Answers to questions in

Lab 3: Image segmentation

Name: \_\_\_\_Simon\_Lagerqvist\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Program: \_\_\_\_\_\_\_\_\_\_TIEDB\_\_\_\_\_\_\_\_\_\_

**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 started by using the colors from randomly selected pixels, to initialize the K values. But I quickly discovered that which seeds that were used for the random selection effected the result greatly. It became clear to me that the initialization process is crucial for obtaining a good result. The problems are that the algorithm gets stuck because none of the colors selected are better than the other. My theory is that the colors selected in the beginning is colors that doesn’t really represent the picture. I therefor tried out a K means ++ initialization process which gave more stable results between runs. The reason for this is that K means ++ prioritizes more extreme colors in relation to previously selected ones. The K:s are therefore better distributed and more representative for the picture.

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

**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:

Convergence happens at L = 40 for orange.jpg using k=8

Convergence happens at L = 68 for tiger1.jpg using k=8

But there is typically very little change happening after 10.

The convergence also depends on the value of k.

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

**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:

L = 100 was used to make sure that the convergence was 100%. The result become K = 5



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

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

Answers:

We need a higher L to reach convergence. A higher K is also beneficial since the picture contains more colors. We can increase blur to smoothen out details and get less corruptions in super pixels

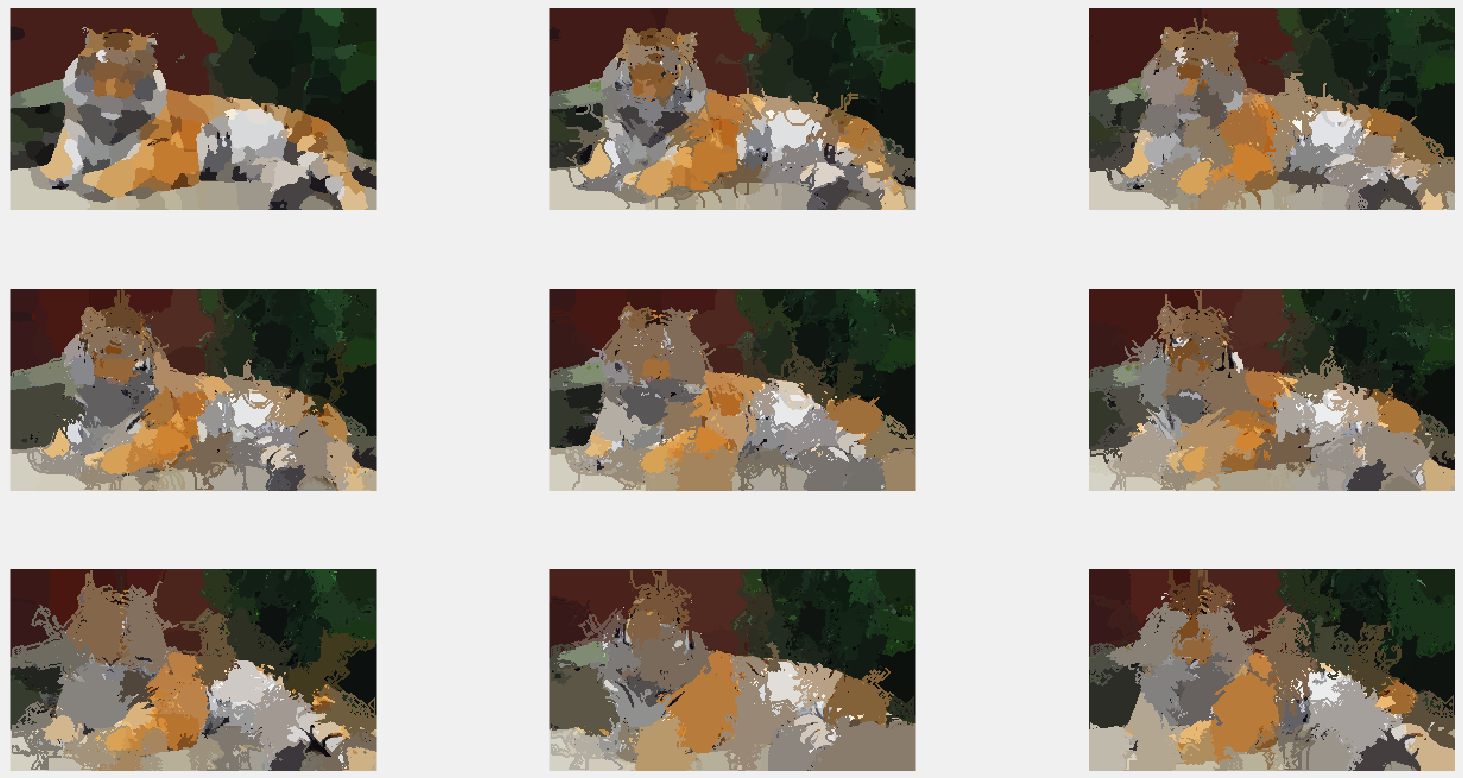
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

Spatial bandwidth lower to higher.

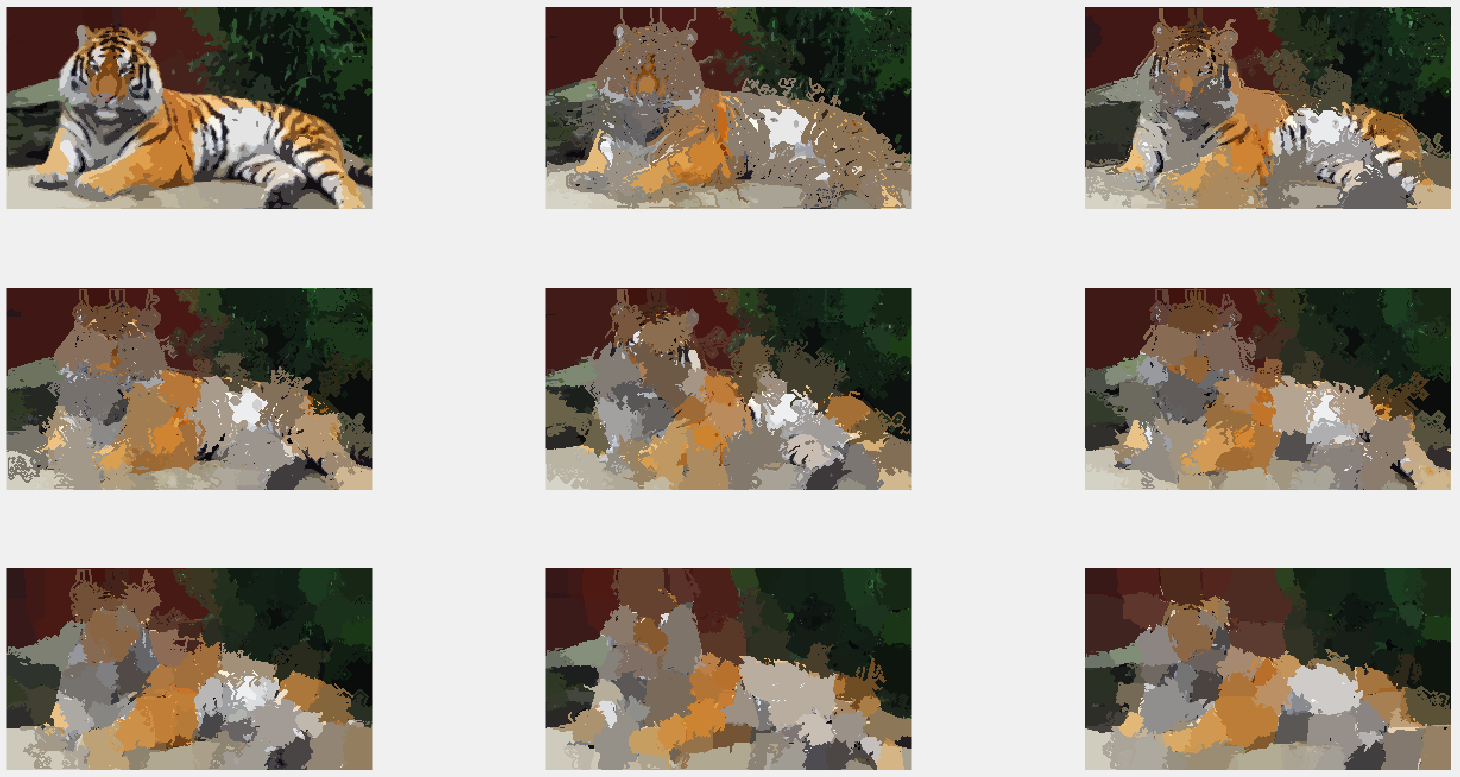
Increasing the spatial bandwidth makes the regions bigger and decrease “groups”:



This is since the spatial bandwidth controls the size of the kernel used along the spatial axis, so more pixels will be included

Color bandwidth, lower to higher.

Increasing the color bandwidth will result in more colors being mapped to regions:



This is since the color bandwidth controls the size of the kernel used along the color axis.

Blur can be used to control how many details we want to “find” however using I high blur will also result in that the locations of the edges of the segments will be less precise.

Which settings that are suitable will depend on what we aim to cover by the segments.

With suitable settings. (getting the contour right were prioritized in this run.)



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

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

Answers:

Similarities

* Both uses a probability distribution to determine color clusters

Differences

* K-means doesn’t care about the spatial information of the image. (only takes color information into account)
* We can specify the amount of cluster we want to find in K-means. Mean shift is non-parametric and clusters data by finding the densest regions. Therefore the amount of clusters found will depend on the “modes” or cluster centers that the method find.
* Menshift preforms better at segmenting images that consists of similar colors.

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

**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:

**min\_area**, parameter will depend on how many segments we want and how many segments the picture will contain.

**Deph**, will also depend on how many segments we want to find, but also on the shape of the segments since normalized cut tends to make straight cuts.

**ncuts\_thresh**, decides if the cut should be included. I will need to be tweaked deepening on the number of incorrect cuts. This is useful since depth is limited to doubling the number of lines found due to recursion. This will depend on how many cuts we need to find/make and the contrast in between colors in the image.

**Radius** increases the segment size. This is since we take less details into consideration since we look at more pixels when calculating weights.

**Reasons why the parameters will differ:**

* Object size
* The number of segments we want to find
* The contrast in the image
* The number of details in the picture
* Number of colors in the object
* The complexity of the background

|  |  |  |  |
| --- | --- | --- | --- |
| picture | paramters | segments | Segments in picture |
| tiger1 | colour\_bandwidth = 16.0  radius = 7  ncuts\_thresh = 0.040  min\_area = 105  max\_depth = 10  scale\_factor = 0.4  image\_sigma = 2.1 |  | En bild som visar hund, djur, inomhus  Automatiskt genererad beskrivning |
| tiger2 | colour\_bandwidth = 8.0  radius = 4  ncuts\_thresh = 0.04  min\_area = 100  max\_depth = 8  scale\_factor = 0.4  image\_sigma = 2.0 |  |  |
| orange | colour\_bandwidth = 18.0  radius = 8  ncuts\_thresh = 0.06  min\_area = 100  max\_depth = 7  scale\_factor = 0.4  image\_sigma = 2.0 |  | En bild som visar orange, frukt, citrus, apelsiner  Automatiskt genererad beskrivning |

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

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

Answers:

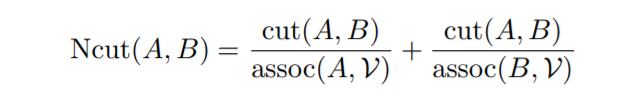
Decreasing Depth since in determine the number of divisions, increase threshold to take les subdivisions into account, increase min\_area to make room for less divisions and increase depth to remove details

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

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

Answers:

This is since we try to minimize the function:



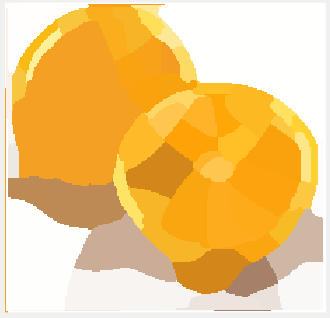
The function will have a small value when assoc(A,V) and assoc(B,V) have large values. The assoc function counts the number of edges connected to any A ore any B. This means that graphs containing 50 % A,B, there will be a large amount of edges connecting in between A:s and B:s, resulting in that these edges will increase both assoc(A,V) and assoc(B,V). If instead we have a large amount of connections inbetwen A:s, the same edges will only increase assoc(A,V) leading to a smaller assoc(B,V) and a bigger result. In other words when the amount A:s and B:s are equal, the amount of edges counted for both assoc A and assoc B are the largest, resulting in the smallest result.

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

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

Answers:

Increased radius lead to longer computation time, larger segments and slightly less defined colors. We won’t get as many details however the contour of the object will get more pronounced when radius increases.



First image, radius = 2, second radius=9.

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

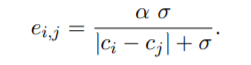
**Question 11**: Does the ideal choice of *alpha* and *sigma* vary a lot between different images? Illustrate with an example image with the parameters you prefer.

Answers:

Alpha controls the maximum penalty that an edge can get.

Sigma controls the difference between the maximum penalty and the minimum penalty.

This can be seen by the equation:

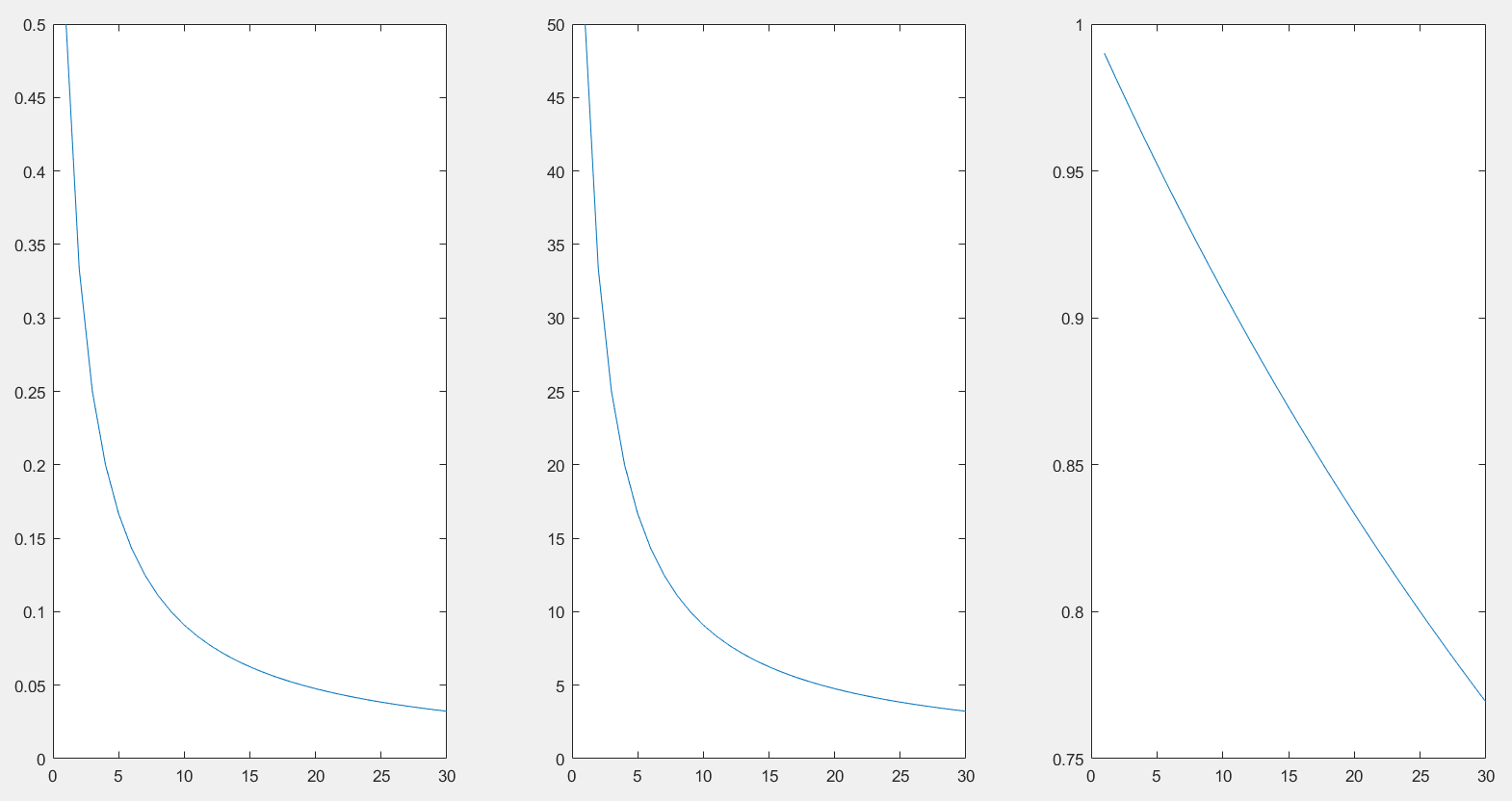


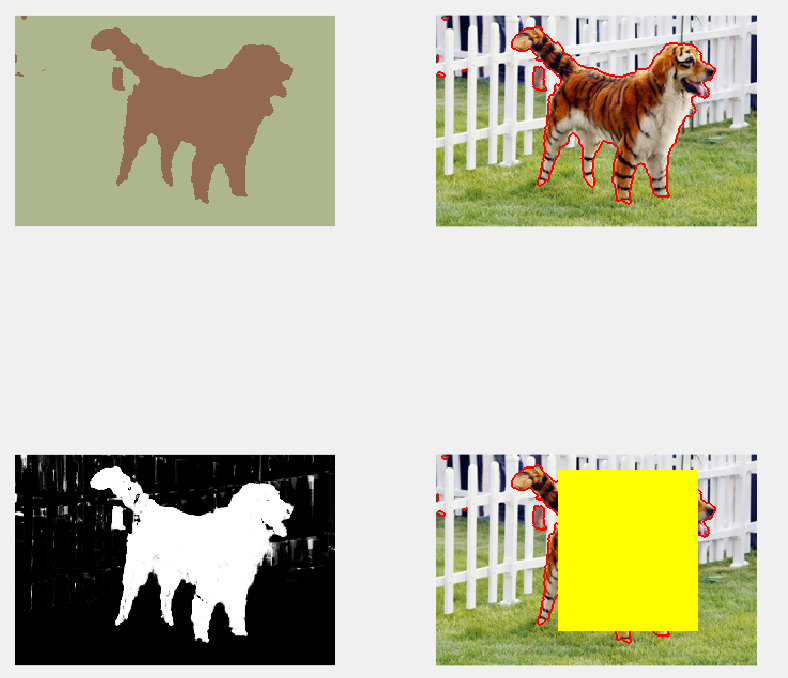
The values that works well will depend on the contrast, and color differences in the image used, and the selected k value. I also depend on the size of the mask.

For example, tiger2 needs a higher alpha and lower sigma due to the low contrast.

The following gave good results for tiger3, however it also yielded in some miss segmentation.

X = diff , Y = cost. (alpha 1, sigma 1) , (alpha 100, sigma 1), (alpha 1, sigma 100)





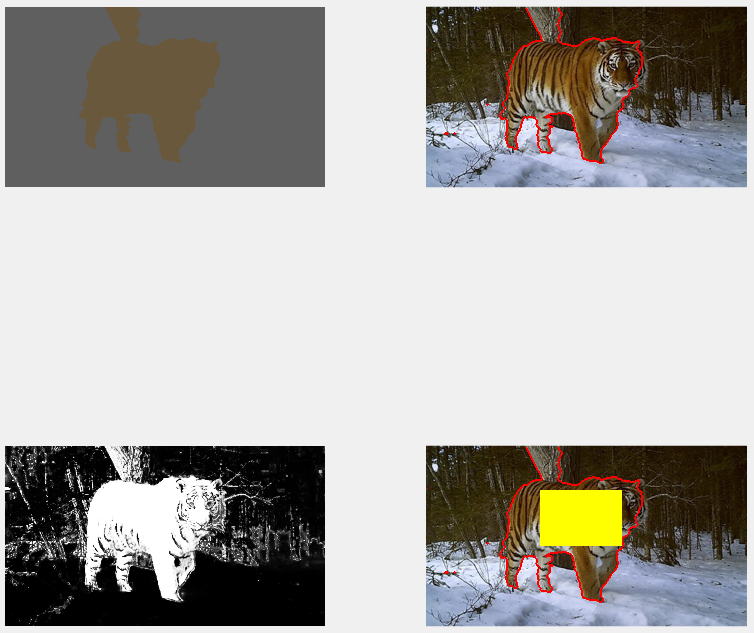
Params:

K = 10;

alpha = 15;

sigma = 2.0;

Result for tiger2:



Params:

K = 80;

alpha = 45;

sigma = 2.0;

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

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

Answers:

We can go very low, as long as we adjust sigma and alpha accordingly.

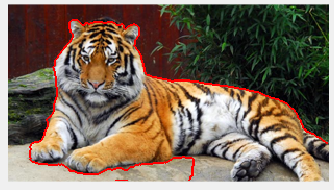
|  |  |  |  |
| --- | --- | --- | --- |
| **Image** | **K** | **Comment** | **Example** |
| ‘tiger1.jpg’ | 3 | It gets very har to separate the stone and the bottom part of the tiger with low Ks. |  |
| ‘tiger2.jpg’ | 2 | K can be lowered without rely effecting the results in a negative way, the segmentation of the feet even improved slightly. |  |
| ‘tiger3.jpg’ | 2 | For lower K values, some parts of the background were included, When sigma and alpha was adjusted to avoid that, the contours became a little bit rougher. |  |

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

**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:

The result is significantly improved; however, the use case will determine if it is worth the effort. Very big datasets could require a lot of work; however, this work could be automated using other object detection algorithms.



K-means vs graphcut.

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

**Question 14**: What are the key differences and similarities between the segmentation method

ds (K-means, Mean-shift, Normalized Cut and energy-based segmentation with Graph Cuts) in this lab? Think carefully!!

Answers:

**Similarities:**

* All uses segmentation through clustering.
* Uses statistical model like gaussian distribution.
* Works through an interactive process.
* Converges either to a maxima or a minima.
* Requires adjusting parameters based on the type of image.
* Normalized cut and graph cut are both based on graph theory.

**Differences:**

* Energy based requires human input.
* Soft assignment vs hard assignment.
* Energy based segm, only has one forground and one background.
* K-means does not take spatial information into account.
* Energy based uses probabilitys.

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