16720: Introduction to Computer Vision
HW1

Saudamini Ghatge
sghatge

September 19, 2024

Q1.1.1

Properties of Filter Functions:

- **Gaussian Filters:** It smoothes the image by the same amount, in any direction we want which helps in removinng noise from the image
- **Laplacian of Gaussian:** This is a second-order derivative of the Gaussian filter. It highlights areas of rapid intensity change i.e. edges.
- **Derivative of Gaussain:** These are the first derivative of Gaussian filter. They are used to detect horizontal and vertical edges by applying in either x or y direction.

Multiple Scales of filter responses:

There are required because different features in an image are derived from different filter responses by changing the scale of the filters. Bigger values of sigma give high-level information, i.e. the image will be more blurred. Smaller values of sigma will give finer information.

Q1.1.2

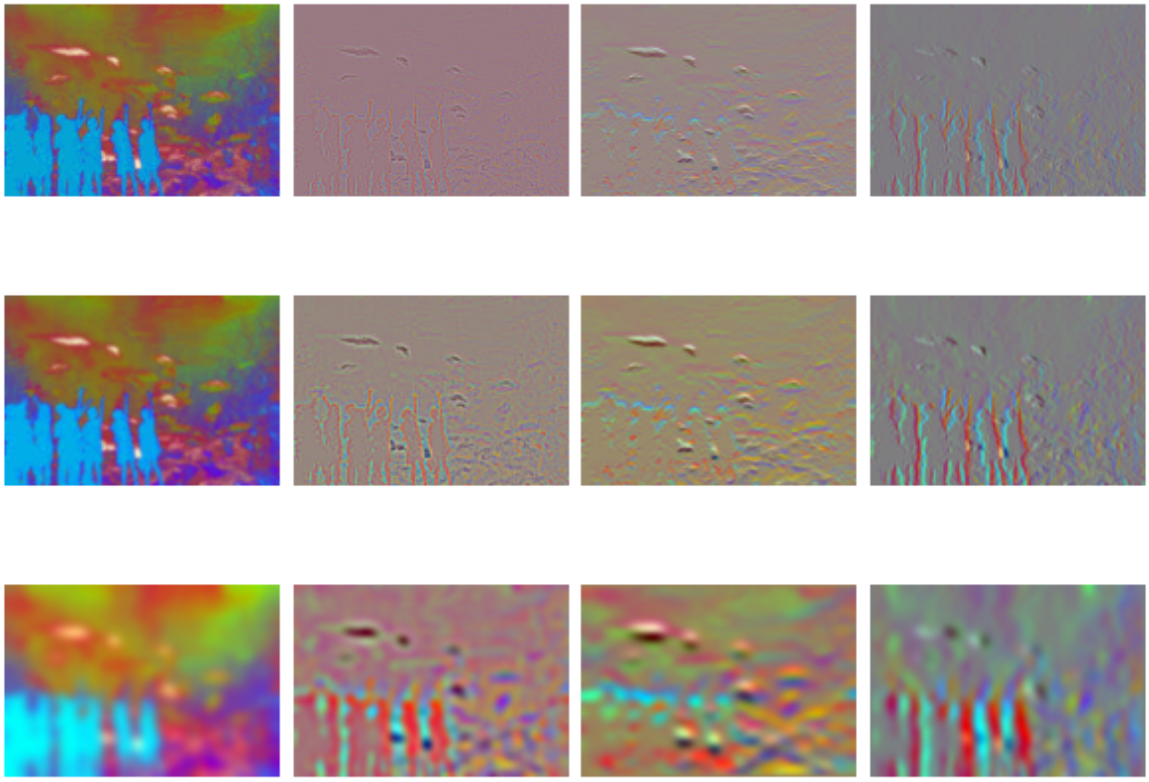

Figure 1: Visualization of filter responses.

Filter Scales used: 1, 2, 8

Q1.3

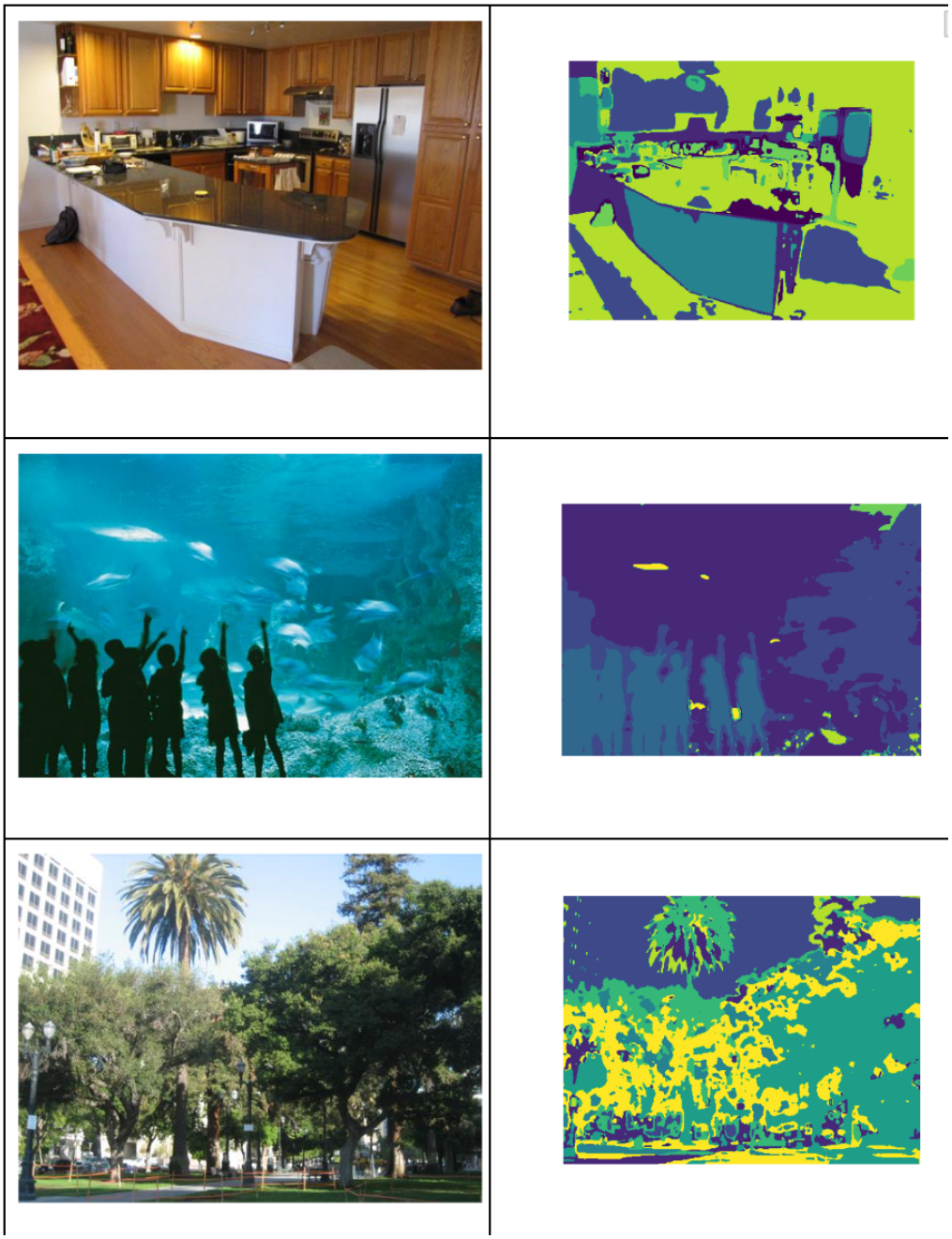

Figure 2: Wordmap

The word boundaries do make sense to me for the second and third image where I can identify kitchen top in the first image, people in the second image and trees in the third image.

Q2.5

The Confusion Matrix is -

```
[[33.  0.  2.  0.  2.  2.  3.  8.]
 [ 1. 25.  6.  9.  2.  0.  3.  4.]
 [ 3.  4. 31.  0.  1.  2.  1.  8.]
 [ 2.  3.  0. 35.  9.  1.  0.  0.]
 [ 1.  1.  2. 17. 16.  4.  5.  4.]
 [ 4.  2.  5.  3.  1. 31.  2.  2.]
 [ 5.  2.  1.  0.  4. 11. 25.  2.]
 [ 2.  3.  7.  2.  3.  7.  3. 23.]]
```

Accuracy: 0.5475 (54.75%)

Q2.6

Some classes/samples were difficult to classify by the model because of the scene containing similar features. In the first example of Kitchen and Laundromat, the two scenes look very similar and it is very easy for the model to confuse between the two. Likewise in the second example, both the scenes have water and greenery which can lead to the model getting confused.

If we compare kitchen with Park, the features are very different for both the scenes thus the model can recognise the scenes correctly.

Q3.1

filter_scales	K	alpha	L	accuracy
[1,2,8,10]	15	30	2	61.25%
[1,2]	5	20	4	45.75%
[1,2,8,10]	20	25	2	61.5%
[1,2,8,10]	10	35	4	50.75%
[1.2,8,10]	25	30	2	65.75%

With the default parameter values of - filter_scales:[1,2], K:10, alpha:25 and L:1 I got 54% accuracy.

1) I first decided to increase all the values of the parameters, which gave me 61.25% accuracy.

2) I then tried to decrease K and alpha as compared to the default parameters and increasing L while keeping filter scale same as default, I got very low accuracy which tells that the classification system does not improve only by splitting the images into more layers, there needs to be more features in the image.

3) Next I increased the number of filter scales to capture more features in the image, increasing K value to get more visual words, kept alpha same as default and number of layers as 2. This resulted in an accuracy of around 62%. This proves the above testcase that adding more features and visual words to the image improves the classifier.

4) I ran a new test case where I use four filter scales, good number of words, greater value of alpha and larger value of L. This resulted in lesser accuracy because the number of words formed are less as compared to the number of clustered pixels (alpha).

5) Lastly I changed K to be almost equal to alpha and around 66% accuracy.

So increasing both the K and alpha values to similar values gives more features to the system to classify from thus resulting in a good accuracy.