<u>Lab 3</u>
<u>Assignment 3</u>
<u>PART1</u>
K-means clustering

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

Answer:

K-means clustering depends very much on initialization of the cluster canters. Choice of a wrong initialization may lead to classification of wrong clustering and also may lead to higher computational cost. There are several different algorithms for good initialization of clusters and in this exercise, I have chosen to do the initialisation through k means ++ with some modifications. Also, according to the image, we can choose the initial clustering by ourselves, and I tried to do it using several standard methods.

<u>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?</u>

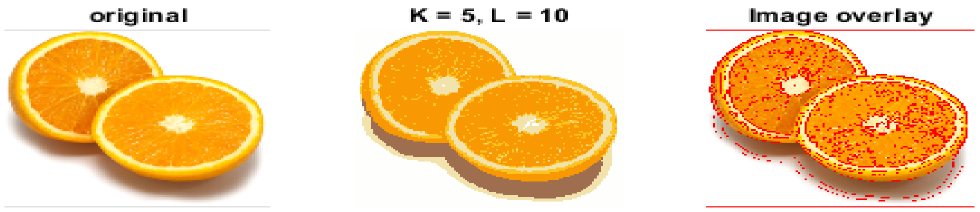
Answer:

This problem of finding the least Iterations L_{least} after which any number of Iterations will show the same result can be solved by thresholding. This assumes the most acceptable threshold as the difference between the precision of the segmentation after successive iteration which can be said termed as acceptable. Thus, we can have a threshold of 0.02 to be acceptable and we find the value of L_{least} for the same for images that we consider. Primarily this value depends on the Image, the number of clusters, the scale factor as well as the standard deviation of gaussian smoothing.

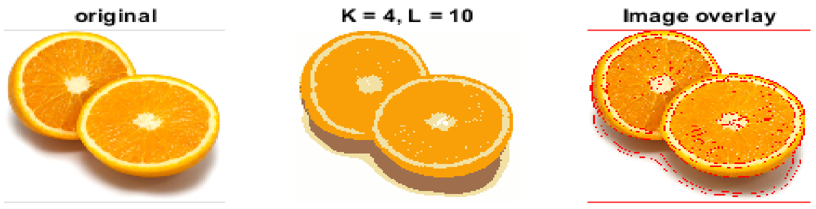
<u>Question 3: What is the minimum value for K that you can use and still get no superpixel that covers parts from both halves of the orange? Illustrate with a figure.</u>

Answer:

When I run the algorithm, I get K=5 as the minimum number of clusters required to show a distinct boundary between oranges and any value below it gives me not so distinct boundary between the two parts of the orange. The conclusions are illustrated from the figures below



Above Figure when the number of clusters is 5 (K=5)



Above Figure when the number of clusters is 4 (K=4)

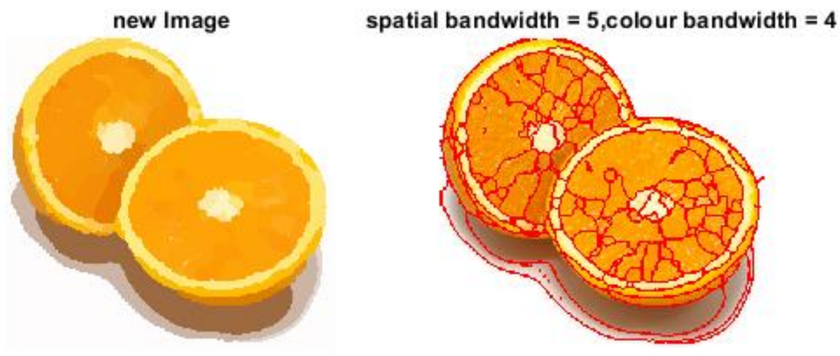
Question 4: What needs to be changed in the parameters to get suitable superpixels for the tiger images as well?

Answer:

Clearly the tiger Images require more segmentation than the orange Images as it has several colours and distinct edges as well therefore we require more number of clusters to showcase these distinct segmentation and thus we will require a greater K value and also with that the number of iterations required to converge will be greater.

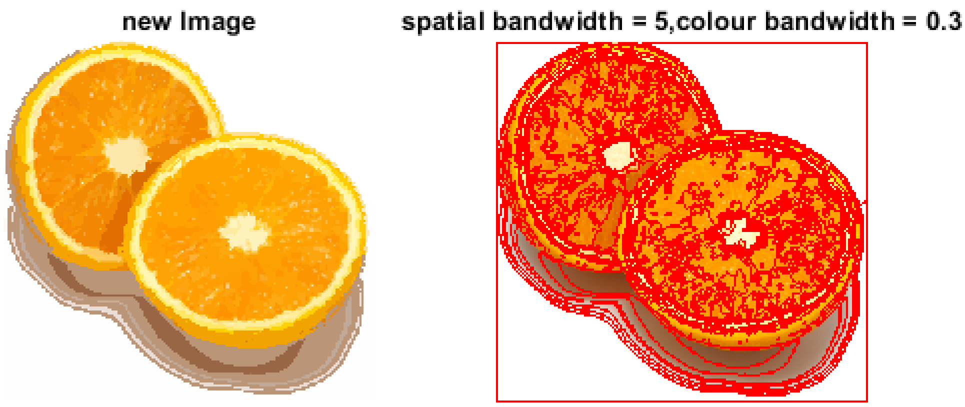
Question 5: How do the results change depending on the bandwidths? What settings did you prefer for the different images? Illustrate with an example image with the parameter that you think are suitable for that image.

Answer:
For the mean shift segmentation, the region of Interest or the window is chosen by the user. Here if the spatial bandwidth is high the region of Interest encapsulates greater number of points and there is a chance of missing modes thus fewer segmentation lines are obtained and if the region of Interest has lower spatial bandwidth, we require more computations but can possibly cover local maxima and minima modes. As the colour bandwidth becomes greater, the weights while considering the centre of Image in the area of Interest becomes greater

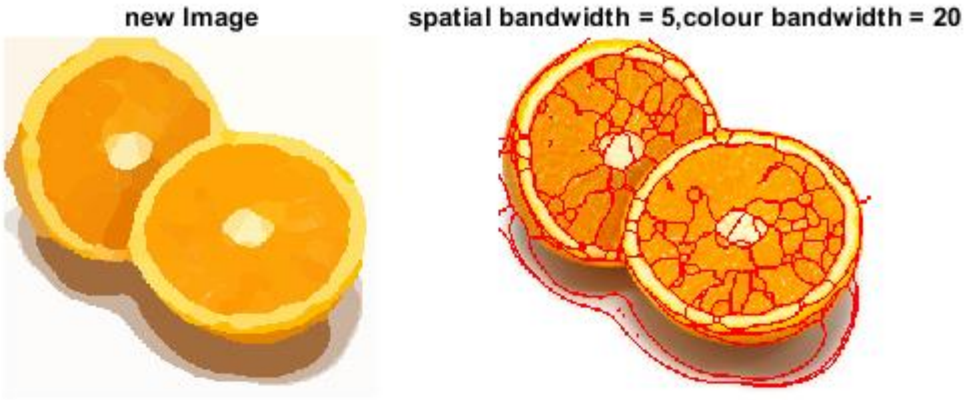


Spatial Bandwidth = 5 and colour Bandwidth = 4

Spatial Bandwidth = 20 and colour Bandwidth = 4



Spatial Bandwidth = 5 and colour Bandwidth = 0.3



Spatial Bandwidth = 5 and colour Bandwidth = 20

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

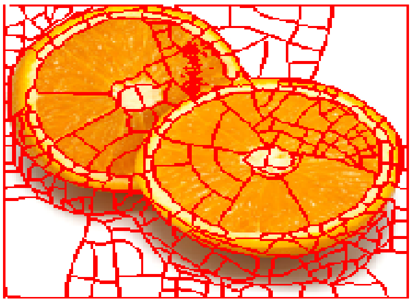
Answer:

The K-means as well as the mean-shift algorithm are both used when the features are mainly colour, position and for the purpose of associating clusters to their respective cluster heads. However, In K-means we have to know before hand the number of clusters which are possible in our data and also K-means is sensitive to both initialization as well as outliers. For the mean shift algorithm, the modes or number of clusters formed are identified by the algorithm itself but size of window (region of Interest) plays a crucial role in accuracy of segmentation as well as Identifying the number of modes that are possible for a given data. Also mean shift algorithm has scalability issues and doesn’t perform very nicely for large scales.

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:



ncuts_thresh = 0.4, min_area = 52, max_depth = 9

ncuts_thresh = 0.4, min_area = 25, max_depth = 9

Thus we see that the parameter settings vary from Image to Image. As mentioned, three factors affect the algorithm namely the minimum area which decides the size of segment required, and the threshold that controls the maximum allowed value of the cost of normalisation cut as well as the maximum depth that limits the depth of recursion. In the above example we see that in the case of tiger the features are more complex and thus our area of Interest or minimum area should be small and increase the n_cut threshold to inculcate the minute changes in colour and shape.

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

Answer:

For a satisfactory segmentation we need to effectively choose the minimum area, normalization cut threshold and the maximum depth allowance.

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

Answer:

The graph cut algorithm has various variations. Most of the algorithms give very less weightage to the sought out isolated points which forms an incorrect segmentation if the feature space is very complex and it has many isolated features which cannot be merged to form segments. The normalisation cut algorithm fairly introduces a scheme where the cost of segmenting such isolated vertices is high as it also depends as can be seen by the cost equation. Also, if we differentiate the normalised cost and make it zero in order to find optimal cut, we get that the cut is preferred to be of an equal size between two distinct features. However, in practise the optimisation problem sometimes becomes an NP- hard Problem and thus it doesn’t happen in practise.

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

Without any loss of generality, we can write,

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

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

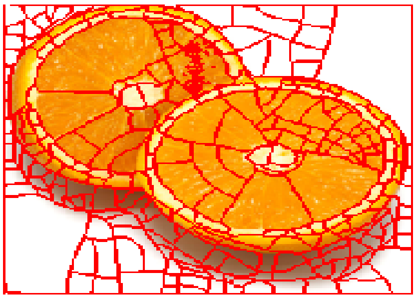
Let us compute the derivative and equate it to zero for optimal solution,

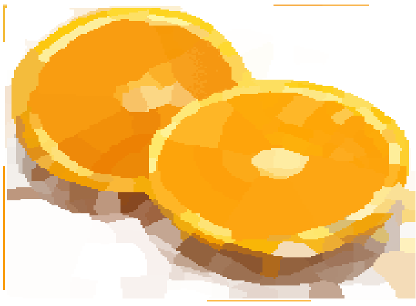
$$\frac{\mathrm{dNcut(A,B)}}{\mathrm{dassoc(A,V)}} = \frac{\mathrm{cut(A,B)}\big(\mathrm{assoc\ V} + \mathrm{cut(A,B)}\big)\big(-2\mathrm{assoc\ (A,V\ )} + \mathrm{assoc\ (V)} + \mathrm{cut(A,B)}\big)}{\mathrm{assoc(\,A,V)}^2\big(\mathrm{assoc\ (V)} - \mathrm{assoc(\,A,V)} + \mathrm{cut(A,B)}\big)^2}$$

Thus we clearly see that $assoc\ (A,V) = assoc\ (B,V\ )$ is an optimal solution for the weight function.

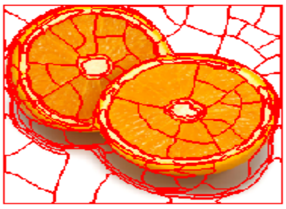
<u>Question 10: Did you manage to increase radius and how did it affect the results?</u>

Answer:





For Radius = 3



For Radius = 10

For a higher radius it takes more time to converge as the calculations are higher and it gives more accuracy for segmentation. However, the colour becomes distorted as we increase the value of radius.

<u>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.</u>

Answer: The images are as follows:

Figure when sigma = 10 and alpha = 8.

Figure when sigma = 30 and alpha = 25

This is a control problem if we see the equation that determines the weights between segments only and not between vertex to source or vertex to drain. This operation of selecting the right sigma and alpha is therefore dependant on the image and the complexity of it's features.

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

Answer:

K is a factor that depends on the image. In the given tiger1 image K=2 also gives a result not accurate enough but still segmentation is possible , however for K=1 a bad result is obtained which is also logically justifiable as K=1 is by common knowledge no segmentation at all.

Figure when K = 2 (We can see that the foreground and the background are merged completely in the bottom of the 1st Image, however that doesn't completely disregard K, by accurate thresholding methods we can possible classify the wrong part with certain accuracy.)

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:

Yes, there are many ways in which we can play around with user's input like the pixels in which user has established as foreground we can give them infinite or very high weight in our algorithm such that they will least preferred for segmentation. Thus, the user's input is highly valuable also it saves us from many calculations and uncertainties unlike K- means where initialisation plays an important rule or the scaling problem like in the mean shift segmentation.

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?

Answer:

K-means is guaranteed to converge however it has several disadvantages such as :
1. It is memory intensive
2. The user needs to give number of clusters that will be formed in advanced
3. It is sensitive to initialization and outliers
4. Generally, fits the data in a "spherical" cluster.
5. Categorical values aren't allowed. (as operations like average cannot be performed on different categories.

The mean shift algorithm has several advantages over K-means which are as follows:
1. Automatically finds the modes we don't have to give the K value or number of clusters
2. Non-parametric method data doesn't have to fit any gaussian or alike distributions.
3. Does not assume clusters to be "spherical", like K- means
4. Can find multiple modes and facilitate thresholding

However, there are disadvantages as well in this method which are as follows:
1. Uses gradient accent and possibility of algorithm stuck on saddle point
2. Have to run the algorithm from different spatial locations to check if the mode is correct and hence extra computation.
3. Selection of window size is the most important aspect. If too larger modes will be skipped if too small noise will also be considered.
4. Does not scale well with dimension of feature space and higher computation time.

The above 2 algorithms are good while considering features like the spatial or colour however they fail when it comes to texture or complex images with minute difference in colour of an image but a distinctive texture.

For graph cut methods we convert the image to vertices and edges where vertices are the image pixels where as the edges are the lines connecting between vertices. We assume affinity of weight or Energy parameter for these edges and try to find a solution where we can minimize the weights. The disadvantage of min cut method is that it takes weight of cut proportional to the number of edges in the cut and hence we have unnecessary cuts for isolated points in the image thus to solve this issue Normalization cuts algorithm was introduced as it accounts all the other connected edges in the cost, it is theoretically a fair algorithm. However in practise we find that sometimes optimization becomes an NP-hard Problem (Non deterministic polynomial time hence it has to use Eigen value approximation to solve the problem). The advantage of Normalization cut are as follows:
1. It has a generic framework and there are several methods of dealing with graph structures.
2. It only requires a distance function and not a model of distribution

The disadvantages are as follows:
1. For highly densed graphs there may be many affinity calculations and time complexity
2. Generally, only preferable for balanced partitions.

The Energy based graph cuts have the following advantages:
1. The probability distribution for the pixels in taken into account by iterating through a matrix and all factors such as mean, covariance are taken into account.

However, the results will ultimately depend on the object as well as the input from the user who creates a window in which the probability of foreground maximizes. There are of course several parameters to tune for this technique to give very good results.