Mean Shift Segmentation

by Sahil Dureja(50168872)

Date: Dec 14th, 2015

1) Introduction

The basic idea is find local maxima from the density distribution of feature vectors of an image. A feature vector is the representation of all possible features of an image into n dimensions. For example an RGB image will have Red, Green, Blue value as well as their x and y coordinate all counted as features. Similarly a gray scale image can be represented by using gray scale value, x coordinate and y coordinate as a 3 dimensional feature vectors.
Now using optimal thresholding based on image features we will find local maximas and then later we will merge those which fall below a threshold

The thresholding process is applied two times.
1) First time to find out a lot number of local maximas and store region of convergence of each local maxima
2) Instead of pixels to smaller 'regions of convergence' we merge these regions of convergence into bigger regions which fall within a threshold.

Applications:

As far as segmentation is concerned mean shift may not be the best solution but in case of object tracking it is really a good solution. By good solution we mean how good is the running time of the algorithm

Literary Review

Technologies Used and Learnt

1) Image into feature vector
2) Optimal Thresholding
3) Point of convergence(Local maximas)
4) Spatial and Range domains
5) Maxima vs Plateaus (False Plateaus as Peaks
6) Defining regions
7) Region Merging
8) Regions boundary
9) Function running complexity of O(n2) of each iteration

2) Approach

Steps followed in the project

-Choose an spacial and range threshold value by experiment and shift size

These values depend upon the size of regions as well as contrast of the images

-Image to feature vectors

Loop over feature vectors array till not empty

-Pick a random feature vector as mean

-Get all feature vectors that fall within spatial and range threshold

-Calculate the mean column wise to get a point of convergence/centroid

Calculate the shift from old mean to new mean

If the shift is low/ local maxima found(False maximas not checked in this project**)**

Store the centroid and the region vectors to be later merged

Remove feature vectors keys of the regions found.

If the shift is high reiterate and find local maxima from new mean

-Choose again spacial and range threshold for the regions to be converged
-Use the stored means and store vectors list of regions.
--If regions are found within threshold :
-Calculate global mean of all the regions to be merged and assign value to each pixel on an empty image.
Remove these keys from maximas list (list where means of regions are stored) so that the remaining regions can be merged
-Ignore where the number of pixels in a region are less than a threshold.

Snippets from paper useful to write algorithm (paper[2])

Mean Shift Algorithm

1. Choose the radius r of the search window.
2. Choose the initial location of the window.
3. Compute the mean shift vector and translate the search window by that amount.
4. Repeat till convergence.

Feature Space Analysis

1. Map the image domain into the feature space.
2. Define an adequate number of search windows at random locations in the space.
3. Find the high density region centers by applying the mean shift algorithm to each window.
4. Validate the extracted centers with image domain constraints to provide the *feature palette*.
5. Allocate, using image domain information, all the feature vectors to the feature palette.

O(n2) is the time complexity we face for one iteration.

Number of iterations is dependent upon

1) Thresholds we set for segmentations
2) Thresholds vary depending upon contrast in features of the image
3) Size of the image
4) Average region size

Cost / time of calculations (paper refered [2])

Segmentation resolution is the most general parameter characterizing a segmentation technique. While this parameter has a continuous scale, three important classes can be distinguished.

Undersegmentation corresponds to the lowest resolution. Homogeneity is defined with a large tolerance margin and only the most significant colors are retained for the feature palette. The region boundaries in a correctly undersegmented image are the dominant edges in the image.

Oversegmentation corresponds to intermediate resolution. The feature palette is rich enough that the image is broken into many small regions from which any sought information can be assembled under knowledge control. Oversegmentation is the recommended class when the goal of the task is object recognition.

Technologies used

Language Python

Python IDe Enthought Canopy

Scientific and Analytic Python Deployment with Integrated Analysis Environment.

3) Outcome and Deviations

Shift of hs hr Window shown in Canopy results window:

-Size of feature vectors array

-Size of shift from old mean to the new

-When shift is within threshold (iter(3 is used in this code)) we mark it as maxima and remove that region vectors from feature vectors list and start to find a local maxima again from any random mean vector

```
Remaing feature vectore size
221952
remove region pixel count
5724
shift dist
20.3648685459
end
Remaing feature vectore size
221952
remove region pixel count
6950
shift dist
9.85350113844
end
Remaing feature vectore size
221952
remove region pixel count
7496
shift dist
5.09712742557
end
Remaing feature vectore size
221952
remove region pixel count
7781
shift dist
2.70824988878
Low shift/ Local Maxima  - Remove this region of convergence
end
Remaing feature vectore size
214171
```

1) **RGB Image: Butterfly**

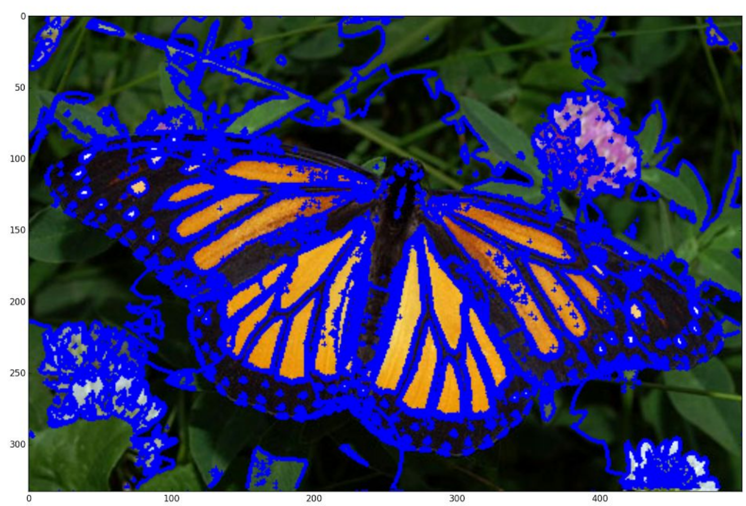
Example1: Mean shift segment preserve filtering hs:40 hr:40 iter:5 merge hs:80 merge hr:70

Bottom 3 images in order

1) **Images with local maxima with regions of convergence having same RGB value as of local maxima**
2) **Regions merged**
3) **Regions segmented**

Regions after being merged

Regions Segmented



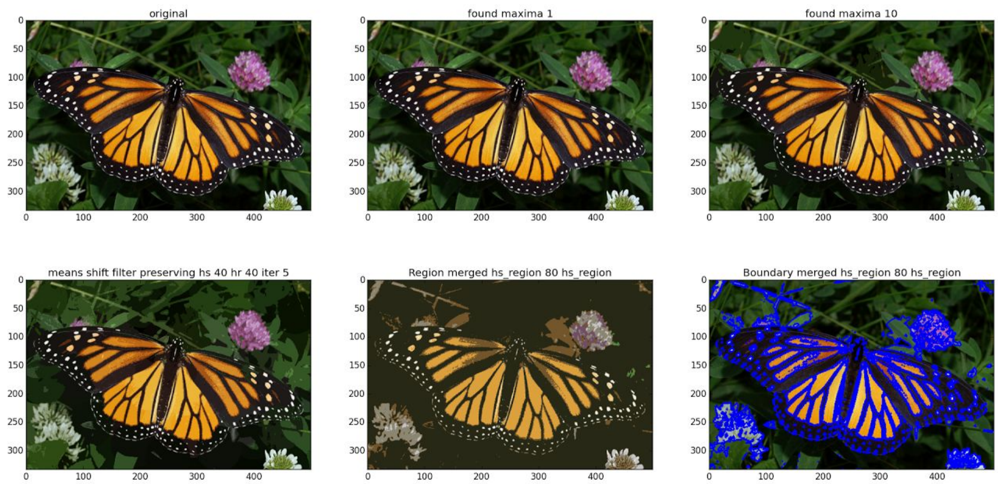
Example 2: Mean shift segment preserve filtertering hs:40 hr:40 iter:5 merge hs:80, No merge hr

Bottom 3 image in order

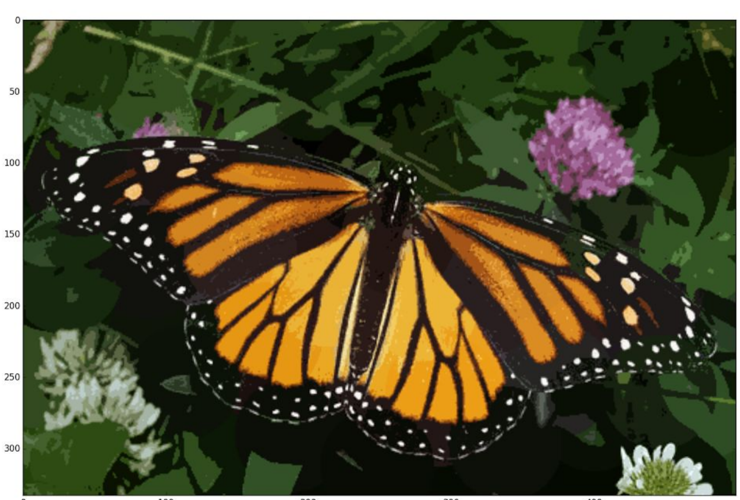
1)Images with local maxima with regions of convergence having same RGB value as of local maxima

2)Regions merged

3)Regions segmented

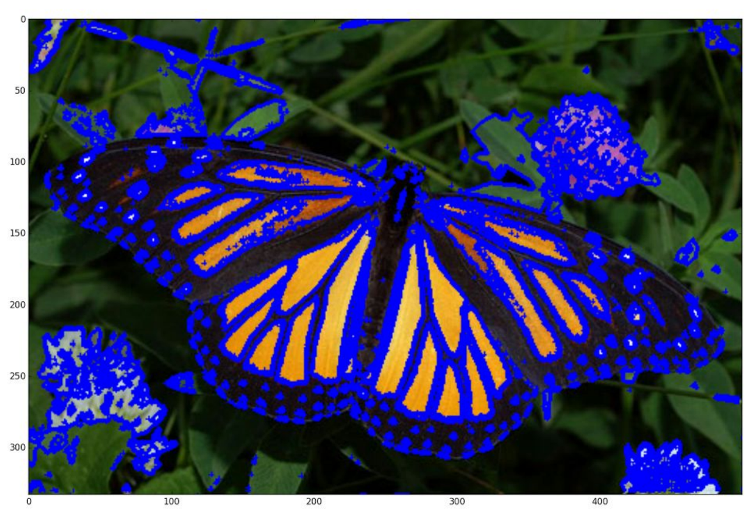


Regions with local maximas filled by RGB of mean value



Regions within hs of 80 are merged

Boundary of the regions marked



2) Gray Scale: House.jpg

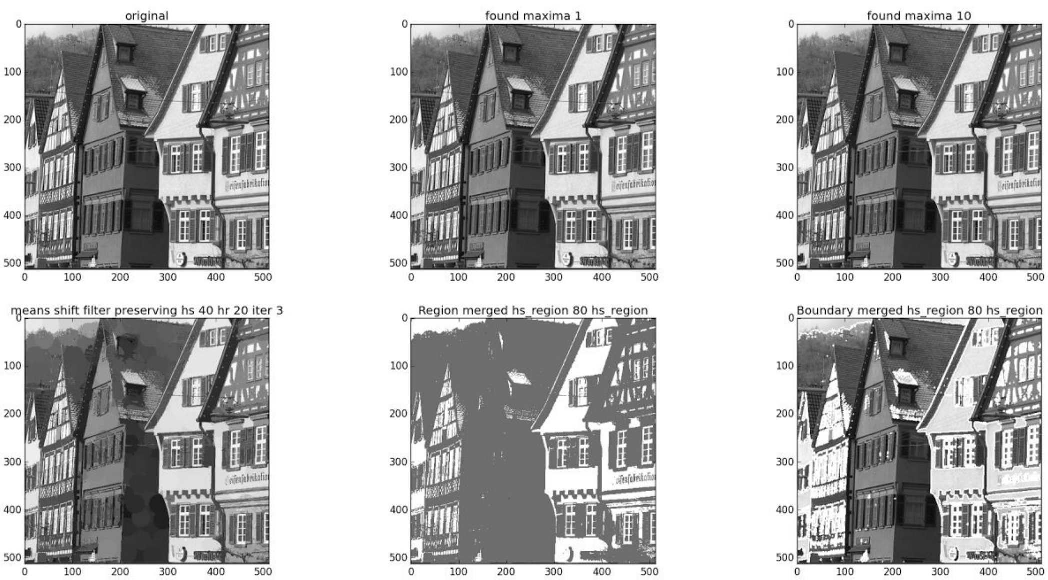
mean shift segment preserve filtertering hs:40 hr:20 iter:3
merge hs:80, No merge hr

Bottom 3 image in order

1)Images with local maxima with regions of convergence having same RGB value as of local maxima

2)Regions merged

3)Regions segmented



Local maximas image

Local maximas merged regions

Regions boundary marked

3)Gray Scale: Hill House.jpg

mean shift segment preserve filtertering hs:40 hr:20 iter:3

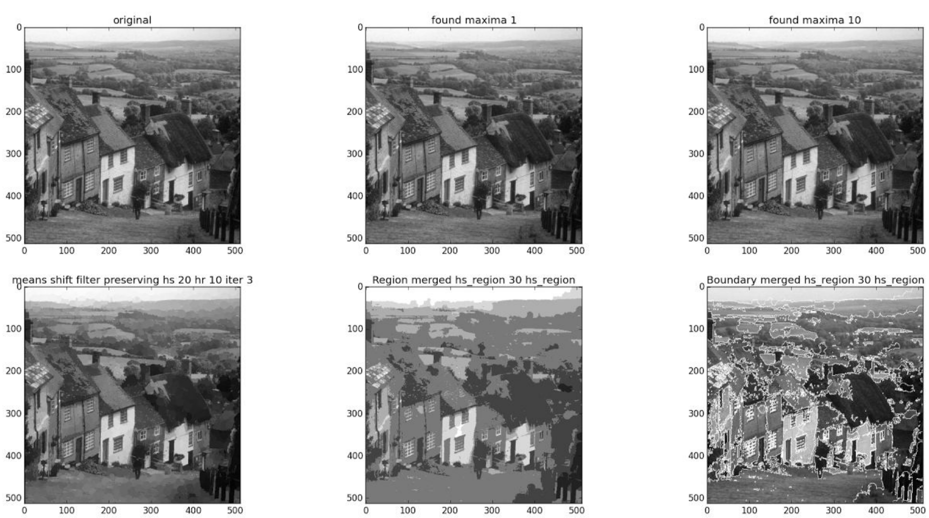
merge hs:30, No merge hr

Bottom 3 image in order

1)Images with local maxima with regions of convergence having same RGB value as of local maxima

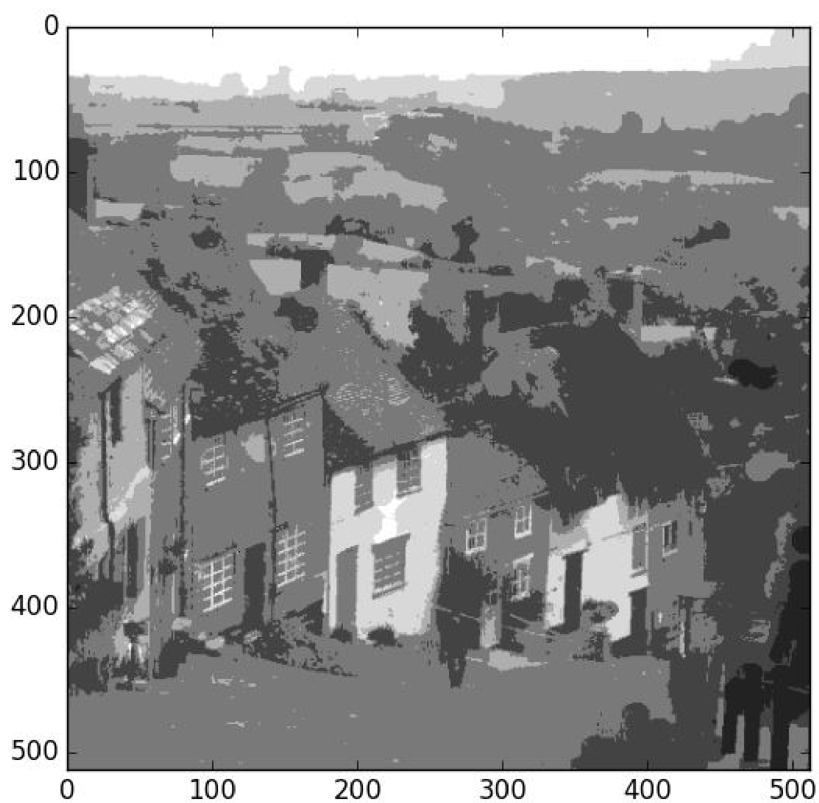
2)Regions merged

3)Regions segmented



Local maximas

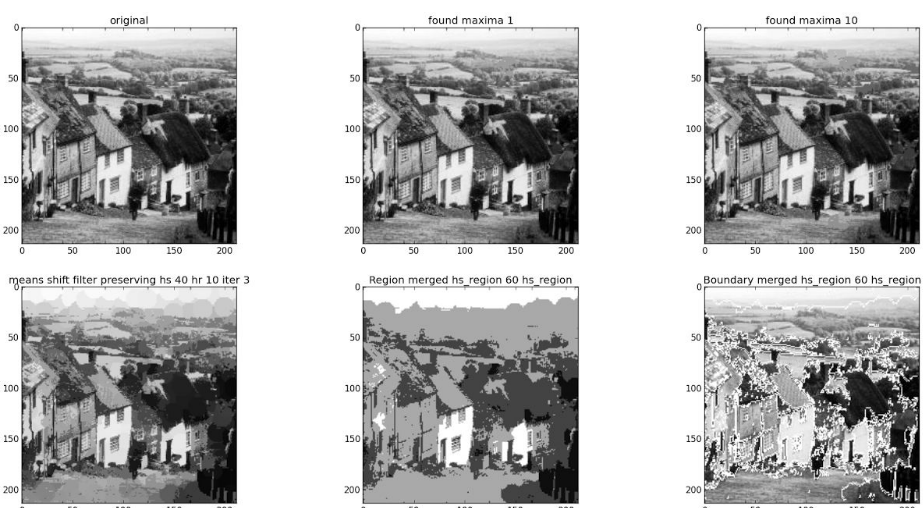
Maximas merged



Segment boundaries

Experiment HIll House Image with Histogram Equalization used below

hs for segmentation value was doubled from 20 to 40 as well as for region merging to achieve the similar result



4) Explanation of Software and Program Development

Libraries Used in the Project

import numpy as np

For all array data structures storing vectors also including mask and means over vertical axis

import matplotlib.pyplot as plt

To plot in Canopy plot window

from scipy.spatial import distance

For calculation euclidean distance

from scipy import misc

Reading of the images

from random import randrange

For random index selection of the vectors space

import cv2

cv2 line function in the bonus part

Summary of Lessons learnt from the project

Optimal Thresholding:

1) Manually selecting various thresholds and then a long wait for the program to run and give results.

2) Only one threshold for the full image didn't seem appropriate.

They should be adaptive as the image changes.

Time Complexity

1) Algorithm complexity and accuracy goes hand in hand.

- Local maximas can be separated on different features of an image within threshold
- Hs and Hr values depend on the contrast and size of the objects of the image.
- Iter is further dependent upon hr and hs selected. For small size window we need very small shift to confirm it is a local maxima.
- O(n2) complexity algorithm is not good to get results.
- Efficient data structure is required.

5) Summary and Discussion

Lessons learnt from this course

Topics Floating top of the mind

-How simple can be the image processing be in frequency domain by changing image from spatial domain to frequency domain for doing all the processing and then back to spacial domain for a user's view
-Playing with images by applying kernels for Smoothing, Edge Detections, Removing Noise etc
-Smoothing of images without losing boundaries
-Sharpening of edges
-Identifying corners
-Compress and store images in data structures
-Erosion and Dilation (Opening Closing: New ways to calculate edges and removal of noise)
-Line thinning
-Edges from Regions and Regions from Edges(Removing weak edges and joining strong edges).
-Global or experimental knowledge necessary for object identification and optimal thresholding
-Object Identification: Image as graphical view with nodes and edges
-Variation from the graphical mean graphs to train and identify facial expressions
-New way to look at the image and find regions by applying semantic knowledge and using adjacency list and probability of objects within an image. Or Probability of a scene with related objects.

The lessons learnt from the course gives me a start to proceed in a direction of any project I would be making in the future. Earlier I had no idea as to how an image should be stored using proper data structures, cleaned of regular noises, segmentations, boundaries, regions. There are different various ways to calculate each. Some requiring Derivatives and fourier transform etc while some require online linear kernel with hit or miss. In the last chapter I learnt to categorize human face image with emotions This information will be helpful to carry out some projects of training models in machine learning I have planned to take next semester.

6) Bonus Presentation and Discussion

The basic idea to track an object over image sequence.

Steps

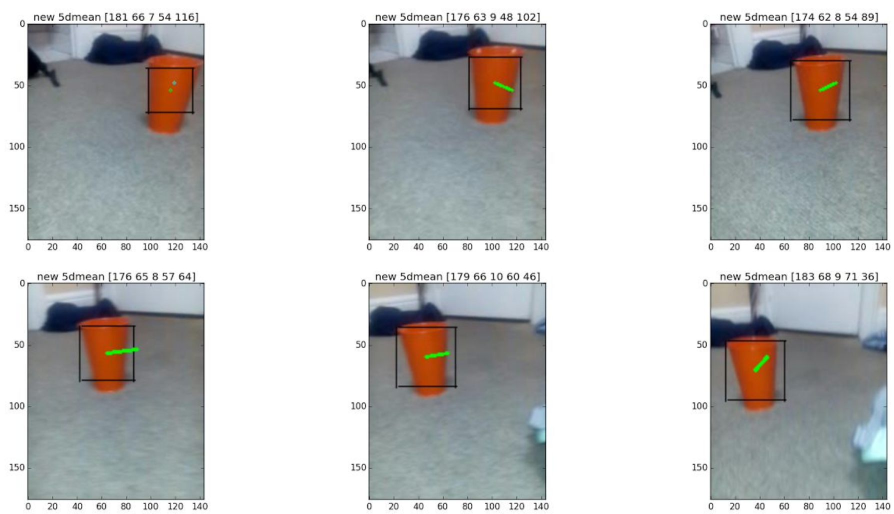
Instead of random point we need to specify an initialization point so that it is considered as a mean around which window using threshold is formed.

As discussed in the algo above the first iteration is kept in never ending loop and the image with that window is shifted as the object in the image shift/convergence point/centroid shifts.

Example 1)

***Square** is formed around new mean / centroid calculated.*

Line marked between old mean and new centroid



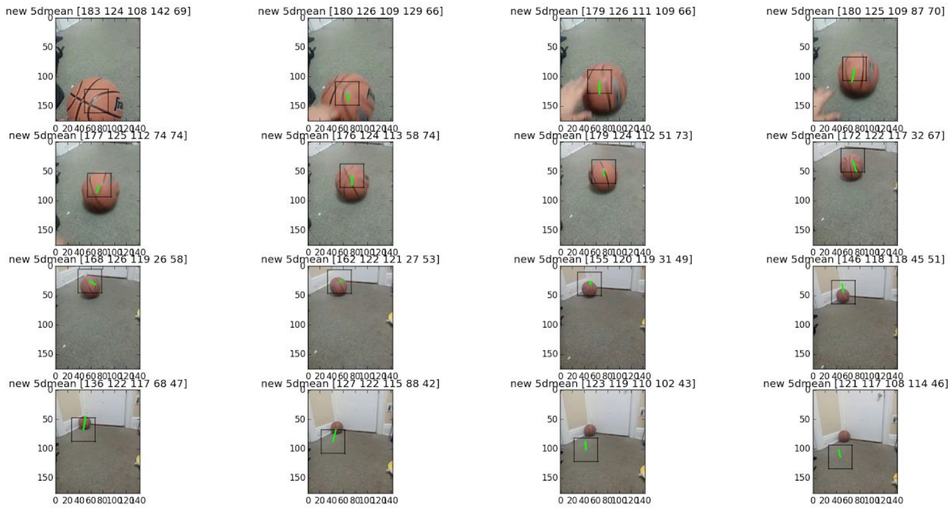
Above window showing shift of points

Example2:

Example of a lesson learnt

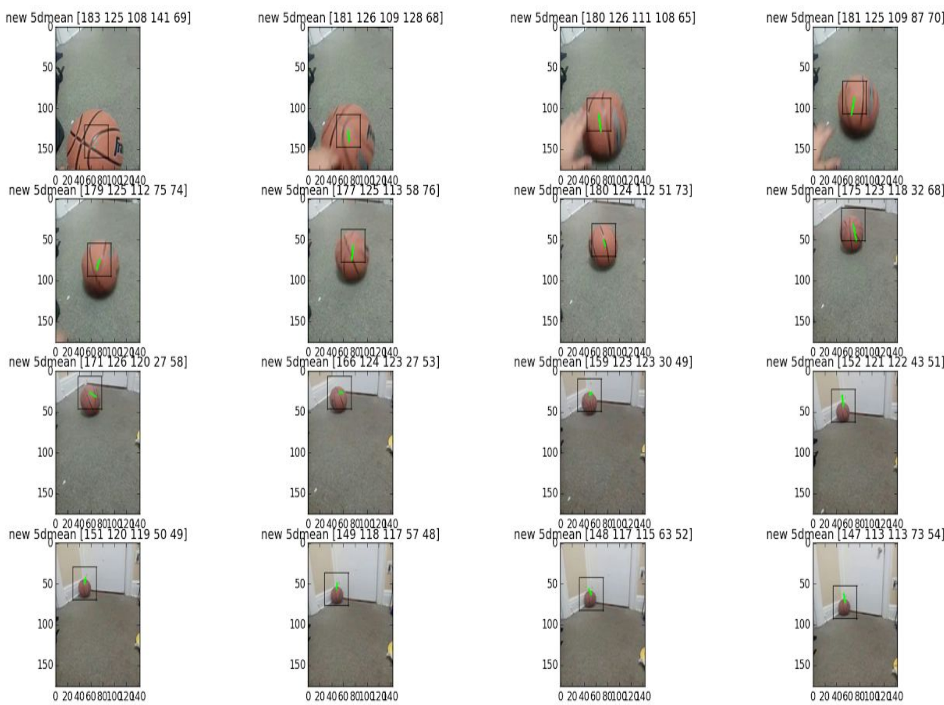
Thresholds hs 20 hr 40

After a while when the ball moves away it becomes darker and its RGB match comes close to the floor RGB. Therefore in the 13th image we lost the centroid.



Example2: Improved

Now hs was taken as 15 and we are able to separate floor with fall in 13th image as well



Results window Shift of Object/ centroid points

start_mean is the initialization point/mean

Calculated mean is the centroid for the window

mean for the new window will work same as the start mean for next iteration

```
Python
In [4]: %run "C:/Septermber2015/cvip/mean/vid.py"
1
Start_mean
[180 128 106 148  66]
Calculated mean
[ 183.14690909  124.35563636  107.59490909  141.58109091   69.16072727]
mean for new window
[ 183.14690909  124.35563636  107.59490909  141.58109091   69.16072727]
2
Calculated mean
[ 180.47452229  125.54564756  109.25796178  128.99522293   66.10828025]
mean for new window
[ 180.47452229  125.54564756  109.25796178  128.99522293   66.10828025]
3
Calculated mean
[ 178.98588235  125.58352941  111.44588235  108.87764706   65.53529412]
mean for new window
[ 178.98588235  125.58352941  111.44588235  108.87764706   65.53529412]
4
Calculated mean
[ 180.49211909  124.86164623  109.22591944   87.13485114   69.61471103]
mean for new window
[ 180.49211909  124.86164623  109.22591944   87.13485114   69.61471103]
5
Calculated mean
[ 177.31929825  124.82982456  112.28947368   74.34561404   74.27719298]
mean for new window
[ 177.31929825  124.82982456  112.28947368   74.34561404   74.27719298]
6
Calculated mean
[ 175.74225122  124.28548124  112.90212072   58.39641109   74.32137031]
```

Lessons learnt from the bonus part

Constraint

Paper reference [5]

The best results can be achieved if the following conditions are fulfilled:

- The target object is mainly composed of one color.
- The target object does not change its color.
- Illumination does not change dramatically.

• There are no other objects in the scene similar to the target object
• The color of the background differs from the target object.
• There is no full occlusion of the target object.

Features of the object should not vary very fast.
Sudden shift in object position can still be tracked back if no other object with similar vector is available but a sudden shift in the object color is impossible to track back after the track is lost

Issue
The frame rate of the image may become low or there can be fast change in object position and object is not found within the threshold window
Resolution:
Object with previous threshold is found again if lost.
You may end up getting another object with similar features mean

Future scope of improvement
hs and hr value should become **adaptive** as the object changes its position its size relative to the image as well as it color due to illumination changes.

7)Reference List:

(Lecture slides, Image analysis and computer vision by milan sonka, Youtube videos and reference papers in the project guidelines)

[1]

Y. Cheng, “Mean shift, mode seeking, and clustering,” IEEE Tans. Pattern Analysis and Machine Intelligence, Vol. 17, No. 8, pp. 790-799, August 1995

Further queries : To reduce the complexity from O(n2) of each iteration to achieve real time processing.

[2]

D. Comaniciu and P. Meer, “Robust analysis of feature spaces: color image segmentation,” Proc. IEEE Computer Vision and Pattern Recognition, pp. 750-755, June 1997

Further queries : What possibly can be feature vectors instead of just RGB and xy coordinate. Example: some gradient information into feature vector.

[3]

D. Comaniciu, V. Ramesh and P. Meer, “Real-time tracking of nonrigid objects using mean shift,” Proc. IEEE Computer Vision and Pattern Recognition, pp. 142-149, June 2000

Further queries: a)How to predict the new location of the object and taking that into account for the mean shift calculation while selecting hr value if any.

b)Instead of loading full image into feature vectors how we take only the prediction window within an image for our hs and hr window

[4]

False-Peaks-Avoiding Mean Shift Method for Unsupervised Peak-Valley Sliding Image Segmentation *(For understanding purpose No implementation in the project)*

[5]

A Comparison of Mean Shift Tracking Methods∗ Nicole M. Artner

Youtube videos

https://www.youtube.com/watch?v=RG5uV_h50b0
https://www.youtube.com/watch?v=74GnbHe7yeY

**There can be some shift from the paper algorithms to the interpretation of the project development*

* Not every part of the papers were read thoroughly.