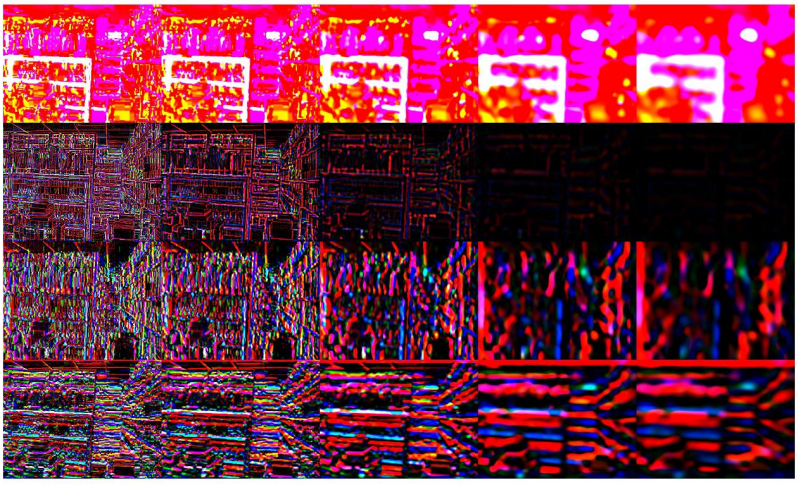
CSE 573 Computer Vision
Homework 1
Ting Zhou #50169020

Part1 Representing the World with Visual Words

Q1.0 What properties do each of the filter functions (see Figure 3) pick up? You should group the filters into broad categories (i.e., all the Gaussians). Answer in your write-up.

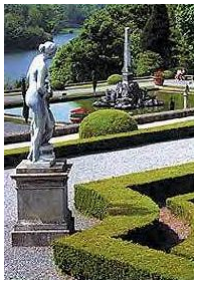
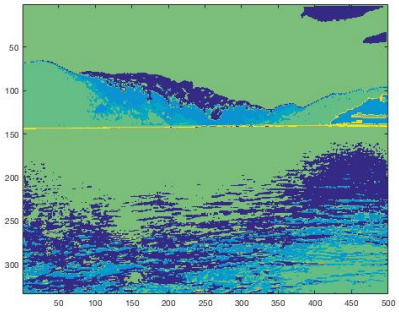
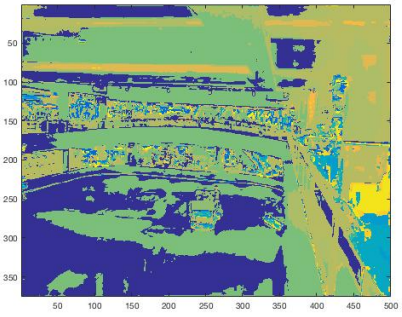
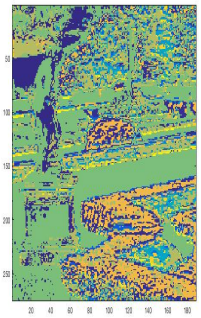
1. Gaussian: Smoothing image and reduce noise.
2. Laplacian of Gaussian (LoG): Second derivative of Gaussian filter to highlight region of rapid intensity change for edge detection.
3. Derivative of Gaussian on X axis: Highlighting the intensity change in horizontal direction.
4. Derivative of Gaussian on Y axis: Highlighting the intensity change in vertical direction.

Q1.1 Apply all 20 filters on a sample image, and visualize as a image collage.



Q1.2 & Q1.3 Visualize three wordmaps of three images from any one of the category and submit in the write-up along with their original RGB image.



Part2 Building a Recognition System

Quantitative Evaluation

When $\alpha = 50$; K = 100 and using Spatial Pyramid Matching to get image features, the test accuracy is **0.5437**. The confusion matrix C is,

C =

9	2	0	0	6	1	1	1
1	6	0	0	8	2	1	2
0	0	19	0	1	0	0	0
1	2	0	12	3	0	0	2
3	4	0	0	12	1	0	0
2	0	2	0	2	8	5	1
0	1	0	0	0	0	19	0
2	0	5	0	3	4	4	2

The accuracy for each image class is,

class	art_gallery	Computer_room	Garden	Ice_skating	Library	Mountain	Ocean	Tennis_court
Accuracy	0.45	0.30	0.95	0.60	0.60	0.40	0.95	0.10

From the table, it is obvious that the images of tennis court and computer room are more difficult to classify than the rest using the bags-of-words approach. It's not a surprise because the objects like computer and tennis nets are relatively hard to be detected when we arbitrarily take 50 pixels from each image to build the visual words. The garden and ocean images are very easy to

classify because the features like green leaf and the sea are distributed in the whole image, thus they are more likely to be extracted.

References

[1] Kristen Grauman and Trevor Darrell. The pyramid match kernel: Discriminative classification with sets of image features. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, volume 2, pages 1458–1465. IEEE, 2005.

[2] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. Beyond bags of features: Spatial pyramid match- ing for recognizing natural scene categories. In Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, volume 2, pages 2169–2178. IEEE, 2006.