[1] Khali Hamuda, "A Comparative study of data clustering technique", Department of system Design Engineering, University of Waterloo, Canada.

[2] H, D, Cheng, X.H. Jiang, Y. Sun and J. Wang," Color image segmentation: advances and prospects," Pattern Recognition, 34: 2259-2281.

[3] Ibrahim A. Almerhag, Idris S. E, Feighi and Ali A Dulla, " Modified k-means clustering algorithm for gray image segmentation", Naseer International University, Tarhoona, Libya.

[4] Nicholas Sia Pik Kong, Haidi Ibrahim and Seng Chun Hoo," A literature Review on Histogram Equalization and its Variation for Digital Image Enhancement", in Internation Journal of Innovation, Management and Technology, vol 4, No. 4, August 2013.

[5] Huiz Zhu, Francis H.Y. Chan and F.K.Lam, " Image Contrast Enhancement by Constrained Local Histogram Equalizatio",in Computer Vision and Image Understanding, Vol. 73, No. 2, Febbruary,, pp. 281-290, 1999..

[6]Sorin Israil," An overview of clustering methods", in With Application to Bioinformatics.

[7] Aimi Salihai, Abdul Yusuff Masor, Zeehaida Mohamed, " Color image segmentation approach for detection of Malaria parasites using various color models and k-means clustering", in Wseas Transaction on Biology and Biomedicine, vol. 10, January 2013.

[8] Khaled Hammouda, A Comparative study of Data Clustering technique. Department of System Design Engineering, University of Waterloo,Canada.

[9] Koheri Arai and Ali Ridho Barakbah, Heirarchical K-means: An algorithm for Centroid initialization for K-means, Saga University, (2007).

[10]D. L. Pham, C. Y. Xu, J. L. Prince, A Survey of Current Methods in Medical Image Segmentation, In Annual Review of Biomedical Engineer,(2000).

[10]Pallavi Purohit and Ritesh Joshi, A New Efficient Approach towards k-means Clustering Algorithm, In International Journal of Computer Applications, (0975-8887), vol. 65, no. 11, March (2013).