Answers to questions in

Lab 3: Image segmentation

Name: Zacharie McCormick_____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.

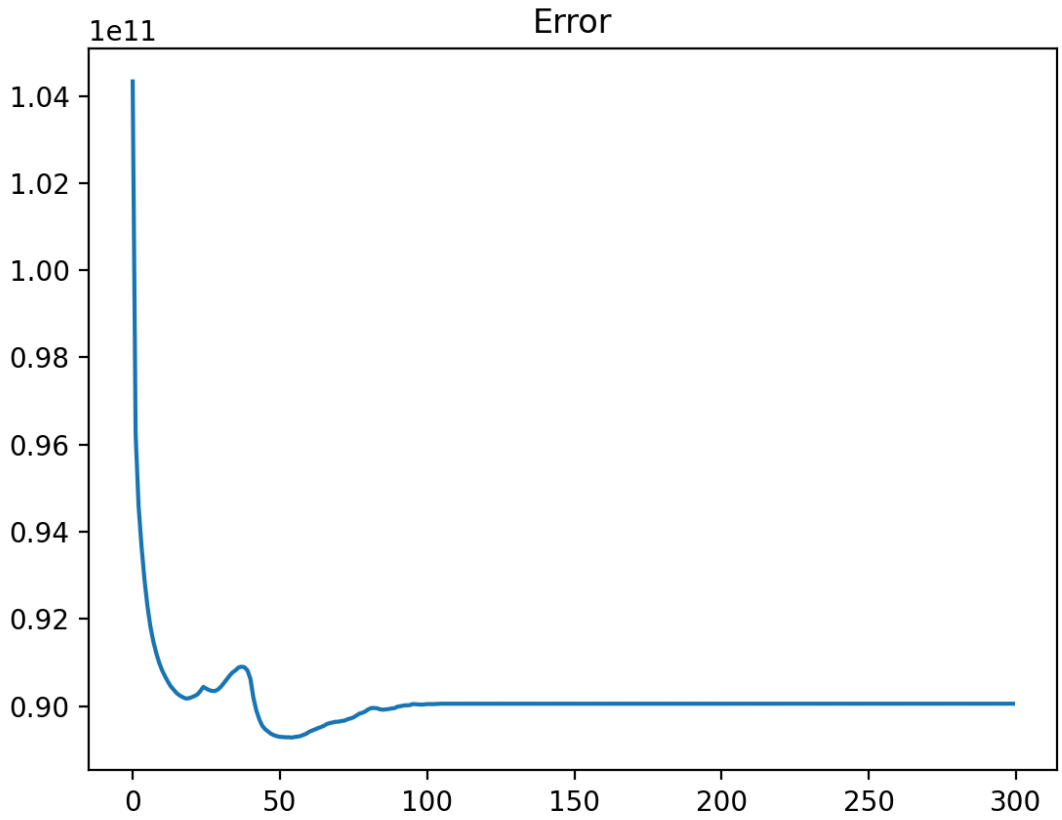
Good luck!

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

Answers:
I used a set to pick pixels from. This way, no cluster is initialized to the same value.

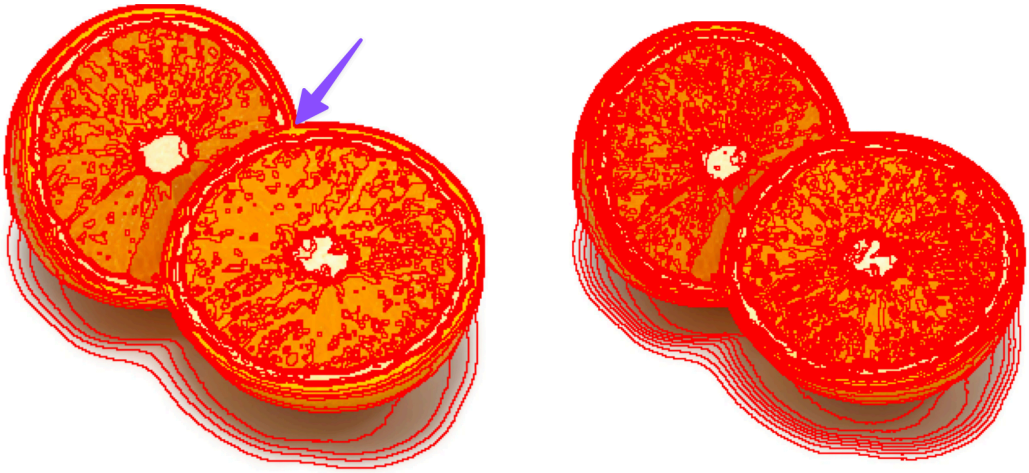
Question 2: How many iterations L do you typically need to reach convergence, that is the point where no additional iterations will affect the end results?

Answers:
I plotted the sum of square of the distances to get an idea on the spread of the distances compared to the clusters values and we see that at around 30-ish but it varies a lot between different runs.



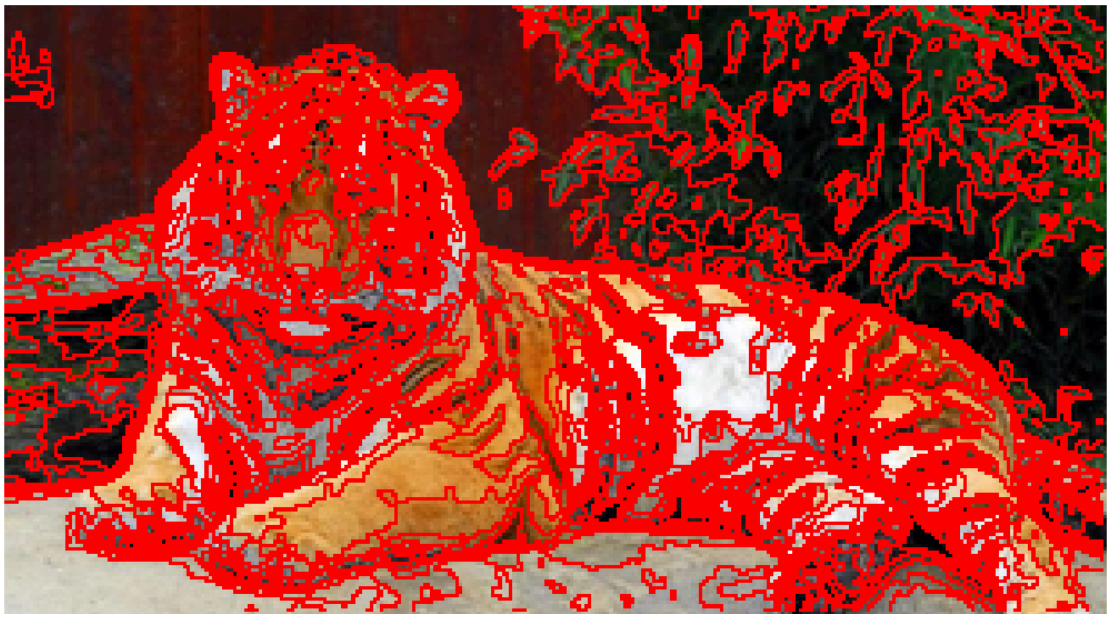
Question 3: What is the minimum value for K that you can use and still get no super-pixel that covers parts from both halves of the orange? Illustrate with a figure.

Answers:
No value of K gave me that because the top and bottom part pointed here in the image are always connected together (although at high value of K (K>30) you simply can't see anymore due to the red line covering these parts.



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

Answers:
K=15 L=50



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:
The higher the color bandwidth the more pixel with distant values are allowed to be grouped together. The higher the spatial bandwidth the further away from the center of the cluster the pixels are allowed to be.

spatial_bandwidth: 5, color_bandwidth: 70

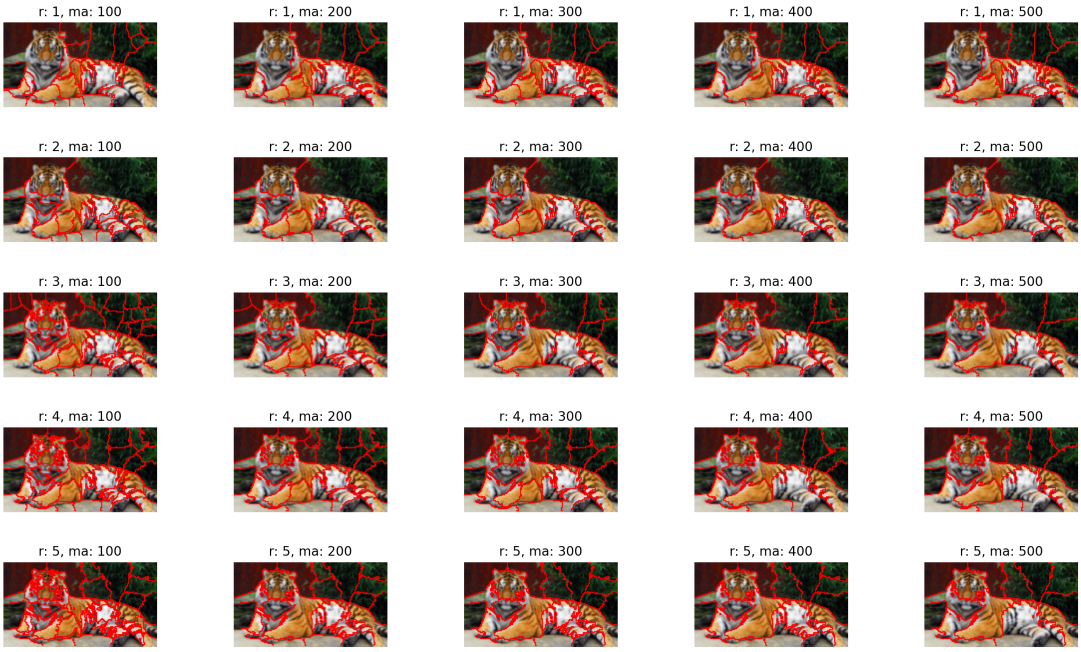
Question 6: What kind of similarities and differences do you see between K-means and mean-shift segmentation?

Answers:
Similiraties:
Both are clustering algorithms.
They both use a vector that contains the RGB value of the pixels (although the mean-shift segmentation uses x and y too)

Differences:
Computation time.
With K-mean you can pick the number of color cluster you want to have whilst, the mean-shift segmentation converges to an unknown number of maximas in it's 5D space. Also, the K-mean method does not need to iterate on every pixel.

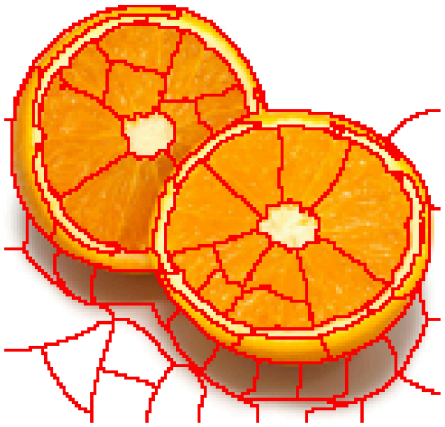
Question 7: Does the ideal parameter setting vary depending on the images? If you look at the images, can you see a reason why the ideal settings might differ? Illustrate with an example image with the parameters you prefer for that image.

Answers:
Here I show the difference that that radius and the maximum area have on the output. We can see that minimum area allows the chunk to be bigger and bigger (this is visible on the second row where the section in between the tigers leg is allowed to grow more and more as the maximum area increases). As for the radius, it's fairly hard to see what is happening and I would say increasing it, does not seem to yield great result.



Coming back to the question, here is the best parameters I found for the tiger:
colour_bandwith=20
radius=3
ncuts_treshold=0.1
min_area=100
max_depth=15

And if I apply these same parameters to the orange, I obtain this:



This is not bad but could probably be improved on. The problem with an algorithm like this is that it apply the same parameters aver the entire image. Thus, if you try to improve the parameters hoping to fix one part, it will undoubtedly affect another part of your image in a way that makes it worse.

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

Answers:
The min_area allows us to increase the area of similar cluster while still being very consistent between different values which made it easier to use than the other parameters to refine the result.

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

Answers:
Based on this equation

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)}$$

If the size of both clusters are not approximately the same, one of the divisor will end up being fairly small due to the sum of the vertex weight not having many links (not many value to sum up) and the other other will end up with a larger value. This will lead to a very small value added with a fairly large one which leads to an big weight for the cut. On the other hand, if the size of both clusters are almost the same, both divisor will be fairly big and lead to a smaller value in the end.

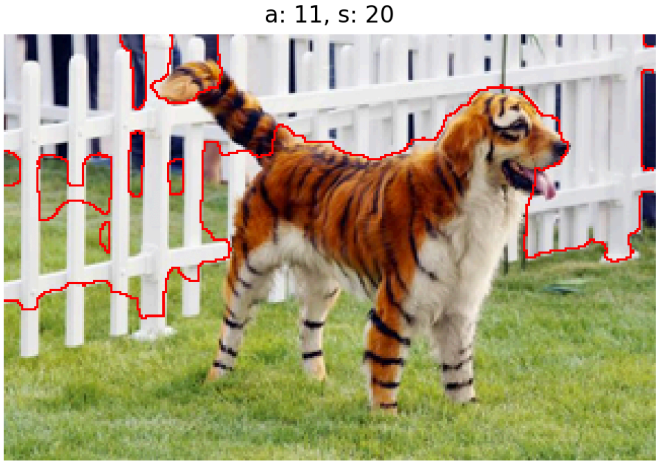
Question 10: Did you manage to increase *radius* and how did it affect the results?

Answers:
Not in a way that enhanced the image.

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:
As can be seen from the image below, even picking a suitable alpha and sigma for a single image is fairly hard (finding a correlation between the parameter and it's effect in the image is quite difficult although some parameters are indeed a lot better than other). The best image in this set in my opinion is the one with alpha=11 and sigma=10 as it follows the contour of the tiger really well.

Applying this same set of parameter (alpha=11 and sigma=10) to the image of the dog tiger yields dubious result (note that I also adjusted the area for this image in order to capture the correct part of the dog).



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

Answers:
As can be seen from the image below (realized with the same ideal parameter), using a K of 12 or less yields pretty bad result.

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:
It depends on the application, since this method requires some prior knowledge to be given (aka. The mask) it cannot be used for real time application as is. On possible way to use this in a realtime application would be to use a neural network to give us the bounding box of the important thing in the image and then pass it to this function. I think that for manual photo editing though this method is pretty neat compared to the other ones.

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:
K-mean is ignorant to what is in the image and wether or not it should be part of the background, it's purpose is to segment the image in similar color patches. Mean-shift does a similar thing but tries to keep spatially correlated pixel in the same cluster. Normalized cut tries to minimize to minimize the weights of the cuts it makes without knowing much about the image itself (mostly relies on the similarities of the neighbouring pixels). Finally, the Graph cuts is similar to the energy based but as some idea as to which pixel should be considered background and foreground from the prior knowledge given by the input mask.
