

Homework 1

16-720A Computer Vision

Matthew Swenson

September 26, 2017

1 Question 1.0

1. The Gaussian filters detect local neighborhood features.
2. The Laplacian of Gaussian filters detect rapid changes in features (edges and corners in both directions).
3. The dx filters detect changes along the x axis of the image.
4. The dy filters detect changes along the y axis of the image.

All of the filters are scaled such that they detect progressively larger features.



Figure 1: The various filter responses of a picture of an auditorium. Rows are Gaussian, LoG, dX, dY. Columns are increasing scale of filter.

2 Question 1.3

Section 1.3 involved the creation of "wordmaps" of each image using the computed dictionary; namely matching the closest texton to each pixel of an image in an attempt to generalize image textures. This was reasonably successful, as depicted in the associated image. I found it somewhat surprising that areas that look extremely visually similar to me are often classified as different textons; such as parts of the sky in the windmill picture, or even more confusingly, parts of the cabinets in the kitchen picture.



Figure 2: Various false-color wordmaps of images. Each hue is the closest texton to that part of the real image.

3 Question 2.5

The overall accuracy of this system is 56.25%, for $\alpha=400$ and $K = 175$.

4 Question 2.6

The single greatest source of error in the system was the misclassification of kitchens as laundromats. Intuitively, this makes sense, as both kitchens and laundromats feature a significant number of similar scale rectangular appliances, counters, and chairs, and often have similar layouts.

Table 1: Confusion Matrix

	<i>auditorium</i>	<i>baseball field</i>	<i>desert</i>	<i>highway</i>	<i>kitchen</i>	<i>laundromat</i>	<i>waterfall</i>	<i>windmill</i>
<i>auditorium</i>	12	2	0	0	4	1	1	0
<i>baseball field</i>	0	11	2	1	1	0	2	0
<i>desert</i>	0	3	9	3	0	0	0	0
<i>highway</i>	0	0	4	12	0	1	1	8
<i>kitchen</i>	3	1	2	0	13	9	0	0
<i>laundromat</i>	3	0	1	0	1	9	1	0
<i>waterfall</i>	1	0	0	1	1	0	14	2
<i>windmill</i>	1	3	2	3	0	0	1	10

Slightly more confusing is the high number of highways that were misclassified as windmills. My guess is that both of these image sets feature lots of sky and grassy areas, with a darker line shaped object running up the middle, and because the system primarily looks for textures, a lot of those images are going to be very similar.