Answers to questions in

Lab 3: Image segmentation

Name: David Dashti Program: CMEDT

**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 randomly initialized the clusters with RGB from 0 – 255. It is possible to set this to the dominant colors of the image, but I kept it random since it was simpler. You might get a performance increase with the other option.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

This is highly dependent on the definition of convergence in this case. I defined the convergence as when the algorithm has reached a point where the difference between old clusters and new ones is below a certain value. In other words, the algorithm has converged when the difference between clusters of different iterations is small.   
  
Another thing that might affect this is the complexity of the image in terms of colors and the number of clusters. Both of these factors increase the number of iterations needed.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

It is pretty hard to find an exact number as it seemed to differ from image to image but if we are using the orange image then K = 8 was the magic number.



Figure 1 - Clustering with K = 7, halves cannot be separated

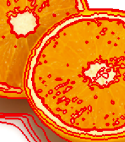


Figure 2 - Clustering with K = 8, halves can be separated by the yellow area on the outer rim of the edges

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Answers:

Well since it is more colors in this image, we can conclude that we are going to need more iterations, as I reasoned in question 2. Furthermore, there is a lot of color shift which means that we are going to need more clusters to compensate.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

If you increase the color bandwidth then that will correspond to a smoothing of the image. On the other hand, if you make the color bandwidth narrower, it will correspond to a sharpening of colors.

If you increase the spatial bandwidth then you are going to increase the number of pixels that get included when doing the calculation. Thus, you will be able handle images where the modes are further separated. On the other hand, if you choose a narrow spatial bandwidth, then you will lose this ability but gain an increased precision in “pixel-mode” allocation.



Figure 3 - Example of mean-shift segmentation done on tiger1.jpg. Spatial- and color bandwidth was set to 7.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Answers:  
  
They both are algorithms that use mean calculation to step in the algorithm and they are used for similar things i.e., image segmentation. However, mean-shift looks at density whilst k-means looks at centroids. Also, k-means does not consider spatial information, only color information.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

This is as the questions entails, heavily dependent on the image itself. The reason why the ideal settings differ is because the complexity of the images varies. If you have a complex image, then you want a high number of cuts and a high max depth. And vice versa for a simpler image. You do not want to go too low though since it might result in a bad segmentation.



Figure 4 - Example of normalized cut segmentation of same image as in Figure 3. Color bandwidth is set to 14, the radius to 12, threshold for number of cuts to 0.4. minimum area of 16 and maximum depth of 7.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Answers: I would say the minimum area, the number of cuts and the maximum depth.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Answers:

It boils down to the math behind the algorithm which favors a proportionality between foreground and background that is 1:1. However, this is not considering the other parameters used in the implementation which will skew this proportionality into something more uneven.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Answers:  
  
Yes, I did but at a cost. There was a heavy toll on the computational time. This is because adjacent pixels are being included.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

Well alpha is the maximum cost of an edge which means that a high alpha will result in a harder time to cut similar vertices. Sigma on the other hand is how much the costs lowers when we have dissimilar vertices. The optimal parameters are again highly dependent on the image complexity but for the image tiger1, alpha of 13 and sigma of 8 achieved good results.



Figure 5 – Prior foreground probabilities of tiger1.jpg.



Figure 6 - Segmentation using graph cuts. Not perfect but good enough. Parameters are described in the question above.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

Answers:

The threshold is definitely at K = 3. Any lower than that and the results looks terrible. It basically gets segmented into two pieces which are unintelligible.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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 really depends on if the background and foreground of the image can conform to a box really. I have also seen implementations of graph cut that does not use a box. So, if the background and foreground can conform to the geometry then yes, it is worth the effort. Otherwise, the seed is useless.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

First of all, might be obvious but it must be mentioned. They are all methods of segmenting images. Now I would say that K-means and Mean-shift are closely related as they are both clustering algorithms. In this implementation, k-means does not include spatial information whilst mean-shift does. In a similar way, both normalized cut and graph cut uses graph theory to model the image. Normalized cuts do not need a priori information, whilst graph cut need it to achieve a good accuracy.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_