Foundations of Graphics

Universität Bonn
Institut für Informatik II
November 19, 2019

Winter term 2019
Prof. Dr. Matthias Hullin

Sheet 6 - Clustering

Please upload your solutions to
`https://uni-bonn.sciebo.de/s/XfRkWhpWDEeR8Y7`
by **Thu, November 28, 2019**.
Make sure to **list all group members** on all pages / source files.
Theoretical solutions must be texed or scanned, **photos will not be accepted**.

Practical Part

Assignment 1) Segmentation via k-means *(3+1+1=5Pts)*

The k-means algorithm can be used for segmenting data points. In this exercise we will consider the applications of mesh segmentation and image segmentation.
Please submit the whole framework with its original folder structure and all files.

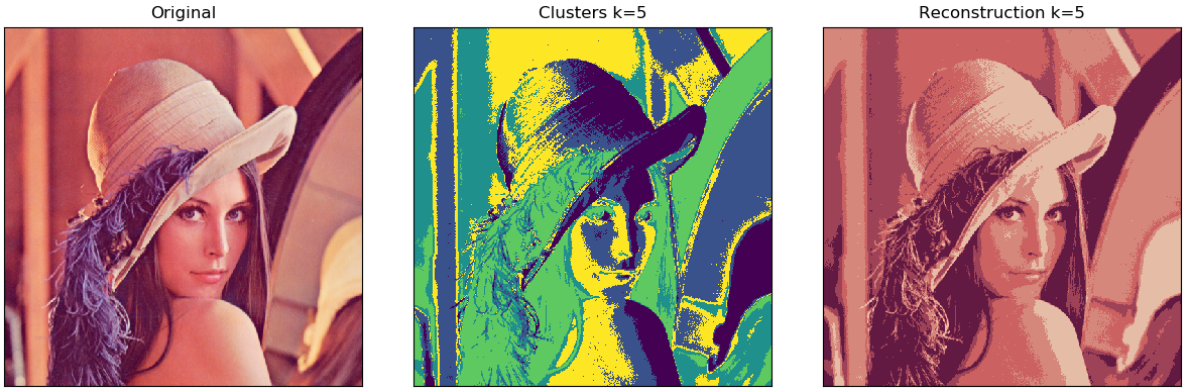

Figure 1: Result for $k = 5$.

a) Implement a function `[index, centers] = my_kmeans(data, k)` which takes as an input the number of clusters `k` and an array of data points `data` (each point in a row). The function should return the cluster center positions (`centers`) and a label (assigned cluster, (`index`)) for each point.

b) Complete '`image_segmentation`' and reconstruct the image `'lenna.png'` with 2, 5, and 15 colors. Plot reconstructions of the image using only the colors of the `k` cluster centers.

c) Consider the application of k-means in mesh segmentation. Therefore we cluster points in the $\mathbb{R}^3$ with their difference of normal vectors.

$$d(n_1, n_2) = \|n_1 - n_2\|^2$$

Complete '`mesh_segmentation`' to partition a mesh into 5 clusters. Display the clusters using multiple colors in the same plot.

Theoretical Part

Assignment 2) Convergence of k-means *(1+1+1 = 3Pts)*

Consider the k-means algorithm discussed in the lecture. To analyze its convergence we define the current assignment error by the sum of the squared distances of the points p_i to their centers $c_{\pi(i)}$:

$$\mu[c_i, \pi] = \sum_i d^2(c_{\pi(i)}, p_i)$$

Show that on a finite number of points, the algorithm must converge after a finite number of iterations. Therefore show that,

a) the error function μ strictly decreases in each step

b) there is only a finite number of point to cluster assignments

c) the algorithm terminates after a finite number of steps

Assignment 3) Expectation Maximization *(1+1+1+1=4Pts)*

To understand the convergencce of the EM-algorithm better, please discuss the following questions:

a) How does the result and run-time of EM change over different initial cluster parameters (positions and variance)? Why is EM a local optimization?

b) How does the result change for a different number of clusters k?

c) What methods can be used to gain better global convergence (i.e. to find the global optimum)?

d) How does the result/run-time change for different samplings of the underlying distribution but the *same* initial starting conditions?

Good luck!