Project 3

Locality-constrained Linear Coding for Scene Classification

Giang Doan, Emily Busche

{gdoan@cs.wisc.edu, ebusche@cs.wisc.edu}

1. Download images (Emily)

The 15-class natural scene dataset was downloaded into a file called **scene_categories(1)** (with the subclasses in separate folders inside this folder) which was in the project folder. You can change the path to these images by updating the **image_dir** variable in **Example.m**.

2. Download baseline and set up your pipeline (Emily)

For the pipeline, I modified the baseline code in the following ways:

a. **Example.m**

   This is the main file used to run the image classification code. It contains the variables for the parameters that are modified in the various experiments that were run.

b. **labelImages.m**

   This file moves the images to a single folder and renames them in a way that makes identification easier.

c. **SVMclass.m**

   This file performs the classification and prints out the confusion matrix for the test images. We use LIBSVM for classification. To get LIBSVM to run, go into libsvm-3.18\matlab in matlab and run make.

3. Implement LLC method by modifying the spatial pyramid code (Giang)

I modified some files when implementing LLC method:

a. **BuildPyramid2.m**

   This file will call new methods to generate SIFT descriptor, calculate dictionary, build histogram with LLC and compile pyramid.

b. **GenerateSiftDescriptors2.m**

This file is a modified version of **GenerateSiftDescriptor.m**. I modified it to extract proper features for LLC method.

c. **CalculateDictionary2.m**

This file is a modified version of **CalculateDictionary.m**. Here I replaced *sp_kmeans* method with *FastKMean* to train a k means cluster model faster.

d. **BuildHistograms2.m**

This file is a modified version of **BuildHistogram.m.** In this file, I implemented LLC method (function *CalculateLLC*) as in the paper which uses locality constrained linear coding.

e. **CompilePyramid2.m**

This file is a modified version of **CompilePyramid.m**. LLC uses max pooling to form the histogram feature.

4. Evaluate your implementation on the 15-class natural scene dataset (Emily)

For testing, we used 1,500 test images (15 classes, 100 images per class), gridSpacing = 8, patchSize = 16, dictionary size 200 and 2 pyramid levels. The unedited spatial pyramid code gave a mean accuracy of 69.02%, and the LLC code gave a mean accuracy of 73.59%.

Spatial Pyramid Confusion Matrix

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	125	0	3	0	2	2	0	2	2	0	0	2	0	1	2	141	88.65%
coast	1	193	1	19	0	7	37	0	0	0	2	0	0	0	0	260	74.23%
forest	3	0	214	0	0	5	3	2	0	0	0	0	0	0	1	228	93.86%
highway	0	15	0	116	1	6	11	3	0	0	0	6	0	0	2	160	72.50%
insidecity	7	0	0	3	134	0	2	17	4	1	1	16	13	2	8	208	64.42%
mountain	1	6	10	5	0	233	14	1	1	0	0	2	0	0	1	274	85.04%
country	7	48	18	1	0	33	196	6	0	0	0	0	0	0	1	310	63.23%
street	1	1	0	0	4	3	0	169	4	0	0	7	0	1	2	192	88.02%
building	0	0	1	0	16	1	0	4	199	0	0	23	2	3	7	256	77.73%
office	0	1	0	0	4	0	0	0	0	76	8	3	16	4	3	115	66.09%

bedroom	0	1	0	1	0	2	0	2	1	6	57	6	11	26	3	116	49.14%
industrial	7	3	1	4	21	8	2	4	30	5	5	92	6	2	21	211	43.60%
kitchen	0	0	0	0	8	1	0	0	1	5	11	0	62	18	4	110	56.36%
livingroom	2	0	0	0	3	2	0	2	1	10	31	6	15	107	10	189	56.61%
store	0	0	5	1	31	5	1	4	13	2	0	11	12	10	120	215	55.81%
Total	154	268	253	150	224	308	266	216	256	105	115	174	137	174	185		69.02%

LLC Confusion Matrix

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	129	0	1	0	0	0	0	0	0	0	0	5	1	0	5	141	91.49%
coast	0	215	2	4	0	9	29	0	0	1	0	0	0	0	0	260	82.69%
forest	0	0	208	0	0	15	2	1	0	0	0	1	0	0	1	228	91.23%
highway	0	14	0	133	2	3	3	4	0	0	1	0	0	0	0	160	83.13%
insidecity	1	0	0	1	166	0	1	25	7	3	0	2	1	0	1	208	79.81%
mountain	0	7	11	3	0	231	18	2	1	0	0	0	0	0	1	274	84.31%
country	3	54	15	3	0	33	194	6	0	0	0	0	0	0	2	310	62.58%
street	0	0	0	2	4	2	1	177	4	0	1	1	0	0	0	192	92.19%
building	0	0	0	0	24	1	0	4	216	0	2	3	1	0	5	256	84.38%
office	0	0	0	0	0	0	0	0	0	89	2	3	15	6	0	115	77.39%
bedroom	1	0	0	0	1	0	0	1	0	12	52	8	8	30	3	116	44.83%
industrial	11	2	0	0	11	0	0	1	13	2	9	116	3	6	37	211	54.98%
kitchen	0	0	0	0	1	0	0	0	0	6	17	5	57	22	2	110	51.82%
livingroom	2	0	0	0	1	1	0	0	0	6	32	7	24	105	11	189	55.56%
store	1	0	2	0	12	0	0	1	3	2	2	21	10	16	145	215	67.44%
Total	148	292	239	146	222	295	248	222	244	121	118	172	120	185	213		73.59%

5. Bonus

a. Optimize codebook:

- **Use fast k-mean for faster clustering (Giang):**
  As mentioned in (3c), I used a new algorithm fast k-mean to train a k means cluster model faster than the original k-mean algorithm in *sp_kmeans*.

- **Build optimized codebook (Emily)**:

I implemented the codebook optimization code described in algorithm 4.1 in the file **CodebookOpt.m**. It is called by **CalculateDictionary2.m**, but because the optimization did not show much improvement that line has been commented out (see **Try different parameters and experimental settings** section).

b. **Try different parameters and experimental settings:**

Several tests were performed by trying different values for the parameters. For all testing, we used 1,500 test images (15 classes, 100 images per class), gridSpacing = 8, and patchSize = 16.

Changing the Dictionary size

We tested the LLC code with 2 pyramid and varied the dictionary size. As we increased the dictionary size, the accuracy increased. Increasing the dictionary size by 100 lead to about a 2.5% increase in accuracy.

Confusion Matrix with 300 Dictionary Size (mean accuracy 76.49%)

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	135	0	1	0	0	0	0	0	0	0	0	1	0	1	3	141	95.74%
coast	0	222	2	1	0	4	31	0	0	0	0	0	0	0	0	260	85.38%
forest	0	0	211	0	0	9	7	0	0	0	0	0	0	0	1	228	92.54%
highway	0	12	0	134	2	3	6	3	0	0	0	0	0	0	0	160	83.75%
insidecity	2	0	0	4	169	0	0	24	4	0	0	2	0	1	2	208	81.25%
mountain	0	9	11	1	0	242	10	0	1	0	0	0	0	0	0	274	88.32%
country	1	44	21	3	0	35	202	4	0	0	0	0	0	0	0	310	65.16%
street	0	0	1	2	6	1	0	176	5	0	0	1	0	0	0	192	91.67%
building	0	0	1	1	24	1	0	2	221	0	0	3	1	0	2	256	86.33%
office	0	0	0	0	0	0	0	0	0	94	2	2	13	4	0	115	81.74%
bedroom	1	0	0	0	1	1	0	0	0	7	52	7	10	31	6	116	44.83%
industrial	8	2	0	1	9	1	1	2	12	2	7	122	5	6	33	211	57.82%
kitchen	0	0	0	0	0	0	0	0	0	6	7	4	62	29	2	110	56.36%
livingroom	0	0	0	0	0	1	0	0	0	7	18	3	26	126	8	189	66.67%
store	1	0	2	0	12	2	0	1	4	2	3	15	8	15	150	215	69.77%
Total	148	289	250	147	223	300	257	212	247	118	89	160	125	213	207		76.49%

Confusion Matrix with 400 Dictionary Size (mean accuracy 78.03%)

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	138	0	1	0	0	0	0	0	0	0	0	1	0	0	1	141	97.87%
coast	0	219	2	3	0	5	31	0	0	0	0	0	0	0	0	260	84.23%
forest	0	0	213	0	0	10	4	1	0	0	0	0	0	0	0	228	93.42%
highway	0	13	1	135	2	1	5	3	0	0	0	0	0	0	0	160	84.38%
insidecity	1	0	0	2	175	0	0	22	5	1	0	0	0	0	2	208	84.13%
mountain	0	4	14	2	0	239	13	0	2	0	0	0	0	0	0	274	87.23%
country	0	40	15	4	0	29	217	5	0	0	0	0	0	0	0	310	70.00%
street	0	0	0	0	7	2	0	177	6	0	0	0	0	0	0	192	92.19%
building	0	0	0	1	24	3	0	2	216	0	1	1	1	0	7	256	84.38%
office	0	0	0	0	0	0	0	0	0	98	5	3	7	2	0	115	85.22%
bedroom	1	0	0	0	0	1	0	1	0	2	59	4	13	30	5	116	50.86%
industrial	6	2	0	0	6	1	1	2	9	3	6	130	4	4	37	211	61.61%
kitchen	0	0	0	0	0	0	0	0	0	4	8	1	62	30	5	110	56.36%
livingroom	0	0	0	0	0	1	0	0	0	6	23	8	19	123	9	189	65.08%
store	0	0	1	0	10	0	0	1	4	1	2	13	9	16	158	215	73.49%
Total	146	278	247	147	224	292	271	214	242	##	##	161	##	205	224		78.03%

Changing the Pyramid Levels

We tested the LLC code with a dictionary size of 200 and varied the number of pyramid levels. When the pyramid levels increase from 2 to 3 levels, the mean accuracy increased from 73.59% to 79.47%. However, increasing the number of levels from 3 to 4 showed a slight decreasing in accuracy (to 79.22%). Increasing the number of pyramid levels can improve performance, but there is a limit to how many levels produce an increase in performance.

Confusion Matrix with 3 Pyramid Levels (mean accuracy 79.47%)

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	140	0	0	0	0	0	0	0	0	0	0	1	0	0	0	141	99.29%
coast	0	217	2	8	0	2	31	0	0	0	0	0	0	0	0	260	83.46%
forest	0	0	209	0	0	13	4	1	0	0	0	0	0	0	1	228	91.67%
highway	0	14	0	130	2	3	7	4	0	0	0	0	0	0	0	160	81.25%

insidecity	0	1	0	1	172	0	0	26	6	0	0	0	0	1	1	208	82.69%
mountain	0	5	12	6	0	238	10	0	3	0	0	0	0	0	0	274	86.86%
country	0	46	13	3	0	28	216	4	0	0	0	0	0	0	0	310	69.68%
street	0	0	0	1	6	2	0	176	6	0	0	0	0	0	1	192	91.67%
building	0	0	1	0	24	3	0	1	221	0	0	1	1	0	4	256	86.33%
office	0	0	0	0	0	0	0	0	0	96	3	3	10	3	0	115	83.48%
bedroom	0	0	0	0	0	0	0	1	0	3	62	8	7	31	4	116	53.45%
industrial	2	0	0	1	2	0	1	1	7	0	4	151	2	4	36	211	71.56%
kitchen	0	0	0	0	0	0	0	0	0	4	9	1	70	22	4	110	63.64%
livingroom	0	0	0	0	0	1	0	1	0	5	20	7	12	132	11	189	69.84%
store	0	0	1	0	3	0	0	2	3	5	2	7	7	19	166	215	77.21%
Total	142	283	238	150	209	290	269	217	246	113	100	179	109	212	228		79.47%

Confusion Matrix with 4 Pyramid Levels (mean accuracy 79.22%)

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	138	0	0	0	0	0	0	0	0	0	0	2	0	0	1	141	97.87%
coast	0	221	2	6	0	2	29	0	0	0	0	0	0	0	0	260	85.00%
forest	0	0	208	0	0	16	2	1	0	0	0	0	0	0	1	228	91.23%
highway	0	14	0	128	2	5	7	4	0	0	0	0	0	0	0	160	80.00%
insidecity	0	2	0	1	167	0	0	30	6	0	1	0	0	0	1	208	80.29%
mountain	0	6	13	5	0	233	13	1	3	0	0	0	0	0	0	274	85.04%
country	0	43	12	3	0	30	218	4	0	0	0	0	0	0	0	310	70.32%
street	0	0	0	1	8	2	1	174	5	0	0	1	0	0	0	192	90.63%
building	0	0	1	0	26	2	0	0	224	0	0	0	1	0	2	256	87.50%
office	0	0	0	0	0	0	0	0	0	96	1	3	11	4	0	115	83.48%
bedroom	0	0	0	0	0	0	0	1	0	3	63	8	8	28	5	116	54.31%
industrial	3	0	0	0	1	0	1	1	6	0	3	153	2	5	36	211	72.51%
kitchen	0	0	0	0	0	0	0	0	0	2	8	2	65	28	5	110	59.09%
livingroom	0	0	0	0	0	1	0	1	0	4	14	8	11	137	13	189	72.49%
store	0	0	1	0	4	1	0	1	3	3	2	1	9	21	169	215	78.60%
Total	141	286	237	144	208	292	271	218	247	##	92	178	107	223	233		79.22%

Combined test

Combining the knowledge we learned from the previous tests, we increased the pyramid levels to 3 and the dictionary size to 300. This produced a mean accuracy of 79.72% an improvement over all previous tests.

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	138	0	0	0	0	0	0	0	0	0	0	2	0	1	0	141	97.87%
coast	0	227	2	2	0	2	27	0	0	0	0	0	0	0	0	260	87.31%
forest	0	0	210	0	0	14	4	0	0	0	0	0	0	0	0	228	92.11%
highway	0	13	0	130	2	3	7	5	0	0	0	0	0	0	0	160	81.25%
insidecity	0	1	0	2	174	0	0	24	6	1	0	0	0	0	0	208	83.65%
mountain	0	6	13	3	0	239	10	0	3	0	0	0	0	0	0	274	87.23%
country	0	40	13	3	0	30	220	4	0	0	0	0	0	0	0	310	70.97%
street	0	0	0	3	7	1	0	176	4	0	0	1	0	0	0	192	91.67%
building	0	0	1	0	22	1	0	2	226	0	0	0	1	0	3	256	88.28%
office	0	0	0	0	0	0	0	0	0	102	1	1	9	2	0	115	88.70%
bedroom	0	0	0	0	0	1	0	1	0	5	59	6	9	31	4	116	50.86%
industrial	3	1	0	0	8	1	1	3	10	1	3	145	3	3	29	211	68.72%
kitchen	0	0	0	0	0	0	0	0	0	6	7	5	62	24	6	110	56.36%
livingroom	0	0	0	0	0	1	0	0	0	4	17	6	11	141	9	189	74.60%
store	0	0	2	0	7	1	0	1	5	3	2	4	6	20	164	215	76.28%
Total	141	288	241	143	220	294	269	216	254	122	89	170	101	222	215		79.72%

Codebook optimization

We implemented the codebook optimization algorithm in the LLC paper. To compare it to our initial LLC test, we used a dictionary size of 200 and 2 pyramid levels to compare to spatial pyramids. We used the entire training set for our X and used $\lambda = 500$ and $\sigma = 100$. The optimized codebook version of LLC performed only slightly better than our original test with a mean accuracy of 73.83%.

	suburb	coast	forest	highway	insidecity	mountain	country	street	building	office	bedroom	industrial	kitchen	livingroom	store	Total	Mean accuracy
suburb	133	0	1	0	0	0	0	0	0	0	0	5	0	0	2	141	94.33%
coast	0	216	2	2	0	7	33	0	0	0	0	0	0	0	0	260	83.08%
forest	0	0	208	0	0	9	7	2	1	0	0	0	0	0	1	228	91.23%
highway	0	21	1	122	2	4	5	5	0	0	0	0	0	0	0	160	76.25%
insidecity	1	0	0	1	166	0	0	26	5	3	1	2	0	2	1	208	79.81%
mountain	0	8	11	5	0	235	13	0	2	0	0	0	0	0	0	274	85.77%

country	2	49	19	2	0	23	209	6	0	0	0	0	0	0	0	310	67.42%
street	0	0	1	0	5	2	0	174	9	0	0	1	0	0	0	192	90.63%
building	0	0	1	0	24	2	0	2	218	0	1	3	1	0	4	256	85.16%
office	0	0	0	0	0	0	0	0	0	91	5	2	16	1	0	115	79.13%
bedroom	0	0	0	0	1	1	0	0	0	6	50	8	8	36	6	116	43.10%
industrial	9	0	0	1	11	0	1	4	14	2	5	126	4	4	30	211	59.72%
kitchen	0	0	0	0	1	0	0	0	0	8	10	3	55	28	5	110	50.00%
livingroom	0	0	0	0	0	1	0	0	0	9	29	10	27	101	12	189	53.44%
store	2	0	2	0	11	0	0	1	2	2	2	20	8	18	147	215	68.37%
Total	147	294	246	133	221	284	268	220	251	121	103	180	119	190	208		73.83%