BIL 415 - Image Processing Practicum

HACETTEPE UNIVERSITY
Department of Computer Engineering

Problem Set 4
Fall '2012-2013
Dr. Erkut Erdem
Levent Karacan

K-means Clustering for Quantization and Segmentation

Due Date: 03:00pm on Wednesday, January 02nd, 2013

k-means clustering example k=2

Image from Berkeley Segmentation Dataset

Background

Clustering can be defined as grouping a given set of data points so that points similar to each other are assigned to same clusters or groups. In image domain clustering is used for various problems like quantization and segmentation. A good clustering algorithm must group data points to homogeneous subsets as it can. How the similarity is defined is the most critical step in a clustering algorithm that it determines how the clustering algorithm groups data points. K-means clustering is one of the most popular clustering algorithms that group data to k disjoint clusters. K-means clustering is an unsupervised learning algorithm for clustering problem where the main idea is to define k centroids one for each cluster. These centroids are randomly sampled from data space for first iteration. In the next step for each data point distances to these centroids are calculated and data points are assigned to the nearest centroids as cluster elements. Then for each cluster, new k centroids are calculated from k clusters. These steps go on until clusters centroids stabilize.

Overview

In this assignment, you will implement k-means clustering algorithm for the following image processing problems:

1. Image Quantization
2. Image Segmentation

Details

Your program will take a color or gray-level image as input and produce a quantized image or segmentation result.

1. **Image Quantization**

   Image quantization is a discretization process that reduces the number of observed image intensity values. A color image is represented by 24-bit and this means 2^{24} different color can be seen an image. You will implement a image quantization function that will reduce color number to less number by using uniform quantization or k-means approach. In your implementation you can use the following prototype definition:

   f=quantize(g,method,k);
   g:input image
   method: uniform or k-means
   f:output quantized image

2. **Image Segmentation**

   Segmentation can be defined as clustering image pixels into meaningful image regions. Each region may be an object(car, cup, mouse, etc.), a part of an object(car tire, human body parts,) or a surface(sky,sea,lake). You will implement a k-means clustering function that segments an image into meaningful parts. Your segmentation function must take the total number of segments as a parameter.

   f=segment(g,segmentNumber)
   g:input image
   f:segmentation result

Grading

The assignment will be graded out of 4:

- 0 *(no submission)*, 20 *(an attempt at a solution)*, 40 *(a partially correct solution)*, 60 *(a mostly correct solution)*, 80 *(a correct solution)*, 100 *(a particularly creative or insightful solution)*.

What to Hand In

You are required to submit all your code along with a short webpage in *HTML*. For that purpose, prepare a folder containing

- `README.txt` *(text file containing details about your project)*
- `code/` *(directory containing all your code)*
- `HTML/` *(directory containing all your documents, including your images)*
- `HTML/data/` *(including your data images)*
- `HTML/result/` *(including your result images)*
- `HTML/report.pdf` *(pdf report)*
- `HTML/index.html` *(html report)*

Archive this folder as `pset4.zip` and submit to `Department Submit System`.

In this assignment you will write detailed report which contains a brief overview of the problem, details of your implementation and the results on the least 8 images with your comments. You must play with the number of clusters (k parameter). For image segmentation you are required to use similarities based on RGB and Lab color space and the effect of including spatial constraints.All results you obtain must be put to your report and all observations must be specified clearly. Analyzing the results with appropriate

comments you will do is important for this assignment. Your reports must be prepared with a word processing program(latex,MS Word,etc.) in pdf file format.

If your algorithm failed to give a satisfactory result on a particular image, provide a brief explanation of the reason(s).