Part 3: Scene recognition with bag of words

Date: 2019-6-27

All the extra functions are in the /ec.

- Q1.1 Extract Filter Responses

1   2

3   4

For the filters of Laplacian of Gaussian(image 3), the pixels are zero when they are away from the edge, and on the edge itself, pixels are also zero, and obviously this filter can be used to detect edge.

For the X(Y) Gradient of Gaussian, we can see that it can responds strongly to vertical(horizontal) edges, because only the pixels that have large difference at X(Y) direction can be observed.

CIE Lab color space:

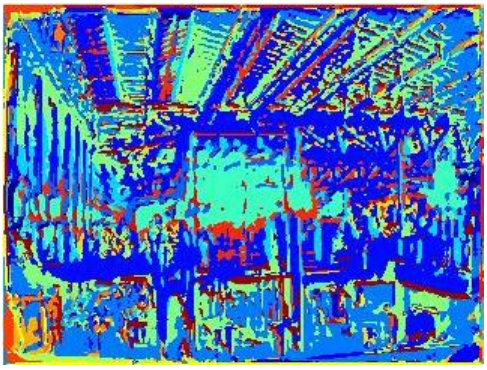
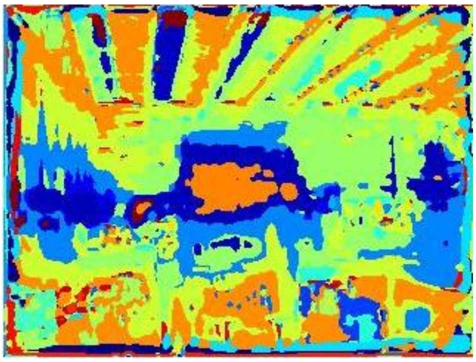
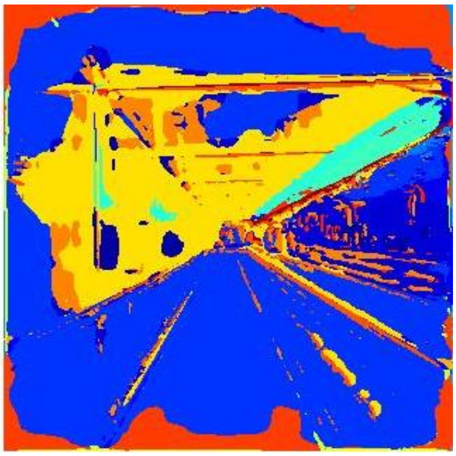
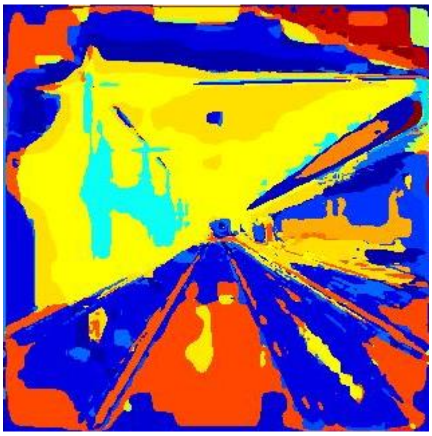
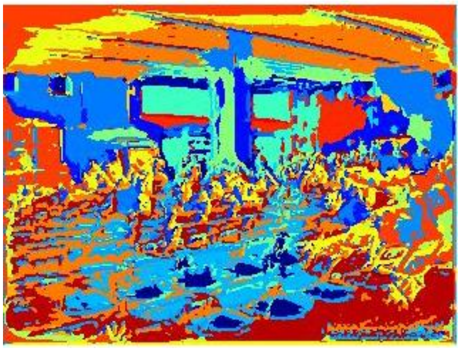
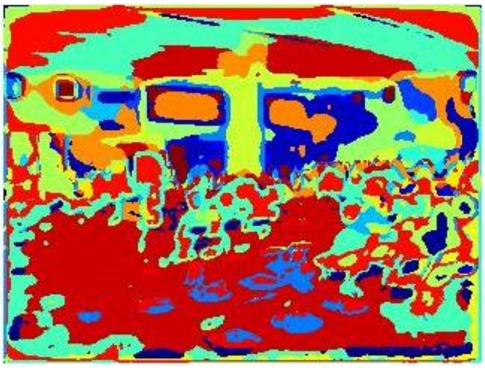
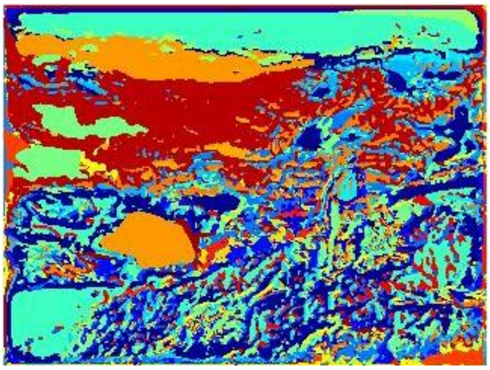
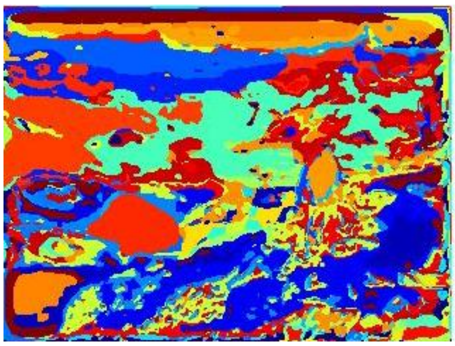
This color space compensates for the drawback of RGB color mode, it is device-independent, and based on physiological features. L (light) and two color channels(a,b)

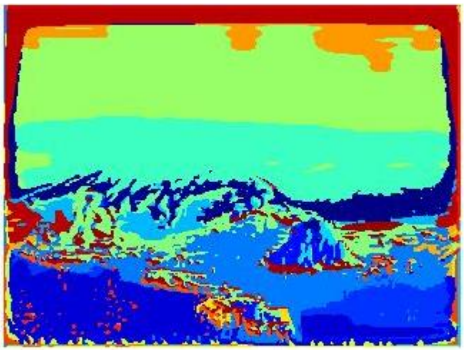
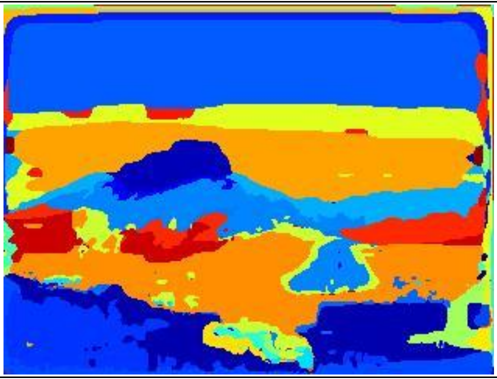
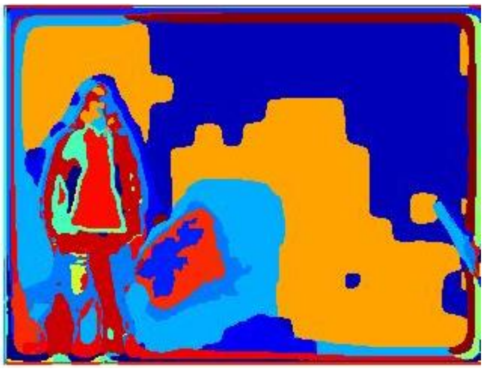
- Q1.2 Collect sample of points from image

Random Harris

- ## Q2.1 Convert image to word map

In this section, since compute the pdist2 of each pixels with each words cost too much time, I used the matrix to compute Euclidean distance, and it is more effectively, only has the time complexity O(K).

class	Harris	Random
airport		
airport		
airport		
landscape		

landscape		
landscape		

the visual words capture semantic meanings, because we can see that in the result images, the pixels that need to be sorted as same classes are connected to same words, so that the words can correctly sorts the pixels. The harris dictionary is better, because in the result of harris dictionary, we can see that the classifications is more detailed(more details can be observed in the result images). Because Harris corner means the edges of the objects and the corners of the objects, so that it do better in express features. But when we choose the points randomly, the points probably be the ones that in the “flat area”, and some pixels that is not related to the word may be mapped to it .

- ## Q3.2 Evaluate Recognition System - NN and kNN

  ## NN:

  In this part, I saved all the confusion metrics and the confusion matrixes in the Random_euclidean_NN.mat, Harris_euclidean_NN.mat, Random_chi2_NN.mat, \Harris_chi2_NN.mat, and output them in the main3_2.m.

Random - Euclidean

```
Random_euclidean
metric =

    0.4188


c =

     9     2     4     0     0     1     0     4
     5     9     4     0     1     1     0     0
     6     3     8     0     3     0     0     0
     2     0     4     7     2     1     4     0
     0     4     2     2     8     1     3     0
     2     2     1     2     1     5     3     4
     4     1     3     2     1     0     7     2
     3     0     1     2     0     0     0    14
```

Random- chi2

```
Random_chi2
metric =

    0.5250


c =

    12     2     4     0     0     0     0     2
     3    12     3     0     1     1     0     0
     7     3    10     0     0     0     0     0
     1     1     3     9     1     0     5     0
     0     3     2     1    12     0     2     0
     2     2     3     0     1     6     2     4
     5     0     3     2     0     0     9     1
     3     0     0     3     0     0     0    14
```

Harris - Euclidean

```
Harris_euclidean
metric =

    0.4188


c =

     9     2     2     2     0     0     3     2
     4    10     2     0     1     1     2     0
     5     4     7     1     3     0     0     0
     1     2     0    10     2     2     0     3
     1     2     3     2     8     1     3     0
     1     4     3     3     0     4     1     4
     2     2     1     4     4     0     6     1
     5     1     0     1     0     0     0    13
```

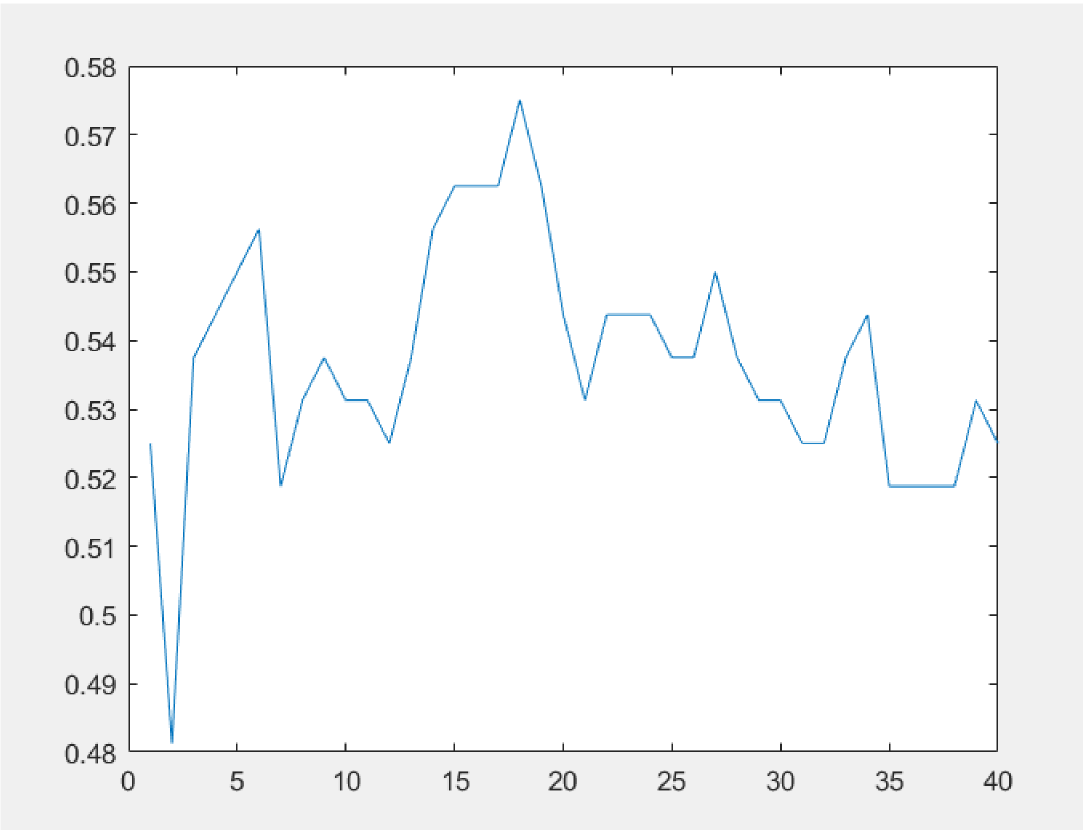
Harris - chi2

```
Harris_chi2
metric =

    0.4500


C =

     8     2     4     2     0     0     2     2
     6    10     2     0     1     1     0     0
     4     4     9     0     3     0     0     0
     1     2     1     9     2     3     2     0
     1     1     3     0    11     0     4     0
     2     3     2     1     1     5     3     3
     1     0     2     6     3     0     6     2
     3     0     1     2     0     0     0    14
```

From the results, we can see that the performances of the two dictionaries have a little difference, it is not surprising, because the Harris corner detection can detect the harris corner that have nothing to do with the rotation of the corner.

And the chi2 distance performs better, because the chi2 distance is more insensitive to outliers, some points that have negative effect on the result can be ignored by the chi2 distance.

KNN: (Random- chi2)



Confusion matrix:

```
ans =

    15     1     2     0     0     0     0     2
     3    12     4     0     0     0     1     0
     3     3    13     0     1     0     0     0
     1     1     2     7     1     0     6     2
     0     3     1     0    15     0     1     0
     2     3     2     3     2     4     2     2
     6     0     3     0     2     0     9     0
     2     0     1     0     0     0     0    17
```

When k=18, the maximum accuracy is 0.5750, from the plot of the accuracies from k=1 to k=40, larger k is not always better,   because when k is too large, the more likely the image is to be classified into the wrong classification(for the more nearby points to be considered)