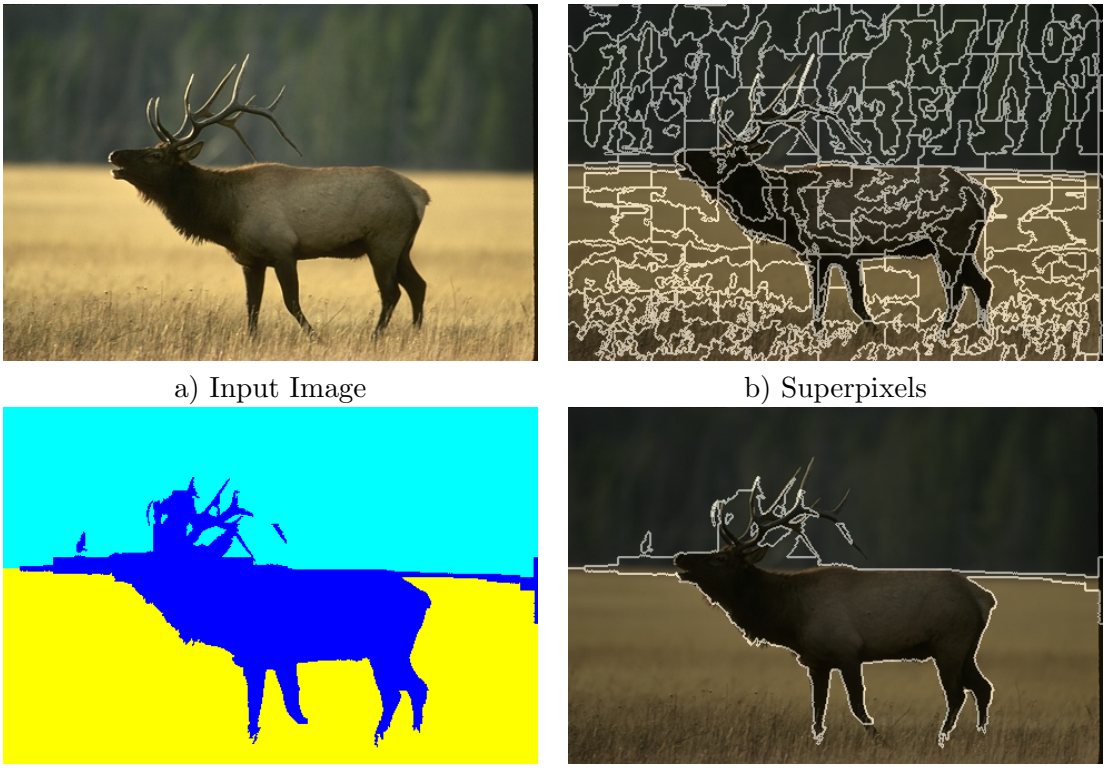
BIL 415 - Image Processing Practicum

HACETTEPE UNIVERSITY
Department of Computer Engineering

Problem Set 5
Fall '2014-2015
Dr. Erkut Erdem
TAs.Levent Karacan

1 K-Means Clustering for Image Segmentation

Due Date: 23:59pm on Monday, January 5th, 2015



a) Input Image b) Superpixels

c) Image Labels d) Image Segments

Figure 1: K-Means Image Segmentation using Superpixels

Background

Clustering is a process that groups data with respect to data similarity so that similar data take part same cluster. In image domain clustering is used for various types of problem e.g. *image quantization*, *image segmentation*. A good clustering algorithm must group data to homogeneous subsets as possibble. Similarity is most critical step in a clustering algorithm that determine how the clustering algorithm groups data. K-means clustering is one of the most popular clustering algorithm that groups data to k dissimilar clusters. It is an unsupervised learning algorithm for clustering problem and the main idea is to define k centroids one of each cluster. These centroids is randomly assigned to data space for first iteration. In the next steps, for each data point, distance to these centroids is calculated and data points are assigned to nearest centoids as cluster elements. Then for each cluster, new k centroids are calculated from k clusters. This steps go on until clusters centroids unchange.

Overview

In this assignment, you will use k-means clustering algorithm for image segmentation by using superpixel representation of an input image. For this purpose you must carry out the following steps:

- Extracts Superpixels

  You will use SLIC Superpixel work to extract superpixel labels. You can find the code in http://ivrg.epfl.ch/research/superpixels

- Feature Representation of Each Superpixel

  You will define a feature vector to represent eachs superpixel. This feature vector may contain various type of features such as color,intensity, location or 1st order statistics of these features in superpixels e.g. *means of intensities*. You will build up your own feature representation.

- K-Means Clustering of Superpixels

  You can use Matlab built-in function for k-means clustering. You will give feature points obtained from each superpixel and cluster number to k-means algorithm so that k-means algorithm will produce cluster labels for each superpixel.

Details

Your Matlab program will take an color or gray level image and cluster number as input and produce segmentation result like in Figure 1-c,1-d.

-You must take results on the least 5 images. You can find lots of image in Berkeley Segmentation Dataset[2].

-You must play with cluster number according to the image content to obtain good results.

-You must use superpixels with different region size and make observation that how the superpixels and region size of superpixels influence clustering results or not.

-You must show your results like in Figure 1-c,1-d. You can use *imagesc* function in Matlab for this purpose.

-You must comment your results why they are satisfactory or not.

2 (Bonus)-Spectral Clustering for Image Segmentation

You will use spectral clustering algorithm for image segmentation in this part of your assignment. It is not mandatory but students who do this part will earn bonus points. You will follow similar way(superpixels and features) as in part one except clustering algorithm.To do this you should carry out the following steps:

- You build a graph structure in which N superpixels are vertices. In this case you will generate a affinity matrix $W : N \times N$ for graph edges.

  $W_{ij} = exp(\frac{-||f_i - f_j||^2}{\sigma^2})$

  $D(i,i) = \sum_j W_{ij}$ is the diagonal degree matrix.

  $L = I - D^{\frac{-1}{2}} W D^{\frac{-1}{2}}$ is the graph Laplacian.(I:unit matrix)

- Calculate eigenvectors and find K smallest eigenvalues of the graph Laplacian L
- Select $N \times K$ matrix of the corresponding eigenvectors of the graph Laplacian.
- Feel free to choose appropriate values for σ and K in defining similarities for intensity and number of neighbors, respectively.
- Use k-means clustering algorithm on new data representation $N \times K$.

Details

Your Matlab program will take an color or gray level image and cluster number as input and produce segmentation result like in Figure 1-c,1-d.

-You must play with cluster number according to the image content to obtain good results.

-You must use superpixels with different region size and make observation that how the superpixels and region size of superpixels influence clustering results or not.

-You must show your results like in Figure 1-c,1-d. You can use *imagesc* function in Matlab for this purpose.

-You must comment your results why they are satisfactory or not.

-You must compare your results with k-means clustering algorithm and comment.

What to Hand In

You are required to submit all your report 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)*
- `report/data/` *(including your data images)*
- `report/result/` *(including your result images)*
- `report/pset5.pdf` *(PDF report)*

Archive this folder as `pset5.zip` and submit to 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 with your comments. You must play with related parameters. .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.

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

References

[1] http://ivrg.epfl.ch/research/superpixels

[2] https://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/.