

CS4495/6495

# Introduction to Computer Vision

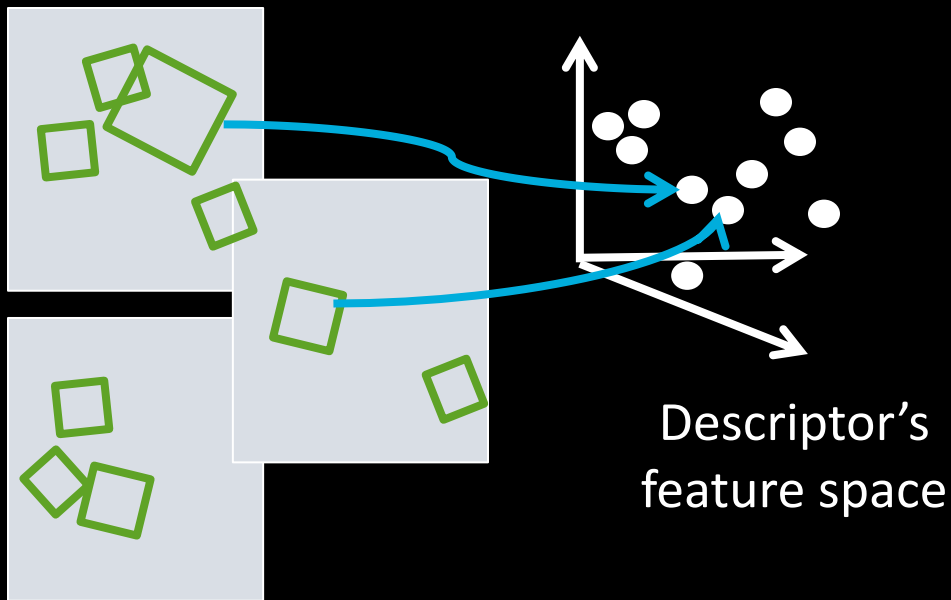
---

8C-L4 *Bag of visual words*



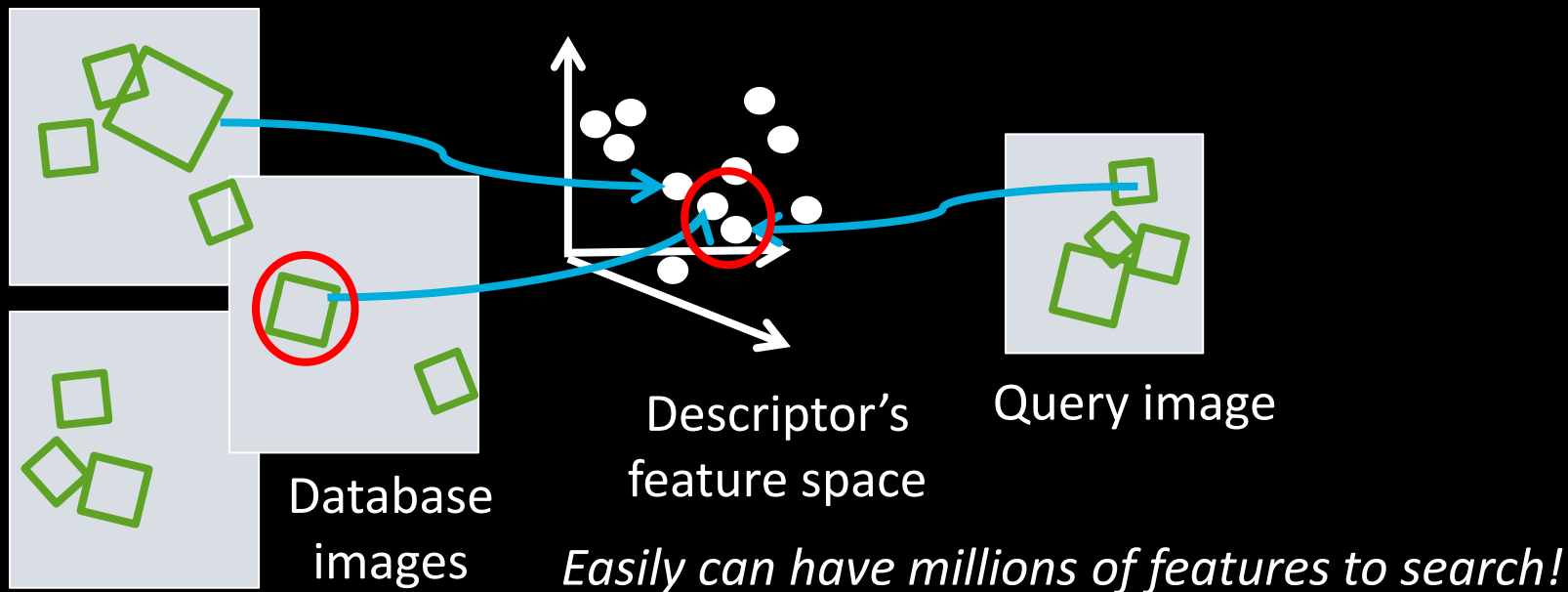
# Indexing local features

- Each patch / region has a descriptor, which is a point in some high-dimensional feature space (e.g., SIFT)



# Indexing local features

- When we see close points in feature space, we have similar descriptors, which indicates similar local content.



# Indexing local features

- With potentially thousands of features per image, and hundreds to millions of images to search, how to efficiently find those that are relevant to a new image?

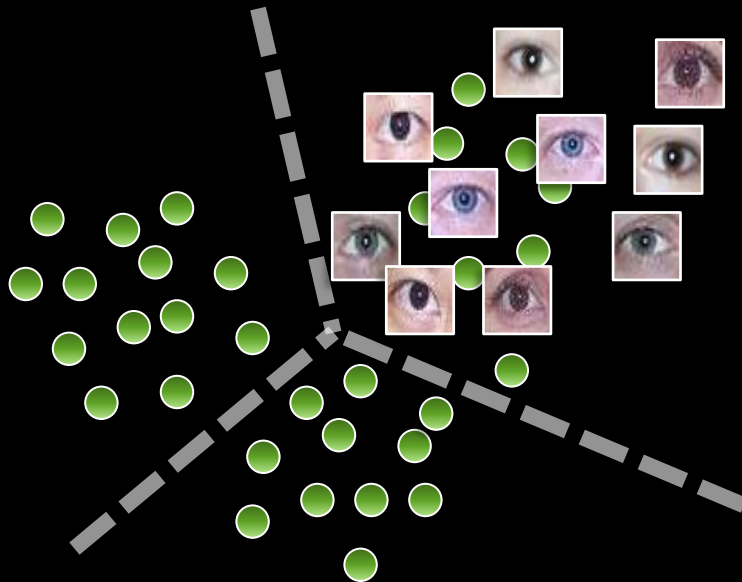
# Indexing local features: inverted file index

Index		
"Along I-75," From Detroit to Florida; <i>inside back cover</i>	Butterfly Center, McGuire; 134	Driving Lanes; 85
"Drive I-95," From Boston to Florida; <i>inside back cover</i>	CAA (see AAA)	Duval County; 163
1929 Spanish Trail Roadway; 101-102, 104	CCC, The; 111, 113, 115, 135, 142	Eau Gallie; 175
511 Traffic Information; 83	Ca d'Zan; 147	Edison, Thomas; 152
A1A (Barrier Is) - I-95 Access; 86	Caloosahatchee River; 152	Eglin AFB; 116-118
AAA (and CAA); 83	Name; 150	Eight Reals; 176
AAA National Office; 88	Canaveral Natl Seashore; 173	Ellenton; 144-145
Abbreviations,	Cannon Creek Airport; 130	Emanuel Point Wreck; 120
Colored 25 mile Maps; cover	Canopy Road; 106, 169	Emergency Callboxes; 83
Exit Services; 196	Cape Canaveral; 174	Epiphytes; 142, 148, 157, 159
Travelogue; 85	Casillo San Marcos; 169	Escambia Bay; 119
Africa; 177	Cave Diving; 131	Bridge (I-10); 119
Agricultural Inspection Stns; 126	Cayo Costa, Name; 150	County; 120
Ah-Tah-Thi-Ki Museum; 180	Celebration; 93	Estero; 153
Air Conditioning, First; 112	Charlotte County; 149	Everglade; 90, 95, 139-140, 154-160
Alabama; 124	Charlotte Harbor; 150	Draining of; 156, 181
Alachua; 132	Chautauqua; 116	Wildlife MA; 160
County; 131	Chipley; 114	Wonder Gardens; 154
Alafia River; 143	Name; 115	Falling Waters SP; 115
Alapaha, Name; 126	Choctawhatchee, Name; 115	Fantasy of Flight; 95
Alfred B Macley Gardens; 106	Circus Museum, Ringling; 147	Fayer Dykes SP; 171
Alligator Alley; 154-155	Citrus; 88, 97, 130, 136, 140, 180	Fires, Forest; 166
Alligator Farm, St Augustine; 169	CityPlace, W Palm Beach; 180	Fires, Prescribed ; 148
Alligator Hole (definition); 157	City Maps,	Fisherman's Village; 151
Alligator, Buddy; 155	Fl Lauderdale Expwys; 194-195	Flagler County; 171
Alligators; 100, 135, 138, 147, 156	Jacksonville; 163	Flagler, Henry; 97, 165, 167, 171
Anastasia Island; 170	Kissimmee Expwys; 192-193	Florida Aquarium; 186
Anhaica; 108-109, 146	Miami Expressways; 194-195	Florida,
Apalachicola River; 112	Orlando Expressways; 192-193	12,000 years ago; 187
Appleton Mus of Art; 136	Pensacola; 26	Cavern SP; 114
Aquifer; 102	Tallahassee; 191	Map of all Expressways; 2-3
Arabian Nights; 94	Tampa-St. Petersburg; 63	Mus of Natural History; 134
Art Museum, Ringling; 147	St. Augustine; 191	National Cemetery ; 141
Aruba Beach Cafe; 183	Civil War; 100, 108, 127, 138, 141	Part of Africa; 177
Aucilla River Project; 106	Cleanwater Marine Aquarium; 187	Platform; 187
Babcock-Web WMA; 151	Collier County; 154	Sheriff's Boys Camp; 126
Bahia Mar Marina; 184	Collier, Barron; 152	Sports Hall of Fame; 130
Baker County; 99	Colonial Spanish Quarters; 168	Sun 'n Fun Museum; 97
Barfoot Mailman; 182	Columbia County; 101, 128	Supreme Court; 107
Barge Canal; 137	Coquina Building Material; 165	Florida's Turnpike (FTP); 178, 189
Bee Line Expy; 80	Corkscrew Swamp, Name; 165	25 mile Strip Maps; 66
Beltz Outlet Mall; 89	Cowboys; 95	Administration; 189
Bernard Castro; 136	Crab Trap II; 144	Coin System; 190
Big "I"; 165	Cracker, Florida; 88, 95, 132	Exit Services; 189
Big Cypress; 155, 158	Crosetown Expy; 11, 35, 98, 143	HEFT; 76, 161, 190
Big Foot Monster; 105	Cuban Bread; 184	History; 189
	Dade Battlefield; 140	Names; 189
	Dade, Maj. Francis; 139-140, 161	Service Plazas; 190
	Dania Beach Hurricane; 184	Spur SR91; 76

- For text documents, an efficient way to find all *pages* on which a *word* occurs is to use an index...
- We want to find all *images* in which a *feature* occurs.
- To use this idea, we'll need to map our features to "visual words"

# Visual words (discretization)

- Words are discrete, visual features are typically continuous...



Discretization via clustering/vector quantization

# Visual words

Example: each group of patches belongs to the same visual word

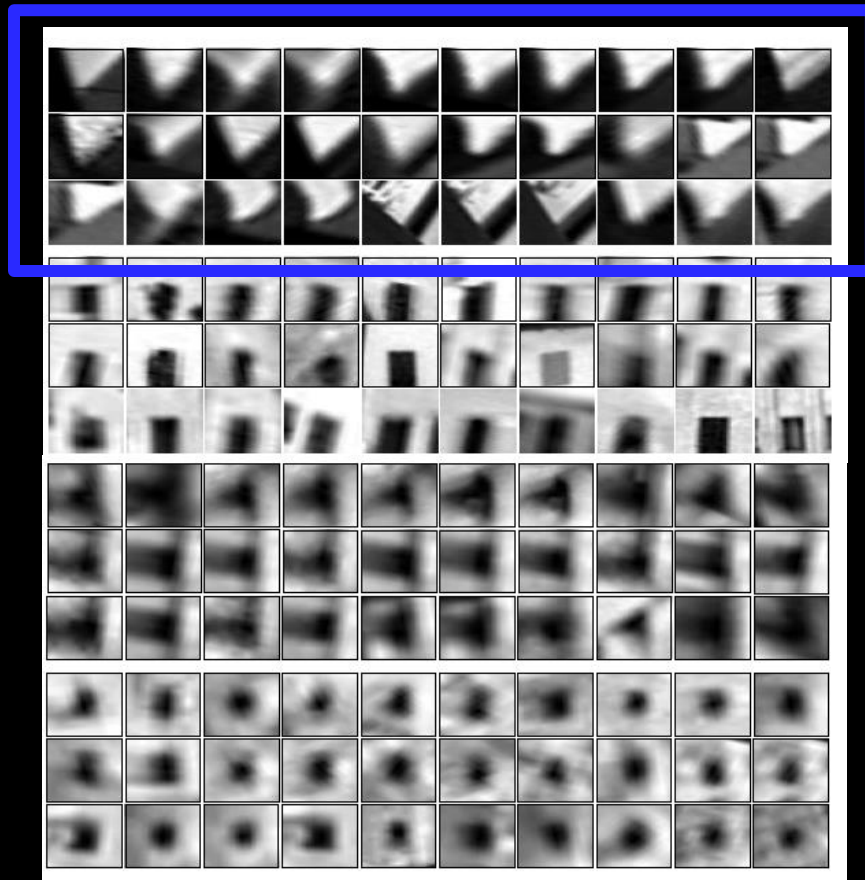
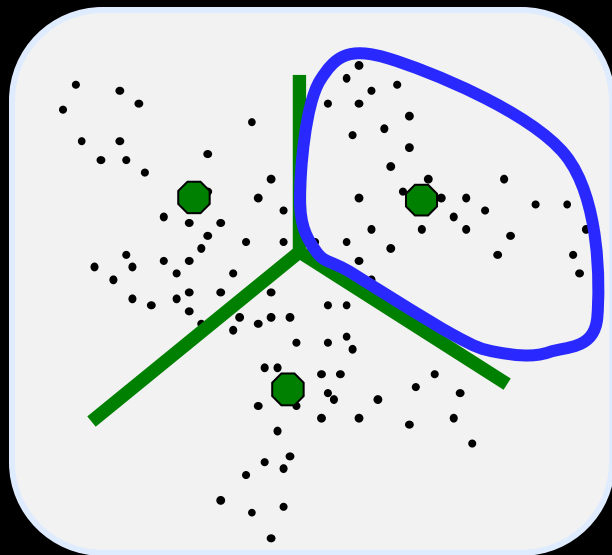
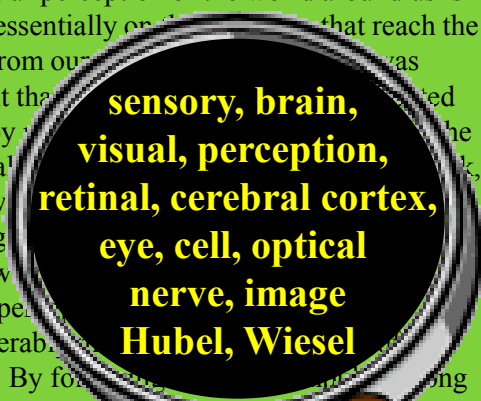


Figure from Sivic & Zisserman, ICCV 2003

# Analogy to documents

Of all the sensory impressions proceeding to the brain, the visual experiences are the dominant ones. Our perception of the world around us is based essentially on the visual impressions that reach the brain from our eyes. It was thought that the visual point by which the cerebral cortex is upon which. Through we now we now visual perception is considerably events. By following their path to the visual cortex, Hubel and Wiesel demonstrate that the *message about the image falling on the retina undergoes a step-by-step analysis in a system of nerve cells stored in columns. In this system each cell has its specific function and is responsible for a specific detail in the pattern of the retinal image.*



**sensory, brain,  
visual, perception,  
retinal, cerebral cortex,  
eye, cell, optical  
nerve, image  
Hubel, Wiesel**

China is forecasting a trade surplus of \$90bn (£51bn) to \$100bn this year, a threefold increase on 2004's \$32bn. The Commerce Ministry said the surplus would be created by a predicted 30% jump in exports compared with a 18% rise in imports. It is likely to be argued that the surplus is only a temporary phenomenon. Zhou Xiaochuan, governor of the People's Bank of China, said the yuan should do more to support the growth of goods stayed in the country. He increased the value of the yuan by the dollar by 2.1% in July. The trade within a narrow band, but the US wants the yuan to be allowed to trade freely. However, Beijing has made it clear that it will take time and tread carefully before allowing the yuan to rise further in value.



**China, trade,  
surplus, commerce,  
exports, imports, US,  
yuan, bank, domestic,  
foreign, increase,  
trade, value**

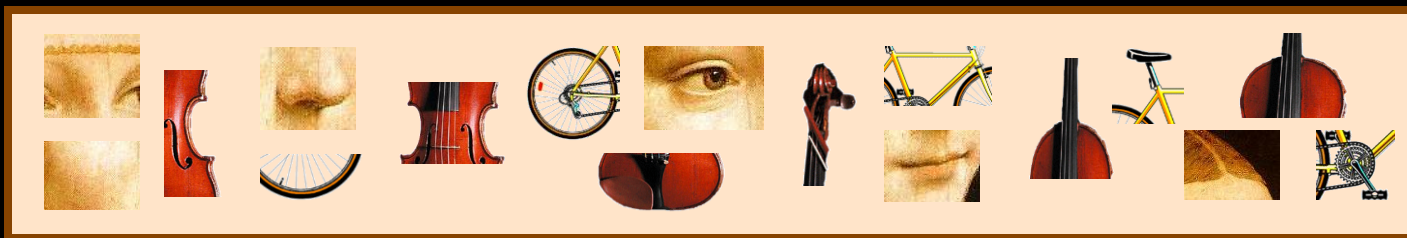
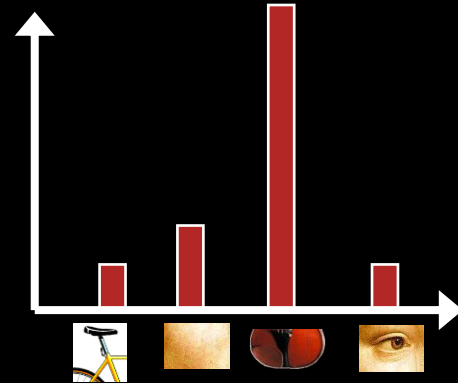
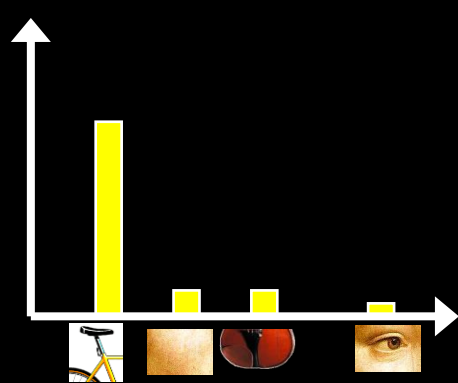
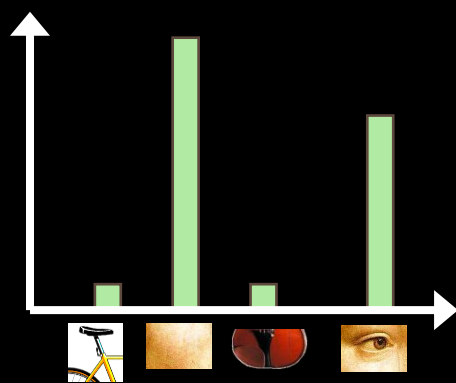


**Object**



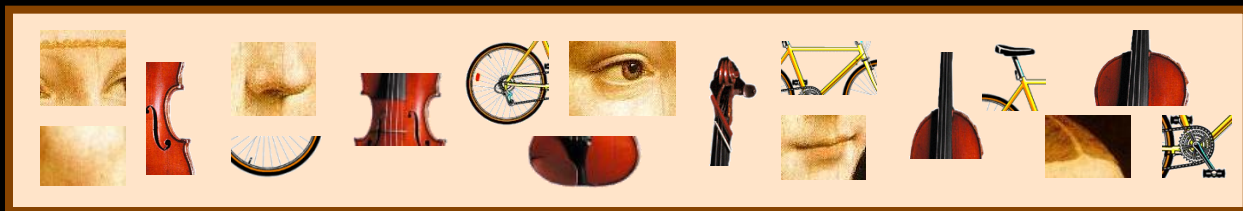
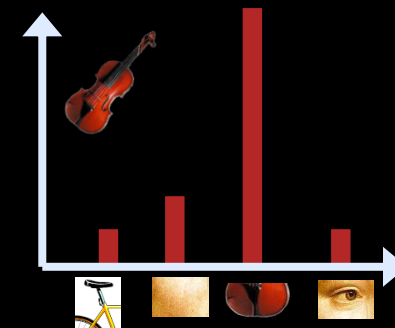
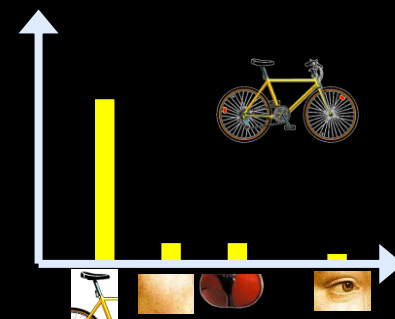
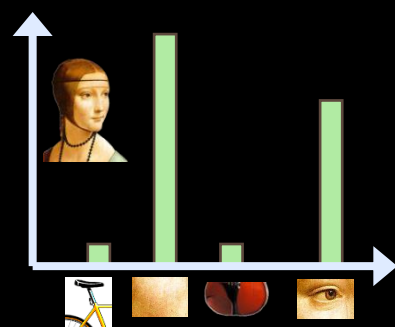
**Bag of 'words'**





# Bags of visual words

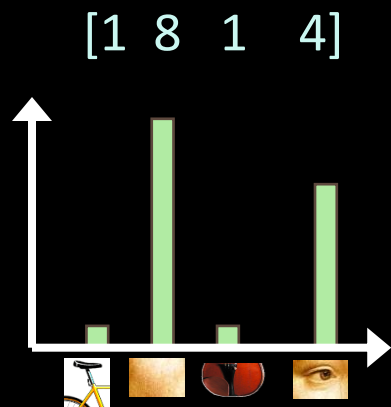
- Summarize entire image based on its distribution (histogram) of word occurrences.
- Analogous to bag of words representation commonly used for documents.



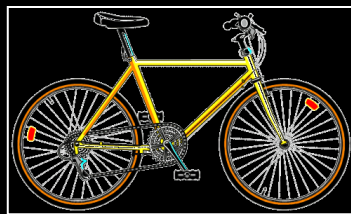
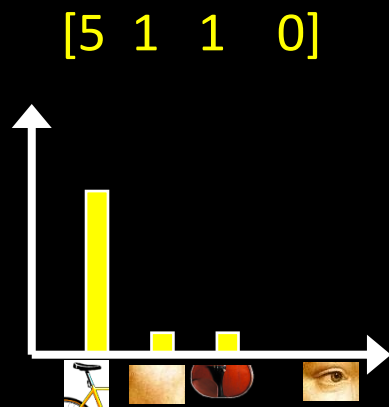
# Comparing bags of words

- Rank by normalized scalar product between their (possibly weighted) occurrence counts-- nearest neighbor search for similar images.

# Comparing bags of words



$d_j$



$q$

$$\begin{aligned} \text{sim}(d_j, q) &= \frac{\langle d_j, q \rangle}{\|d_j\| \|q\|} \\ &= \frac{\sum_{i=1}^V d_j(i) * q(i)}{\sqrt{\sum_{i=1}^V d_j(i)^2} * \sqrt{\sum_{i=1}^V q(i)^2}} \end{aligned}$$

for vocabulary of  $V$  words

# Object classification with bag of words

- Performance on Caltech 101 dataset with linear SVM on bag-of-word vectors:



True classes →	faces (frontal)	airplanes (side)	cars (rear)	cars (side)	motorbikes (side)
faces(frontal)	<b>94</b>	0.4	0.7	0	1.4
airplanes (side)	1.5	<b>96.3</b>	0.2	0.1	2.7
cars (rear)	1.9	0.5	<b>97.7</b>	0	0.9
cars(side)	1.7	1.9	0.5	<b>99.6</b>	2.3
motorbikes (side)	0.9	0.9	0.9	0.3	<b>92.7</b>
Mean ranks	1.07	1.04	1.03	1.01	1.09

[Csurka et al., '04]