Pyramids of Features
For Categorization

Greg Griffin and Will Coulter
(see Lazebnik et al., CVPR 2006, too)
Intuition:
Approximates optimal partial matching

Intuition [cont’d]:
Combine bags of features with spatial information
Example Pyramid Comparison
Disadvantages

- Same objects in different quadrants
- Objects sliced by bins
Possible Solutions

• Flipping / rotating image
• Sliding / shuffling histogram bins
Possible Solutions [cont’d]

• Split histogram in powers of three
Implementation Overview

Image and Feature Extraction (SIFT on regular grid)

Vocabulary Translation (200 words (k-means))

Histogram Generation (flips, slides, arbitrary mixing)

Matching (full or partial pyramid)

Decision (best match, voting, SVM)
Sanity Check 1
Graphical mini-confusion matrix (and rot. invariance)
Scene Database
Scene Database

81.1% vs 67.4%

(100)
Caltech 101 64.6% vs. 33.1% (30, 16)
<table>
<thead>
<tr>
<th>Category</th>
<th>Image 1</th>
<th>Image 2</th>
<th>Image 3</th>
<th>Image 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>minaret</td>
<td><img src="minaret_image1" alt="Image" /></td>
<td><img src="minaret_image2" alt="Image" /></td>
<td><img src="minaret_image3" alt="Image" /></td>
<td><img src="minaret_image4" alt="Image" /></td>
</tr>
<tr>
<td>windsor chair</td>
<td><img src="windsor_chair_image1" alt="Image" /></td>
<td><img src="windsor_chair_image2" alt="Image" /></td>
<td><img src="windsor_chair_image3" alt="Image" /></td>
<td><img src="windsor_chair_image4" alt="Image" /></td>
</tr>
<tr>
<td>joshua tree</td>
<td><img src="joshua_tree_image1" alt="Image" /></td>
<td><img src="joshua_tree_image2" alt="Image" /></td>
<td><img src="joshua_tree_image3" alt="Image" /></td>
<td><img src="joshua_tree_image4" alt="Image" /></td>
</tr>
<tr>
<td>okapi</td>
<td><img src="okapi_image1" alt="Image" /></td>
<td><img src="okapi_image2" alt="Image" /></td>
<td><img src="okapi_image3" alt="Image" /></td>
<td><img src="okapi_image4" alt="Image" /></td>
</tr>
<tr>
<td>cougar body</td>
<td><img src="cougar_body_image1" alt="Image" /></td>
<td><img src="cougar_body_image2" alt="Image" /></td>
<td><img src="cougar_body_image3" alt="Image" /></td>
<td><img src="cougar_body_image4" alt="Image" /></td>
</tr>
<tr>
<td>beaver</td>
<td><img src="beaver_image1" alt="Image" /></td>
<td><img src="beaver_image2" alt="Image" /></td>
<td><img src="beaver_image3" alt="Image" /></td>
<td><img src="beaver_image4" alt="Image" /></td>
</tr>
<tr>
<td>crocodile</td>
<td><img src="crocodile_image1" alt="Image" /></td>
<td><img src="crocodile_image2" alt="Image" /></td>
<td><img src="crocodile_image3" alt="Image" /></td>
<td><img src="crocodile_image4" alt="Image" /></td>
</tr>
<tr>
<td>ant</td>
<td><img src="ant_image1" alt="Image" /></td>
<td><img src="ant_image2" alt="Image" /></td>
<td><img src="ant_image3" alt="Image" /></td>
<td><img src="ant_image4" alt="Image" /></td>
</tr>
</tbody>
</table>

Caltech 101

(30, 16)
Work in Progress

• 256 Performance
  – 64 times more work scene database
  – 6.4 times more work than 101

• SVM
  – one-vs-all weighting issues
  – speed it up?
  – improve performance

• Improvements
  – Flip, Slide, Arbitrary

• Powers-of-3 histogram bins
Open Questions

• Performance of arbitrary match bins
  – Try random sampling?
  – Allow multiple best matches?
• Chess/pattern example
• Grid example
• Optimal kernel level weights
Implementation Details

- [block diagram]
- Images \((288^2, b&w, \text{squished})\) and feature extraction (-weak, -pca, +sift)
- Vocab generation (200 words, 20,000-small)
- Histogram\((\text{fliplr, flipud, slide, arbitrary, bag of features})\)
- match (full&partial pyramids)
- Decision (best match, voting, SVM)