CV Assignment3

Kshitiz Jain 2016051

March 2019

1   Question 1

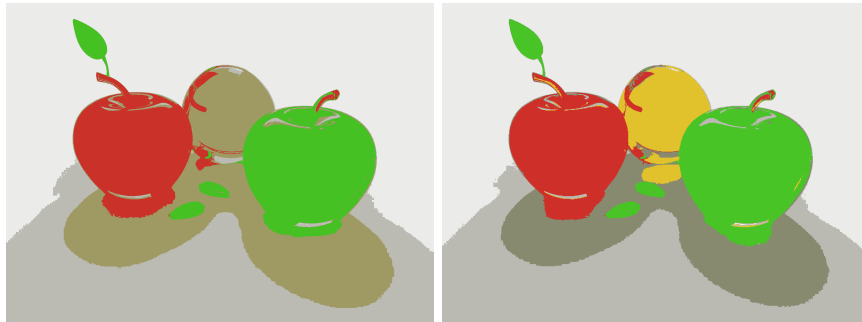Different type of Clustering is performed in the following images :

Similar number of clusters are used in both 3D and 5D clustering so as to compare their results.
Different values of clusters are used for different images so as to variation when number of clusters changes.
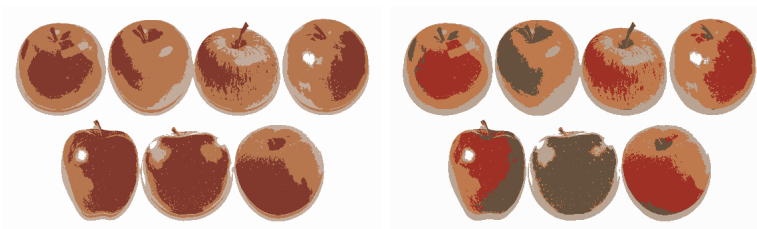
1.1 3D Clusters
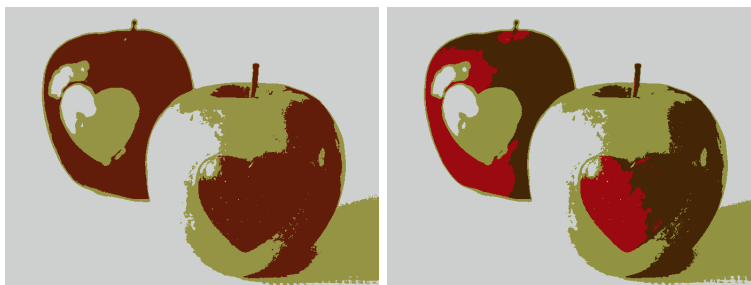
With clusters 5 and 6 respectively.

With clusters 3 and 4 respectively.

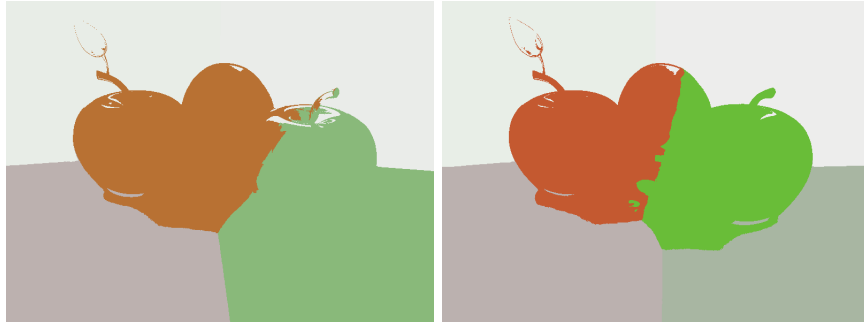With clusters 4 and 5 respectively.
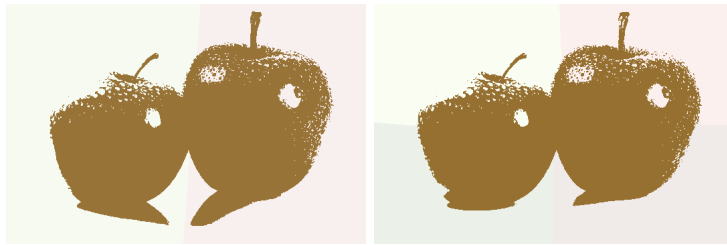

With clusters 3 and 4 respectively.
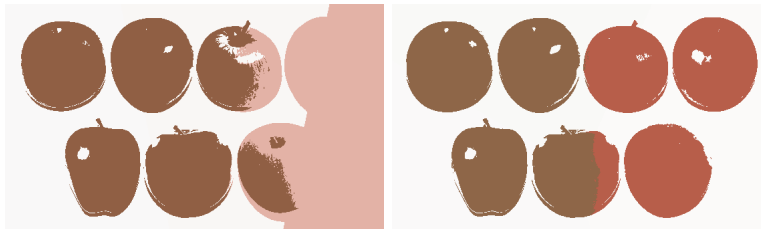
1.2   5D Clusters
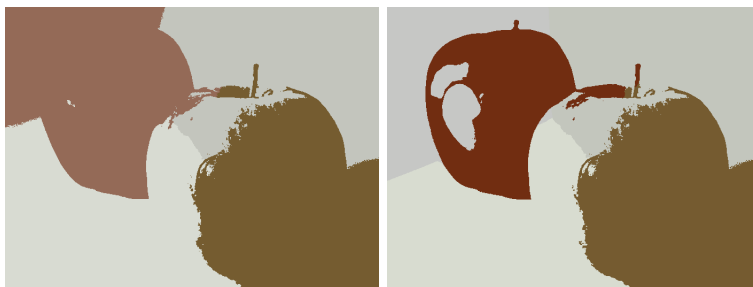
With clusters 5 and 6 respectively.

With clusters 3 and 5 respectively.

With clusters 4 and 5 respectively.



With clusters 4 and 5 respectively.

It is observed that the colors/contrast of clusters have different values for 3D and 5D clustering. In 3D clustering the color closely resembles the original colors of the image as compared to the images from 5D clustering.

This is because in 5D clustering the values are also influenced by the pixel position in the image, whereas in 3D clustering only the colors are clustered and hence preserves more original colors.

In 5D clustering spatial information is also incurred, hence we can see in 3rd image we get whole apples as clusters whereas in 3D clustering only the colors are clustered hence for each cluster it can be observed that there are patches of the same cluster all over the image.

In 5D clustering due to addition of spatial information it happens that the background are separated by different clusters.
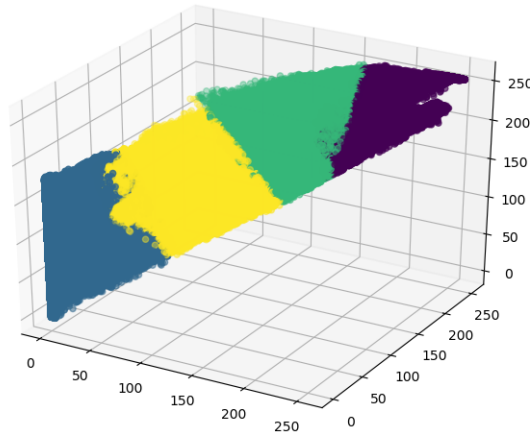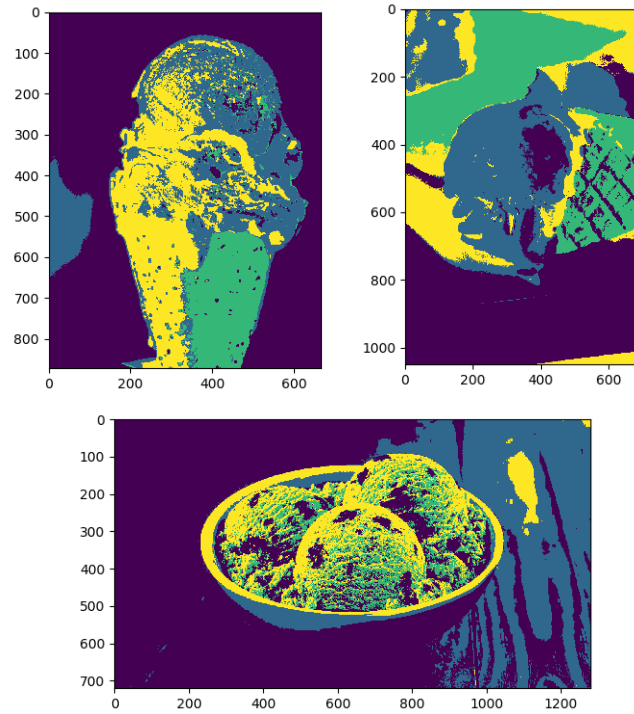

2   Question 2

Mean Shift clustering is done on the images given. Mean shift clustering Python library SKlearn is used in the process. Parameters for the respective library are the following :

- Bandwidth : Bandwidth represent the size of the circle/window in which the algorithm proceeds.

- Quantile : Quantile helps in smoothening the image. Hence it can be used to increase/decrease the number of clusters, lesser the value of quantile greater are the number of clusters.

- Number of samples : Number of samples taken friom image.

- Binary seeding : As the name suggest, binary seeding helps us get minimum number of clusters possible and hence speeding up the process.

For the following images quantile is set to be 0.15 and number of samples are 300.

3   Question 3

3.1   Skin Color Thresholding

The images were converted in the HUE space and then thresholding on all the pixels are done. That is if the pixel color is in the range then is considered white else black.

- Lower Threshold = (0,40,80)

- Upper Threshold = (25,250,250)

3.2   Seeding Segmentation

Queue based seeded segmentation is done, the neighbour pixel is considered in related to seeded region if the distance between their pixel values and distance between the average pixel value to the neighbour pixel is less than a threshold.

Different distance methods are implemented in the code, but the images shown are calculated using Manhattan distance. For every image seed in manually chosen so as it is in the skin region of the face.

4   Question 4

Classification of CIFAR dataset is proceeded as follows :

1. Training dataset is divided into 2 random equal parts with 25 thousand images each. One set is used for creating bag of visual words and second dataset is used for training of the classifier. The testing batch of the dataset is reserved for testing.

2. Creation of Bag of Words : First the each image (32x32) is divided into 8x8 patches, giving 16 patches from each image. For each patch we create histogram features. Now the obtained features are clustered into 10 clusters using KMeans Clustering. Now we have created the BOW of size 10 by using centroids of each cluster.

3. Training : Similar as above first each image is divided into 16 patches and for each patch histogram feature is generated. The obtained 16 features are now mapped with the centroids (BOWs) and a new feature vector is created corresponding to the number of times a centroid selected. Note - Each index of new feature vector represents a centroid. So if the 1st patch is nearest to the 2nd centroid/word the value at 2nd index is incremented by 1.

4. Testing : Similar to training we calculate feature vector of each testing image. Now the training features vectors are passed in the inbuilt classifier and testing feature vectors are predicted and accuracy is repoerted.
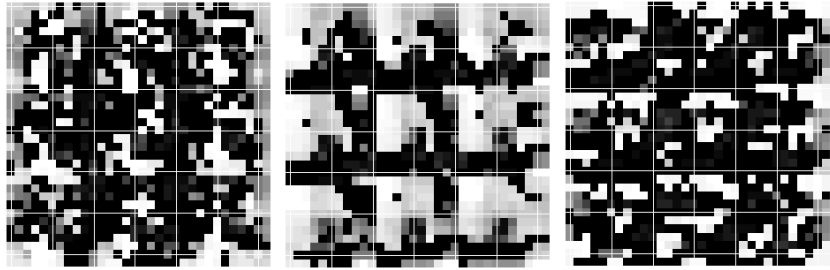
4.1   HOG Features

- Trianing Accuracy : 29.64%

- Testing Accuracy : 16.42%
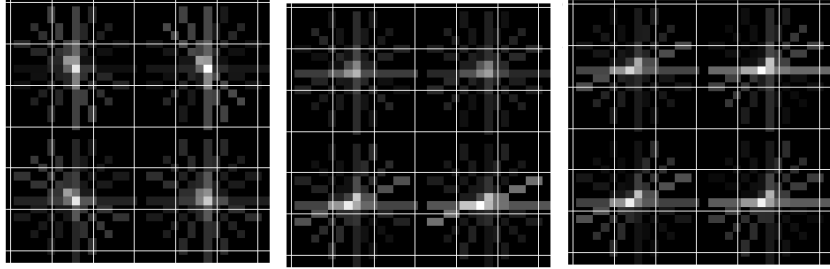
4.2   LBP Features

- Trianing Accuracy : 24.69%

- Testing Accuracy : 18.57%

LBP features are calculated by thresholding of pixels with respect to their neighbour pixel values, whereas HOG features are calculated using the gradient (both magnitude and direction) features of the image. Since there is information loss due to thresholding hence less accuracy is observed in case of LBP features.