Abhishek Bamotra
Andrew ID: abamotra

Homework 1

16-720 A
February 7, 2020

Problem 1

Q 1.1.1

Gaussian:

These filters removes high frequencies from the image and let pass the low frequencies resulting in smoothening/blurring of the image. Therefore, these filters are also called low-pass filters/ smoothening filters.

Laplacian of Gaussian:

These filters are used to detect the change in intensities across the image. Therefore, it gives value only where the intensity is changing and zero for areas with constant intensity.

Derivative of Gaussian in x direction:

These filters are used to detect vertical edges in the image.

Derivative of Gaussian in y direction:

These filters are used to detect horizontal edges in the image.

Why do we need multiple scales of filter responses?

We need multiple scales of the filter to see the different responses of the images with a filter with different sigma values. Sigma value decides the level of smoothening by the filter. The change in sigma value changes smoothening for all the filters as all the above listed filters are some how derived from Gaussian filter which is also known as smoothening filter. Also, Smoothening is directly proportional to the sigma value.

Q 1.1.2

In figure 1, the responses of the original image, shown in Figure 1 (top), are shown as collage. The responses are for four different filters: (1) Gaussian, (2) Laplacian of Gaussian, (3) derivative of Gaussian in the x direction, and (4) derivative of Gaussian in the y direction.

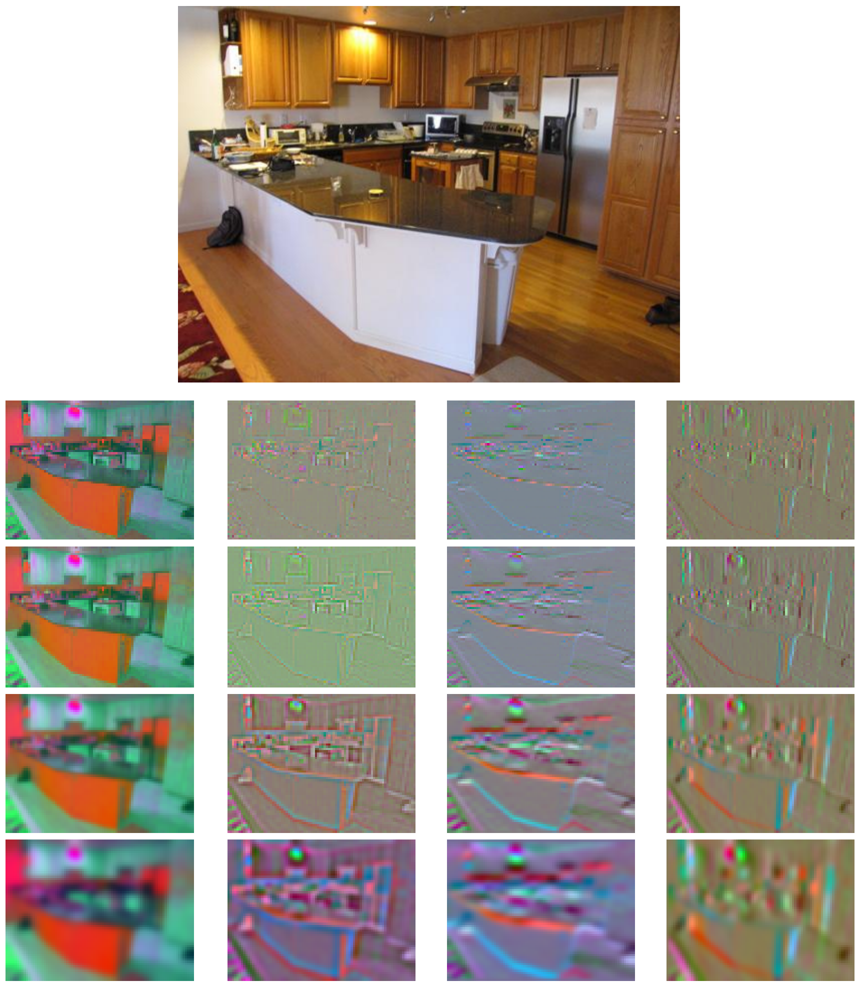
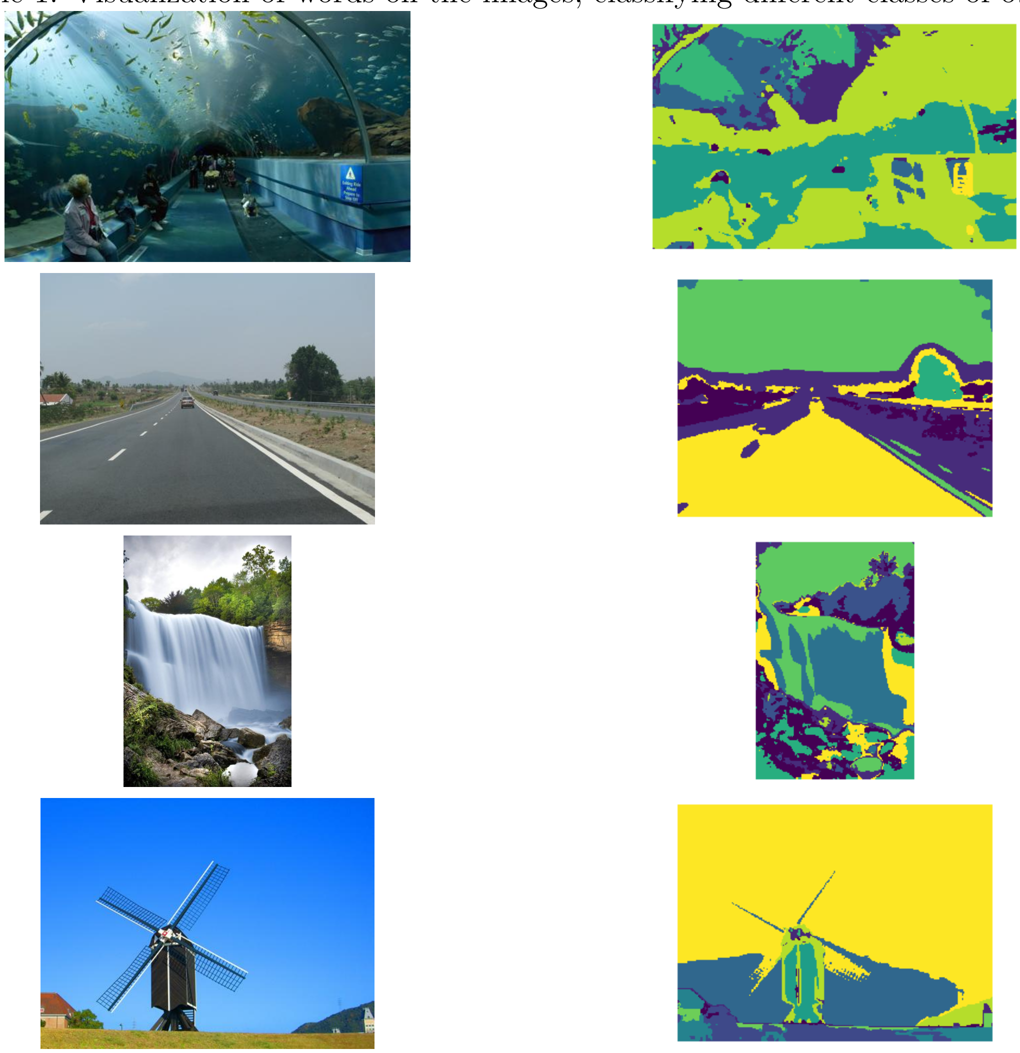


Figure 1: An input image and filter responses for all of the filters in the filter bank. (a) The input image (b) The filter responses in Lab colorization, corresponding to the filters

Q 1.3

The visualization of the word map of 4 different images are shown below in Table 1. The left column of the table 1 shows the original input images and the visualization of the ”word” is shown in right column. The word boundaries make sense, as in the image of the highway and waterfall make it easier to identify the boundaries when compared to the original image. It is clear in the highway image, how road, tree, side path and even white marks on the road are classified as different class. The algorithm is able to cluster similar pixels together into one class. Similar results can be seen in other images shown.

Table 1: Visualization of words on the images, classifying different classes of object.



Problem 2

Q 2.5

Confusion matrix for default parameters as: alpha = 25; L = 1; K = 10;

```
[[29.  0.  4.  1.  2.  1.  5.  8.]
 [ 1. 19.  5. 15.  3.  1.  1.  5.]
 [ 1.  3. 17.  0.  4.  1. 12. 12.]
 [ 4.  1.  0. 33.  7.  1.  3.  1.]
 [ 2.  2.  2. 15. 24.  1.  4.  0.]
 [ 1.  0.  4.  3.  4. 28.  7.  3.]
 [ 5.  1.  7.  3.  6.  5. 21.  2.]
 [ 1.  2.  8.  1.  4.  4.  6. 24.]]
```

Accuracy of the Algorithm = **48.75 %**

Q 2.6

There are various classes in the data-set which are very often miss classified by the system. Some of the difficult classes to classify are: (Kitchen, Laundromat), (Highway, Windmill) and (Highway, Waterfall). We can confirm these pairs from the confusion matrix above.

(Kitchen, Laundromat) $\longrightarrow$ **Figure 2**

Figure 2: Similarity between Kitchen and Laundromat Images

(Highway, Windmill) $\longrightarrow$ **Figure 3**

Figure 3: Similarity between Highway and Windmill Images

(Highway, Waterfall) $\longrightarrow$ **Figure 4**

Figure 4: Similarity between Highway and Waterfall Images

These classes show a lot of difficulty in classification as these pairs of classes contains images which contain similar data which makes the clustering difficult. Also, the K-means used in this case has clusters set to 10 which is very low and results in generalization of some data information. Also, as we are considering alpha as 25, I think the 25 pixels are less to classify the image. Therefore, misclassifies some classes.

Problem 3

Q 3.1

Run no.	Scale	L	K	alpha	Accuracy (%)
1	[1, 2, 5, 9.81]	1	10	50	47.75
2	[1, 2, 5, 9.81]	3	10	50	57.5
3	[1, 2, 5, 9.81]	3	100	50	***65.5***

Run 1:

From the alpha value of 25, I thought the only 25 pixels are low to actually classify the image as it passes a very small information data so, I increased the alpha value to 50. Accuracy was 47.75 which was lower then the default.

Run 2:

From Run 1, the accuracy dropped, So, to improve it I thought the importance should be given to spatial structures of the image. I increased the number of layers, L, to 3. The accuracy shooted to 57.5.

Run 3:

The increase in accuracy in Run 2 was significant, and to further it I figured maybe the number of features i.e K is low. So, I increased K to 100. and boom... accuracy came out to be 65.5. The increase in number of features for the training helped in improving the model and resulted in better accuracy.

On a side note:

As k-means takes different initialization of the clusters on each run, the results might vary a bit.

Q 3.2

Q. What you did?

For improvement to the algorithm, I think improving the weight distribution of the SPM will improve accuracy of the model. Because I think giving more weight to the spatical structures will give more importance to the detailed information in the image and therefore adding with other layers will result in higher performance.

Currently, for L = 2 i.e 3 levels (0, 1, 2), the weights are 1/4 for L = 0,1 and 1/2 for L = 2.

Changing to, for L = 2 i.e 3 levels (0, 1, 2), the weights are 1/8 for L = 0,1 and 3/4 for L = 2.

Q. What you expected would happen?

I expect changing the weights of the spatial pyramid will give more importance to the finest layer of the pyramid. This change will help in improving the accuracy as this will make a better classification model.

Q. What actually happened?

The accuracy dropped as against the expectation. Accuracy comes out to 62.75. Visual response for an image is shown in the figure below:

Figure 5: Visualization of the words of the image

I think the following can be the possible reasons of drop:

1. I have assigned less weights to the 0th and 1st level of the pyramid.
2. There should have been other combination of weights to make it perform better.