IMAGE CATEGORIZATION USING CODEBOOKS BUILT FROM SCORED AND SELECTED LOCAL FEATURES

By
Bala S. Divakaruni and Dr. Jie Zhou
Northern Illinois University
OUTLINE

- Background.
- Method.
- Results.
- Discussion.
- Conclusion.
AIM

Classify an entire image to a known category using sparse local features.
RELATED WORK

- Bags of keypoints (G. Csurka ECCV 2004).
- Codeword binary weighting (J. Winn ICCV 2005).
Codebook (dictionary, vocabulary) is often generated to facilitate image representation and the subsequent classification.
K-Means clustering.
- Favoring dense areas.
Large codebooks.
Codebook Uncertainty and plausibility.
METHOD HIGHLIGHTS

1. Scored codes – Most representative Features
2. Local features, feature selection – no aligning or segmentation required, not based on shapes.
ALGORITHM FLOW

1. Extracting Reliable Features
2. M & C Scoring
   - M-score Calculation
   - C-score Calculation
3. Score Sorting, Code Selection
4. Image Representation
5. Classifier
SIFT features are distinctive scale-invariant keypoints defined as maxima and minima of Difference of Gaussian (DoG) at various scales.

A raw SIFT feature extractor consists of 2D location, scale and orientation, and the keypoint descriptor.
To reduce the computational complexity of feature scoring as well as the inference of some background clutters and outliers, we apply Hough transform to the raw SIFT features.
The setting is similar as in image retrieval, with features to be matched in the query image, and the remaining images viewed as reference images.
The M-score reflects the level of matching of the feature with different features in other images of the same category.

\[ M(\mathbf{f}_{ij}^x) = 1 - \text{fd}(\mathbf{f}_{ij}^x) \times \text{rod}(\mathbf{f}_{ij}^x) \]
Consensus is the process of determining how representative the matched features are for a particular class.

C-score is consensus weighted by level of matching.

\[
C(f_i^x) = \frac{1}{L-1} \sum_{j=1, i-1, i+1, L}^{} \delta(M(f_{ij}^x) - t)M(f_{ij}^x)
\]
SELECTED CODES
K150 DATASET

Sample image, Class 1

Sample image, Class 2

Sample image, Class 3

Sample image, Class 4
BINARY DATASETS

Graz 01 - Bike

Graz 01 - Person

Caltech6 - Faces(front)

Caltech6 - Motorcycle(side)
CODEBOOK GENERATION

Raw Sift features

Class A

35 selected Codes for A Class

Class B

Scores of Codewords in the Codebook
CODEBOOK REPRESENTATION OF AN IMAGE

Histogram representation of images

K150, class1 (a278)
CODEBOOK REPRESENTATION OF AN IMAGE

K150 class2 (a150)

The Histogram of the image

K150 class3 (ato)

The Histogram of the image
## RESULTS

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Graz Bike</th>
<th>Graz People</th>
<th>Caltech Motorcycle</th>
<th>Caltech Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our result</td>
<td>88.0</td>
<td>81.0</td>
<td>96.2</td>
<td>100</td>
</tr>
<tr>
<td>A. Opelt [PAMI 2006]</td>
<td>86.5</td>
<td>80.8</td>
<td>94.3</td>
<td>100</td>
</tr>
<tr>
<td>R. Fergus [CVPR 2003]</td>
<td>-</td>
<td>-</td>
<td>92.5</td>
<td>96.3</td>
</tr>
<tr>
<td>S. Lazebnik [CVPR 2006]</td>
<td>86.3</td>
<td>82.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>D. Crandall [ECCV 2004]</td>
<td>79.0 - 84.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>J. Zhang [IJCV 2007]</td>
<td>92.0</td>
<td>88.0</td>
<td>98.5</td>
<td>100</td>
</tr>
<tr>
<td>H. Cai [CVPR 2010]</td>
<td>83.3 - 86.7</td>
<td>80.7 - 84.0</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Graz Bike</th>
<th>Graz People</th>
<th>Caltech Motorcycle</th>
<th>Caltech Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>0.6</td>
<td>0.52</td>
<td>0.55</td>
<td>0.6</td>
</tr>
<tr>
<td>Size</td>
<td>112</td>
<td>167</td>
<td>424</td>
<td>40</td>
</tr>
</tbody>
</table>
CODEBOOK SIZE VERSUS PERFORMANCE

Trends of performance versus codebook size per class.

K150
Caltech Motorcycle

Graz01 Bike
CONCLUSION

- Local feature scoring and selection.
- Smaller number of codes.
- Distance metric based definitions for matching and scoring.
- Avoided possible bottleneck of computational complexity

(cont..)
CONCLUSION

(cont..)

- More representative Codebooks.
- Very satisfactory results.
THANK YOU
QUESTIONS??