# Computer vision, Lab 3: Image Segmentation

Kildo Alias
Oskar Erlandsson

November 2019

## 1 K-Means Clustering

**Question 1:** How did you initialize the clustering process and why do you believe this was a good method of doing it?
**Answer:** We decided to split the image in to four equally sized areas in order to draw pixels in a random matter but with some bias added, we thought that this was a good choice because we get pixels that aren't to close to each other. Having pixels that are close to each other could make the picture unrecognizable if K is small and you're unlucky.

**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?
**Answer:** For lower values of K (eg. around 3-6) the number of iterations needed for convergence is approximately 20. If K is higher then more iterations are required.

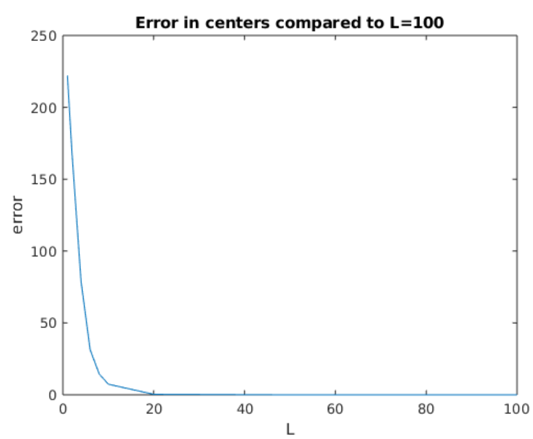


Figure 1: The error is compared to L=100 because at that point convergance is assumed, based on previous tests.

**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. Try using parameters suitable for orange.jpg and see how these affect the tiger images.
**Answer:** The minimum value of K is 7. This can be seen in Figure 2

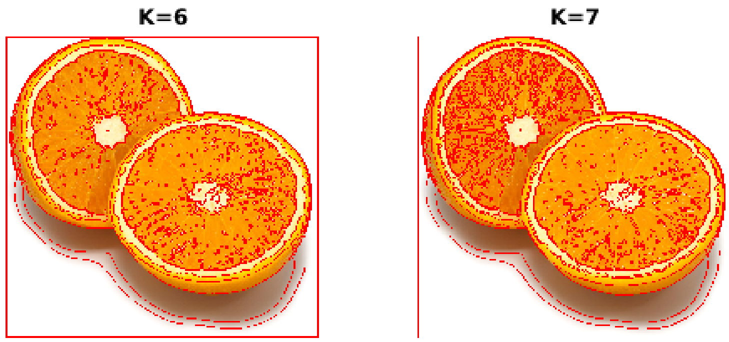


Figure 2: K-means segmentation with K=6 and K=7

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


Figure 3: image tiger1 with K=7 and K=10

Increasing the K-Value in will give more distinct super-pixels, as can be seen in Fig. 3 when increasing the K value to 10 we get a distinct superpixel covering the background and the tree.

## 2 Mean-shift segmentation

**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.
**Answer:** Changing the spatial bandwidth will increase the window size in the spatial domain for which where the mean is taken, changing the color bandwidth will do the same but in the color domain.

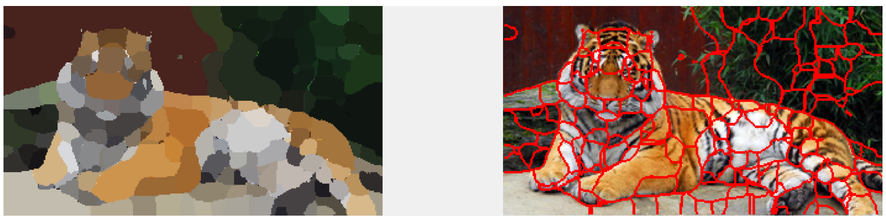

Figure 4: Tiger with spatial_bandwidth = 6.0 and colour_bandwidth = 10

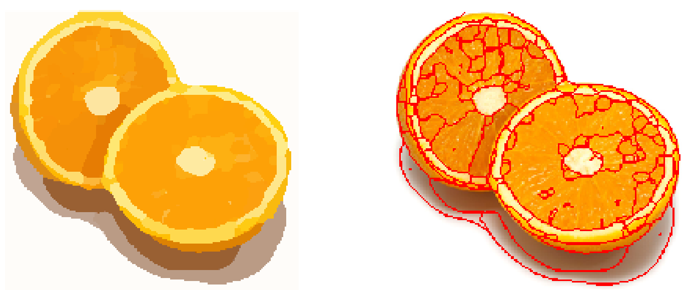

Figure 5: Orange with spatial_bandwidth = 4.0 and colour_bandwidth = 30.0

We decided to try for two pictures with distinct differences, the tiger picture where there is a lot of colors going on and for the orange where there isn't that many colors rather different shades of one color.
For the orange you want a higher value for the color domain because you want to merge similar

shades of orange in to one color in order to make larger segments.
For the tiger you need to be more careful with the window sizes as larger values can make the tigers color merge with the background, making it hard to recognize the tiger.

**Question 5:** What kind of similarities and differences do you see between K-means and mean-shift segmentation?
**Answer:** Both K-means and mean-shift find and group pixels with similiar color values, however where the mean-shift algorithm does this based on a window size in both color and spatial domain the k-means algorithm does this globally covering the whole picture.

## 3 Normalized Cut

**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 using the parameters you prefer for that image.
**Answer:**

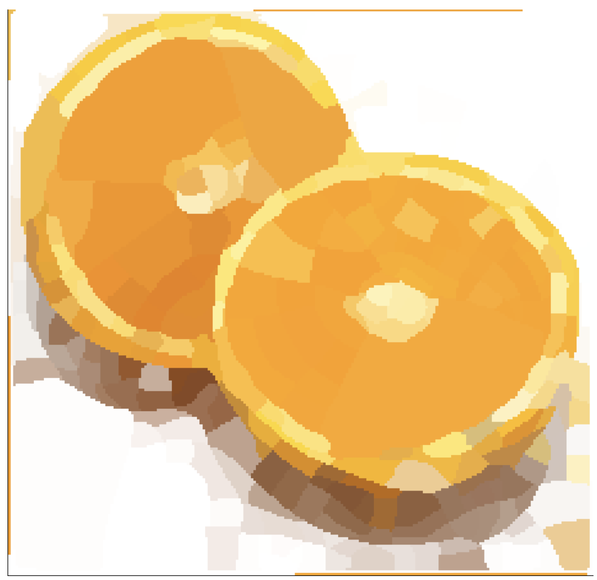

Figure 6: colour_bandwith=20, ncuts_thresh=0.5, min_area=60, max_depth=10

Figure 7: colour_bandwith=20, ncuts_thresh=0.2, min_area=15, max_depth=10

The tuning of the parameters varies depending on how many colors there is in the picture,how many changes there are, and the size of features. The bigger the features the bigger the min area. ncuts threshold determines if there is a cut or not depending on the change between pixels. If an image contains features with similar colors as the background the threshold should be set to a higher value.

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

**Answer:** Varying the parameters, min_area, max_depth and ncuts_thresh is the most effective. Where we would say max_depth was the most efficient, it decides how many times the function will be used recursively.

**Question 9:** Why does Normalized Cut prefer cuts of approximately equal size? Does this happen in practice?
**Answer:** The *normalized cut* is given by

$$\mathrm{Ncut}(A,B) = \frac{\mathrm{cut}(A,B)}{\mathrm{assoc}(A,V)} + \frac{\mathrm{cut}(A,B)}{\mathrm{assoc}(B,V)} \tag{1}$$

where $\mathrm{assoc}(A,V)$ and $\mathrm{assoc}(B,V)$ are the total connection from nodes in A to all other nodes and connection from nodes in B to all other nodes.
A graph can be partitioned into two disjoint sets, $A, B, A \cup B = V, A \cap B = \emptyset$, by simply removing edges connecting the two parts, ie by removing the $\mathrm{cut}(A,B)$, [1]. Equation 2 shows this partitioning.

$$\mathrm{assoc}(V) = \mathrm{assoc}(A,V) + \mathrm{assoc}(B,V) - \mathrm{cut}(A,B) \tag{2}$$

By rewriting equation 2 and solving for $\mathrm{assoc}(B,V)$ and using equation 1 the following equation is given

$$\mathrm{Ncut}(A,B) = \frac{\mathrm{cut}(A,B)}{\mathrm{assoc}(A,V)} + \frac{\mathrm{cut}(A,B)}{\mathrm{assoc}(V) - \mathrm{assoc}(A,V) - \mathrm{cut}(A,B)} \tag{3}$$

The optimal partitioning of a graph is the one that minimizes the $\mathrm{Ncut}(A,B)$. Minimizing the disassociation between A and B is identical to maximizing the association in the groups. Therefore differentiation is done with respect to $\mathrm{assoc}(A, V)$ instead as it is easier to compute. By differentiating equation 3 with respect to $\mathrm{assoc}(A, B)$ we get the equation below.

$$\frac{d\,\mathrm{Ncut}(A,B)}{d\,\mathrm{assoc}(A,V)} = \frac{\mathrm{cut}(A,B)(\mathrm{assoc}(V)+\mathrm{cut}(A,B))(-2\,\mathrm{assoc}(A,V)+\mathrm{assoc}(V)+\mathrm{cut}(A,B))}{\mathrm{assoc}(A,V)^2(\mathrm{assoc}(V)-\mathrm{assoc}(A,V)+\mathrm{cut}(A,B))^2} = 0 \quad (4)$$

By simplifying equation 4 equation 5 is given.

$$-2\,\mathrm{assoc}(A,V)+\mathrm{assoc}(V)+\mathrm{cut}(A,B) = 0 \quad (5)$$

By inserting equation 2 in equation 5 we get

$$-\,\mathrm{assoc}(A,V)+\mathrm{assoc}(B,V) = 0 \quad (6)$$

then $\mathrm{assoc}(A, V) = \mathrm{assoc}(B, V)$ must hold. In theory they are equal in size but because minimizing the normalized cut exactly is NP-complete a approximate discrete solution is used instead [1].

**Pre question 10:** Try to increase the radius to include neighbouring pixels that are a bit further away from each other. This usually leads to a better segmentation, but at the cost of slower computations.

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

(a) radius=5 (b) radius=10

Figure 8: Comparison between oranges for different radiuses, other parameters are the same.

(a) radius=3 (b) A radius=10

Figure 9: Comparison between tiger3 for different radiuses, other parameters are the same.

The computational time increased a lot for higher radiuses, however the result gets better as can be seen in Fig. 8 and Fig. 9, the segmentations gets larger which will lead to less variation in color.

# 4 Segmentation using graph cuts

**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.
**Answer:**

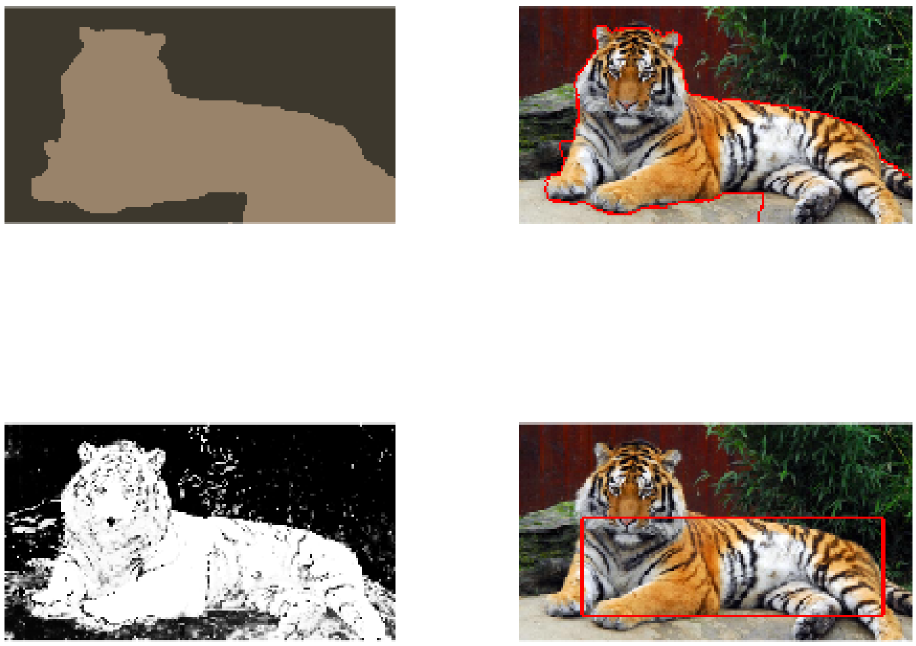

Figure 10: Tiger1 with the parameters K=16, $\alpha$ = 20, and $\sigma$ = 10

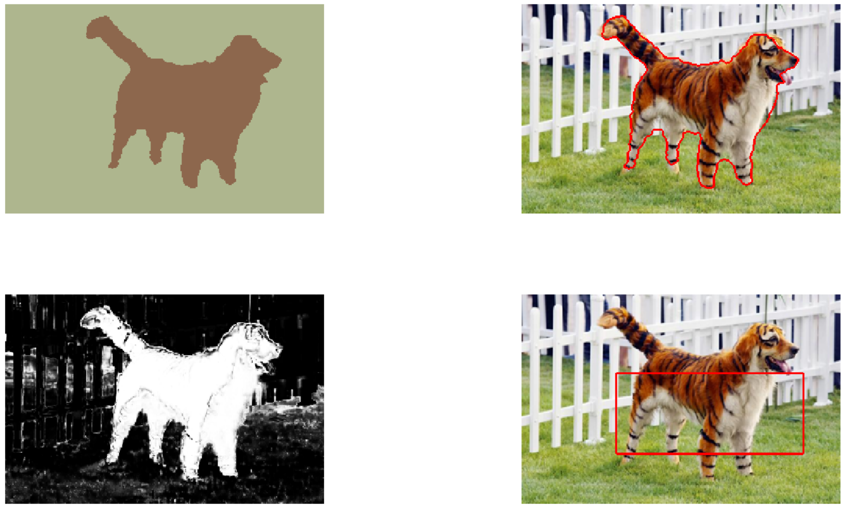

Figure 11: Tiger1 with the parameters K=16, $\alpha$ = 20, and $\sigma$ = 10

If it's more detail in the picture a higher value of $\alpha$ will increase the edge costs and will therefor make it non-beneficial to perform cuts that that will not result in good segmentations.

For similar images like tiger1 and tiger3 the same parameters will work well. They both have approximately the same amount of detail and variance in colors. If an image is different in details or colors then the same parameters will not be optimal.

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

Figure 12: K=4

Figure 13: K=3

If we keep the same parameters for the area, $\sigma$ and $\alpha$ and just change the K parameter , when having $K = 4$ we get a fairly good result, Fig. 12 and if we change $K = 3$ the result gets way worse, Fig. 13.

**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!

**Answer:** The segmentation is much better than in previous methods. The effort is well worth it but depending on the application user input might not be available.

**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!!

**Answer:** The difference are that k-means and mean-shift are not graph based like normalized cut and energy-based segmentation. The difference between k-means and mean-shift is that k-means does not account for spatial distance. The difference between normalized cut and energy-based segmentation is that normalized cut does not need prior information of what is foreground and what is background. The similarities is that all methods try to group similar points and segment them from each other.

# References


[1] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8):888–905, August 2000.