Answers to questions in

Lab 3: Image segmentation

Name: Engel Baptiste, Stefanos Marie-Ange Program: Msc Machine Learning

**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:

We started with a random initialization. However, it often happens that one “cluster center” has no associated point, leading to a hazardous behavior. To tackle this issue, we decided to start by taking a random point of the image. This is a good method to be certain that each cluster center will have at least one point associated.

**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:

The number of iterations is very dependent of the image. For the orange, a very simple image, it usually takes less than 10 iterations before convergence for K=3. However, for a complex image such as tiger2, it takes around 30 iterations.

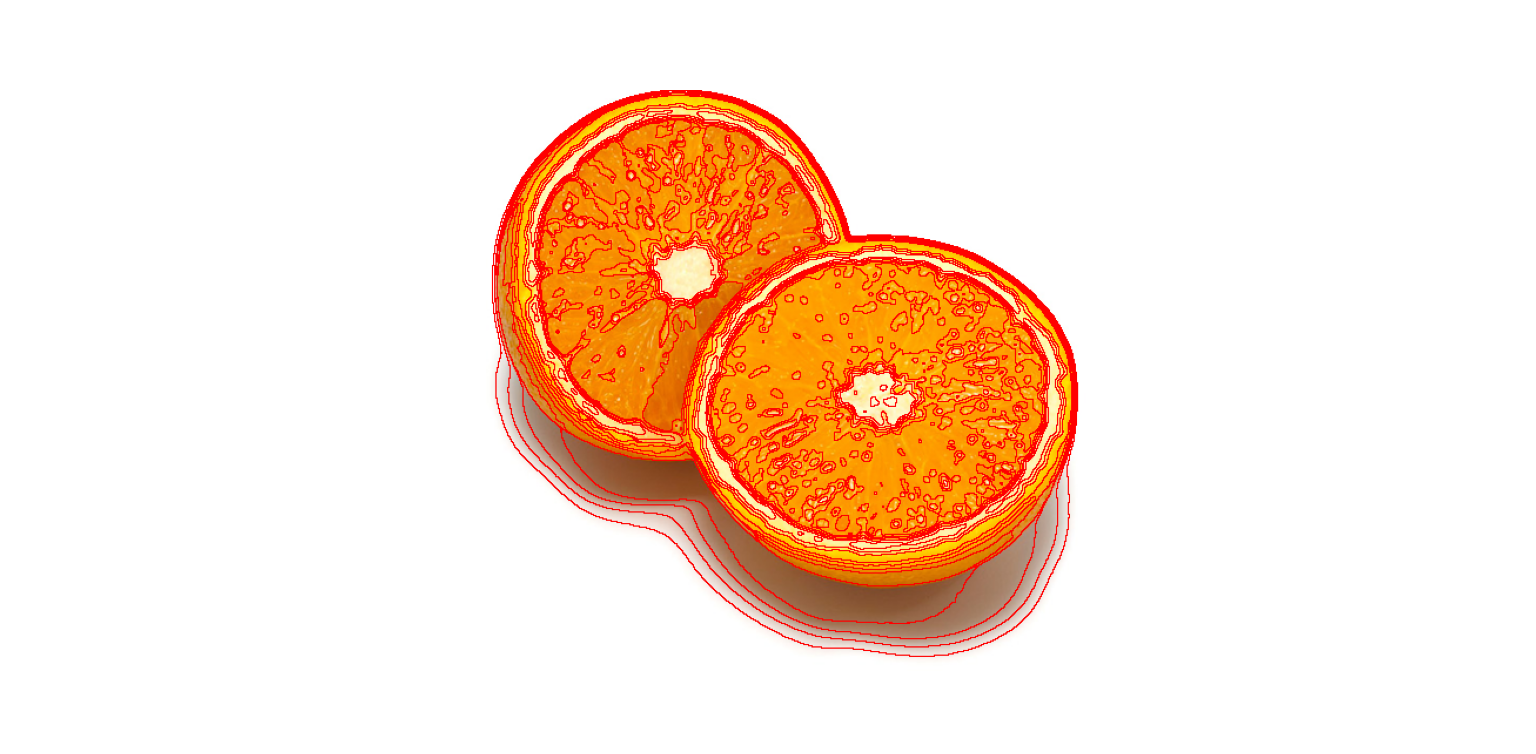
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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:

Une image contenant orange

Description générée automatiquementUne image contenant orange

Description générée automatiquement

K = 5 K = 10 K = 15

Figure 1 : k-means algorithm results applied to orange.jpg image

With k parameter = 5, 10 and 15 (from the left to the right)

And image\_sigma = 0.9

According to several tests we did, we can say that the minimumm value of K to have the orange correctly separated in two halves is K = 10. Indeed, as shows the above figure 1, below K = 10, the number of clusters is to low so the algorithm is not able to make the difference between the 2 halves (see left-hand side image for k=5). From K = 10, wwe can consider that the algorithm is able to detect a boundary between the 2 halves of the orange. Above K =10, it is clear that the value of K enables this separation.

Those results were obtained with an *scale\_factor* value of 0.9, which is the blurring factor.

Reducing the scale\_factor enables to reduce the minimum value of K to have the orange correctly separated in two halves (cf figure 2) where k minimum value seems to be 5 when image\_sigma = 0.5.

Une image contenant orange

Description générée automatiquement

Figure 2 : k-means algorithm results applied to orange.jpg image

With k parameter = 5 and image\_sigma = 0.5

2 :

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

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

Answers:

In order to get suitable superpixels for the tiger images as well, the value of the image\_sigma which is the preblurring scale and the image downscale factor.

The higher the higher the preblurring scale, the bigger the superpixels. This enables to avoid clusters of details or shadows in the leaves for instances.

Moreover, the lower the image downscale factor, the bigger the superpixels as well, what enables to overcome the same issue.

Besides, increasing the number of iterations L can be a good strategy as well to reach a satisfying convergence.

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

**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:

When the color bandwidth grows, the number of segments is reduced. When the spatial bandwidth grows, the number of segments grows.

A good segmentation of the tiger1.jpg image is obtained with a 50 color bandwidth and a 10 spatial bandwidth:

Une image contenant texte, crabe

Description générée automatiquement

Fig ??. Tiger1.jpg segmentation using mean shift.

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

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

Answers:

Kmeans and mean-shift segmentation are two iteratives methods. However, while K-means minimize the distance between in cluster pixels, mean-shift segmentation maximize the given continuous density function.

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

**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:

Yes, the ideal parameter setting vary depending on the images. The ideal settings might differ because all the three parameters (ncuts\_thresh, min\_area, max\_depth) impact the size of the segments, which deeply depends on the image (the size of the object, the colors of the object/background). Moreover, other parameters need to be taken into account, such as the scale\_factor or image\_sigma (which has been set to 0.7) in the below illustration figure 3.

Une image contenant texte

Description générée automatiquement

Figure 3: Result of normalized cuts algorithm on *tiger2.jpg* image

Using the following parameters:

ncuts\_thresh = 0.05

min\_area = 50

max\_depth = 7

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

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

Answers:

Most effective parameter for reducing the subdivision is min\_area (when decreased). However, the one that is also very effective but still result in a satisfactory segmentation is the ncuts threshold.

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

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

Answers:

This is due to theoretical reasons that we can detail by expressing Ncut(A, B) and take its derivative since the goal is to minimize this quantity. Then we can see that the derivative equals zero when the image is cut in 2 cuts of equal size.

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

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

Answers:

Increasing radius not only gives better results (less wrong detected boundaries, larger segments) but it also reduces the computation time.

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

**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:

For the image tiger2, a good result is obtained with the following parameters:

Une image contenant texte, mammifère, tigre, fauve

Description générée automatiquement

Une image contenant texte, herbe, chien, mammifère

Description générée automatiquementIf we keep the same parameters for tiger1 and tiger3, we have the following result:

Une image contenant texte, mammifère, pose, tigre

Description générée automatiquement

We can therefore conclude that the ideal choice of parameter do not vary a lot between different images.

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

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

Answers:

With the same parameters and on tiger3 image, we can see that the results is still good for k=3, but failed for k=2.

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

**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 benefit you get of this does not worth the effort in most cases, even if it can be useful in some cases. More precisely, when the image shows an object in the foreground separated from the background, using a bounding rectangle in order to define the location of the object in the image can be useful for graph cut segmentation. Nevertheless, the user defining a rectangle for an image is not worth for images such as a landscape.

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

**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!!

Four methods for image segmentation have been studied in this lab: k-means, mean-shift, normalized-cut and energy-based segmentation with graph cut.

Differences:

Graph Cut is essentially the only model that need prior information about the expected ratio of foreground/background (in order to produce more accurate results), whereas Normalized cut does not need any such information

Mean-shift (and the Graph-based methods) does take spatial information into account, whereas K-means does not and only looks at the colour dimension

This causes K-means to often result in segments which go across areas that are separated

Similarities:

* Between all the methods: All the methods lean on clustering to group data points and have the same goal, that is grouping parts of the image that are similar in the same cluster, and separating it from the others that are dissimilar.
* Between mean-shift and graph-cut: they both use a Gaussian distribution in order to model the data.
* Between normalized cut and graph-cut: they both see the image as a graph that encompasses vertices and edges, with a defined measure of similarity.

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