MEAN SHIFT SEGMENTATION

Parag Jayant Datar
UBID: 50169787
Date of Submission: 14^{th} Dec 2015
Computer Vision and Image Processing

1. Introduction

Image segmentation is the process of dividing an image into multiple meaningful segments. In real life, images are widely used to capture and store visual data which can be used later. Most of the times we aim to capture certain objects specifically and end up including other things (or background) unintentionally, as we cannot limit our image to the object of interest. This helps us understand that we are mostly concerned with a certain region or object in the image more than the others. This is nothing but what a 'meaningful segment' is. Segmentation helps us identify our areas of interest, separate them from other parts, so that we can specifically concentrate on the required object. Image segmentation is one of the most basic technique needed for many processes and its' application includes areas such as Artificial Intelligence, Computer Vision, Medical imaging, etc.

Broadly, segmentation can be classified as:

- Complete segmentation
- Partial segmentation

In practice, complete segmentation is extremely difficult as great amount of high level knowledge is required.

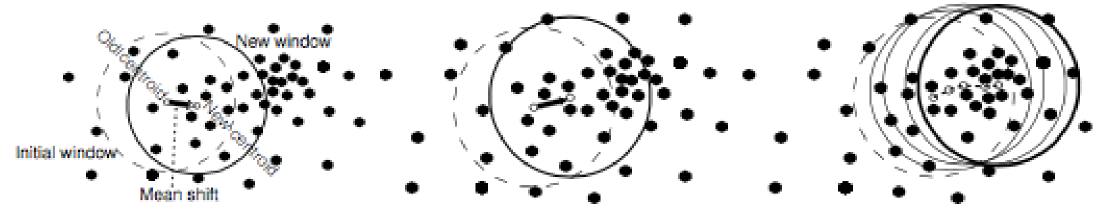
There are three approaches for image segmentation

- Global
- Region based
- Edge based

In this project, we will be using a region based approach known as Mean shift segmentation

2. Literature review

Mean shift technique: It is a non-parametric clustering technique which does not require prior knowledge about clusters such as shape. It is a density gradient estimation using a generalized kernel approach. It falls under region based segmentation approach. Figure below illustrates mean shift concept.



There are two broad level steps in the mean shift algorithm:

- Mean shift discontinuity preserving filtering
- Clustering

Mean shift discontinuity preserving filter:

In this step, we start with aim to modify the intensity of a group of pixels (those that follow the criteria we set up for identifying the neighborhoods) to a value that is the average or 'mean' intensity value of the entire neighborhood.

The algorithm for mean shift discontinuity preserving filtering is as follows:

1. For each image pixel $\mathbf{x}_i$, initialize step $j = 1$ and $\mathbf{y}_{i,1} = \mathbf{x}_i$.
2. Compute $\mathbf{y}_{i,j+1}$ as given in equation (7.12) until convergence $\mathbf{y}_{i,\mathrm{cor}}$.
3. The filtered pixel values are defined as $\mathbf{z}_i = (\mathbf{x}_i^s, \mathbf{y}_{i,\mathrm{con}}^r)$, i.e., the value of the filtered pixel at the location $\mathbf{x}_i^s$ is assigned the image value of the pixel of convergence $\mathbf{y}_{i,\mathrm{con}}^r$.

The complete algorithm i.e. clustering followed by mean shift discontinuity preserving filtering is as follows:

1. Employ the *mean shift discontinuity preserving filtering* and store all information about the d-dimensional convergence points $\mathbf{y}_{i,\mathrm{con}}$.
2. Determine the clusters $\{C_p\}_{p=1,\ldots,m}$ by grouping all $\mathbf{z}_i$, which are closer than h_s in the spatial domain and h_r in the range domain. In other words, merge the *basins of attraction* of these convergence points.
3. Assign $L_i = \{p|\mathbf{z}_i \in \mathbf{C}_p\}$ for each pixel $i = 1, \ldots, n$.
4. If desired, eliminate regions smaller than P pixels as described in Algorithm 6.22.

3. Approach

In our project, we will be using Python as our programming language. The algorithmic steps that we followed while applying the mean shift technique in a programming fashion are as follows:

- Create a Feature vector matrix of the dimension A x 5 where
  - A is the number of pixels in the image.
  - 5 numbers represent the range and spatial information as r,g,b,x,y.
- Till we have rows in our matrix, choose a random row from the matrix and consider it as the initial seed point and initial mean.
- Find neighbors of the seed based on a distance kernel function (we have used a uniform kernel and the distance is calculated using the Euclidean distance. Include only those pixels as neighbor that fall below a certain threshold H.)
- Calculate the mean of all the neighbors and calculate the mean shift (shift related to the initial mean)
- If the shift falls below a certain threshold 'Iter' then mark all the pixels in this neighborhood with the mean value and delete the rows from the matrix. Else, go to the second step.

4. Results

Input image for this section.

Output Image with H = 90 and Iter = 225

Output Image with H = 60 and Iter = 200

Output Image with H = 60 and Iter = 200

As we can see, at a constant Iter value, increased H value merges pixels with a larger intensity spectrum. For the appropriate discontinuity filtering results, the value of H and Iter should be appropriately chosen. Otherwise areas with spatial and range properties over a wider area will be merged and we might lose some information from the object under consideration.

In these results, we can see that the white spots on the wings of the butterfly are left intact as they differ greatly from the surrounding black color. Whereas the leaves have been compromised and only partial leaves with pixels greatly differing in the intensity from neighborhood are preserved.

By analyzing the results for multiple values of H and Iter, we can understand what values are suitable to us. The values purely depend upon the task at hand, amount of information we need to suppress from the surrounding and number of clusters.

5. Problems encountered.

The problems faced during the process were largely related to the run time of the program based on the approaches that were undertaken. Instead of deleting the rows, masking technique was used to mask the pixels which were clustered so that we don't update the matrix for every successful convergence. The biggest improvement was observed when the numpy module was incorporated which greatly sped up the program. As numpy is deployed with heavy optimizations, the runtime was brought down from 1.5 hours to a whopping 2.5 minutes. Once the program was ready where more results could be studied, many results were analyzed and the threshold parameters were exploited.

Other problem was fixating upon the above mentioned H and Iter values.

Some bad values might result into such outputs: (When Hs and Hr are similar or Hs is marginally greater than Hr)

6. Improved algorithm

Another common technique used in the selection of threshold is using 2 threshold values viz. Hs and Hr where Hs is threshold for spatial domain and Hr for the range domain. This pair of threshold is very useful where the range and spatial parameters cannot agree on a single value for a meaningful segmentation result. Also, instead of the normal kernel used for distance calculation, we can use Epanechnikov or Gaussian kernel which gives us different weights for the pixels based on the probability distribution functions governed by these kernels.

Results for Hs and Hr pair

Output Image with Hs = 64, Hr = 42 and Iter = 50

Input image

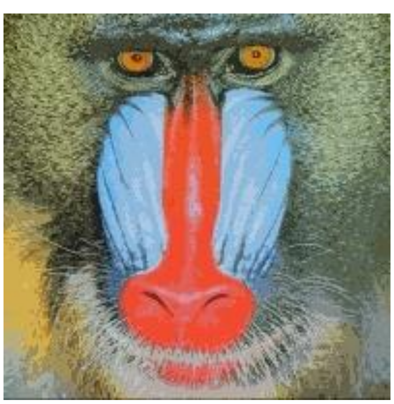
In this Baboon test image, We can see the different outputs for different parameter values. Some points for relating different outputs would be the whiskers, the nose or the hair on his face around the eyes and the nose

Hs = 64, Hr = 50, Iter = 200

Hs = 64, Hr = 50, Iter = 100

Hs = 64, Hr = 50, Iter = 50

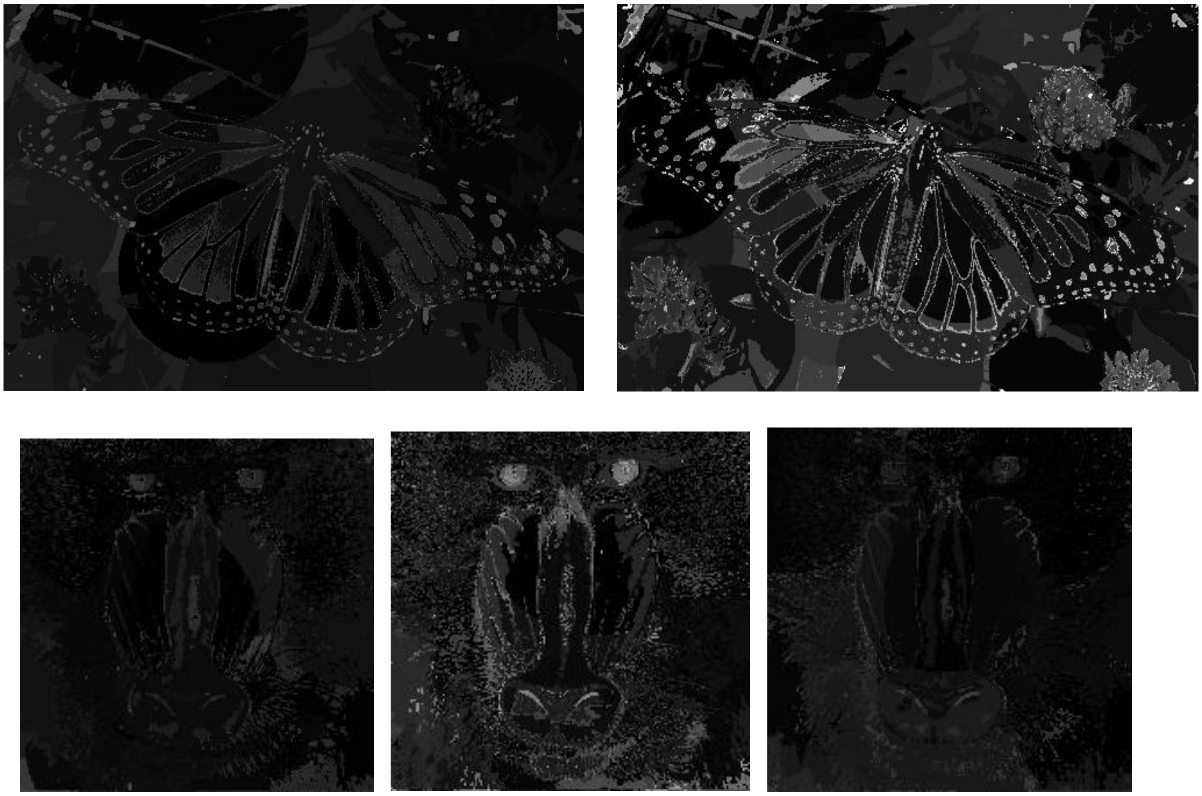
Hs = 64, Hr = 32, Iter = 50



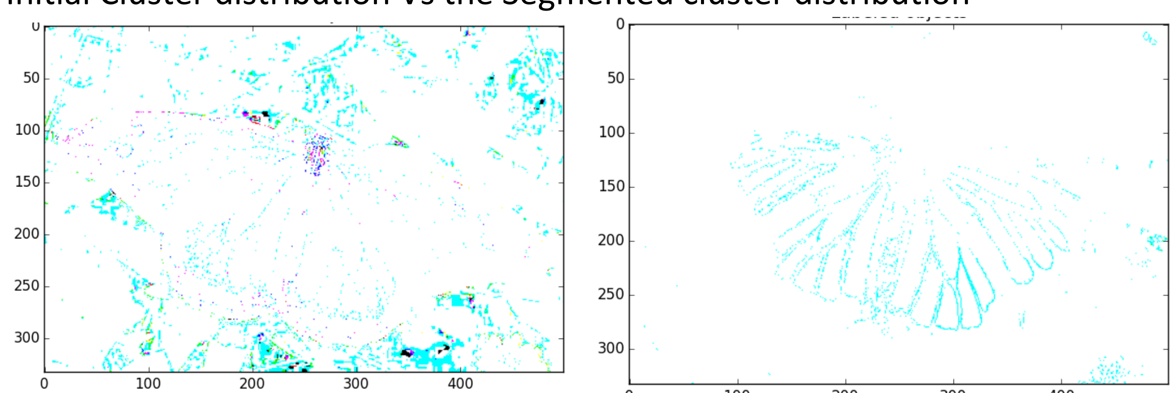
Hs = 64, Hr = 32, Iter = 100

Hs = 64, Hr = 32, Iter = 50

Clustered Images (showing the convergence of neighborhood on the images):



Initial Cluster distribution Vs the Segmented cluster distribution



7. Programming Environment

Python was used as the programming language as mentioned earlier and the Canopy editor by Enthought, Inc was used as the Development environment. No code was adopted and everything was built from scratch with the use of standard libraries such as numpy, opencv for trivial tasks. The mean shift segmentation procedure was created from scratch.

8. Summary

In short, we studied mean shift segmentation in great detail with hands-on practical experience. This project helped me understand the basics behind image segmentation, desired results, the challenges and important points to be considered during the entire process of program development. Mean shift was studied in great detail and exploited to obtain the necessary results.

9. Bonus implementation

The bonus part asked for performing mean shift segmentation over a series of images. The idea is that we have to detect the translation of a cluster (identified by the mean shift algorithm) over a sequence of images. The algorithm for the same is as follows:

a. Perform mean shift segmentation and identify clusters.
b. If we already know the cluster of interest, use only single cluster for consequent images.
c. Else keep a track of the centroid (x_{mid},y_{mid}) of all the clusters.
d. Perform mean shift segmentation for the next image in sequence and find the new centroid of the clusters.
e. If the centroid shifts by a value greater than a threshold, then identify it as moving cluster.
f. Iterate over all the images in the sequence in this fashion.

12. References and Acknowledgements

https://saravananthirumuruganathan.wordpress.com/2010/04/01/introduction-to-mean-shift-algorithm/

http://cmp.felk.cvut.cz/cmp/courses/33DZOzima2007/slidy/meanShiftSeg.pdf

http://wwweb.eecs.umich.edu/vision/teaching/EECS442_2012/lectures/seg_cluster.pdf

https://courses.csail.mit.edu/6.869/handouts/PAMIMeanshift.pdf

http://stackoverflow.com/questions/4831813/image-segmentation-using-mean-shift-explained

http://luthuli.cs.uiuc.edu/~daf/courses/CS-498-DAF-PS/Segmentation.pdf

http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/TUZEL1/MeanShift.pdf

https://en.wikipedia.org/wiki/Mean_shift

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.76.8968&rep=rep1&type=pdf

Image processing, Analysis and machine vision by Milan Sonka, Vaclav Hlavac, Roger Boyle-third edition

Lecture slides by Professor Chang Wen Chen