Answers to questions in
Lab 3: Image segmentation

Name: Mikel Zhobro
Program: _____

Instructions: Complete the lab according to the instructions in the notes and respond to the questions stated below. Keep the answers short and focus on what is essential. Illustrate with figures only when explicitly requested.

Question 1: How did you initialize the clustering process and why do you believe this was a good method of doing it?

Answers:
I initialized the cluster process with random vectors that have integer numbers between 0-255 since R,G,B can only have numbers between 0-255. Concidering the input image before clustering could lead to better strategies such as selecting two of cluster means at colors orange and white, since these are the two predominant colors for orange.

Question 2: How many iterations L do you typically need to reach convergence, that is the point where no additional iterations will affect the end results?

Answers:
As it is seen from the Figure 3. for the orange only 4 iterations are needed. From L=4 to L=5 no significant changes can be noticed.(chekc once more*)
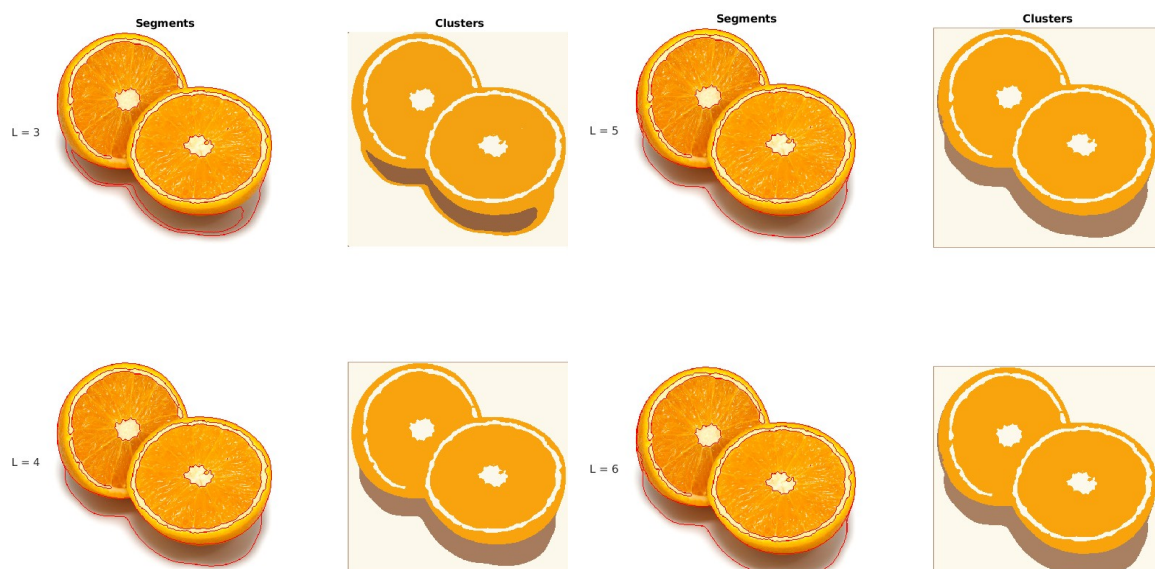

Figure 1. Clustering for K =3 and L = 3, 4,5,6

Question 3: What is the minimum value for K that you can use and still get no superpixel that covers parts from both halves of the orange? Illustrate with a figure.

Answers:
The minimum K for getting superpixels that don't cover parts from both halves of the orange is 8 as can be seen in the Figure 2.
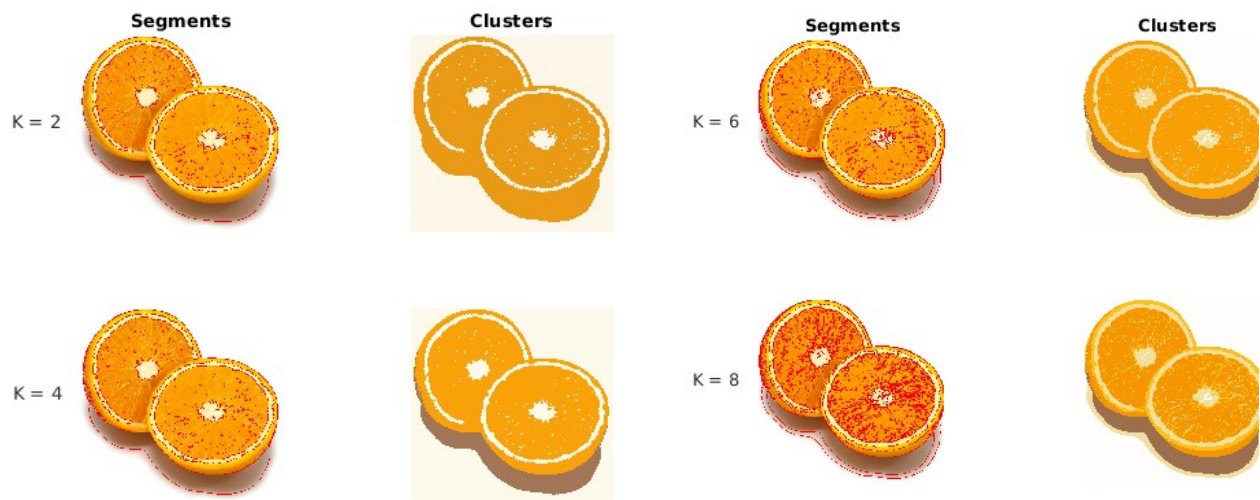

Figure 2. L=12 and K=2,4,6,8

Question 4: What needs to be changed in the parameters to get suitable superpixels for the tiger images as well?

Answers:
For the tiger images more colors are present that is why more clusters are necessary. In the same time we have to increase the number of iterations for the clusters to converge.
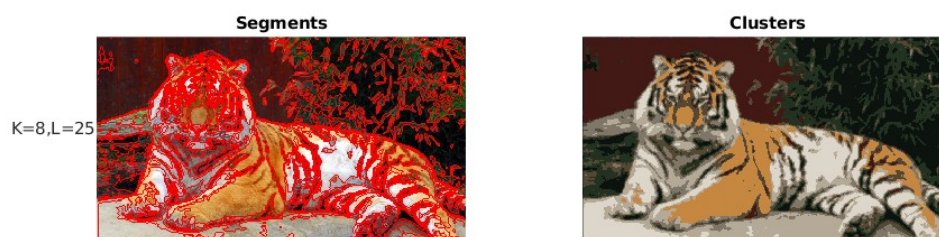

Fgirue3. Clustering for tiger image with K = 8 and L =25

Question 5: How do the results change depending on the bandwidths? What settings did you prefer for the different images? Illustrate with an example image with the parameter that you think are suitable for that image.

Answers:

Bandwidths describe the variance of the kernel around each pixel considered for the corresponding kernels. A big spatial spatial bandwidth corresponds to possibly bigger modes since more pixels are considered while computing the mean. Similar to that increasing the color bandwidth smooths the image out and results in less modes and better color approximation.

For the example of orange image best bandwidths are (4,30) as seen in Figure 4 on the left side.
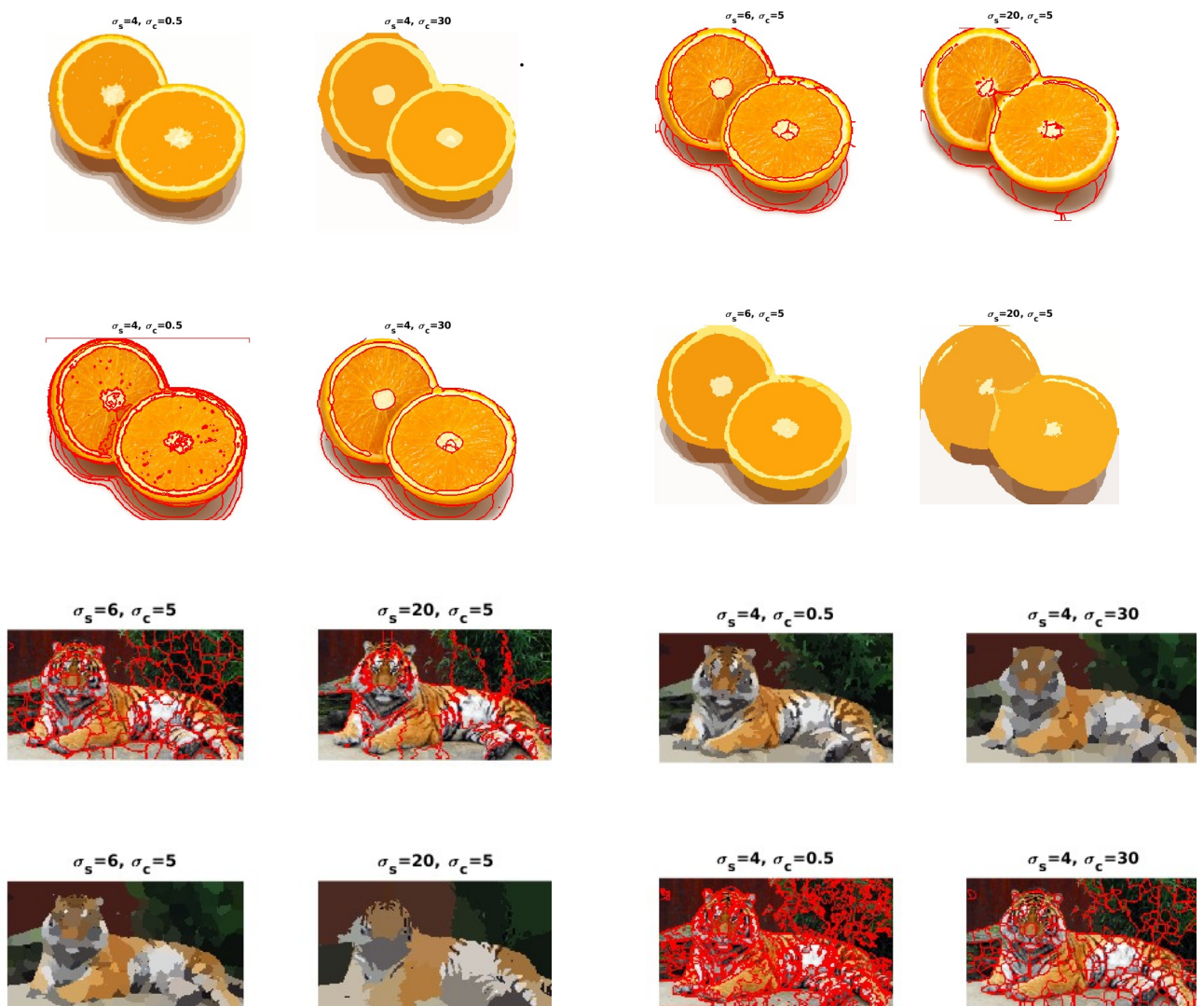

Figure 4 . Different bandwidths, spatial(left) and color(right)

Question 6: What kind of similarities and differences do you see between K-means and mean-shift segmentation?

Answers:
Both K-means and mean-shift are used to create clusters of the image.

Different to K-means, mean-shift takes under consideration the spatial position of the pixels in addition to the color dimensions which explains why clusters don't span over multiple regions.
On the other hand mean-shift doesn't accept a predefined number of clusters like K-means does.

Question 7: Does the ideal parameter setting vary depending on the images? If you look at the images, can you see a reason why the ideal settings might differ? Illustrate with an example image with the parameters you prefer for that image.

Answers:
The parameters depend on the image they are applied too. Depending on the features the image has one has to tune **radius**, **min_area**, **cuts_thresh** and **max_depth**. **Max_depth** limits the depth of rekursion which gives the number of times the graph is separated, the higher the nr gets the smaller the segments. **Min_area** sets the limit on how small the area of segments can be. **Cuts_thresh** sets the maximum cut-value we allow. A higher value would allow more similar graph parts to be separated. **Radius** has a relatively high influence as it gives the size of the neighborhood that will taken in account for each.
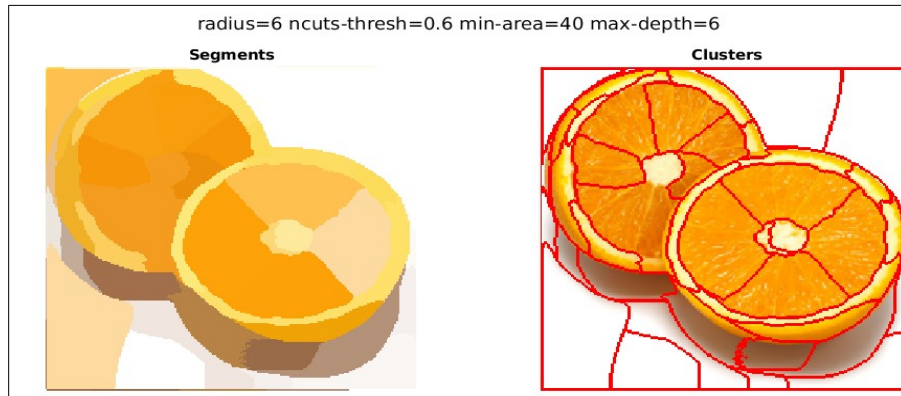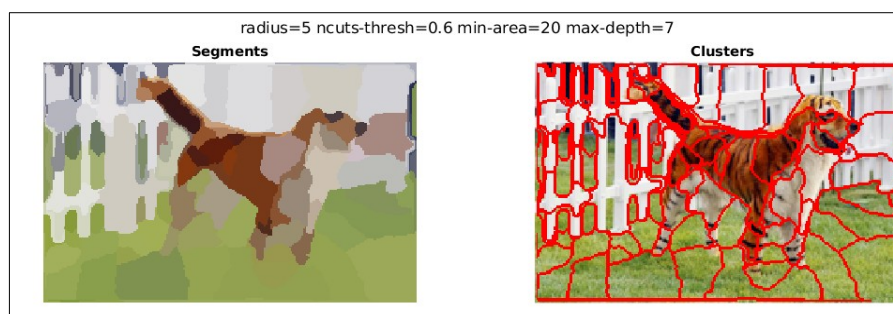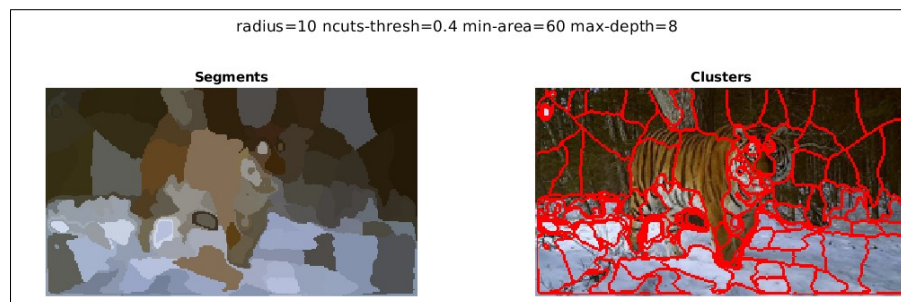
Figure 5.

Figure 6.

Figure 7.

Figure 8.

Question 8: Which parameter(s) was most effective for reducing the subdivision and still result in a satisfactory segmentation?

Answers:
cuts_thresh, min_area, **max depth**

Question 9: Why does Normalized Cut prefer cuts of approximately equal size? Does this happen in practice?

Answers:
In practice this happen if we only look at maximum depth, then the image gets cut into preferably two equally proportional segments. But because of the other parameters the cuts tend to not be proportional.

Question 10: Did you manage to increase *radius* and how did it affect the results?

Answers:
The computational time increases rapidly if we increase radius. That is because we then consider a bigger area of neighbouring pixels into the calculations. The results show better results when it comes to bigger segments but the color association doesn't always work as it can observed in Figure 7.

Question 11: Does the ideal choice of *alpha* and *sigma* vary a lot between different images? Illustrate with an example image with the parameters you prefer.

Answers:
Alpha is the maximum cost of an edge which controls how difficult it is to cut through similar pixels. With higher alpha it is more difficult to cut through similar pixels since the cost is higher.
Sigma regulates how much the cost decays for decreasing similarity between neighbouring pixels. A big sigma would correspond to a low cost even in little changing pixels which makes cutting easier.
Once the alpha and sigma are over 6 the changes with increasing values are negligible.
If we lower both alpha and sigma the algorithm is more sensitive because it is in general easier to make cuts. In conclusion the ideal choice of sigma and alpha do not vary a lot.
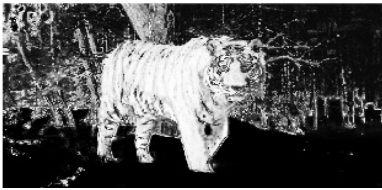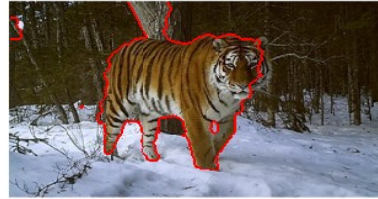
Figure 9. alpha=9, sigma = 10

Figure 10. alpha=9, sigma = 10

Figure 11. alpha=9, sigma = 10

Question 12: How much can you lower K until the results get considerably worse?

Answers:
For the tiger1 figure the result get considerably worse when we choose K lower than 6. But this is relative. It mostly depends on how accurate the mask is. If the mask is very accurate the algorithm gives good results even for K=3 or K =4.

Question 13: Unlike the earlier method Graph Cut segmentation relies on some input from a user for defining a rectangle. Is the benefit you get of this worth the effort? Motivate!

Answers:
Even though the algorithm presents a semi-automatic segmenting method, the benefit is worth the effort. In most cases, e.g. in medicine, the users(doctors) approximately know where the object to be segmentized is located. It allows the user to use extra information which would make the result much better. In addition to that the method is easily expandable in fully automatic by randomly choosing the rectangle or using other methods such as kmeans to decide for the rectangle.

Question 14: What are the key differences and similarities between the segmentation methods (K-means, Mean-shift, Normalized Cut and energy-based segmentation with Graph Cuts) in this lab? Think carefully!!

Answers:
Similiarities
- all methods try to cluster pixels of similar color
- Normalized Cut and Graph Cut are graph-based i.e. they look at the image as a graph, with vertices and edges connecting them which are based on some kind of similarity

Differences
- only Graph-Cut is a semi-automatic method which requires input from the user
- Beside K-means all method consider spatial similarities in addition to color similarities.