

16720 — PS1

Connor W. Colombo (cwcolomb)

September 27th 2020

1 Representing the World with Visual Words

1.1 Extracting Filter Responses

Q1.1.1

- The Gaussian filters pick up / accentuate bright regions, that is the bodies objects, by blurring out the high frequency content of an image.
- The Laplacians of the Gaussians pick up edges of increasing size with increasing scale *sigma* as zero crossings.
- The x-derivatives of Gaussians pick up vertically oriented edges of increasing size with increasing scale *sigma* as peaks in brightness.
- The y-derivatives of Gaussians pick up horizontally oriented edges of increasing size with increasing scale *sigma* as peaks in brightness.

We need multiple scales of filter responses because there's always a trade-off between the smoothing, which gives increased noise resilience with increased filter scale, and localization, which gives better/more fine edge detection with decreased filter scale.

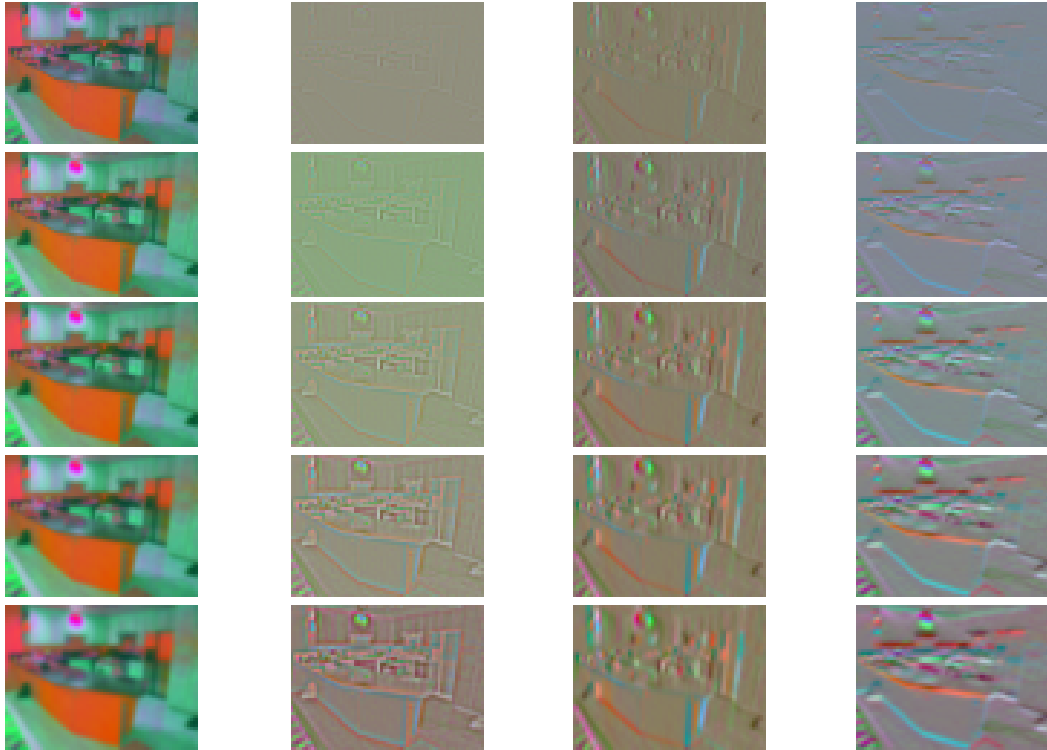


Figure 1: Q1.1.2: Responses of Kitchen Image to Filter Bank as a Collage of all 4 filters (along x axis) at 5 scales (along y axis).

Q1.1.2

1.2 Creating Visual Words

Q1.2

1.3 Computing Visual Words

Q1.3

2 Building a Recognition System

2.1 Extracting Features

Q2.1

2.2 Multi-resolution: Spatial Pyramid Matching

Q2.2

2.3 Comparing Images

Q2.3

2.4 Building A Model of the Visual World

Q2.4

2.5 Quantitative Evaluation

Q2.5

2.6 Find the Failures

Q2.6

3 Improving Performance

3.1 Hyperparameter Tuning

Q3.1

3.2 Further Improvement

3.2.1 Extra Credit