

# BSB663

# Image Processing

Pinar Duygulu

# Revisit Texture

- Texture depicts spatially repeating patterns
- Many natural phenomena are textures



radishes

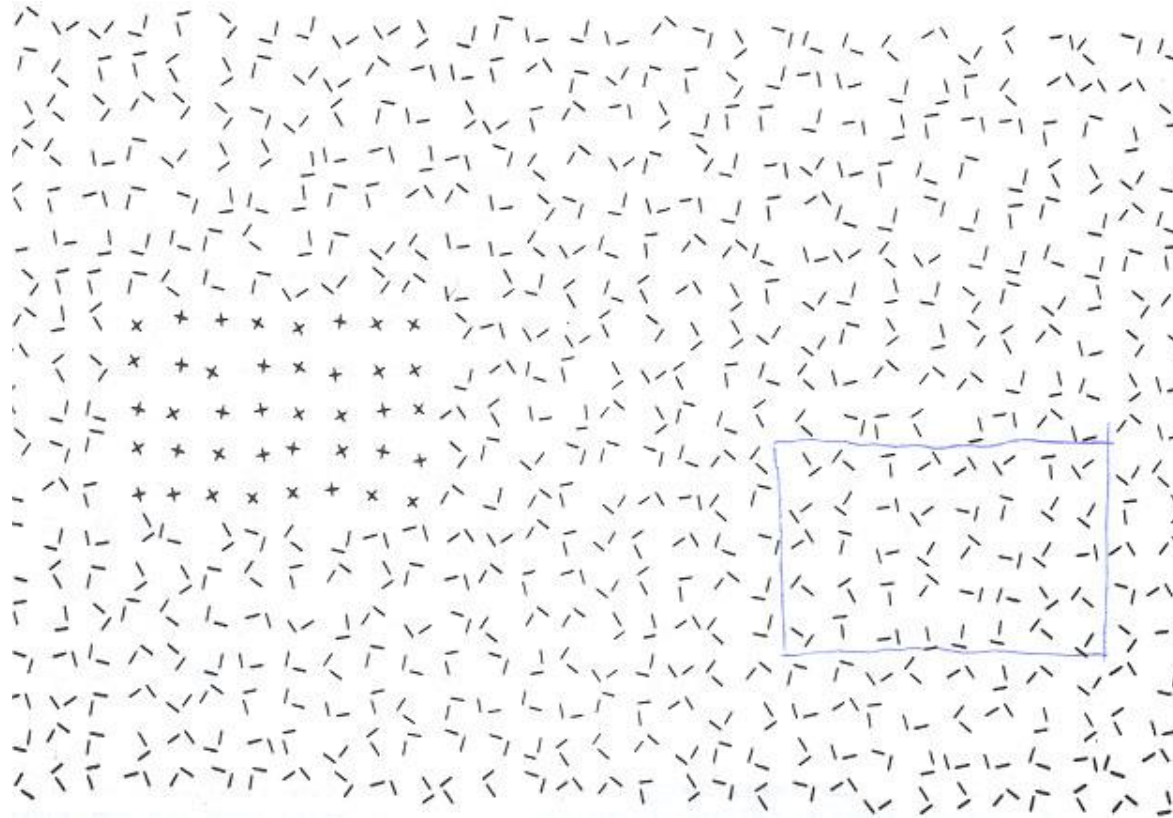


rocks



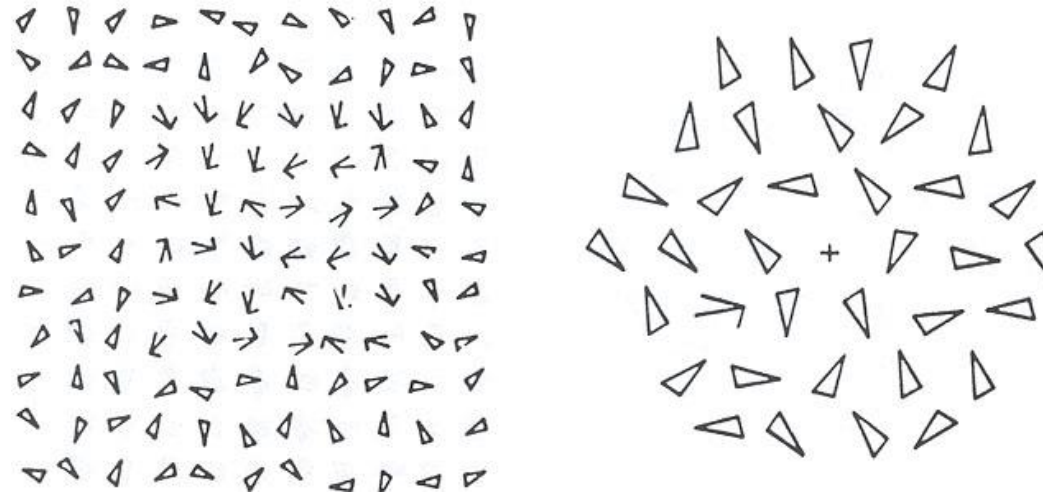
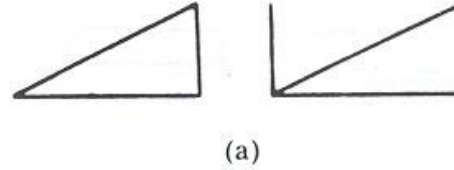
yogurt

# Texton Discrimination (Julesz)



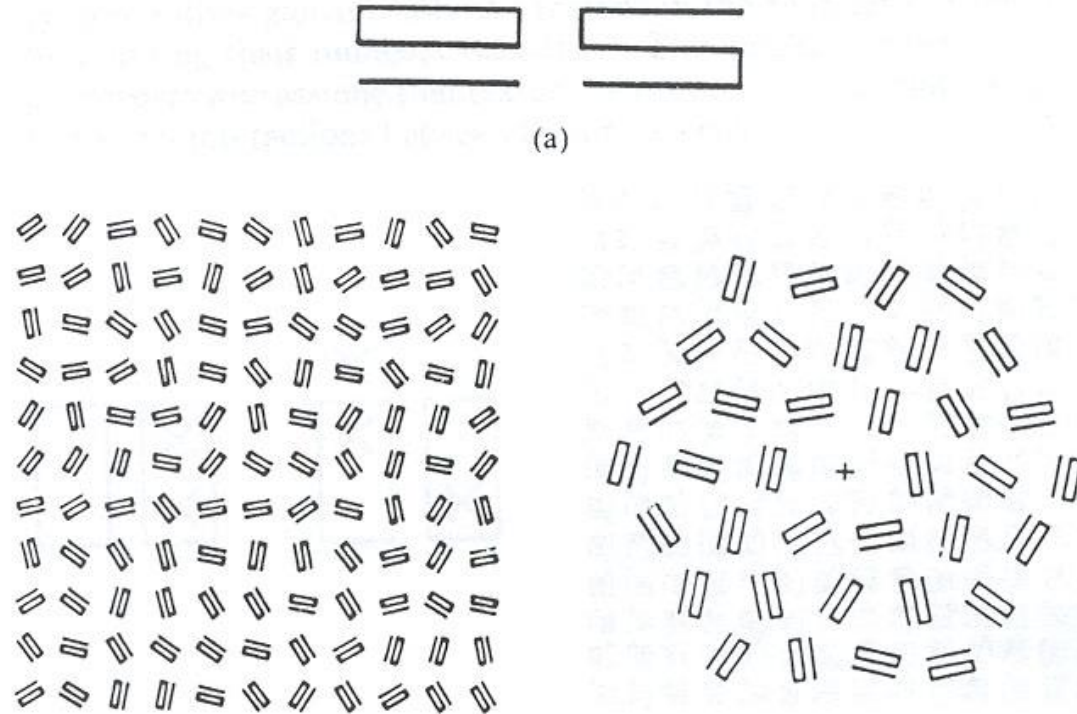
Human vision is sensitive to the difference of some types of elements and appears to be “numb” on other types of differences.

# Search Experiment I



The subject is told to detect a target element in a number of background elements.  
In this example, the detection time is independent of the number of background elements.

# Search Experiment II



In this example, the detection time is proportional to the number of background elements,  
And thus suggests that the subject is doing element-by-element scrutiny.

# Heuristic (Axiom) I

Julesz then conjectured the following axiom:

Human vision operates in two distinct modes:

## 1. Preattentive vision

parallel, instantaneous ( $\sim 100\text{--}200\text{ms}$ ), without scrutiny,  
independent of the number of patterns, covering a large visual field.

## 2. Attentive vision

serial search by focal attention in 50ms steps limited to small aperture.

Then what are the basic elements?

# Heuristic (Axiom) II

Julesz's second heuristic answers this question:

Textons are the fundamental elements in preattentive vision, including

## 1. Elongated blobs

rectangles, ellipses, line segments with attributes  
color, orientation, width, length, flicker rate.

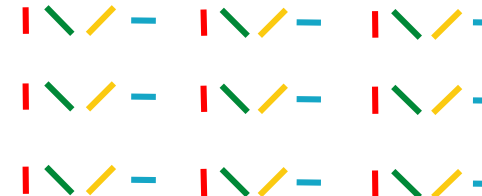
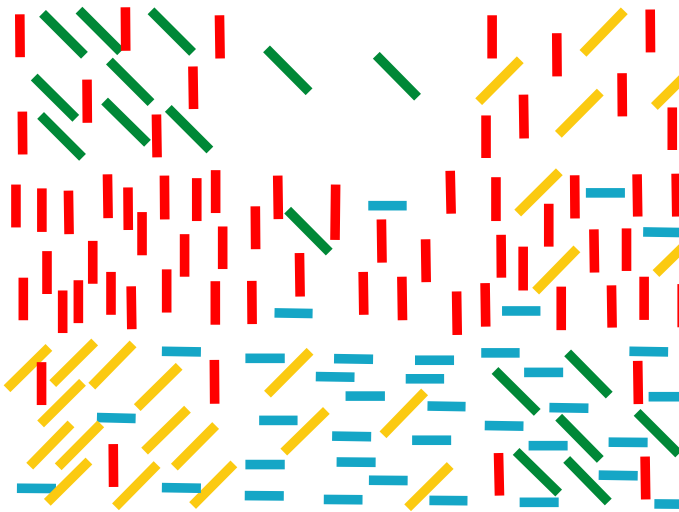
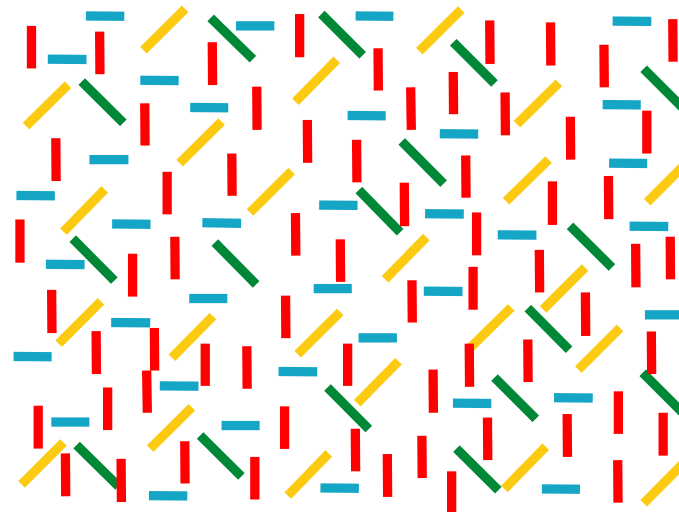
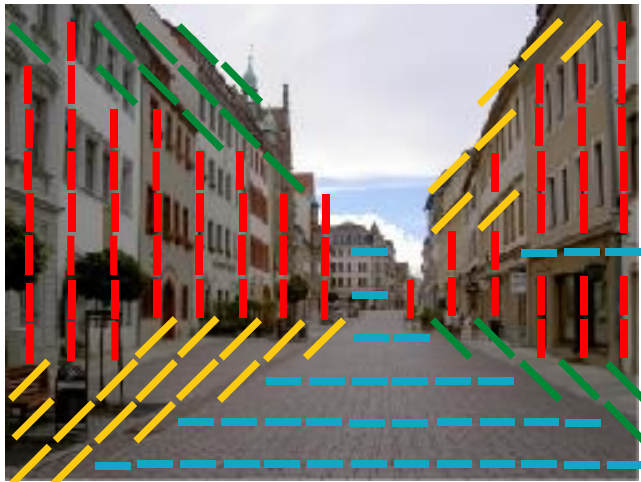
## 2. Terminators

ends of line segments.

## 3. Crossings of line segments.

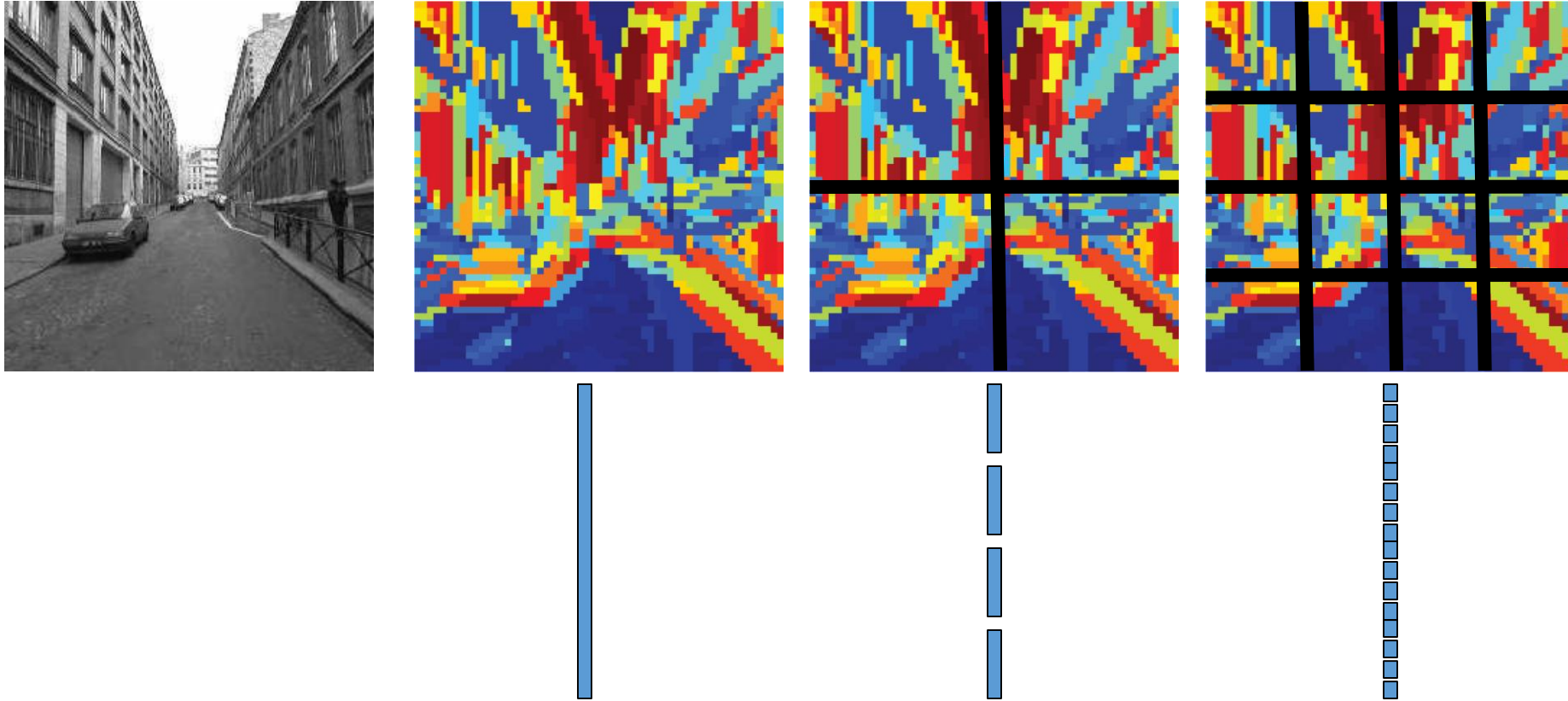
But it is worth noting that Julesz's conclusions are largely based by ensemble of artificial texture patterns. It was infeasible to synthesize natural textures for controlled experiments at that time.

# Bag of words





# Bag of words & spatial pyramid matching



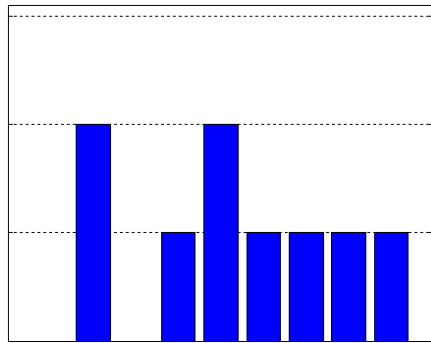
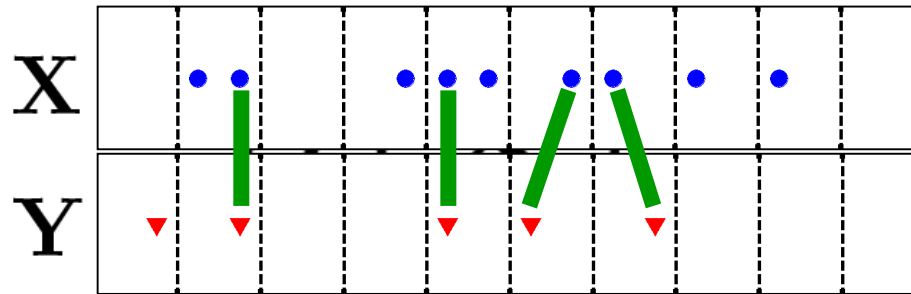
Grauman & Darel,  
S. Lazebnik, et al, CVPR 2006

Torralba, MIT

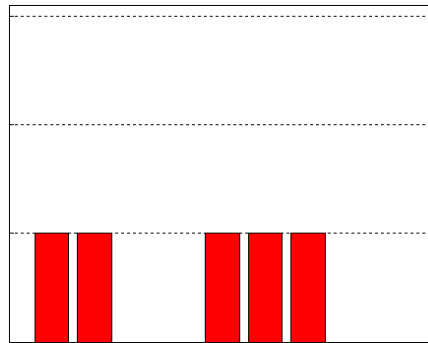
# Histogram Intersection

Histogram intersection

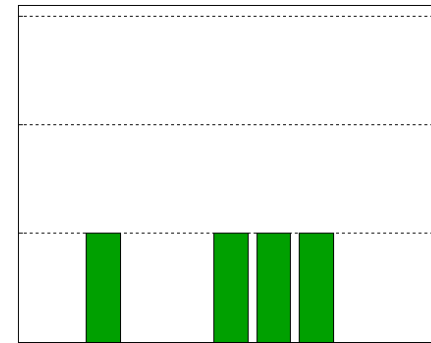
$$\mathcal{I}(H(\mathbf{X}), H(\mathbf{Y})) = \sum_{j=1}^r \min(H(\mathbf{X})_j, H(\mathbf{Y})_j)$$



$H(\mathbf{X})$



$H(\mathbf{Y})$



$$\mathcal{I}(H(\mathbf{X}), H(\mathbf{Y})) = 4$$

# Histogram based distances

Given two histograms:  $h_1, h_2$ , such that  $\text{sum}(h_1)=\text{sum}(h_2)=1$

- Euclidean

$$D(h_1, h_2) = \text{sum} ((h_1 - h_2).^2)$$

- Histogram intersection

$$D(h_1, h_2) = 1 - \text{sum} (\min (h_1, h_2))$$

- $\chi^2$

$$D(h_1, h_2) = \text{sum}((h_1-h_2).^2 ./ (h_1+h_2))$$

(using Matlab notation)

# Capturing the “essence” of texture

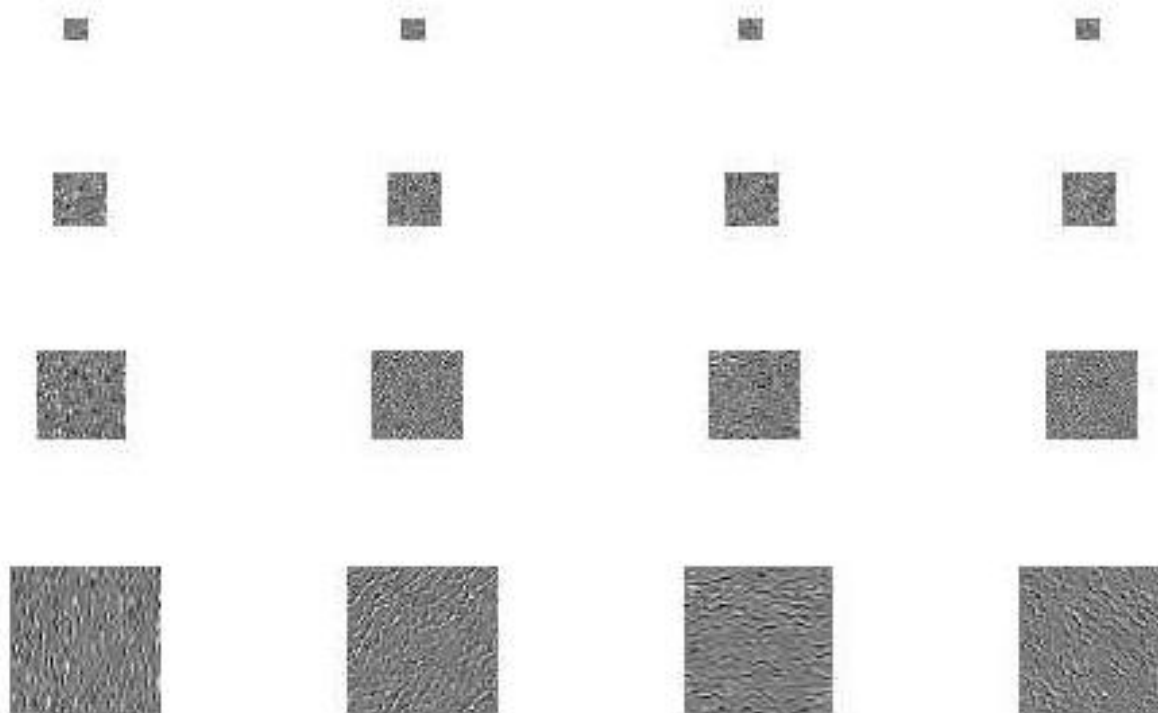
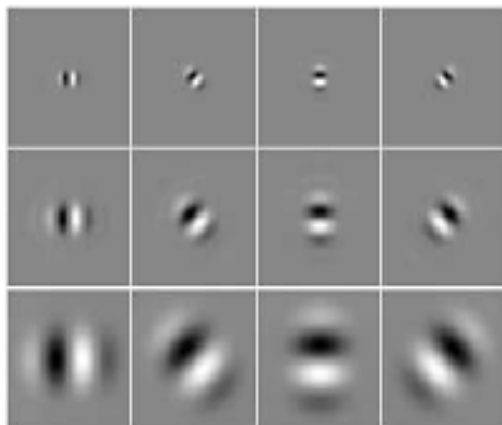
- ...for real images



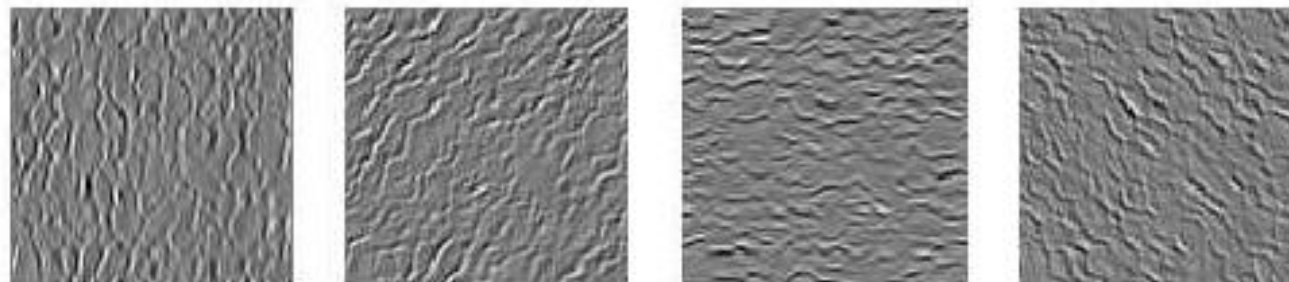
- We don't want an actual texture realization, we want a texture invariant
- What are the tools for capturing statistical properties of some signal?

# Multi-scale filter decomposition

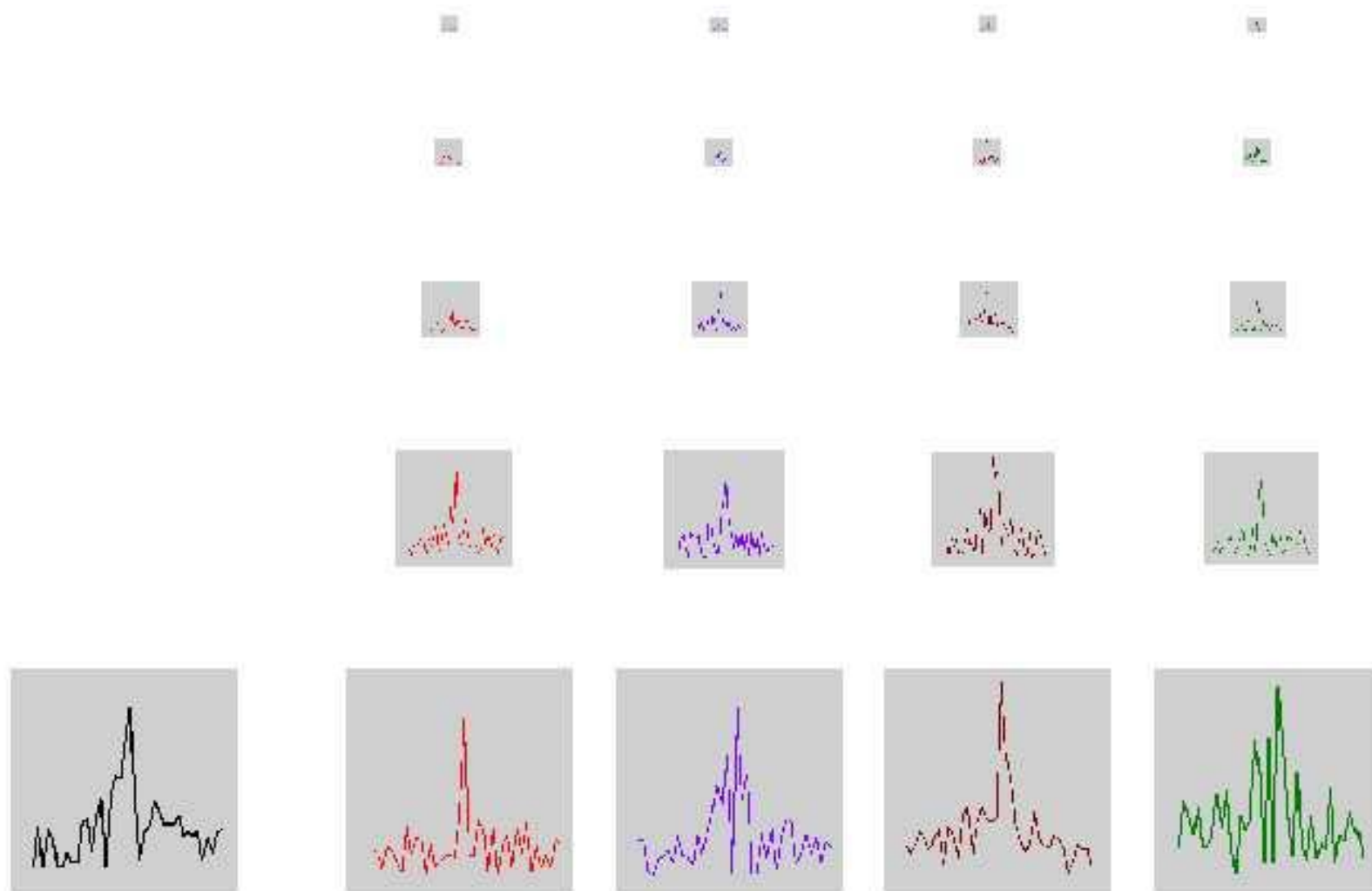
Filter bank



Input image

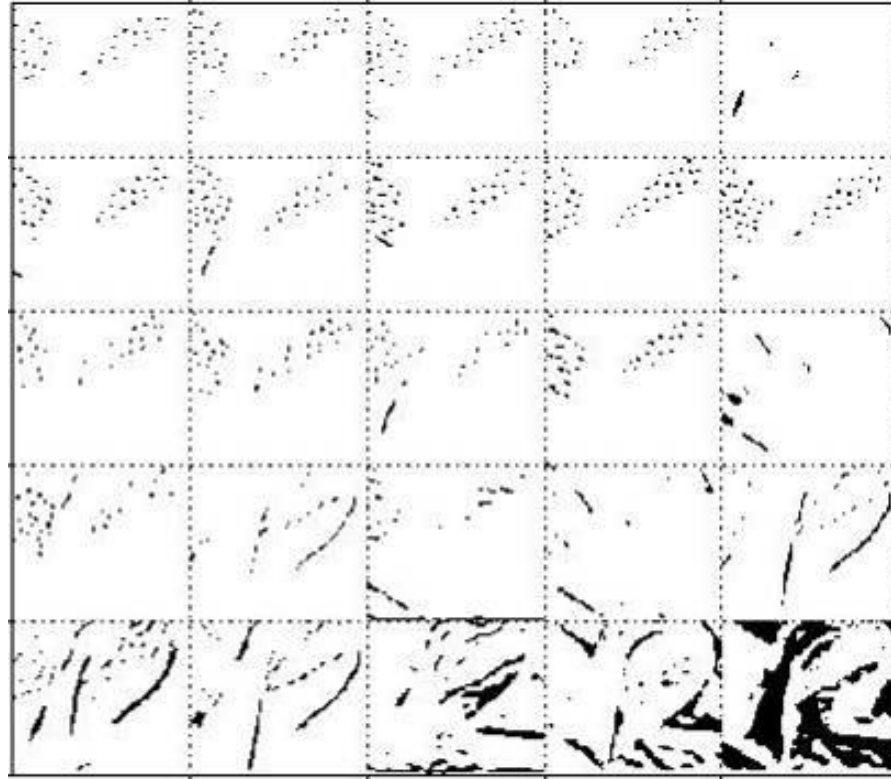
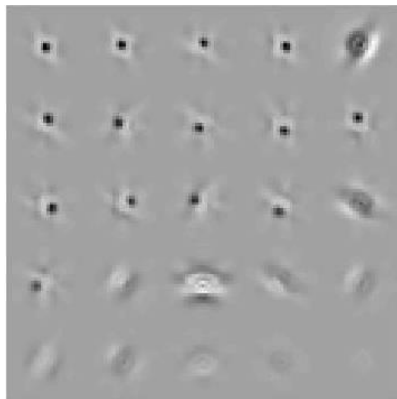
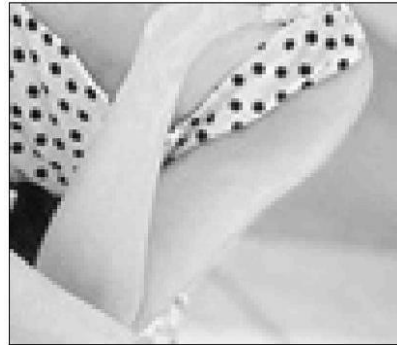


# Filter response histograms

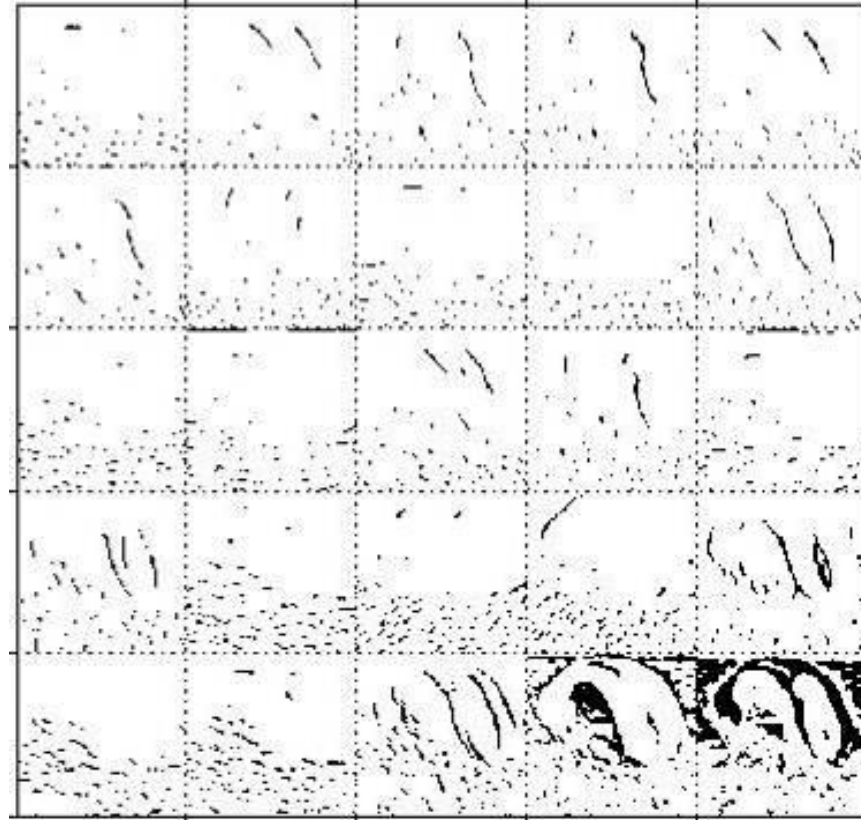
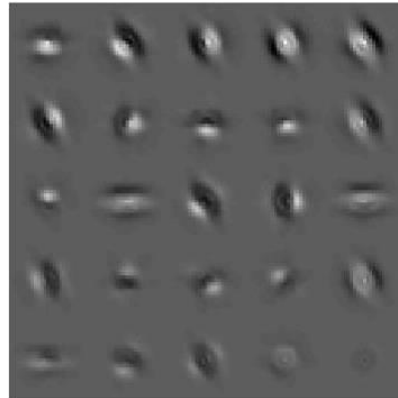


# Textons (Malik et al, IJCV 2001)

- K-means on vectors of filter responses

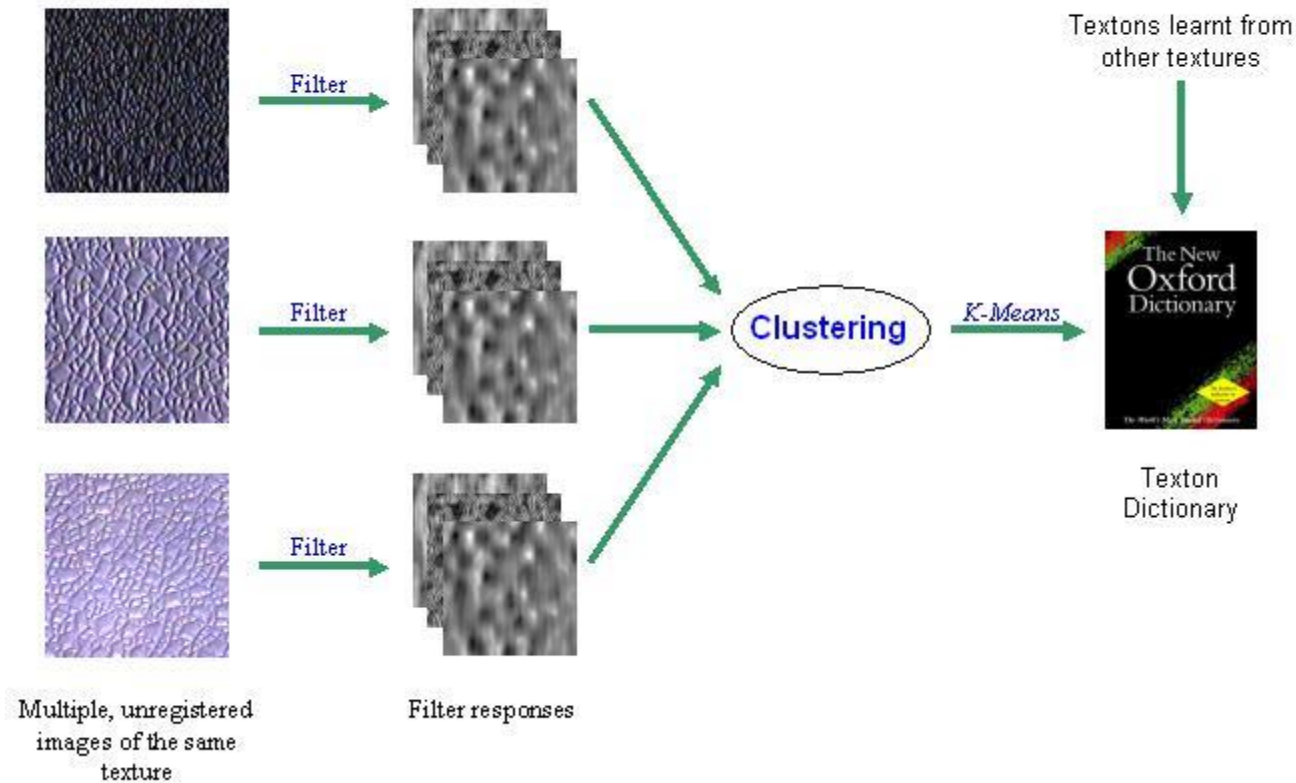


# Textons (cont.)



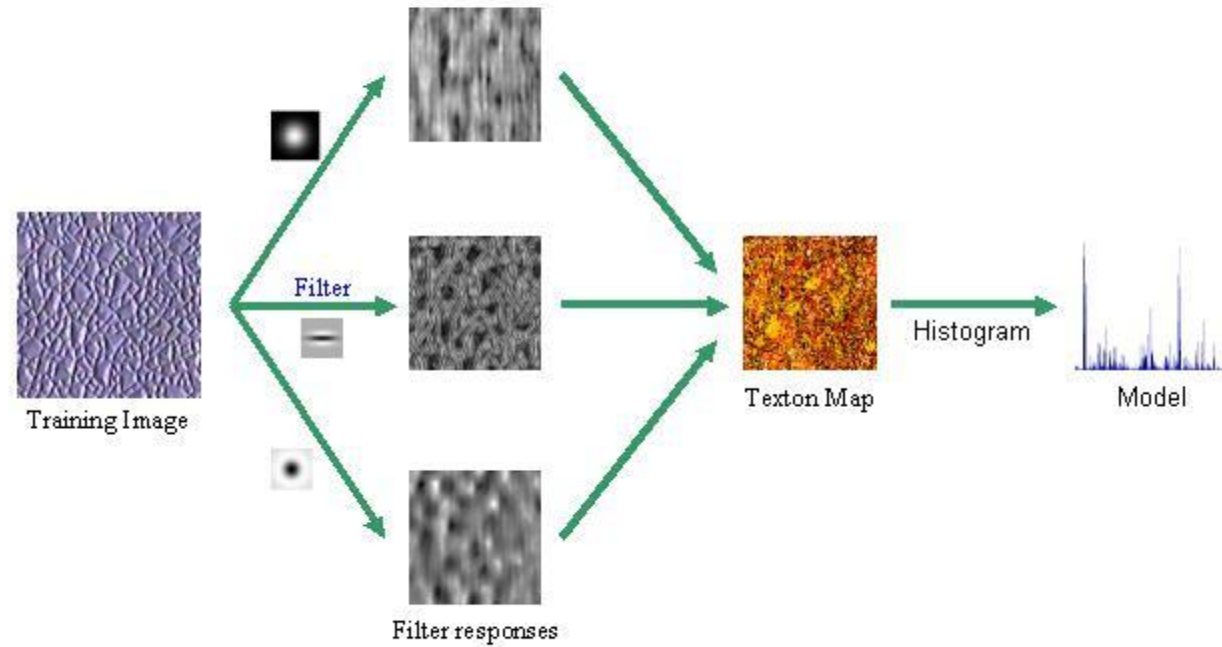


## Modelling I – Learning the Texton Dictionary



## Modelling II – Multiple Models Per Texture

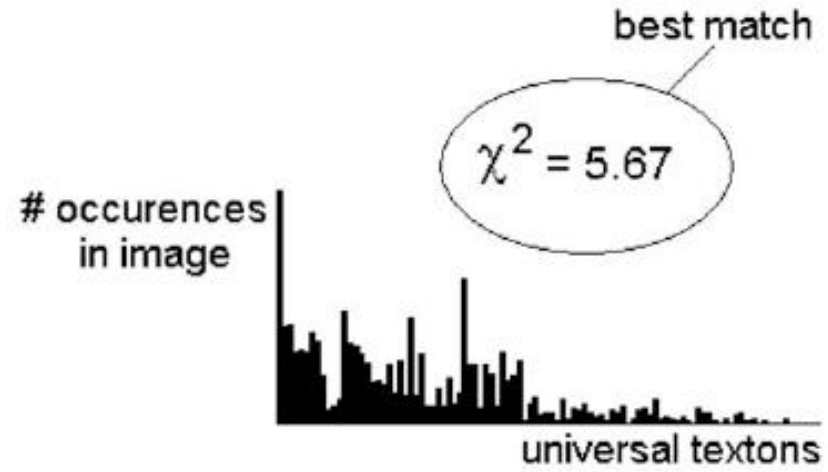
---



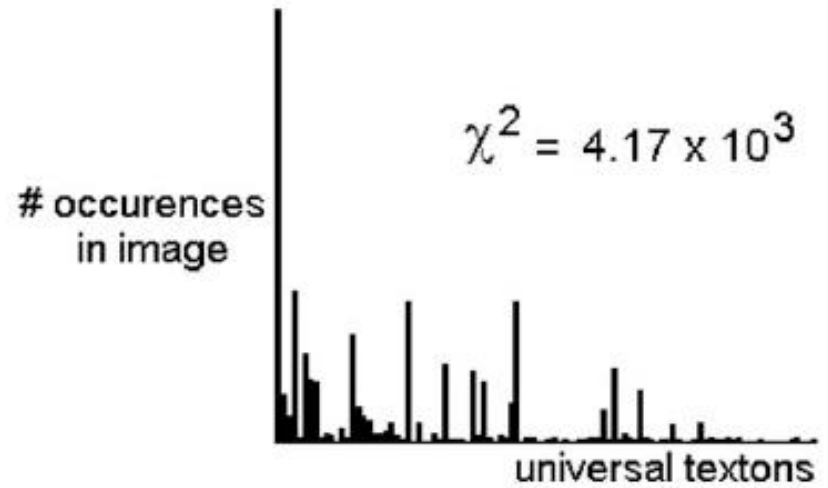
# Textons



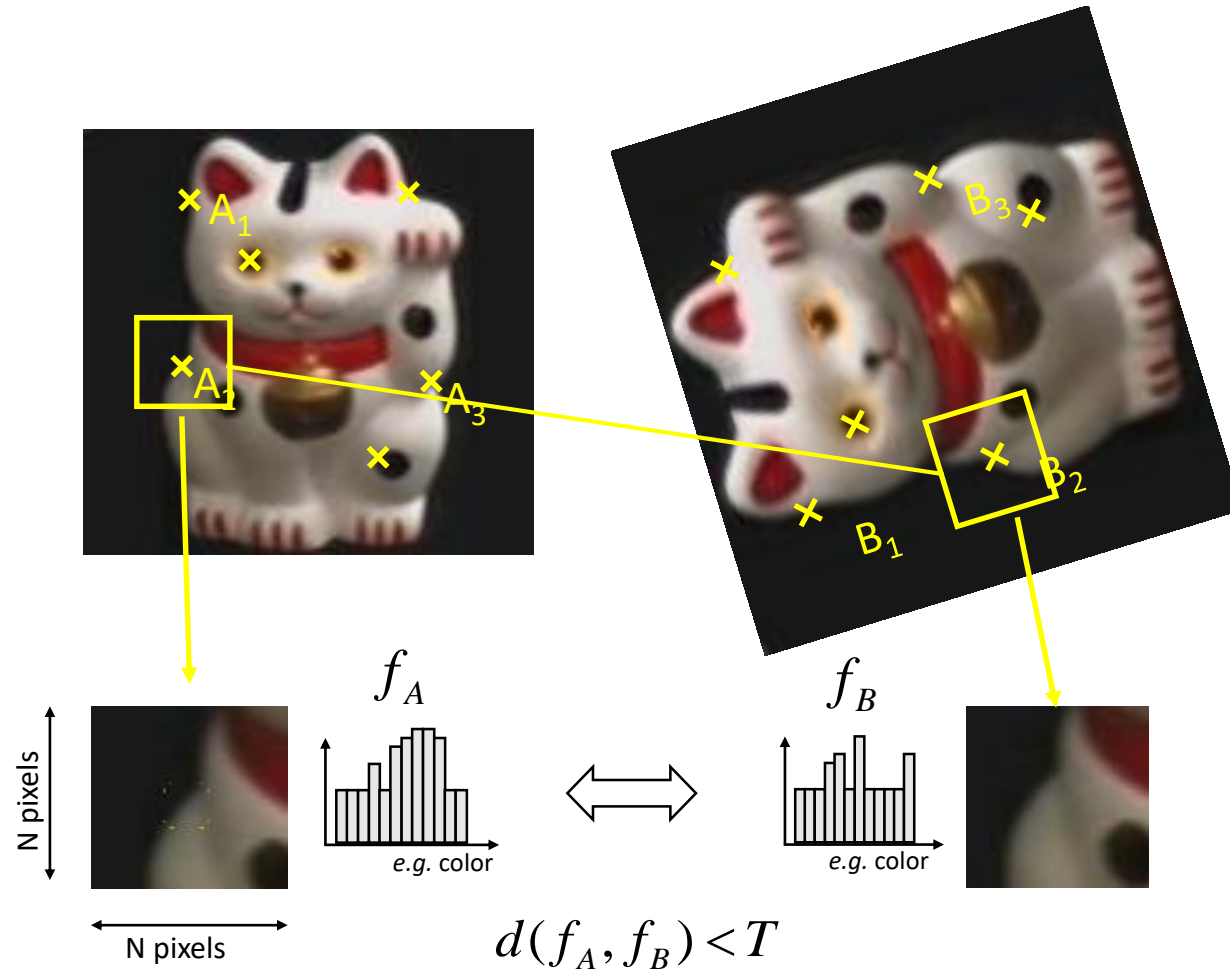
label = bedroom



label = beach

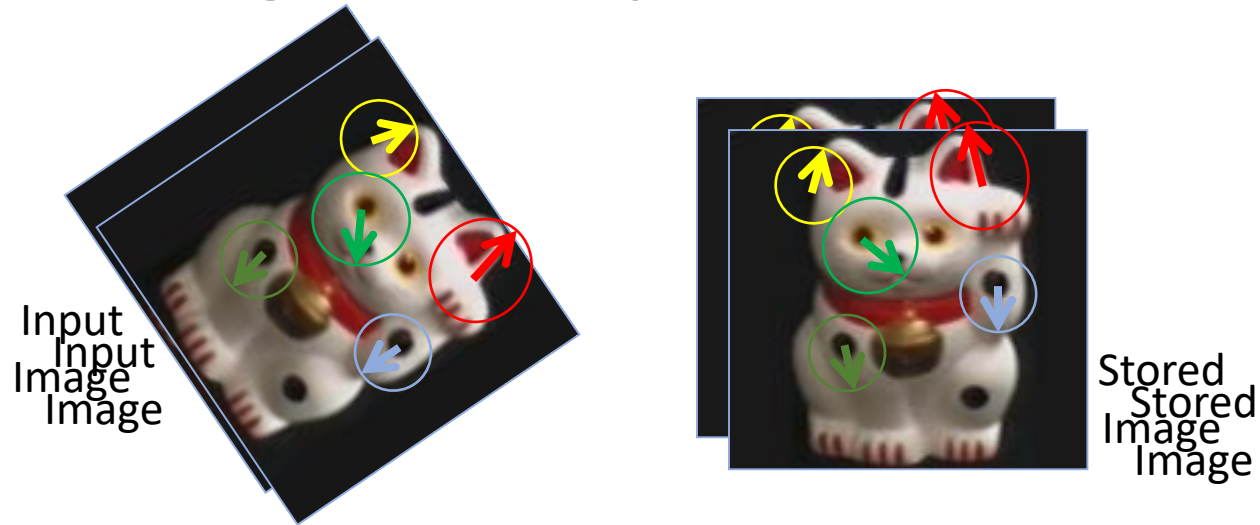


# Revisit Keypoint Matching



1. Find a set of distinctive keypoints
2. Define a region around each keypoint
3. Extract and normalize the region content
4. Compute a local descriptor from the normalized region
5. Match local descriptors

# Finding the objects (overview)



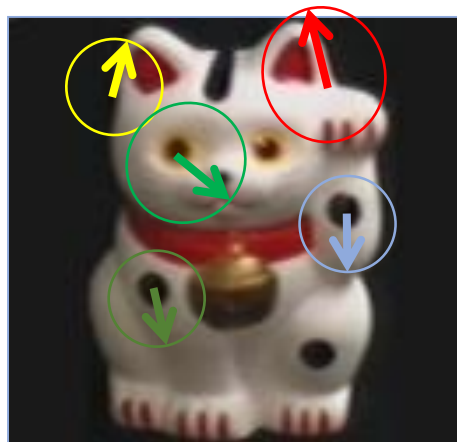
1. Match interest points from input image to database image
2. Matched points vote for rough position/orientation/scale of object
3. Find triplets of position/orientation/scale that have at least three votes
4. Compute affine registration and matches using iterative least squares with outlier check
5. Report object if there are at least  $T$  matched points

# Matching Keypoints

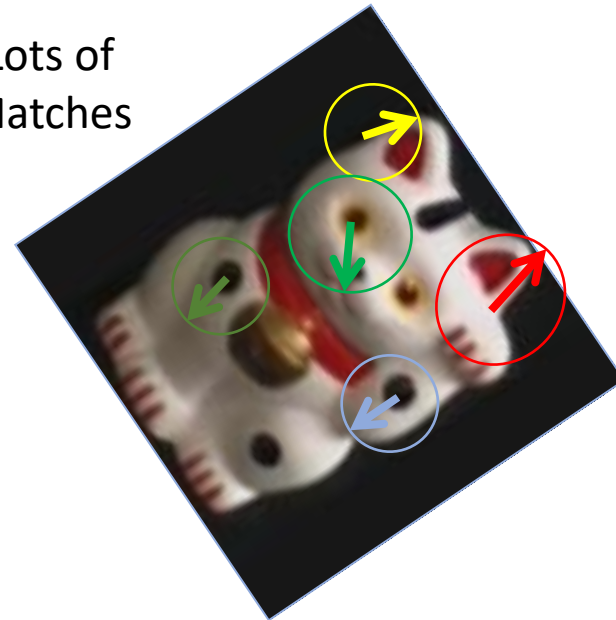
- Want to match keypoints between:
  1. Query image
  2. Stored image containing the object
- Given descriptor  $x_0$ , find two nearest neighbors  $x_1, x_2$  with distances  $d_1, d_2$
- $x_1$  matches  $x_0$  if  $d_1/d_2 < 0.8$ 
  - This gets rid of 90% false matches, 5% of true matches in Lowe's study

# Simple idea

See how many keypoints are close to keypoints in each other image



Lots of Matches



Few or No Matches

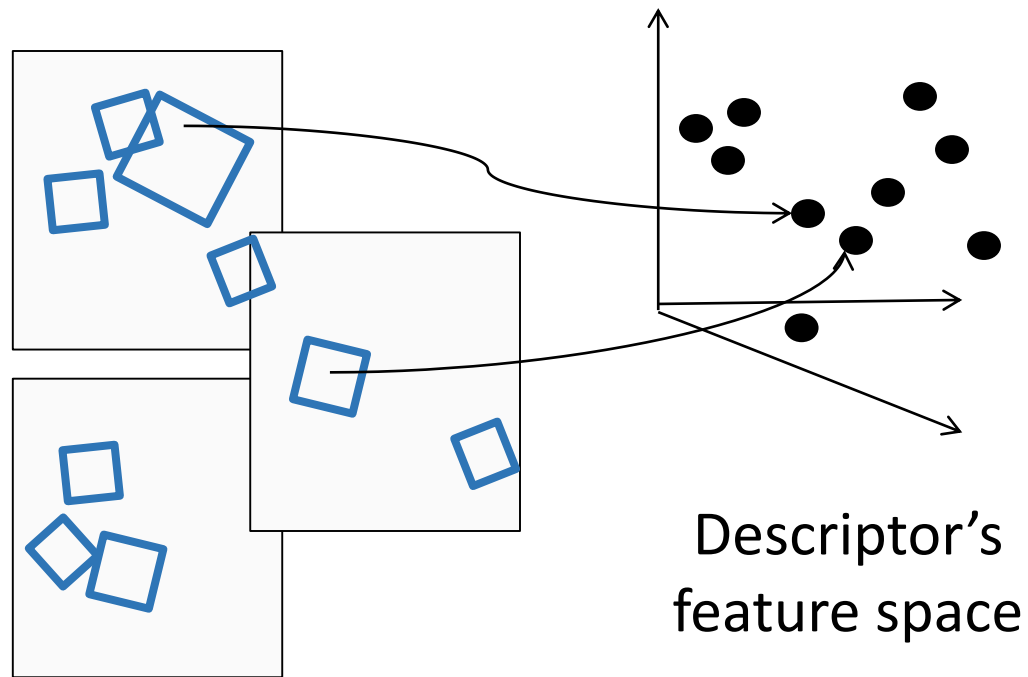


But this will be really, really slow!

Hayes, Brown

# Indexing local features

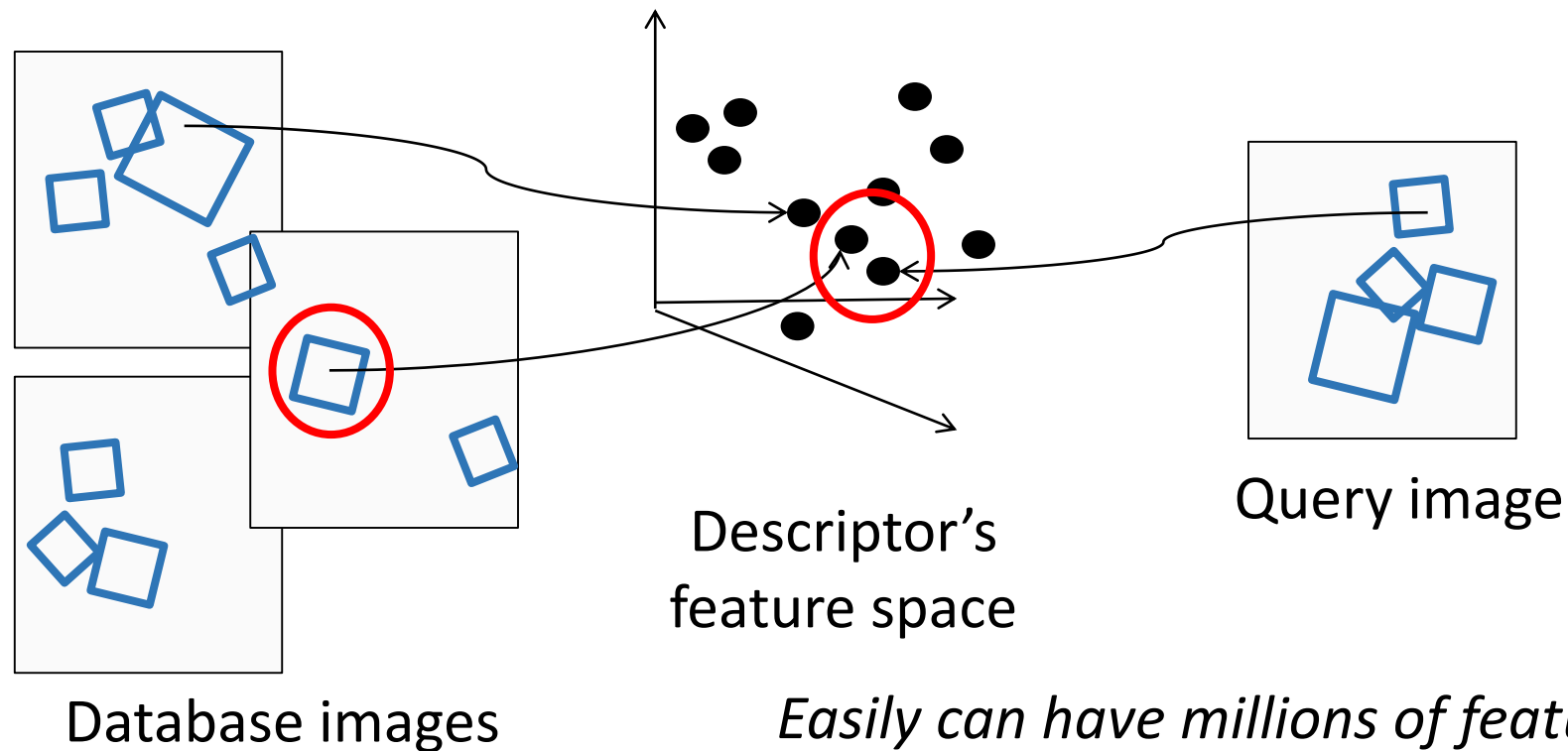
- Each patch / region has a descriptor, which is a point in some high-dimensional feature space (e.g., SIFT)





# Indexing local features

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



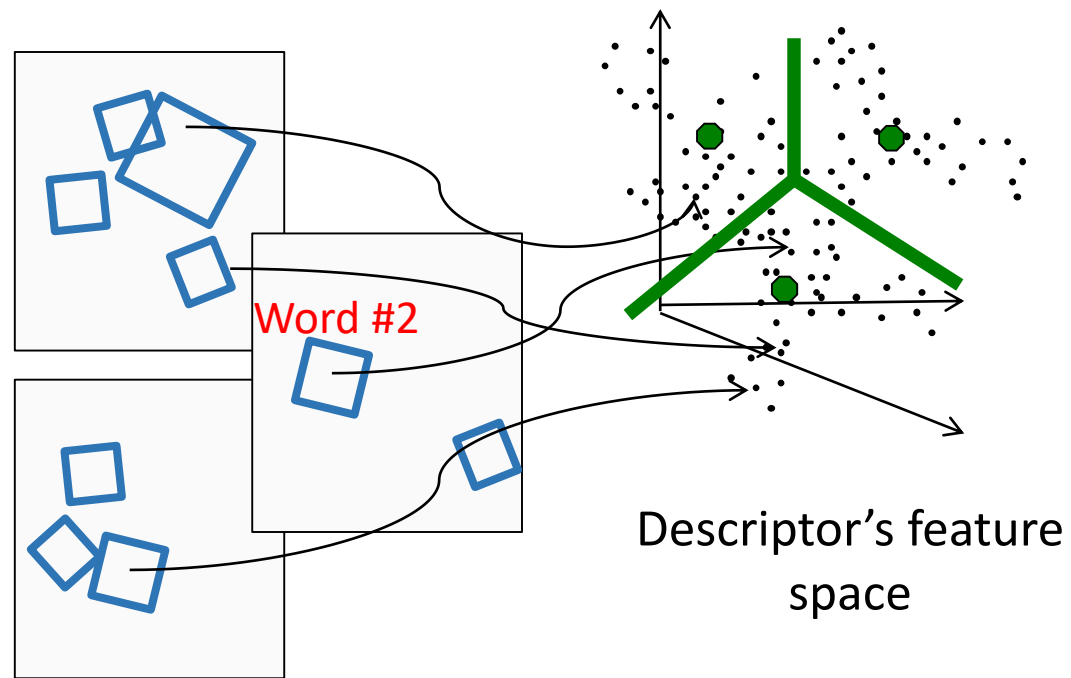
# Indexing local features: inverted file index

Index	
"Along I-75," From Detroit to Florida; <i>inside back cover</i>	Butterfly Center, McGuire; 134
"Drive I-95," From Boston to Florida; <i>inside back cover</i>	CAA (see AAA)
1929 Spanish Trail Roadway; 101-102,104	CCC, The; 111,113,115,135,142
511 Traffic Information; 83	Ca d'Zan; 147
A1A (Barrier Isl) - I-95 Access; 86	Caloosahatchee River; 152
AAA (and CAA); 83	Name; 150
AAA National Office; 88	Canaveral Natnl Seashore; 173
Abbreviations,	Cannon Creek Airpark; 130
Colored 25 mile Maps; cover	Canopy Road; 106,169
Exit Services; 196	Cape Canaveral; 174
Travelogue; 85	Castillo San Marcos; 169
Africa; 177	Cave Diving; 131
Agricultural Inspection Stns; 126	Cayo Costa, Name; 150
Ah-Tah-Thi-Ki Museum; 160	Celebration; 93
Air Conditioning, First; 112	Charlotte County; 149
Alabama; 124	Charlotte Harbor; 150
Alachua; 132	Chautauqua; 116
County; 131	ChIPLEY; 114
Alafia River; 143	Name; 115
Alapaha, Name; 126	Choctawatchee, Name; 115
Alfred B Maclay Gardens; 106	Circus Museum, Ringling; 147
Alligator Alley; 154-155	Citrus; 88,97,130,136,140,180
Alligator Farm, St Augustine; 169	CityPlace, W Palm Beach; 180
Alligator Hole (definition); 157	City Maps,
Alligator, Buddy; 155	Fl Lauderdale Expwys; 194-195
Alligators; 100,135,138,147,156	Jacksonville; 163
Anastasia Island; 170	Kissimmee Expwys; 192-193
Anhaica; 108-109,146	Miami Expressways; 194-195
Apalachicola River; 112	Orlando Expressways; 192-193
Appleton Mus of Art; 136	Pensacola; 26
Aquifer; 102	Tallahassee; 191
Arabian Nights; 94	Tampa-St. Petersburg; 63
Art Museum, Ringling; 147	St. Augustine; 191
Aruba Beach Cafe; 183	Civil War; 100,108,127,138,141
Aucilla River Project; 106	Clearwater Marine Aquarium; 187
Babcock-Web WMA; 151	Collier County; 154
Bahia Mar Marina; 184	Collier, Barron; 152
Baker County; 99	Colonial Spanish Quarters; 168
Barefoot Mailmen; 182	Columbia County; 101,128
Barge Canal; 137	Coquina Building Material; 165
Bee Line Expy; 80	Corkscrew Swamp, Name; 154
Belz Outlet Mall; 89	Cowboys; 85
Bernard Castro; 136	Crab Trap II; 144
Big "I"; 165	Cracker, Florida; 88,95,132
Big Cypress; 155,158	Crosstown Expy; 11,35,98,143
Big Foot Monster; 105	Cuban Bread; 184
Billie Swamp Safari; 160	Dade Battlefield; 140
Blackwater River SP; 117	Dade, Maj. Francis; 139-140,161
Blue Angels	Dania Beach Hurricane; 184
	Daniel Boone, Florida Walk; 117
	Daytona Beach; 172-173
	De Land; 87
	Driving Lanes; 85
	Duval County; 163
	Eau Gallie; 175
	Edison, Thomas; 152
	Eglin AFB; 116-118
	Eight Reale; 176
	Ellenton; 144-145
	Emanuel Point Wreck; 120
	Emergency Callboxes; 83
	Epiphytes; 142,148,157,159
	Escambia Bay; 119
	Bridge (I-10); 119
	County; 120
	Estero; 153
	Everglade,90,95,139-140,154-160
	Draining of; 156,181
	Wildlife MA; 160
	Wonder Gardens; 154
	Falling Waters SP; 115
	Fantasy of Flight; 95
	Fayer Dykes SP; 171
	Fires, Forest; 166
	Fires, Prescribed ; 148
	Fisherman's Village; 151
	Flagler County; 171
	Flagler, Henry; 97,165,167,171
	Florida Aquarium; 186
	Florida,
	12,000 years ago; 187
	Cavern SP; 114
	Map of all Expressways; 2-3
	Mus of Natural History; 134
	National Cemetery ; 141
	Part of Africa; 177
	Platform; 187
	Sheriff's Boys Camp; 126
	Sports Hall of Fame; 130
	Sun 'n Fun Museum; 97
	Supreme Court; 107
	Florida's Turnpike (FTP), 178,189
	25 mile Strip Maps; 66
	Administration; 189
	Coin System; 190
	Exit Services; 189
	HEFT; 76,161,190
	History; 189
	Names; 189
	Service Plazas; 190
	Spur SR91; 76
	Ticket System; 190
	Toll Plazas; 190
	Ford, Henry; 152

- 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

- Map high-dimensional descriptors to tokens/words by quantizing the feature space



- Quantize via clustering, let cluster centers be the prototype "words"
- Determine which word to assign to each new image region by finding the closest cluster center.

# Visual words

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

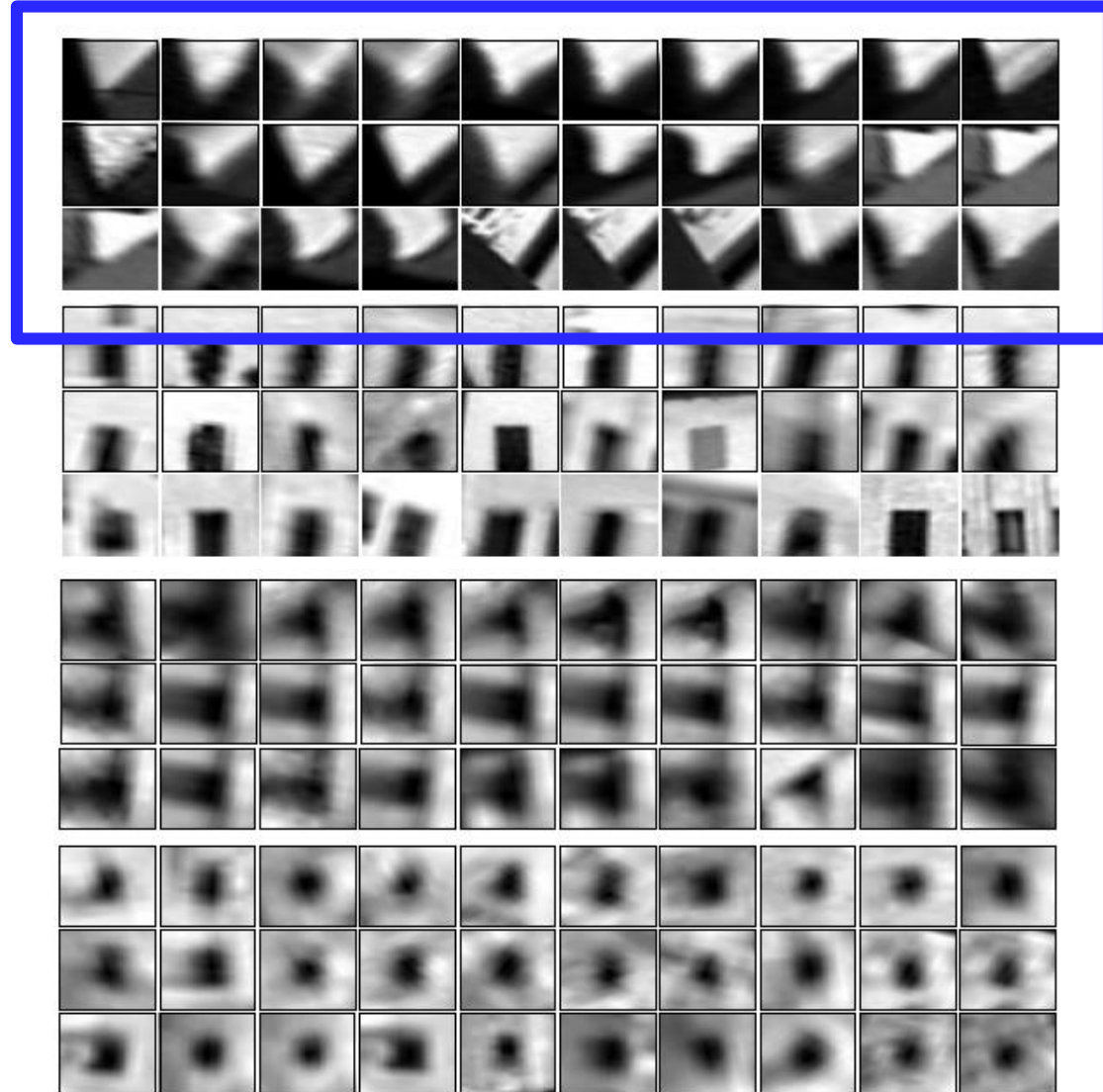
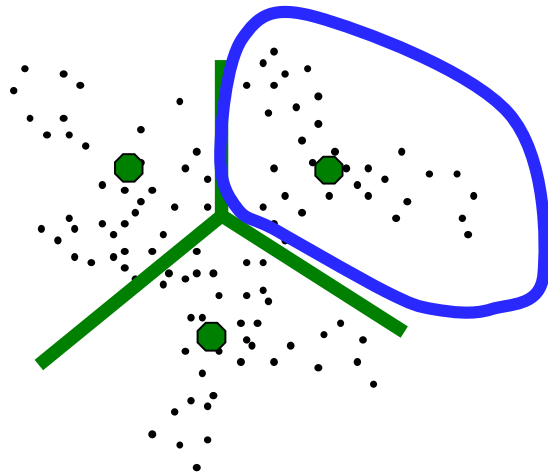


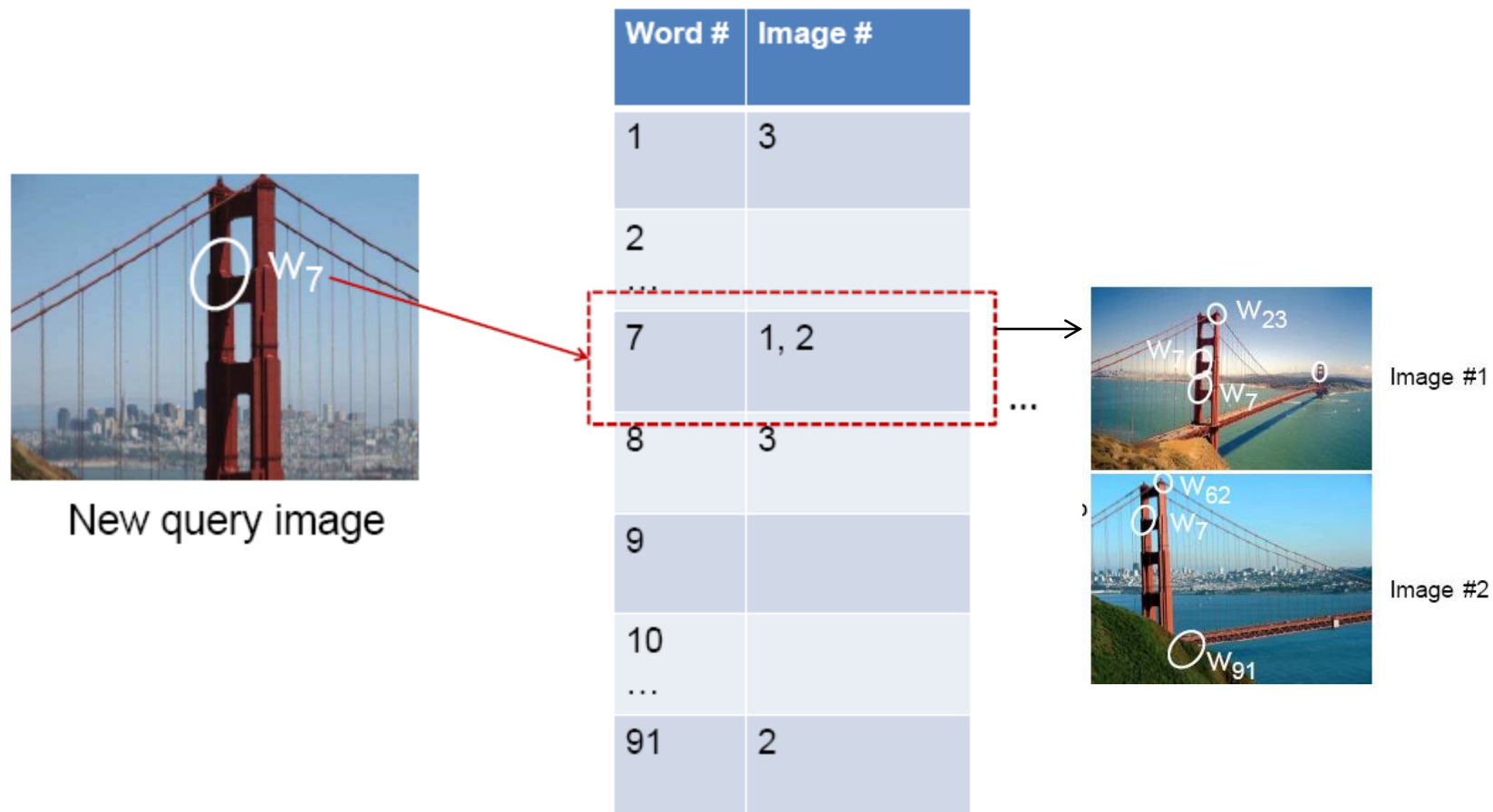
Figure from Sivic & Zisserman, ICCV 2003 Kristen Grauman

# Inverted file index



- Database images are loaded into the index mapping words to image numbers

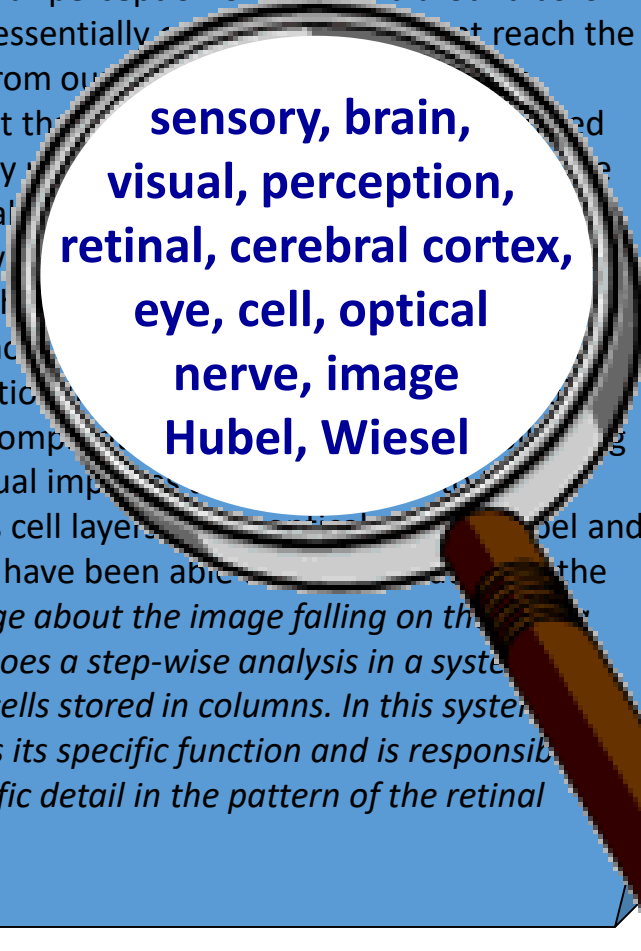
# Inverted file index



- New query image is mapped to indices of database images that share a word.

# 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 visual impressions that reach the brain from our eyes. At the point by which the cerebral cortex is upon which the visual image falls. Through the work of Hubel and Wiesel have been able to show that the message about the image falling on the retina undergoes a step-wise 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 up from \$32bn in 2004 to \$90bn this year, a 30% jump in exports and a 18% rise in imports. The ministry said it will further increase exports to \$100bn. China's government has decided to raise the surplus to \$100bn. One factor is the surging value of the yuan. Xiaochua said the yuan has risen more to be valued at 7.76 yuan per dollar, stayed within the 7.3-7.75 band. The value of the yuan rose 10% in July and permitted it to be valued to trade freely. However, Beijing has made it clear that it will take its time and tread carefully in allowing the yuan to rise further in value.



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

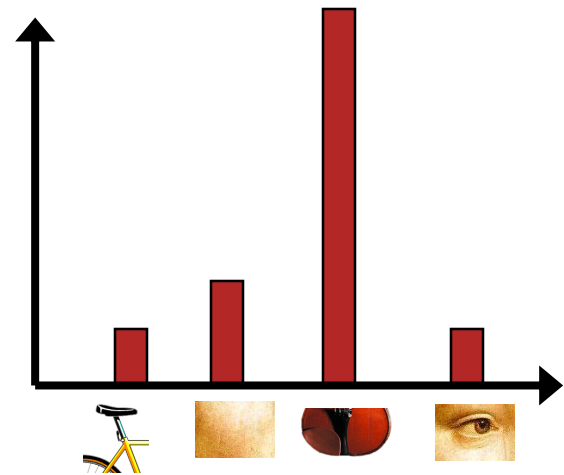
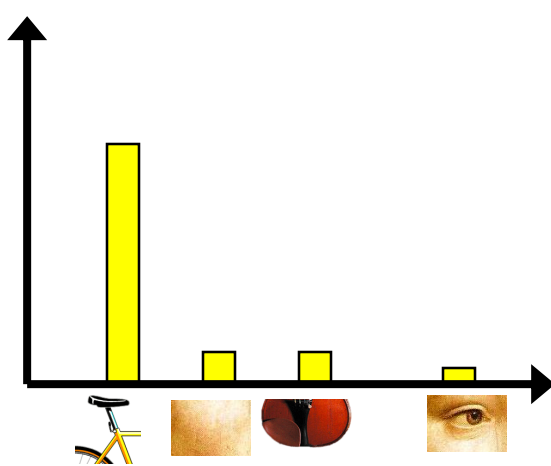
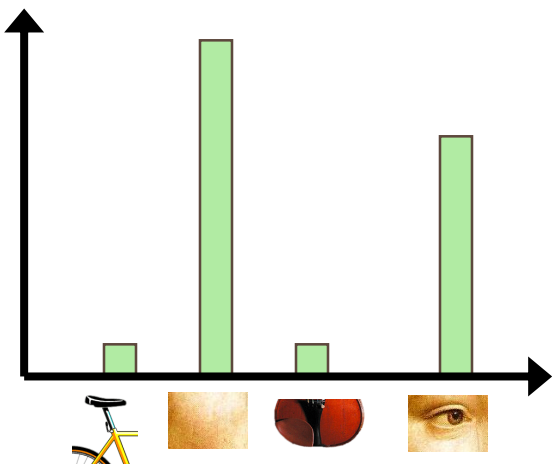
**Object**



**Bag of 'words'**

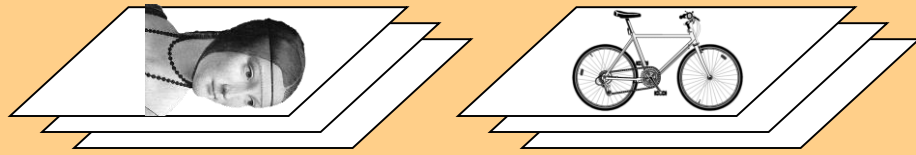




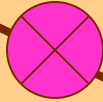


Alyosha Efros, CMU

# learning



feature detection  
& representation



codewords dictionary

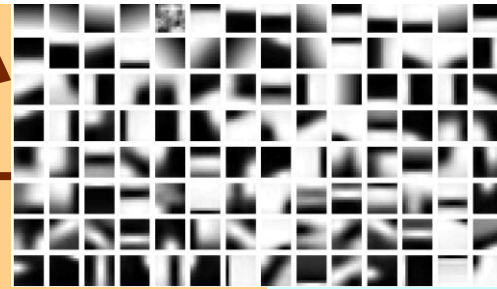
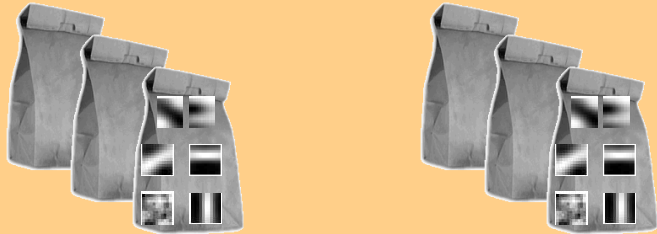
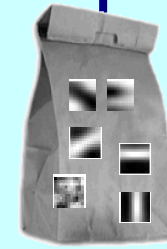
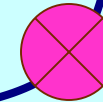


image representation



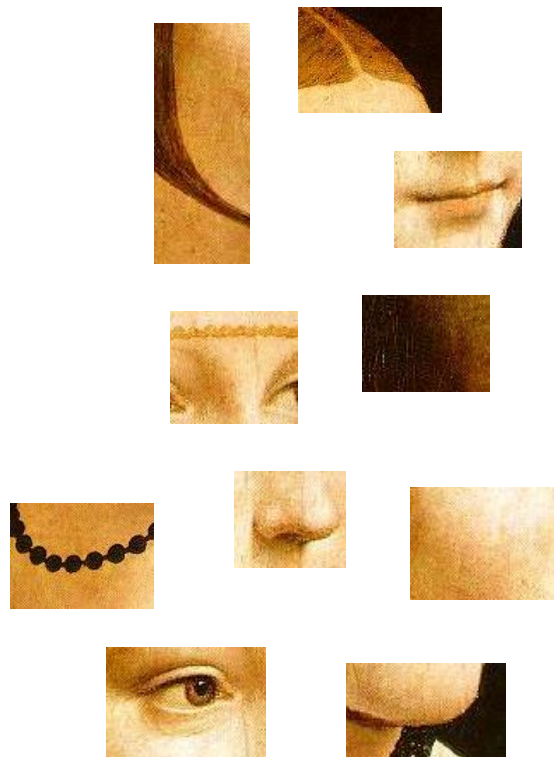
**category models  
(and/or) classifiers**

# recognition



**category  
decision**

# 1. Feature detection and representation



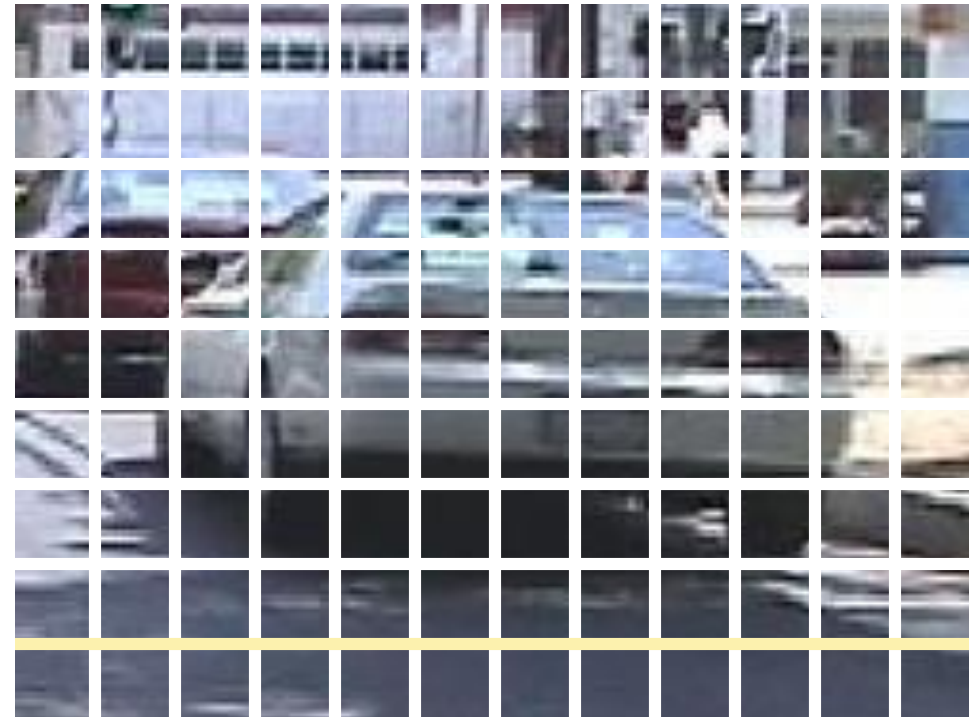
# Feature detection

- Sliding Window
  - Leung et al, 1999
  - Viola et al, 1999
  - Renninger et al 2002



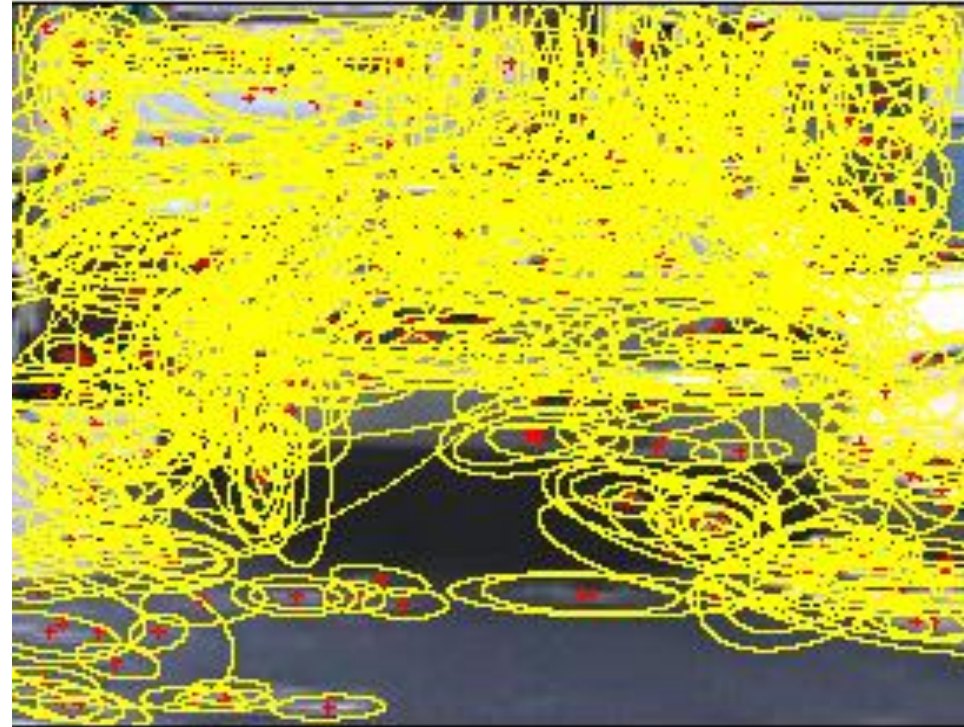
# Feature detection

- Sliding Window
  - Leung et al, 1999
  - Viola et al, 1999
  - Renninger et al 2002
- Regular grid
  - Vogel et al. 2003
  - Fei-Fei et al. 2005



# Feature detection

- Sliding Window
  - Leung et al, 1999
  - Viola et al, 1999
  - Renninger et al 2002
- Regular grid
  - Vogel et al. 2003
  - Fei-Fei et al. 2005
- Interest point detector
  - Csurka et al. 2004
  - Fei-Fei et al. 2005
  - Sivic et al. 2005



# Feature detection

- Sliding Window
  - Leung et al, 1999
  - Viola et al, 1999
  - Renninger et al 2002
- Regular grid
  - Vogel et al. 2003
  - Fei-Fei et al. 2005
- Interest point detector
  - Csurka et al. 2004
  - Fei-Fei et al. 2005
  - Sivic et al. 2005
- Other methods
  - Random sampling (Ullman et al. 2002)
  - Segmentation based patches
    - Barnard et al. 2003, Russell et al 2006, etc.)

# Feature Representation

Visual words, aka textons, aka keypoints:

K-means clustered pieces of the image

- Various Representations:
  - Filter bank responses
  - Image Patches
  - SIFT descriptors

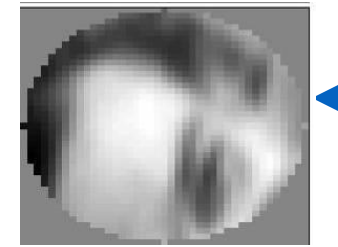
All encode more-or-less the same thing...



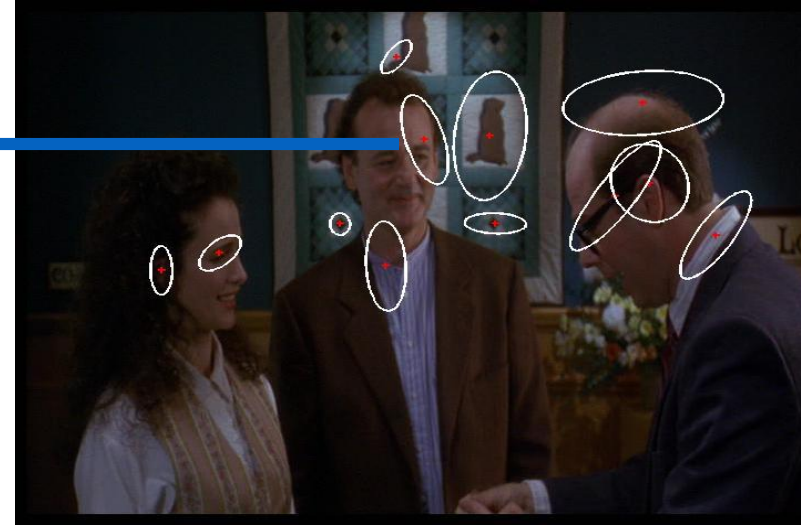
# Interest Point Features



**Compute SIFT  
descriptor**  
[Lowe'99]



**Normalize patch**



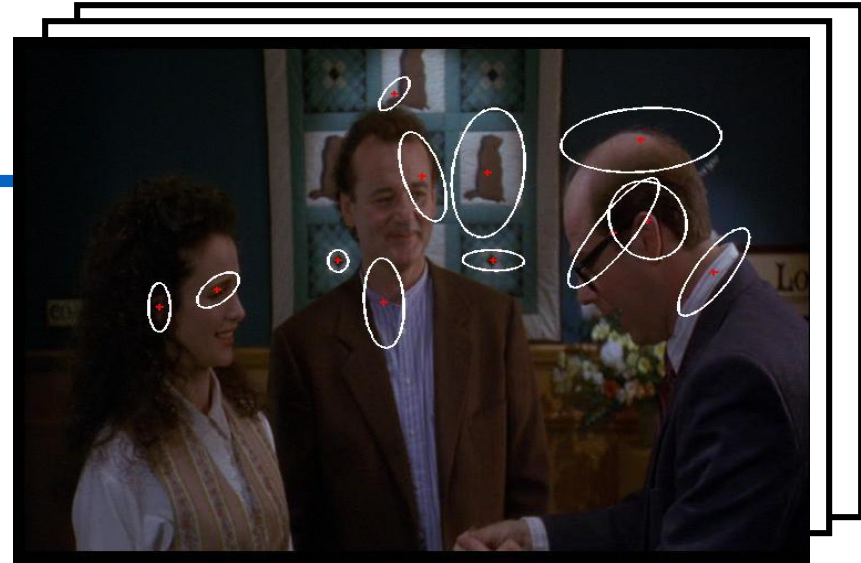
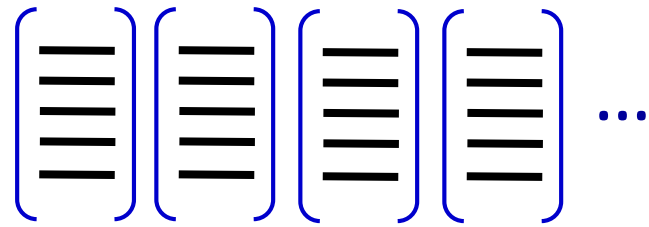
**Detect patches**

[Mikojaczyk and Schmid '02]

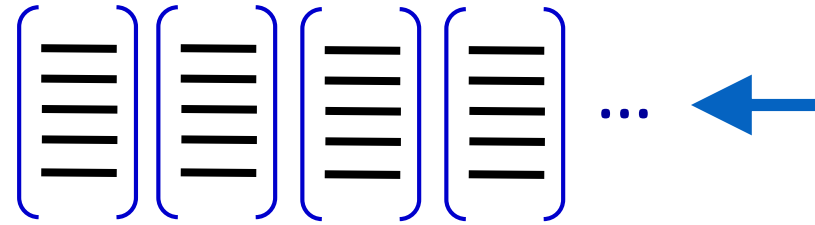
[Matas et al. '02]

[Sivic et al. '03]

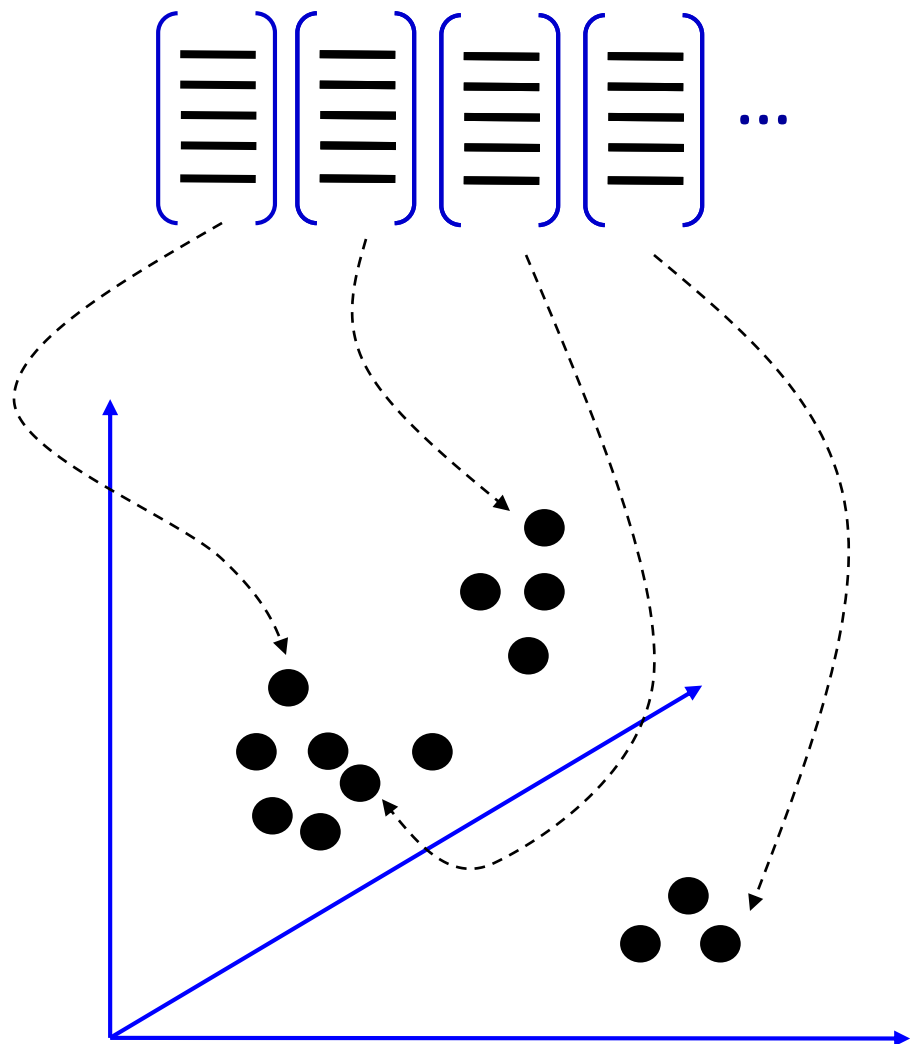
# Interest Point Features



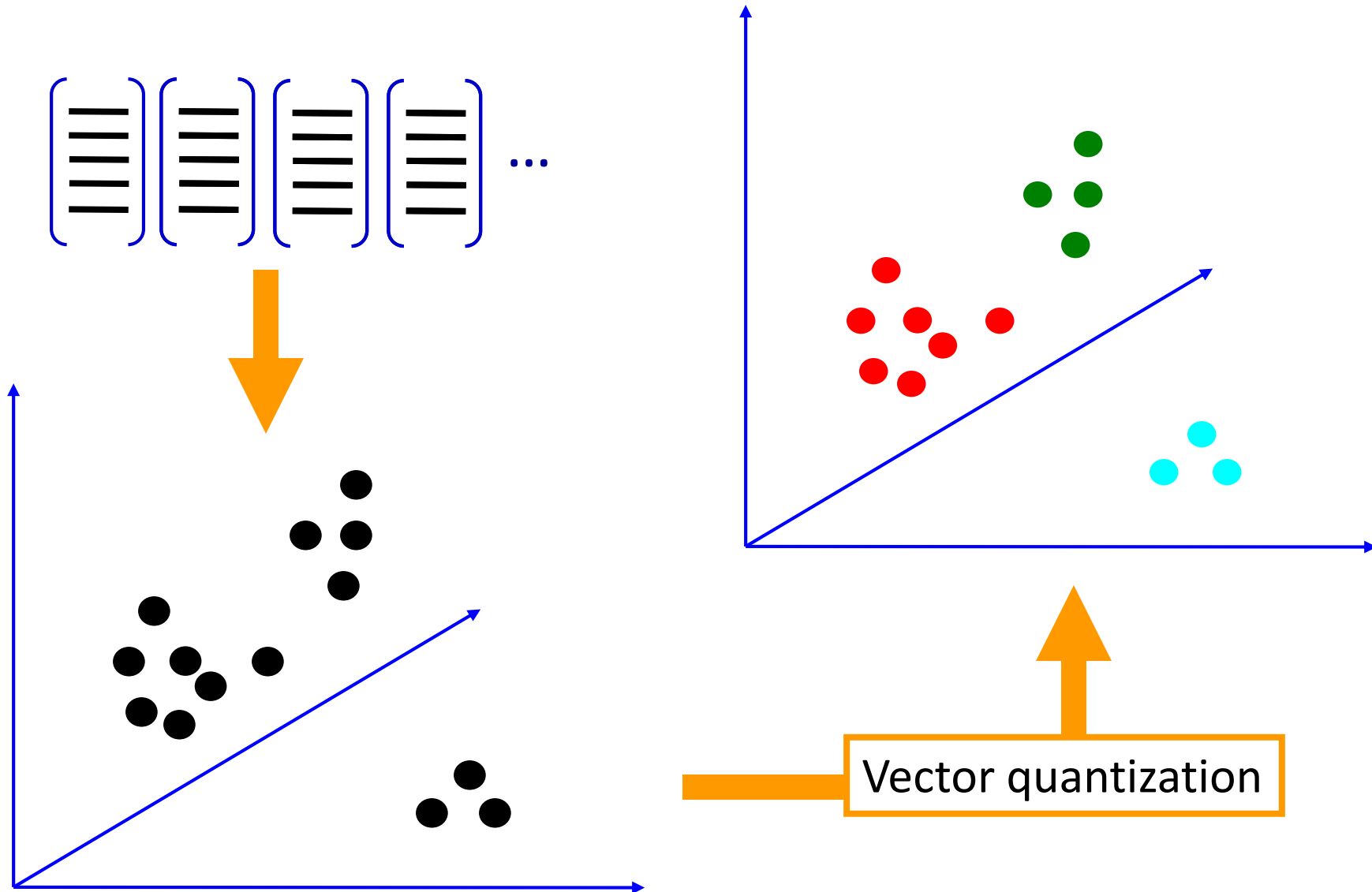
# Patch Features



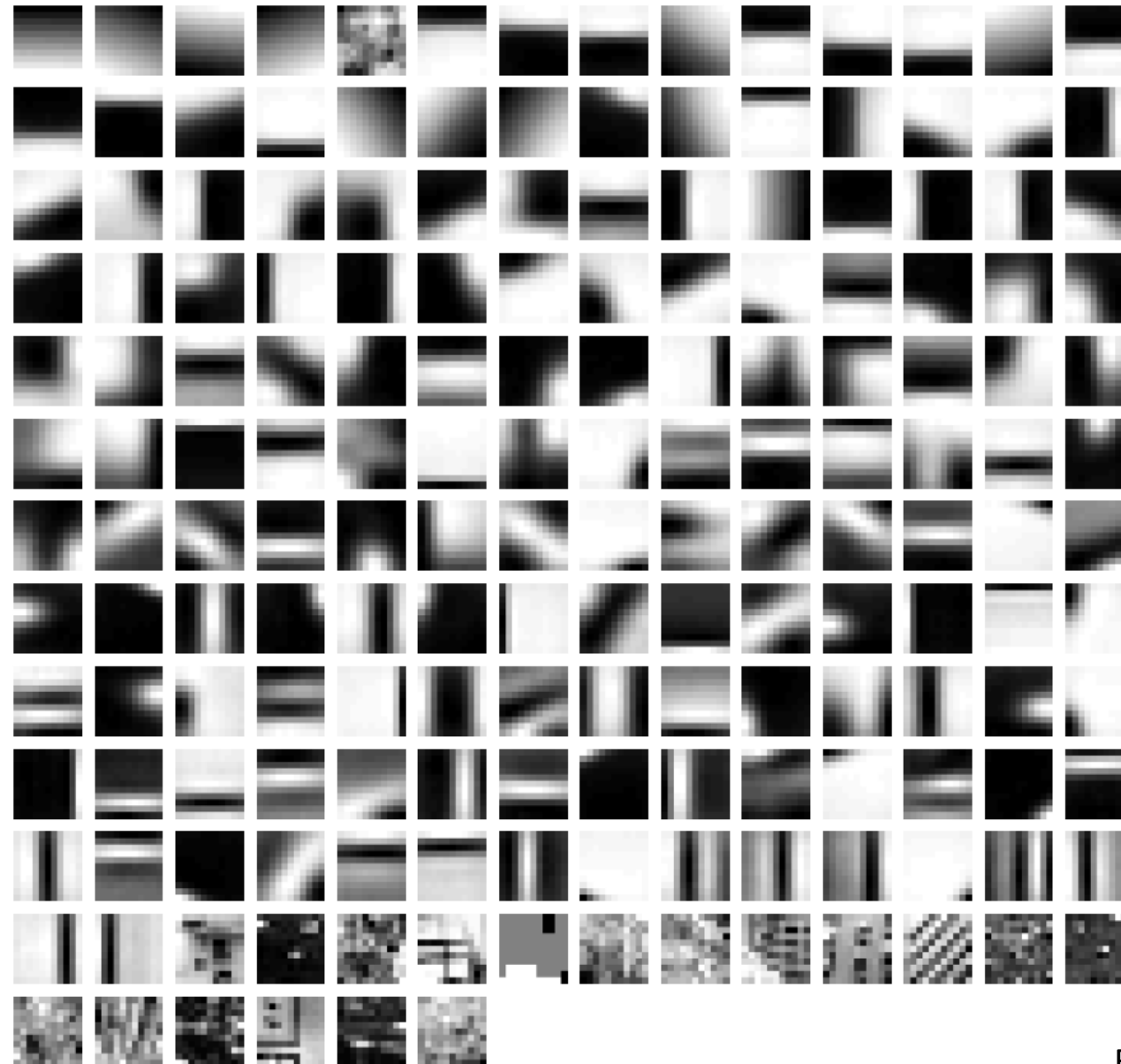
# dictionary formation



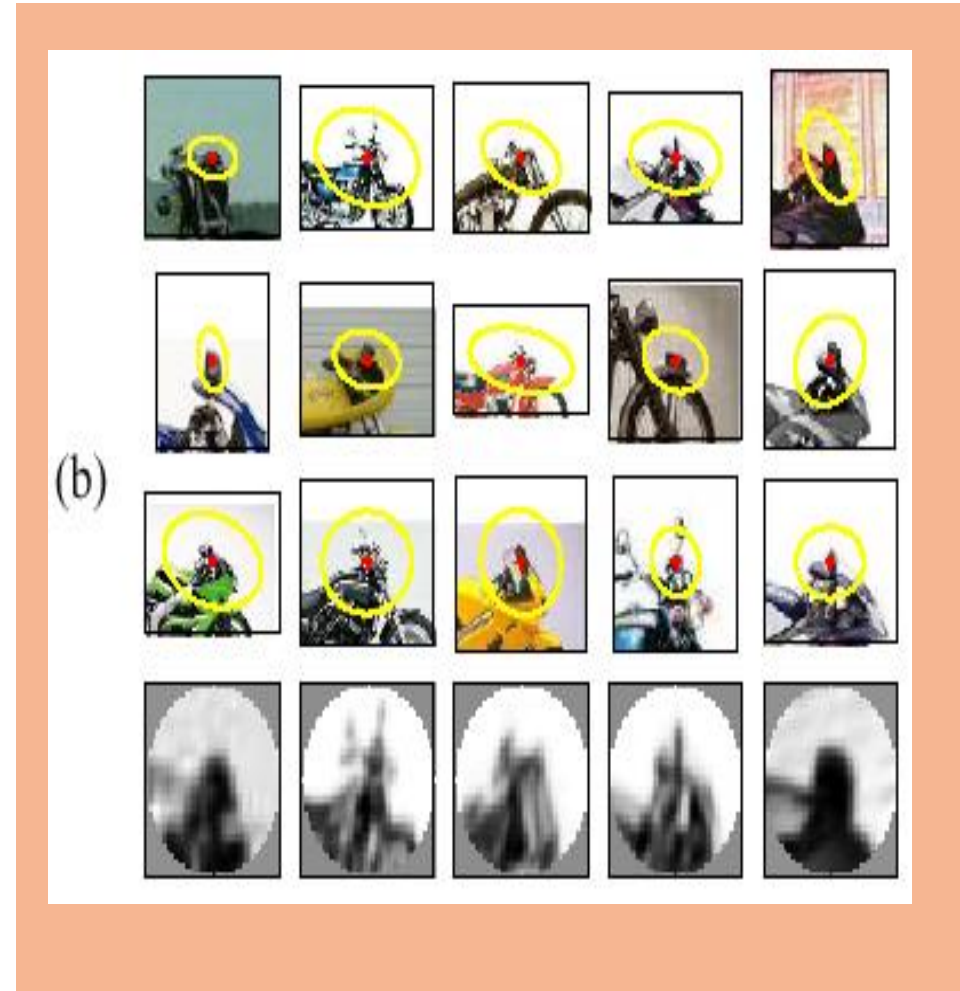
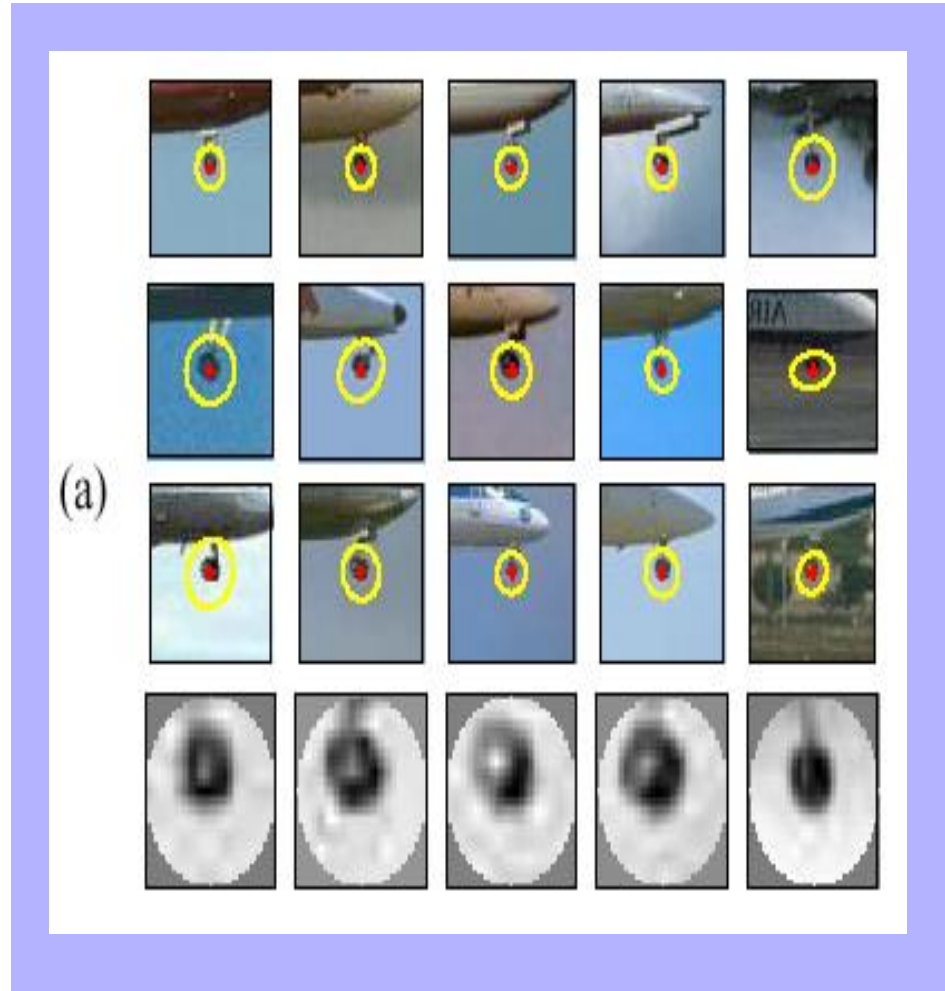
# Clustering (usually k-means)



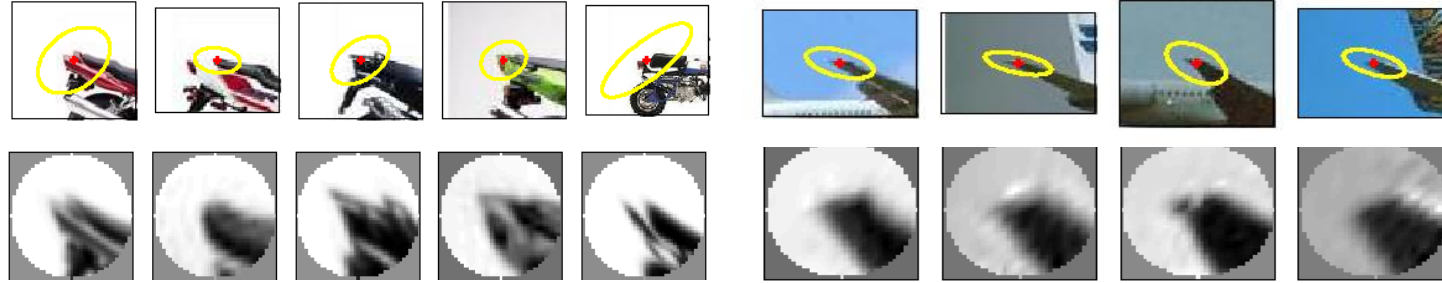
# Clustered Image Patches



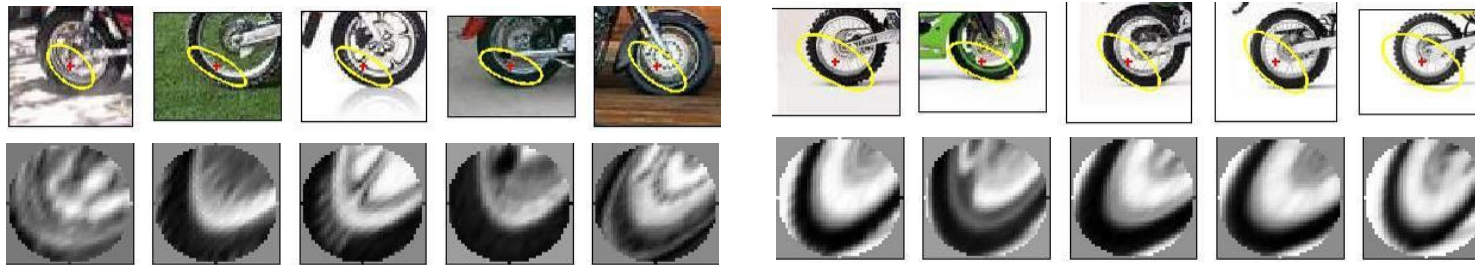
# Image patch examples of codewords



# Visual synonyms and polysemy



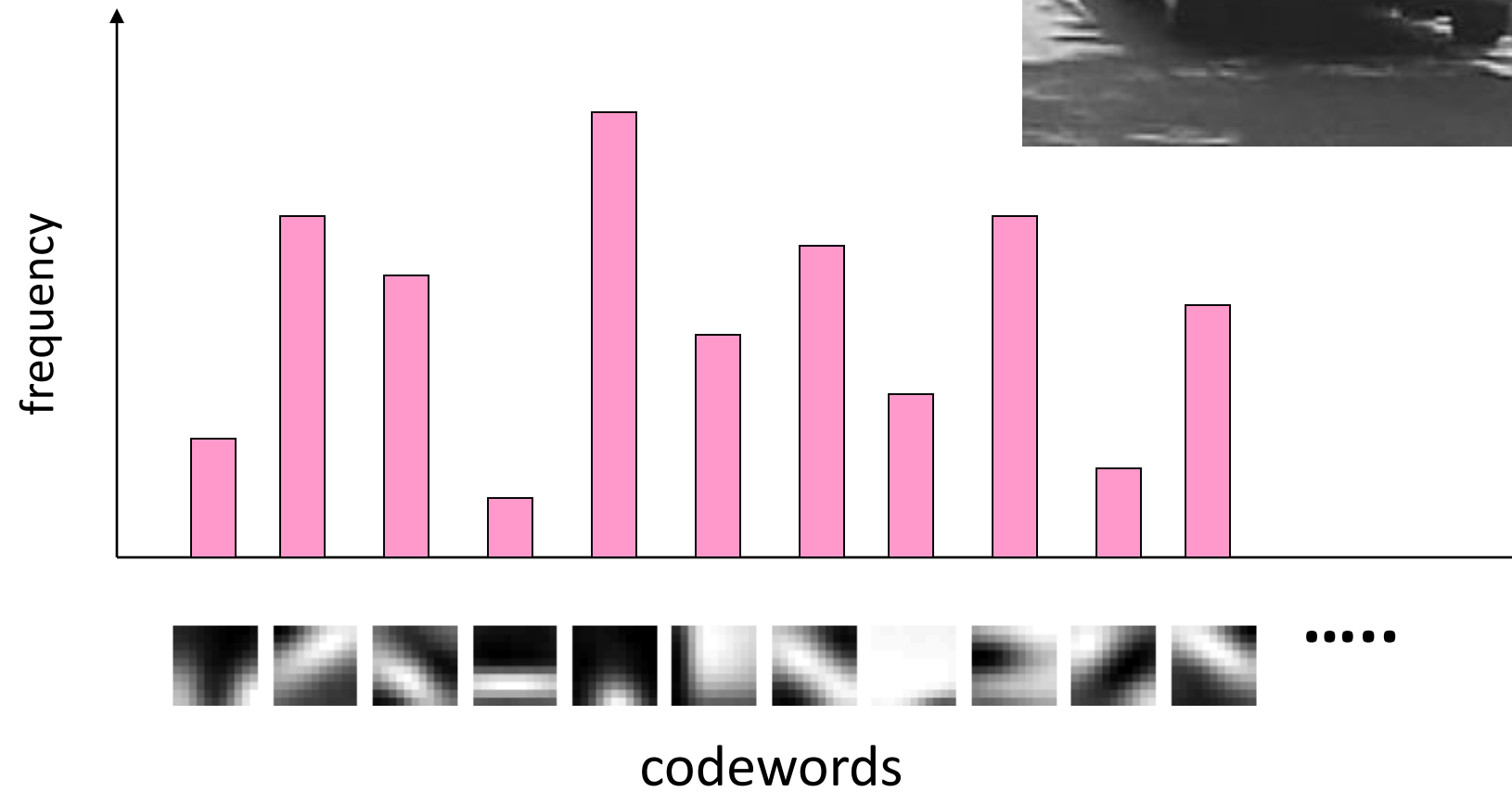
Visual Polysemy. Single visual word occurring on different (but locally similar) parts on different object categories.



Visual Synonyms. Two different visual words representing a similar part of an object (wheel of a motorbike).

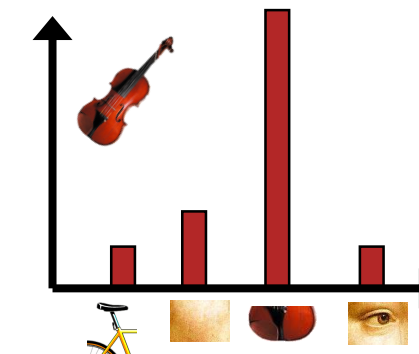
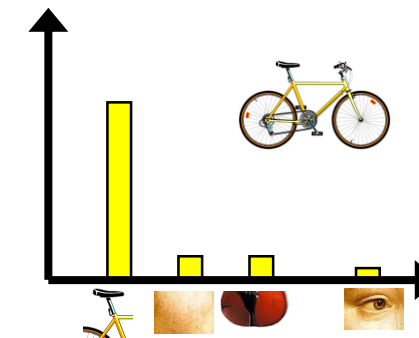
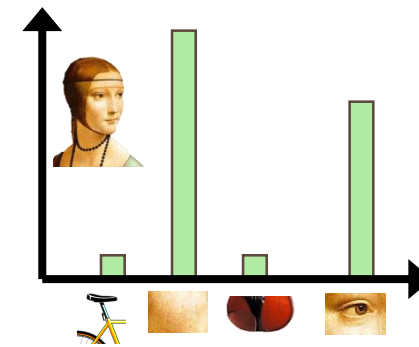


# Image representation



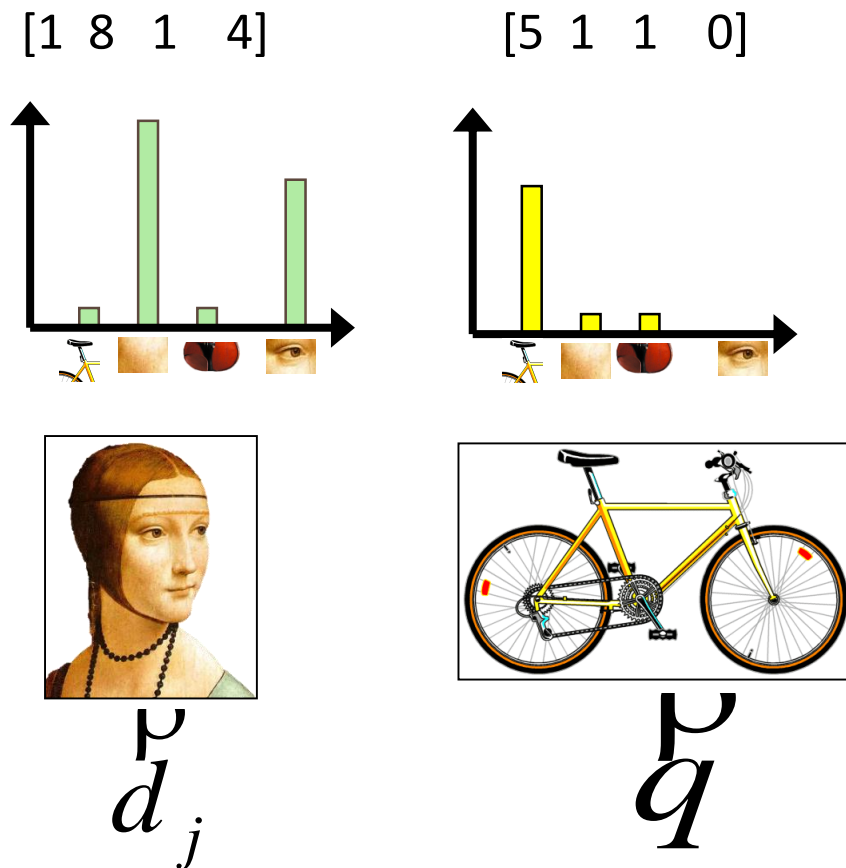
# Bags of visual words

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

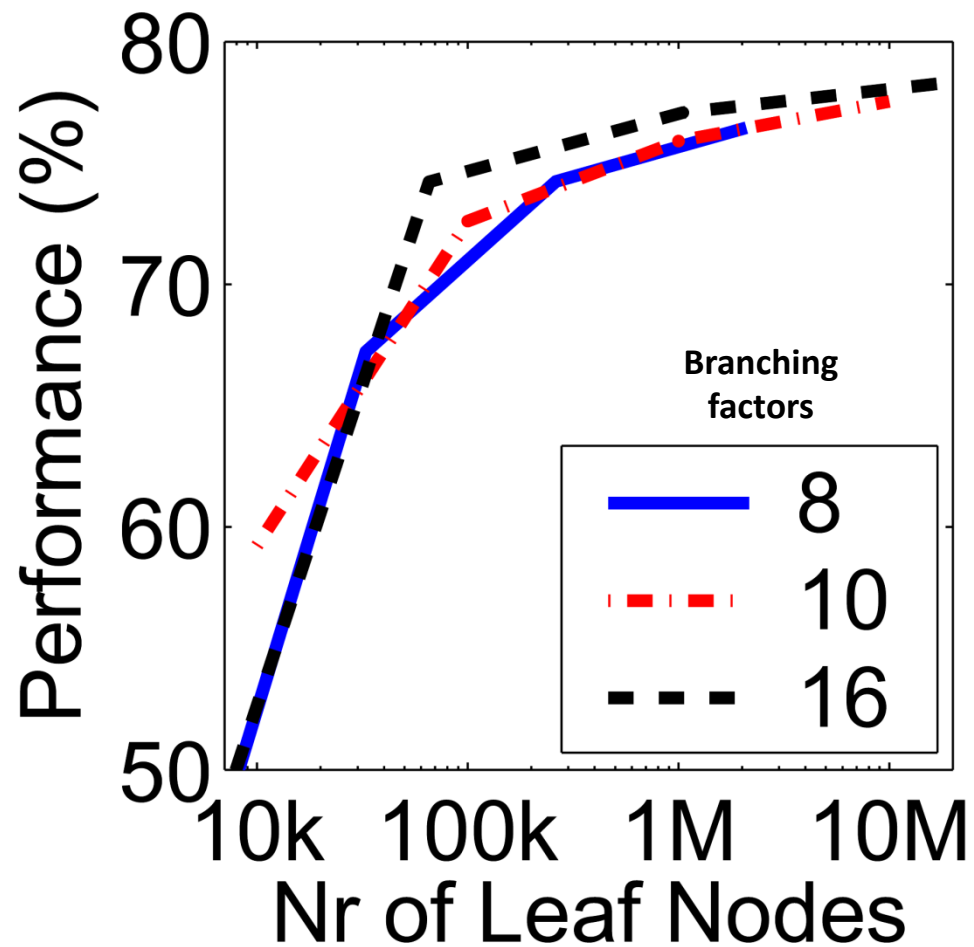


$$\text{sim}(d_j, q) = \frac{\langle d_j, q \rangle}{\|d_j\| \|q\|}$$

$$= \frac{\sum_{i=1}^V d_j(i) * q(i)}{\sqrt{\sum_{i=1}^V d_j(i)^2} * \sqrt{\sum_{i=1}^V q(i)^2}}$$

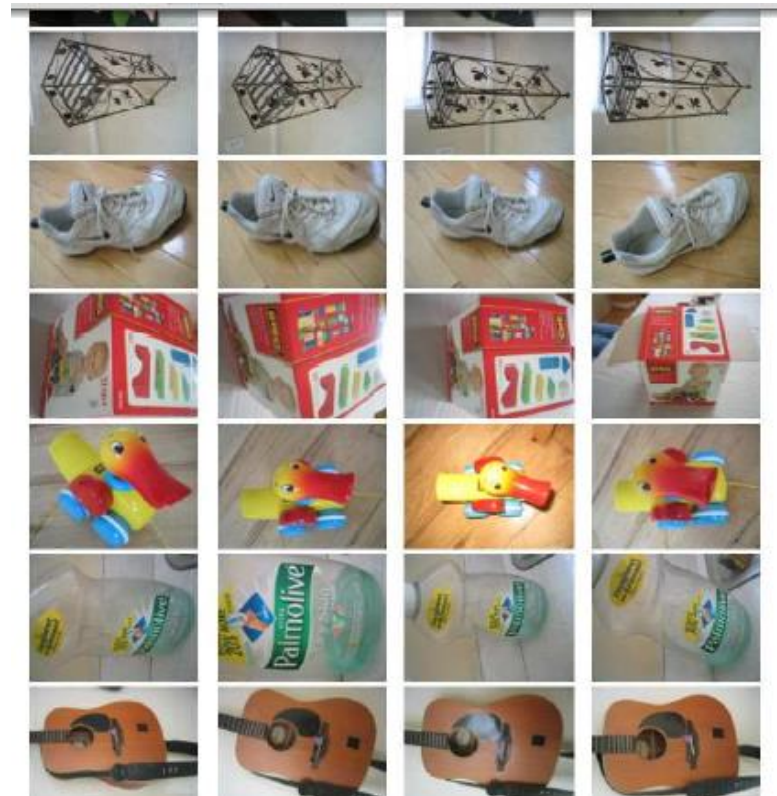
for vocabulary of  $V$  words

# Vocabulary size



*Influence on performance, sparsity*

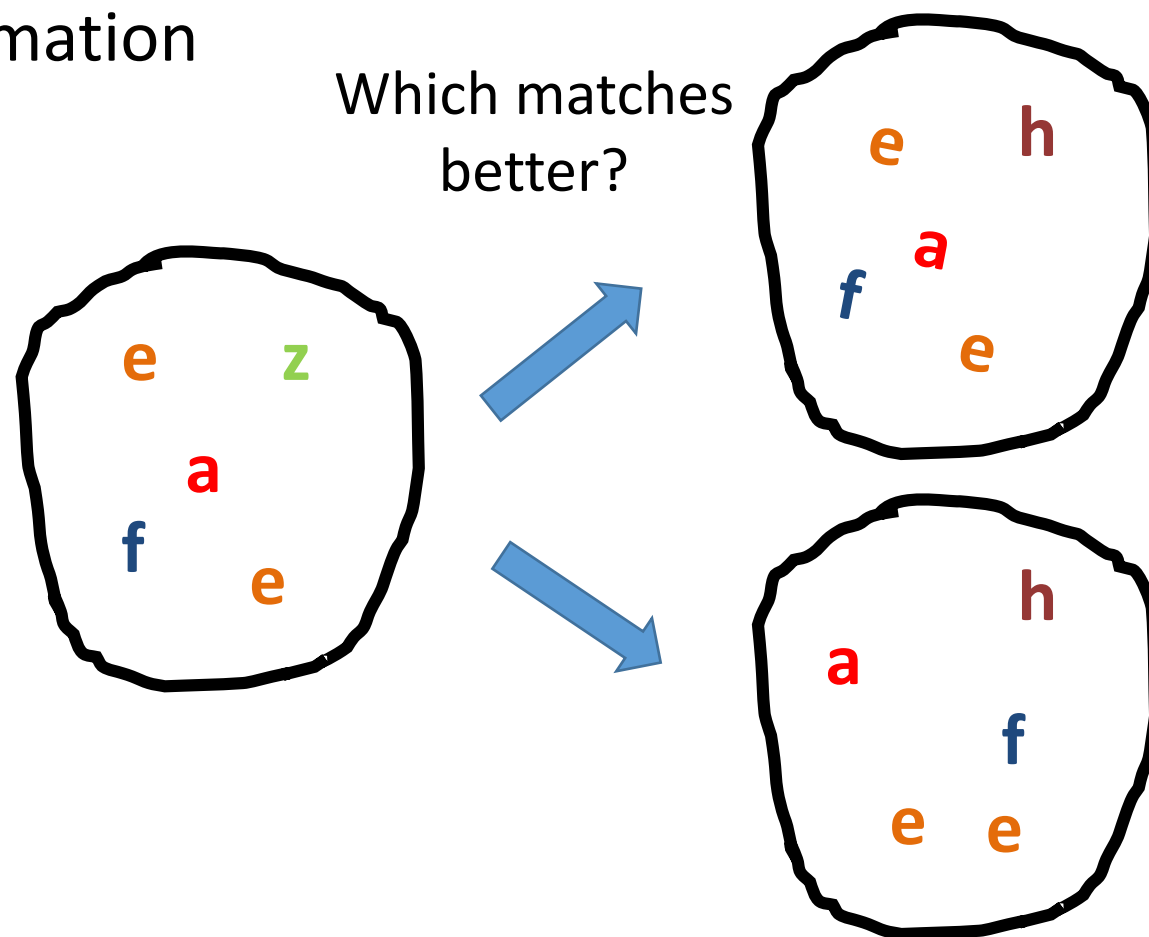
Results for recognition task with 6347 images



Nister & Stewenius, CVPR 2006  
Kristen Grauman

# Can we be more accurate?

So far, we treat each image as containing a “bag of words”, with no spatial information



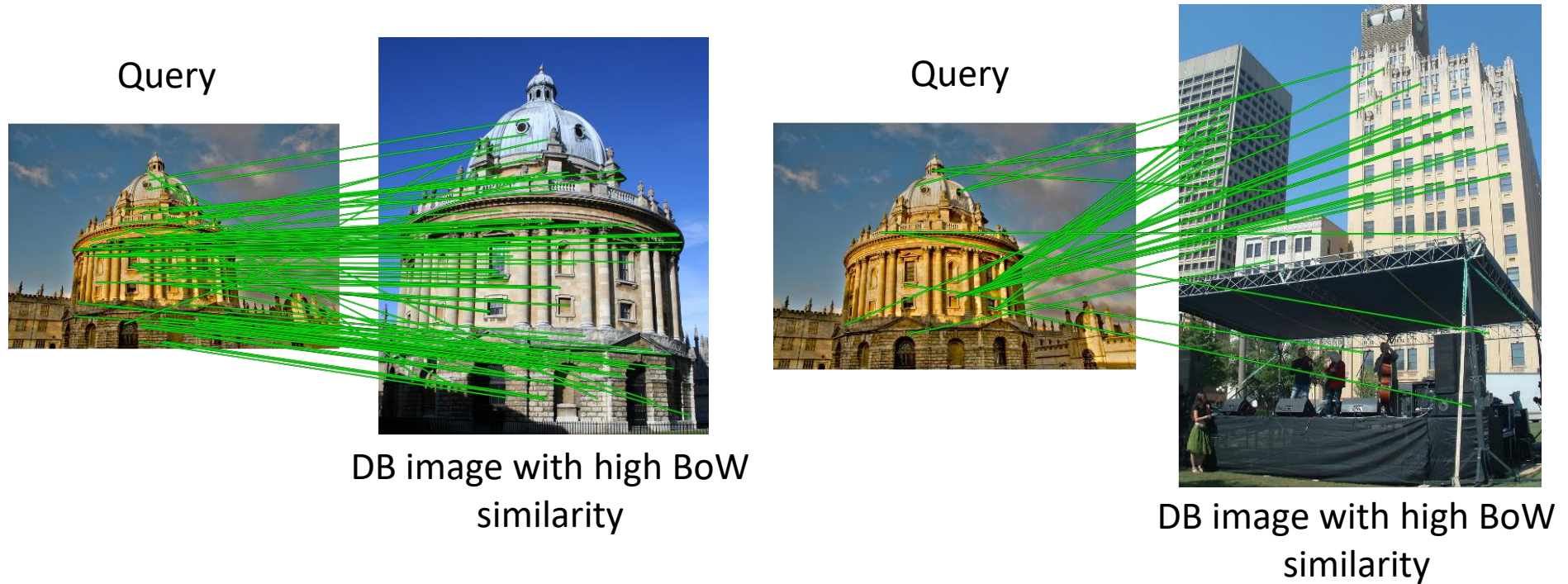
# Can we be more accurate?

So far, we treat each image as containing a “bag of words”, with no spatial information



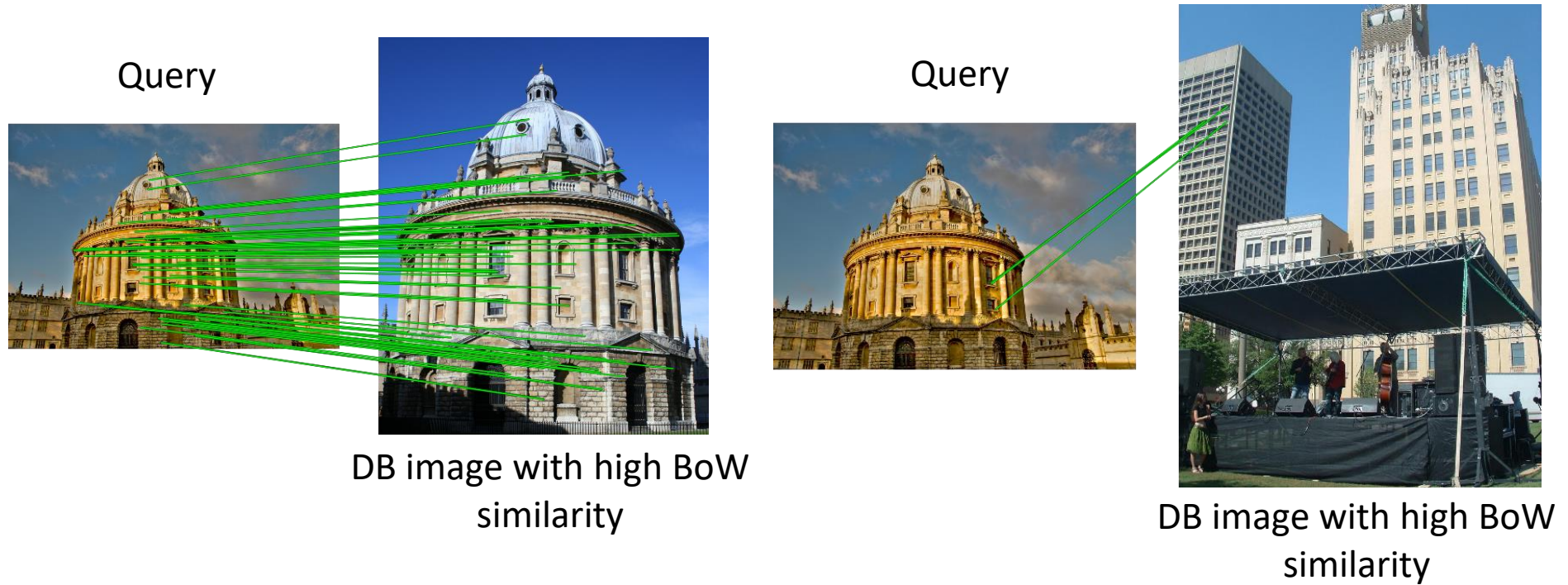
Real objects have consistent geometry

# Spatial Verification



Both image pairs have many visual words in common.

# Spatial Verification



Only some of the matches are mutually consistent



# What else can we borrow from text retrieval?

## Index

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

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 predicted 30% jump in exports with a 18% rise in imports. The yuan is expected to rise further and that China's deliberate policy to the surplus is one factor. Xiaochua more to be stayed within the value of the yuan. July and permitted it to band, but the US wants the yuan to be traded to trade freely. However, Beijing has made that it will take its time and tread carefully allowing the yuan to rise further in value.

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

# *tf-idf* weighting

- **T**erm frequency – **i**nverse **d**ocument frequency
- Describe frame by frequency of each word within it, downweight words that appear often in the database
- (Standard weighting for text retrieval)

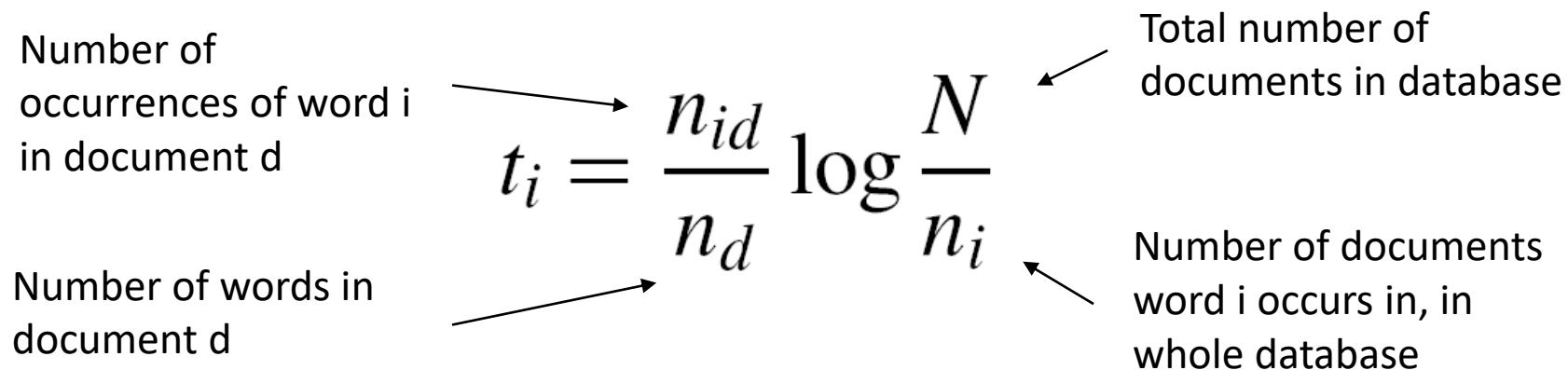
Number of occurrences of word  $i$  in document  $d$

Number of words in document  $d$

$$t_i = \frac{n_{id}}{n_d} \log \frac{N}{n_i}$$

Total number of documents in database

Number of documents word  $i$  occurs in, in whole database

The diagram illustrates the tf-idf formula. On the left, two text labels are connected to the formula by arrows: 'Number of occurrences of word i in document d' points to the numerator  $n_{id}$ , and 'Number of words in document d' points to the denominator  $n_d$ . On the right, two more text labels are connected to the formula by arrows: 'Total number of documents in database' points to the variable  $N$  in the numerator of the log term, and 'Number of documents word i occurs in, in whole database' points to the variable  $n_i$  in the denominator of the log term.

# Query Expansion



Chum, Philbin, Sivic, Isard, Zisserman: Total Recall..., ICCV 2007

Slide credit: Ondrej Chum