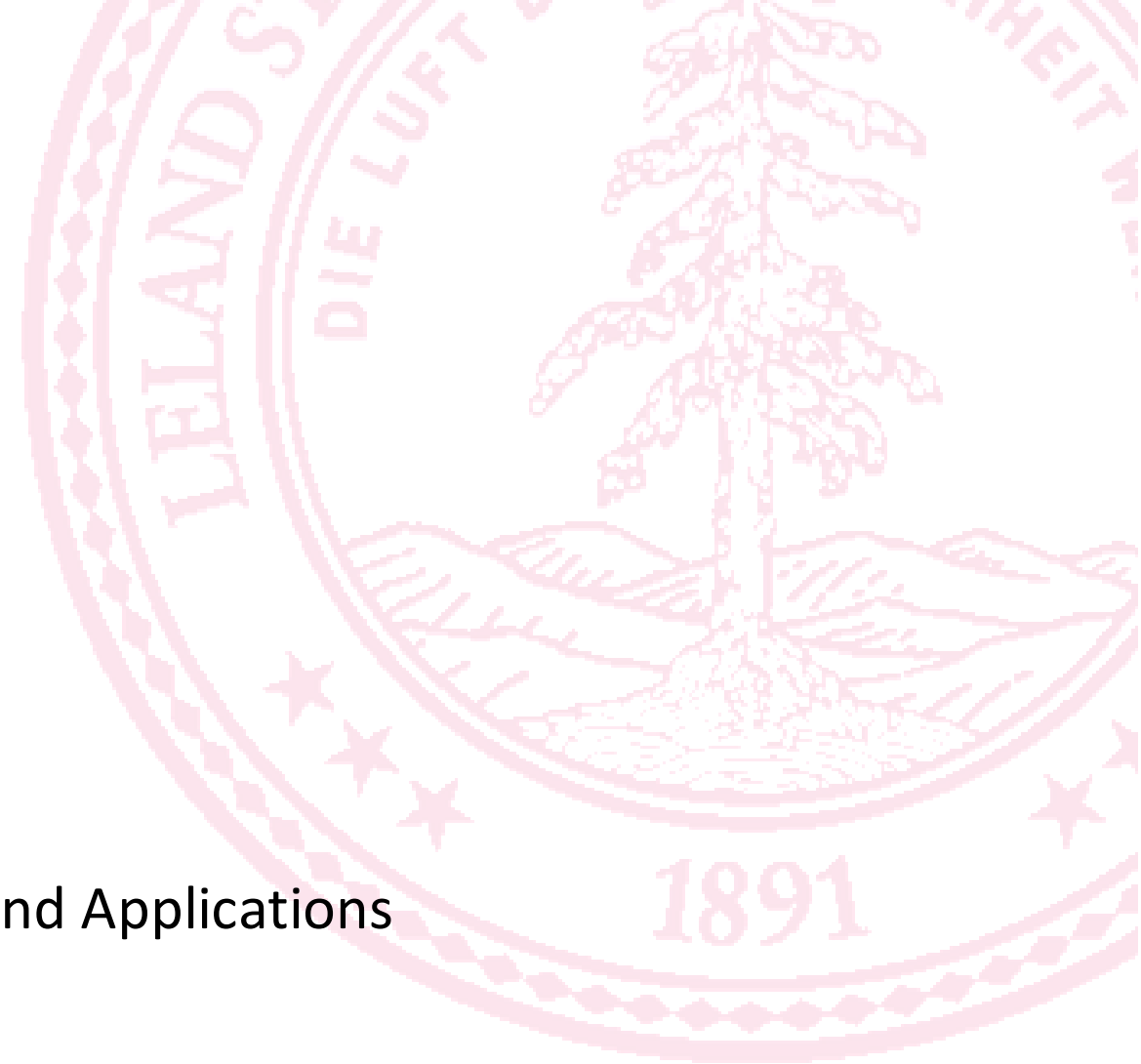


Lecture 13: Visual Bag of Words

Spatial pyramid matching

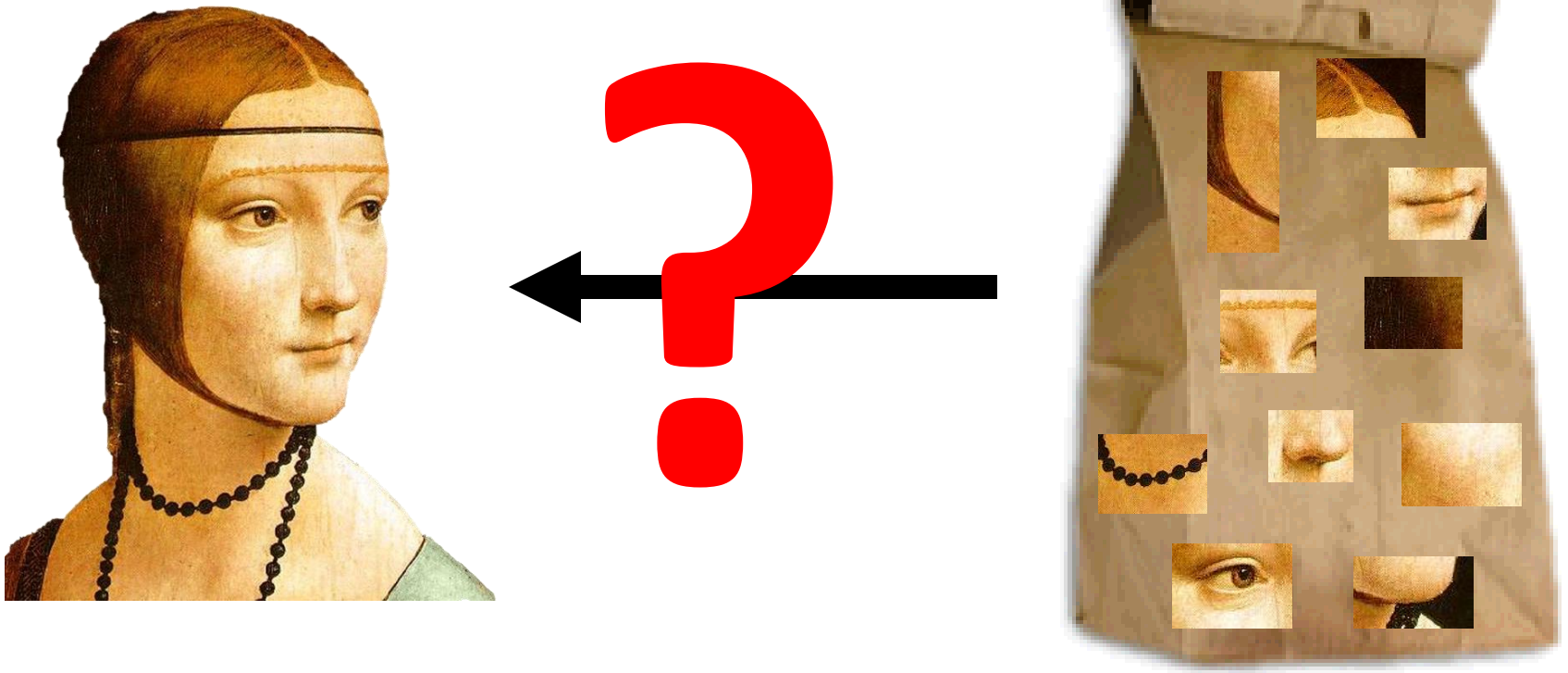
Juan Carlos Niebles and Jiajun Wu

CS131 Computer Vision: Foundations and Applications

What will we learn today?

- Spatial pyramid matching

What about spatial info?



Pyramids

- Very useful for representing images.
- Pyramid is built by using copies of an image at multiple resolutions.
- Each level in the pyramid is 1/2 of the size of previous level.
- The lowest level is of the highest resolution.
- The highest level is of the lowest resolution.

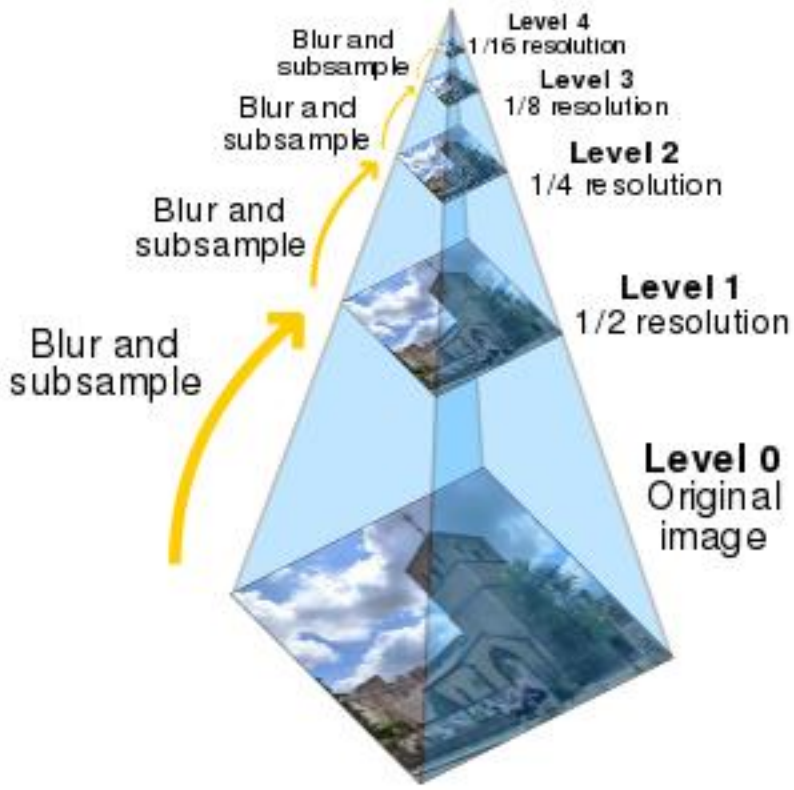
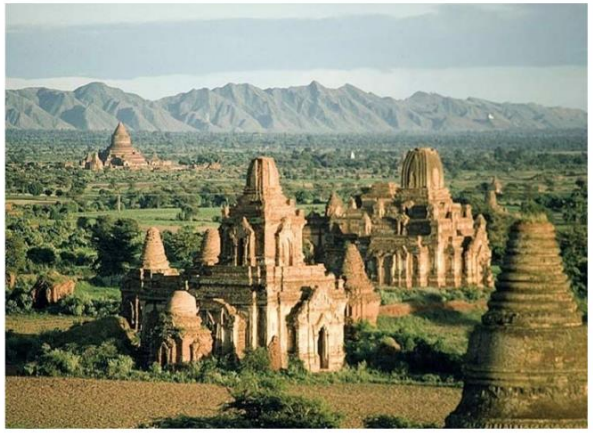
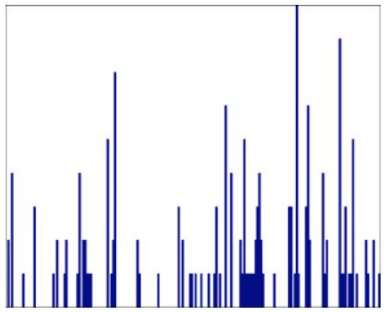


Image source: cmglee at commons.wikimedia.org

Bag of words + pyramids

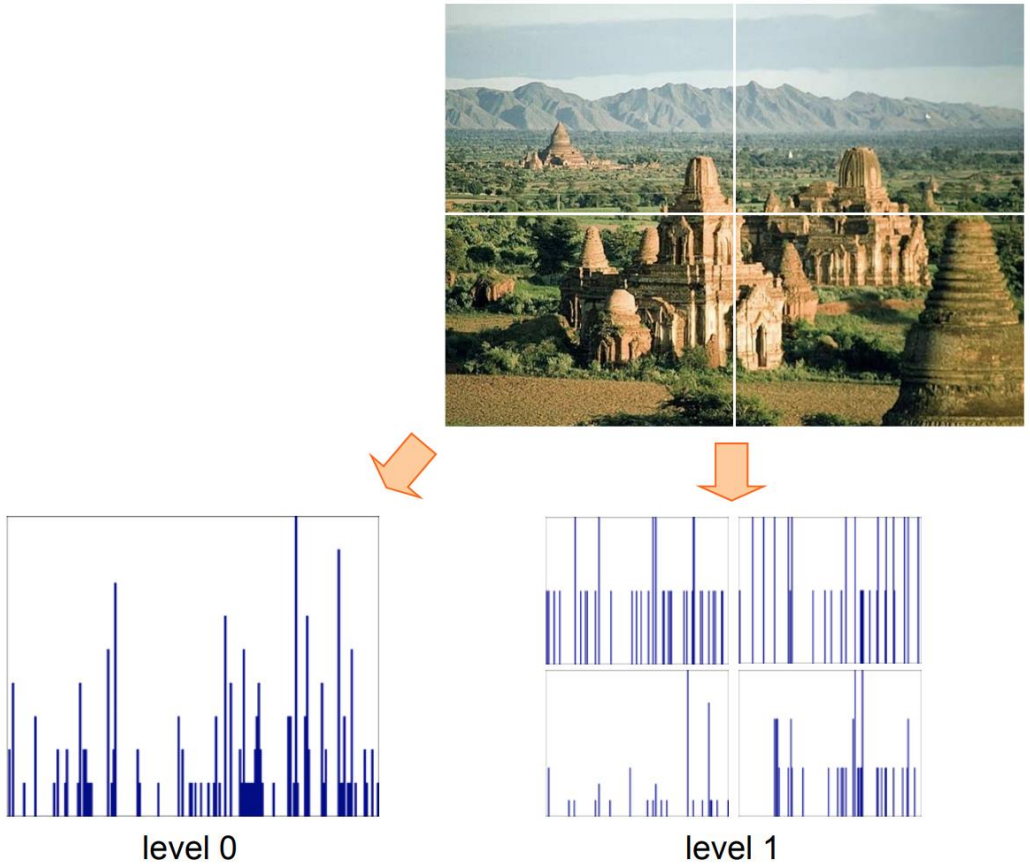




level 0

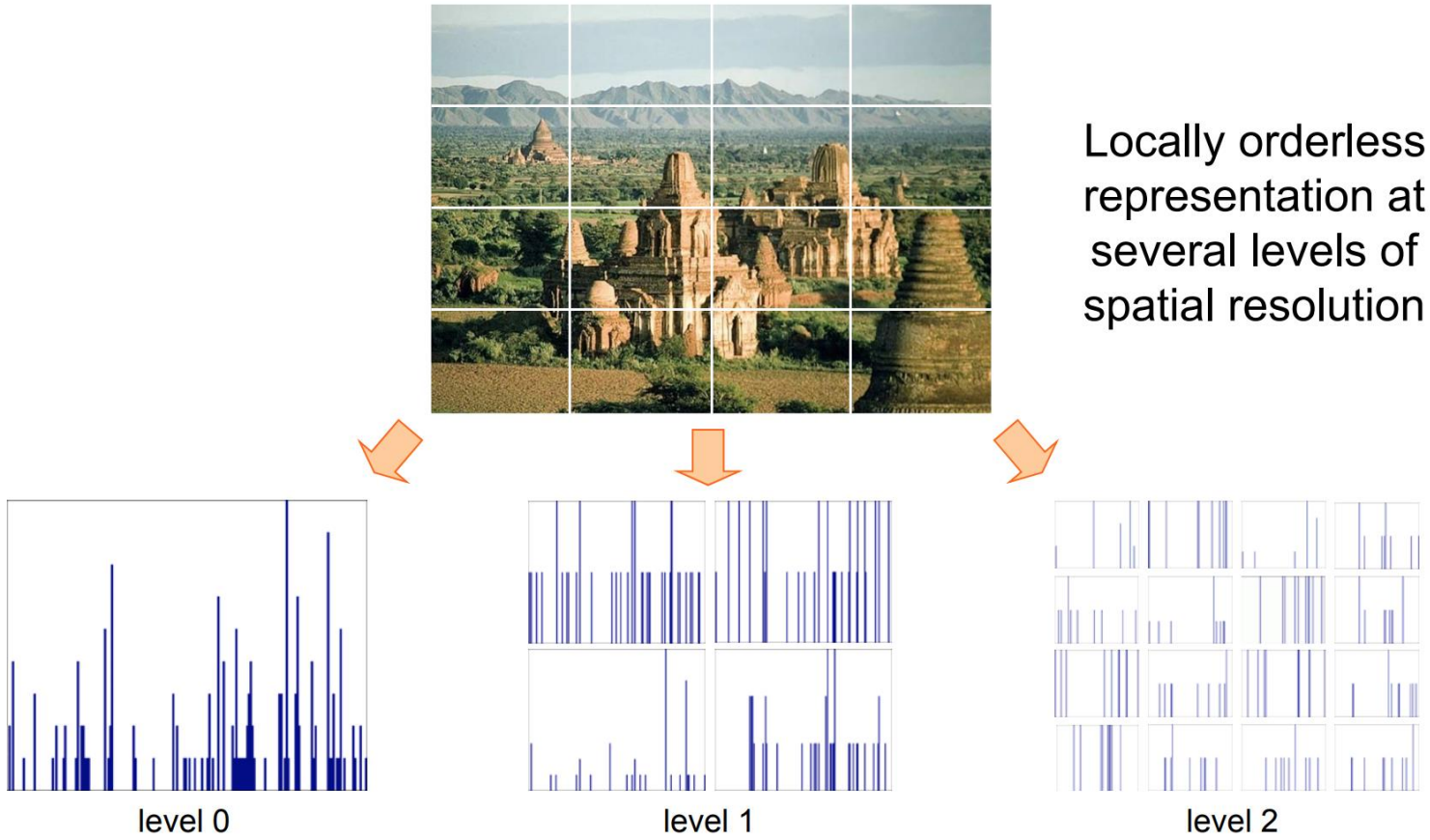
Locally orderless representation at several levels of spatial resolution

Bag of words + pyramids

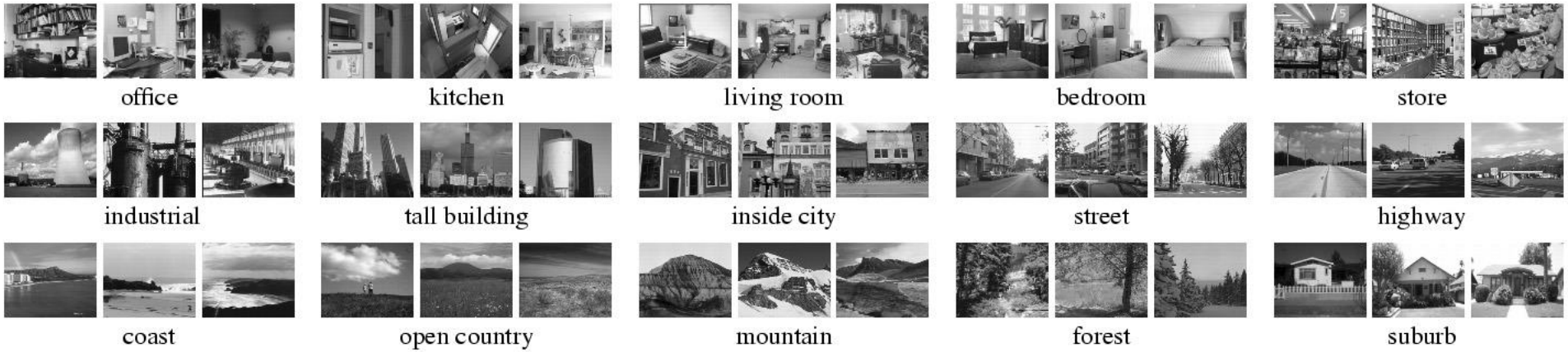


Locally orderless representation at several levels of spatial resolution

Bag of words + pyramids



Scene category dataset



Multi-class classification results
(100 training images per class)

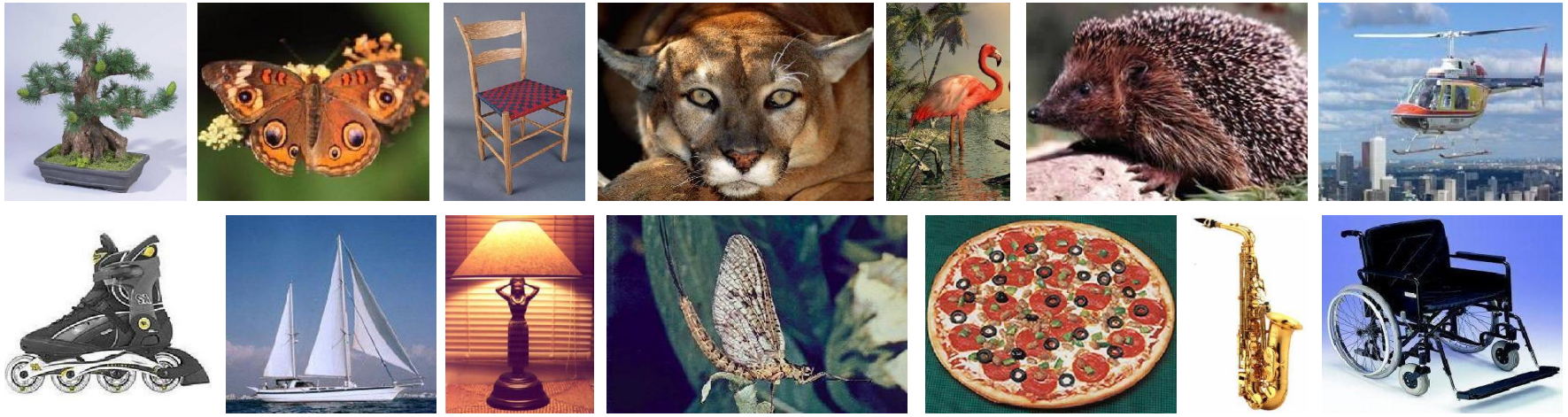
	Weak features (vocabulary size: 16)		Strong features (vocabulary size: 200)	
Level	Single-level	Pyramid	Single-level	Pyramid
0 (1×1)	45.3 ±0.5		72.2 ±0.6	
1 (2×2)	53.6 ±0.3	56.2 ±0.6	77.9 ±0.6	79.0 ±0.5
2 (4×4)	61.7 ±0.6	64.7 ±0.7	79.4 ±0.3	**81.1** ±0.3
3 (8×8)	63.3 ±0.8	**66.8** ±0.6	77.2 ±0.4	80.7 ±0.3

Lazebnik, Schmid & Ponce (CVPR 2006)

Slide credit: Svetlana Lazebnik

Caltech101 dataset

http://www.vision.caltech.edu/Image_Datasets/Caltech101/Caltech101.html



Multi-class classification results (30 training images per class)

	Weak features (16)		Strong features (200)	
Level	Single-level	Pyramid	Single-level	Pyramid
0	15.5 ±0.9		41.2 ±1.2	
1	31.4 ±1.2	32.8 ±1.3	55.9 ±0.9	57.0 ±0.8
2	47.2 ±1.1	49.3 ±1.4	63.6 ±0.9	**64.6** ±0.8
3	52.2 ±0.8	**54.0** ±1.1	60.3 ±0.9	64.6 ±0.7

Slide credit: Svetlana Lazebnik

Summary

- Spatial pyramid matching
  - Motivation
  - Method
  - Results on scene classification