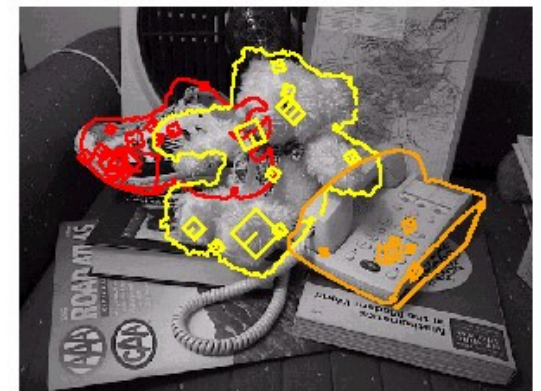




# Matching objects in images

Kristen Grauman



# Recognizing or retrieving specific objects

## Example I: Visual search in feature films

Visually defined query

“Groundhog Day” [Rammis, 1993]

“Find this clock”



“Find this place”



# Recognizing or retrieving specific objects

Example II: Search photos on the web for particular places



Find these landmarks

...in these images and 1M more



# Google Goggles

Use pictures to search the web.

[▶ Watch a video](#)



## Get Google Goggles

**Android (1.6+ required)**

Download from [Android Market](#).

[Send Goggles to Android phone](#)

**New! iPhone (iOS 4.0 required)**

Download [from the App Store](#).

[Send Goggles to iPhone](#)

New!



[Text](#)



[Landmarks](#)



[Books](#)



[Contact Info](#)



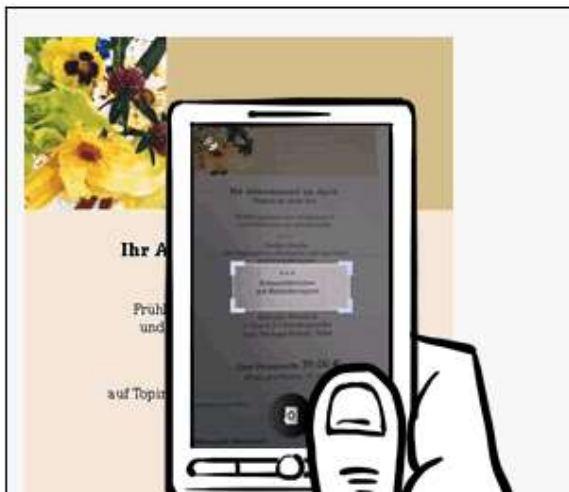
[Artwork](#)



[Wine](#)



[Logos](#)

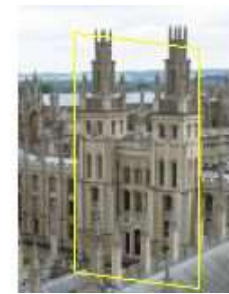
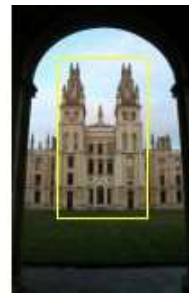


# Why is it difficult?

Want to find the object despite possibly large changes in scale, viewpoint, lighting and partial occlusion



Scale



Viewpoint



Lighting



Occlusion

*We can't expect to match such varied instances with a single global template...*

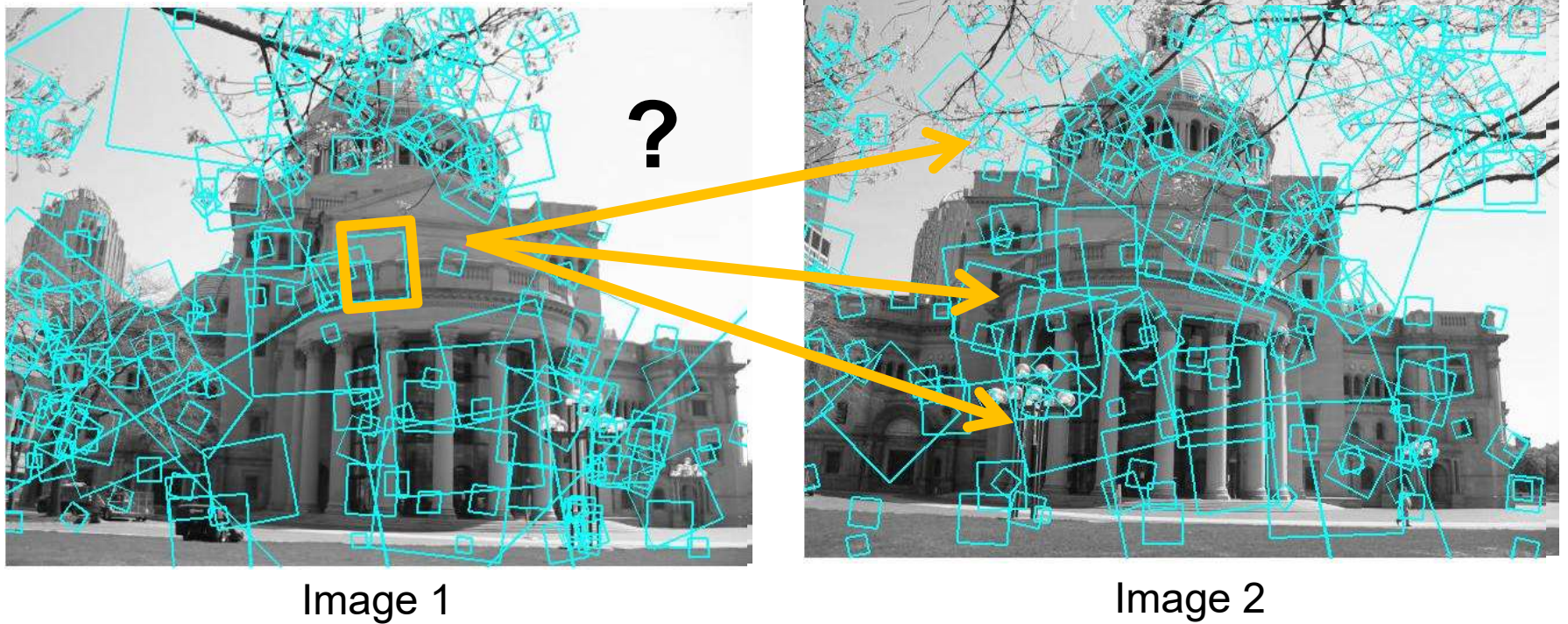
# Text retrieval vs. image search

- What makes the problems similar, different?

# Matching objects in images

- Motivation – visual search
- Visual words
  - quantization, index, bags of words
- Spatial verification
  - affine; RANSAC, Hough
- Other text retrieval tools
  - tf-idf, query expansion
- Example applications

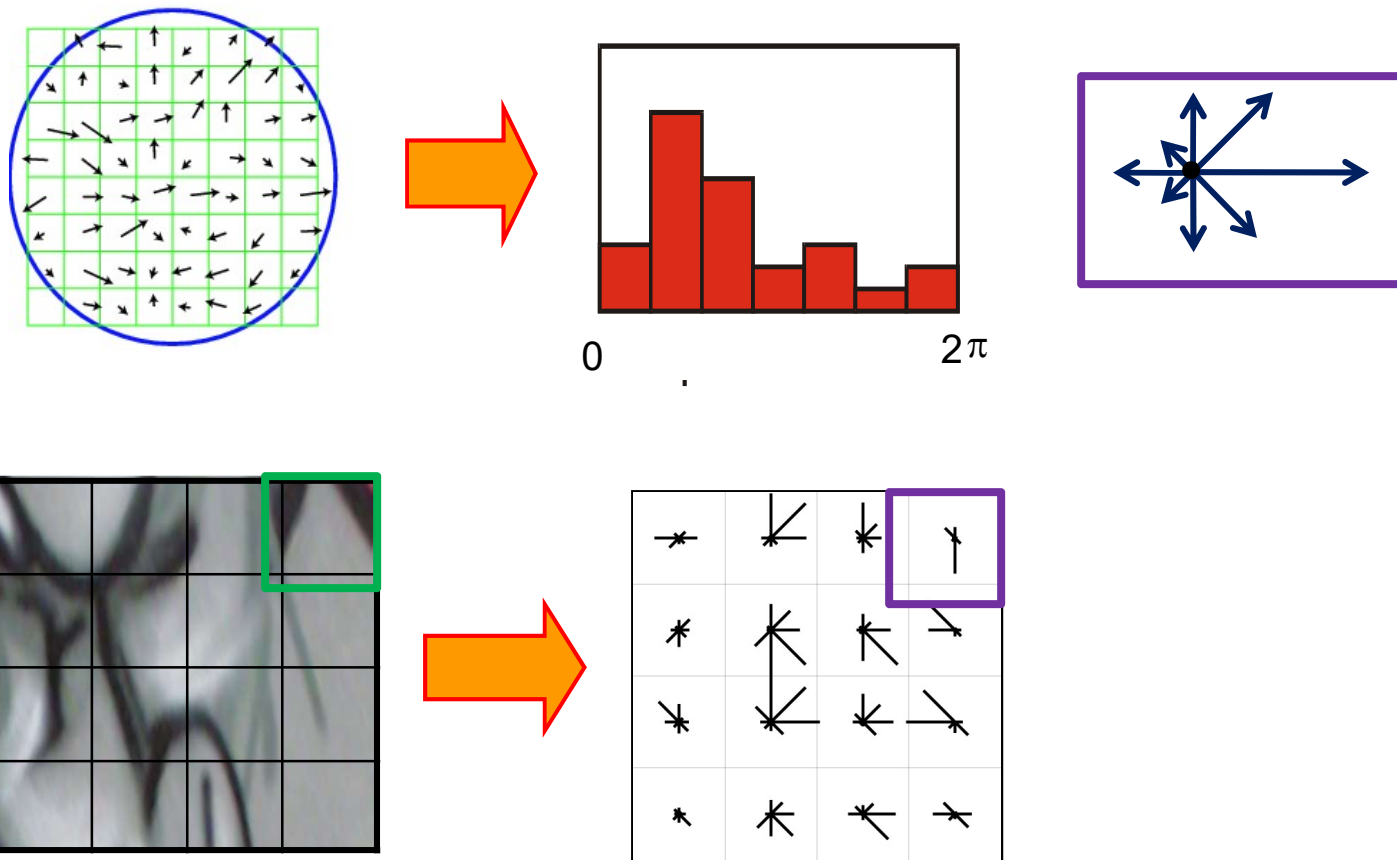
# Invariant local features



We can detect features *independently* per image that are scale, translation, and rotation *invariant*.

# Local feature descriptors

- SIFT [Lowe 2004] : Use histograms to bin pixels within sub-patches according to their orientation.



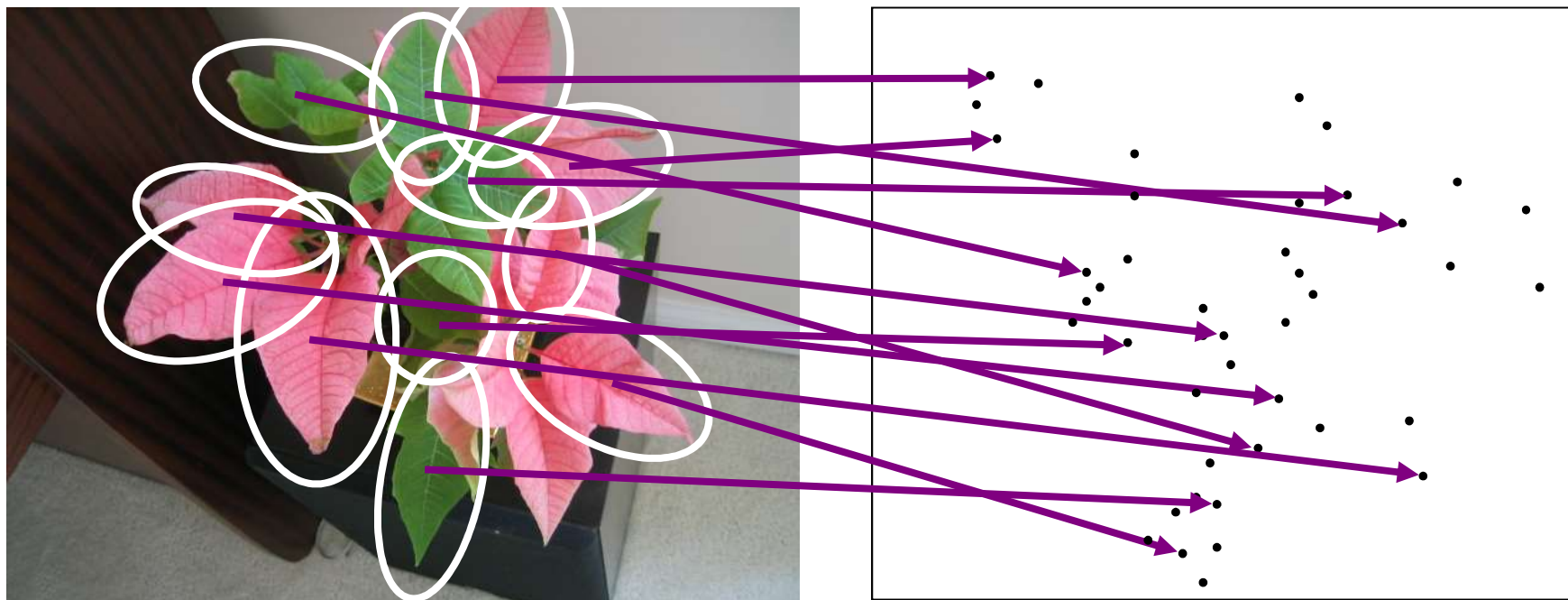
# Visual words: main idea

- Extract some local features from a number of images ...

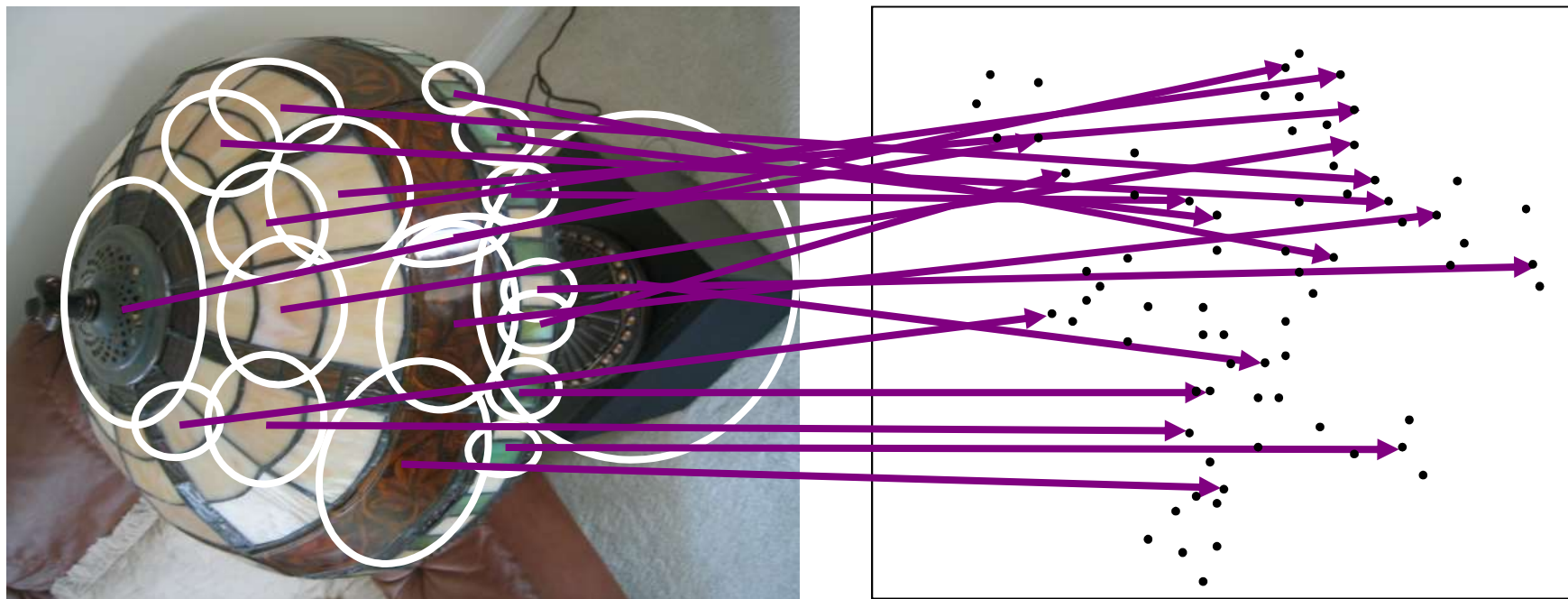


e.g., SIFT descriptor space: each point is 128-dimensional

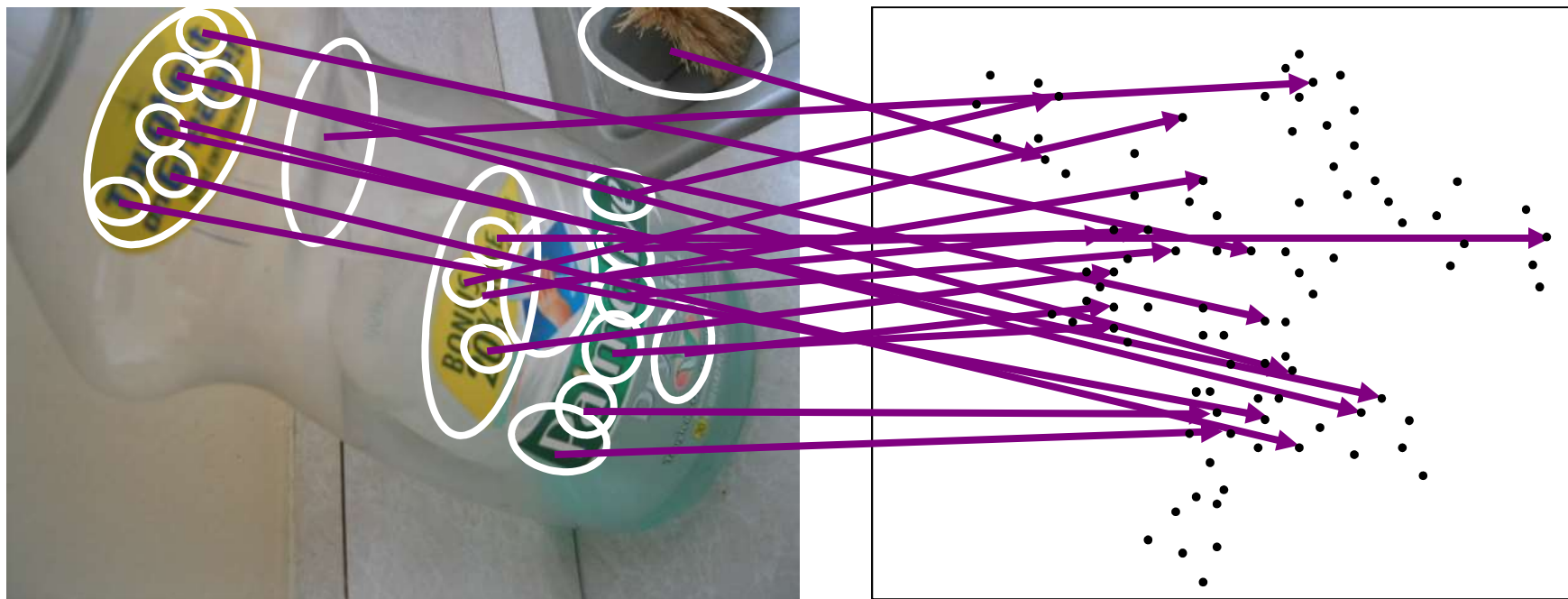
# Visual words: main idea

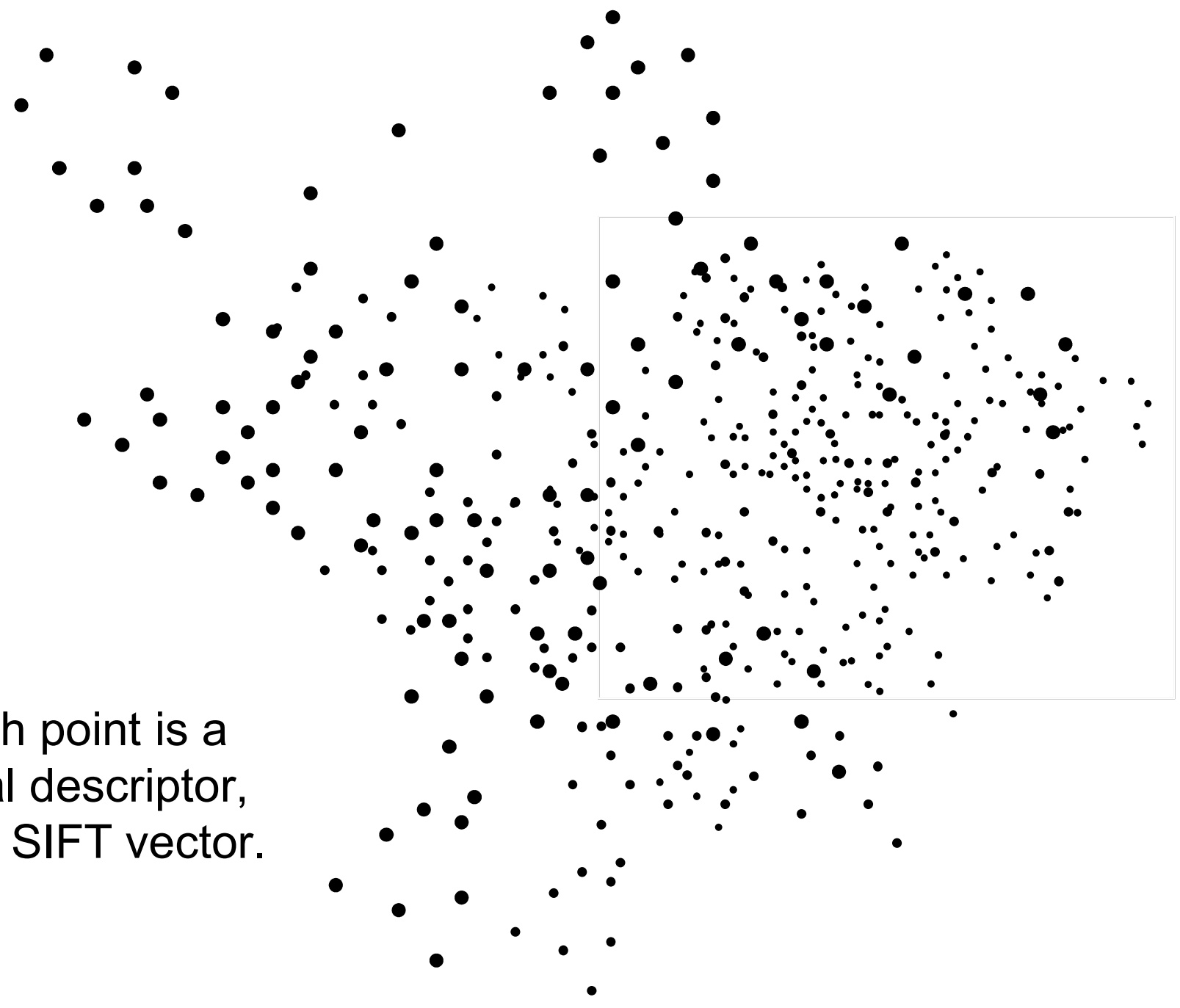


# Visual words: main idea

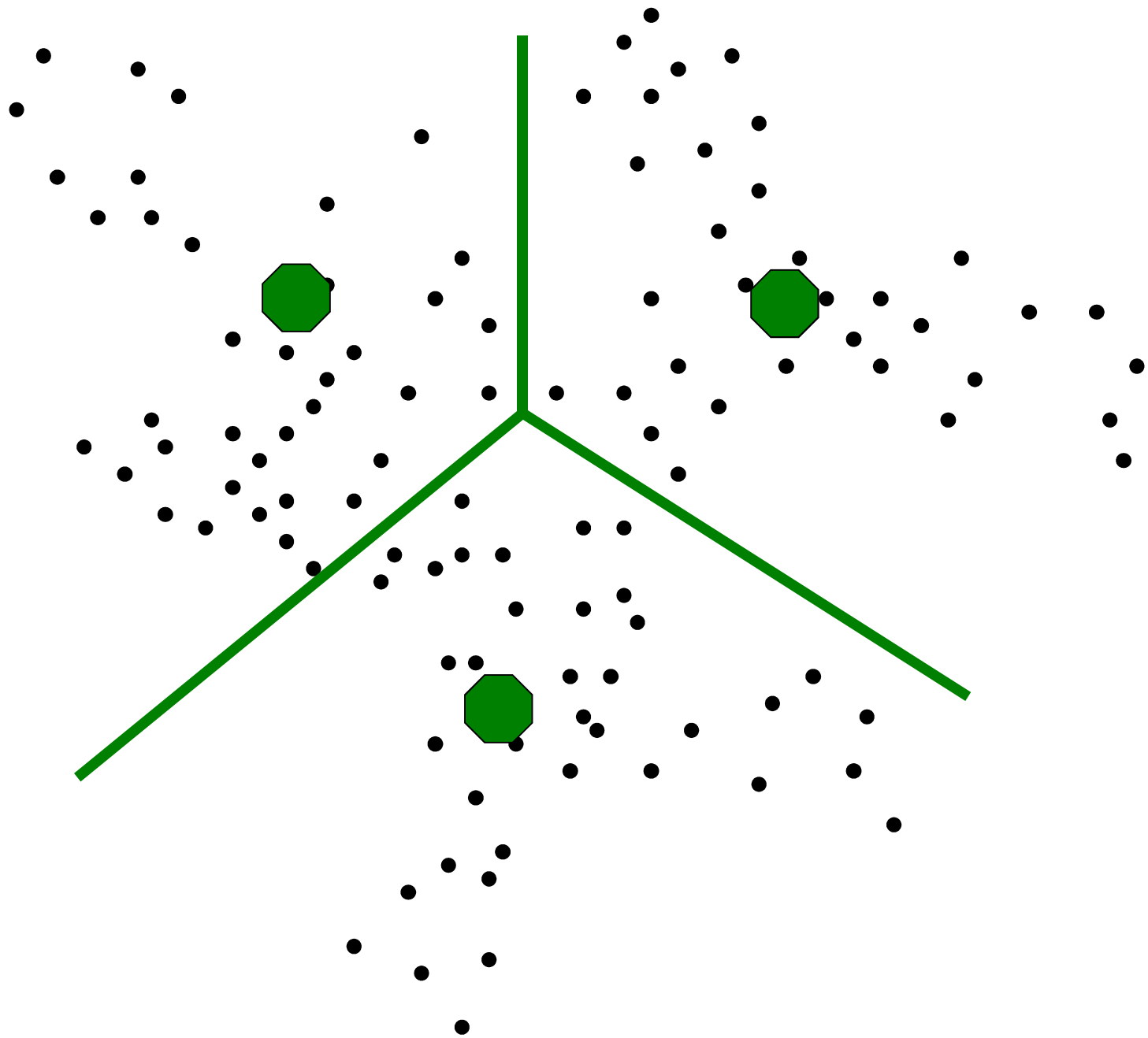


# Visual words: main idea



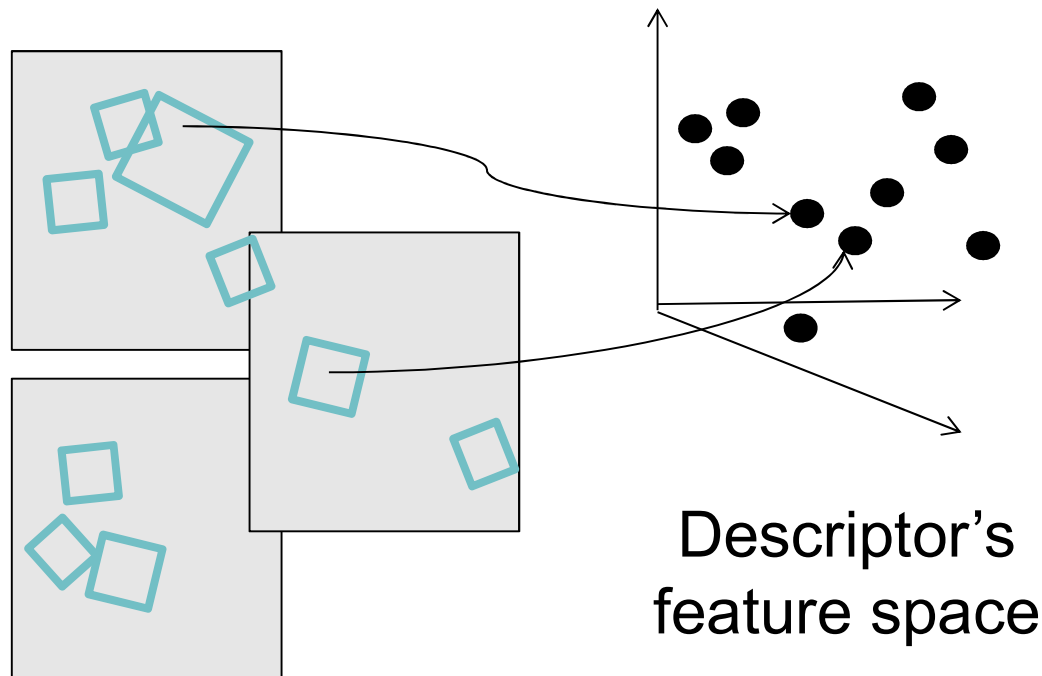


Each point is a  
local descriptor,  
e.g. SIFT vector.



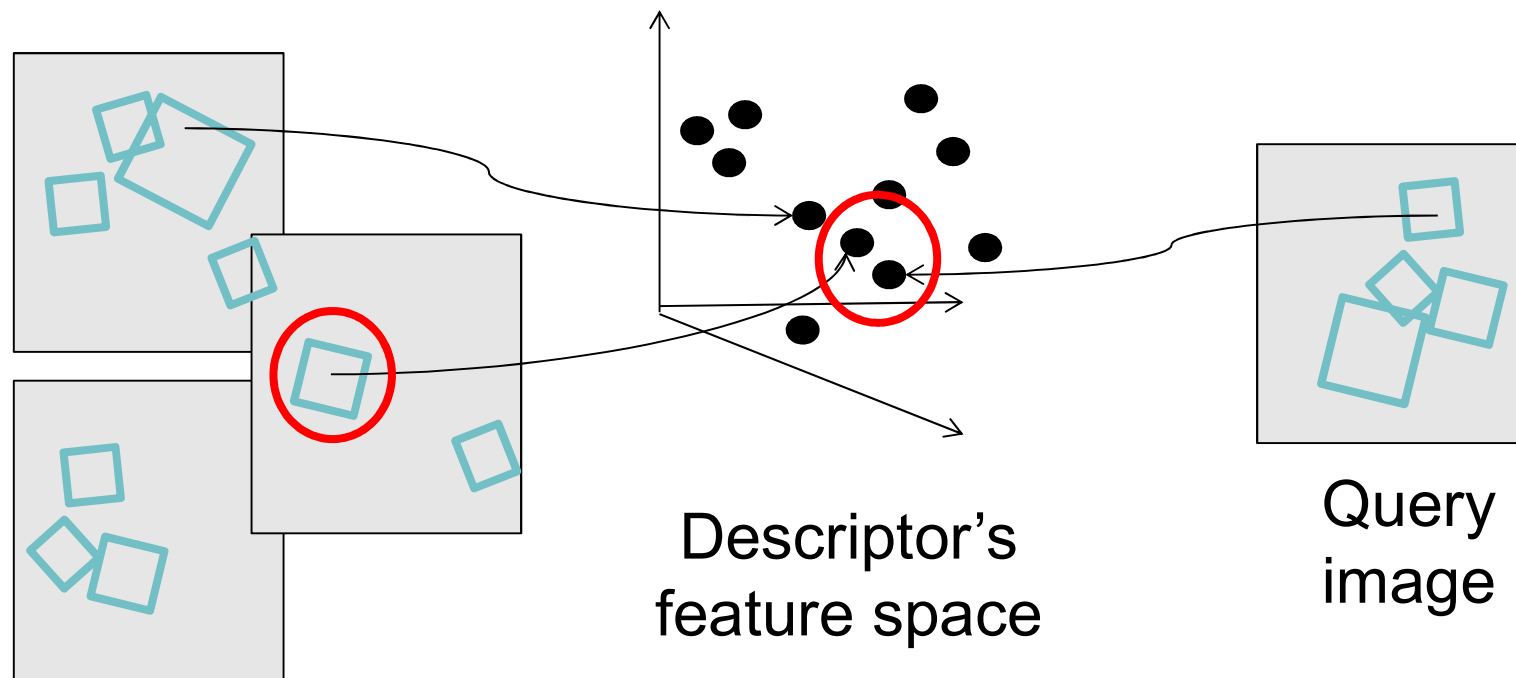
# Indexing local features

- 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.



Database  
images

*Easily can have millions of  
features to search!*

# Indexing local features: inverted file index

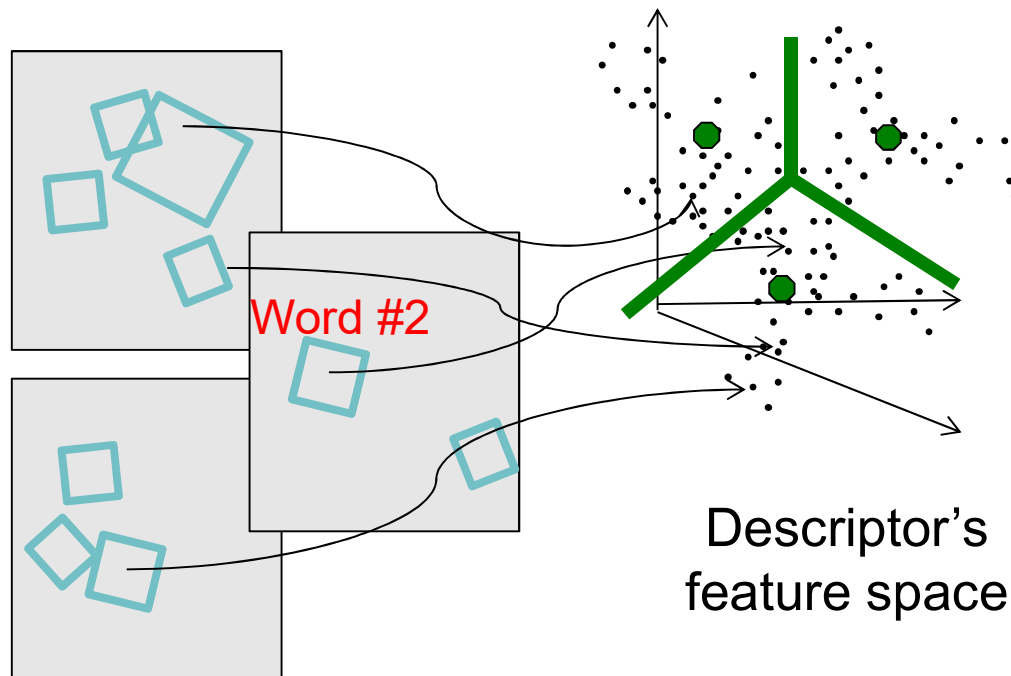
Index		
"Along I-75," From Detroit to Florida; <i>inside back cover</i>	Butterfly Center, McGuire; 134	Driving Lanes; 85
"Drive I-95," From Boston to Florida; <i>inside back cover</i>	CAA (see AAA)	Duval County; 163
1929 Spanish Trail Roadway; 101-102,104	CCC, The; 111,113,115,135,142	Eau Gallie; 175
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 Reale; 176
AAA National Office; 88	Canaveral Natl Seashore; 173	Ellenton; 144-145
Abbreviations,	Cannon Creek Airpark; 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	Castillo 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; 160	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	Chipay; 114	Wonder Gardens; 154
Alafia River; 143	Name; 115	Falling Waters SP; 115
Alapaha, Name; 126	Choctawatchee, Name; 115	Fantasy of Flight; 95
Alfred B MacLay 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	Ft 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; 109-109,146	Miami Expressways; 194-195	Florida,
Apalachicola River; 112	Orlando Expressways; 192-193	12,000 years ago; 167
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	Clearwater 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
Barefoot Mailmen; 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; 154	25 mile Strip Maps; 66
Belz 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	Crosstown Expy; 11,35,98,143	HEFT; 76,161,190
Big Foot Monster; 105	Cuban Bread; 184	History; 189
Camp Safari; 160	Dade Battlefield; 140	Names; 189
Blackwater River SP; 117	Dade, Maj. Francis; 139-140,161	Service Plazas; 190
Blue Angels	Dania Beach Hurricane; 184	Spur SR91; 76
	Daniel Boone, Florida Walk; 117	Ticket System; 190
	Daytona Beach; 172-173	Toll Plazas; 190
	De Land; 87	Ford, Henry; 152

Kristen Grauman

- 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

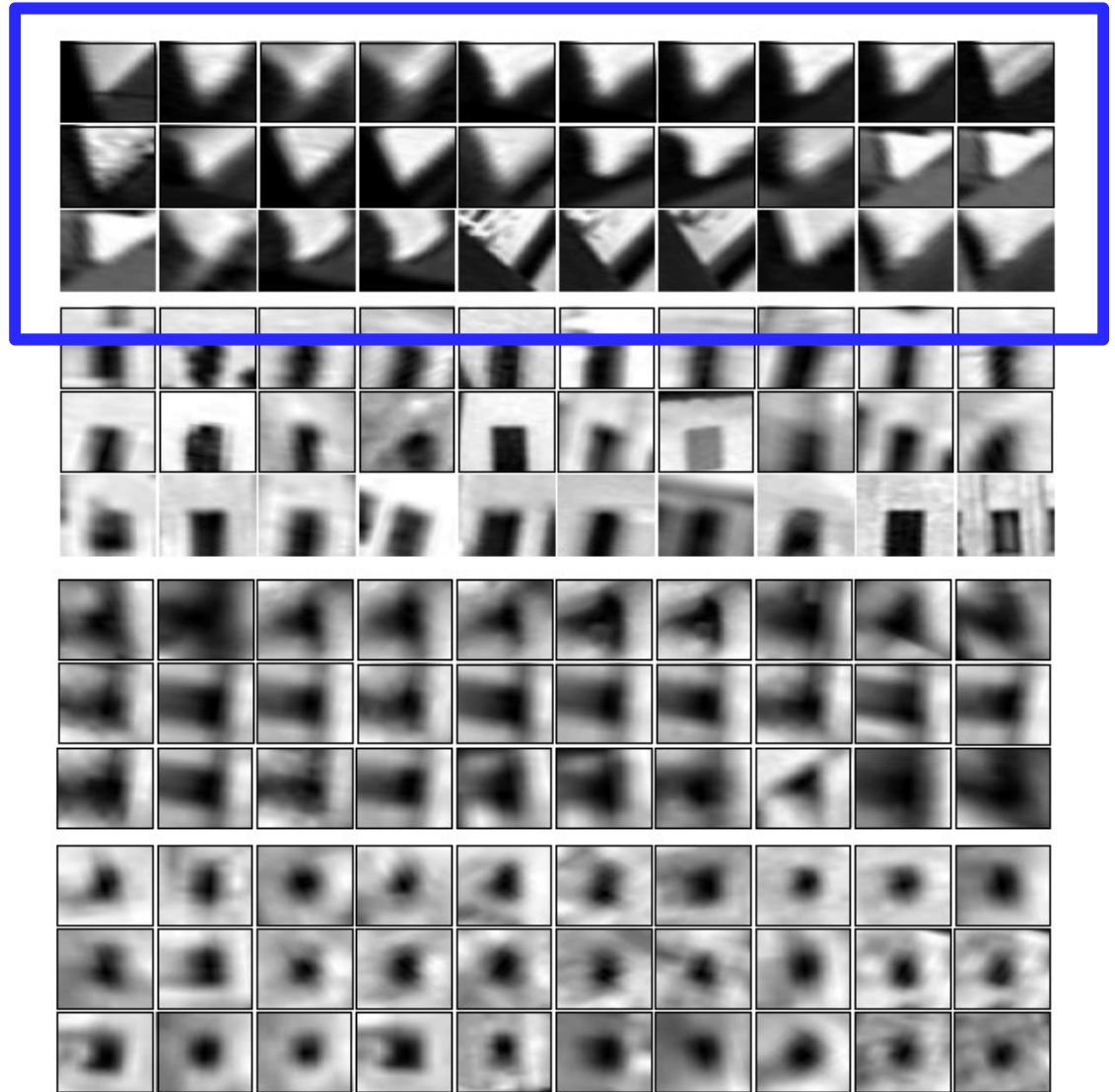
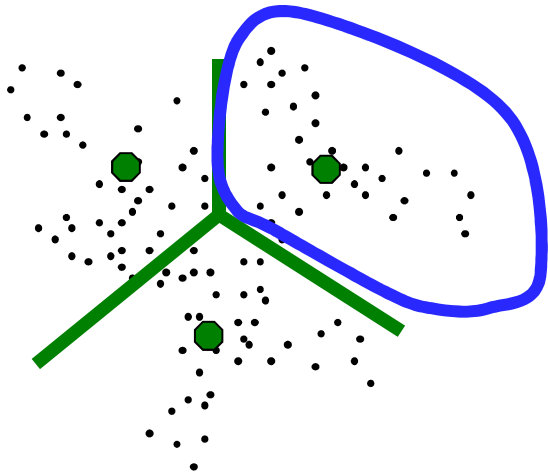
- Map high-dimensional descriptors to tokens/words by quantizing the feature space



- 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

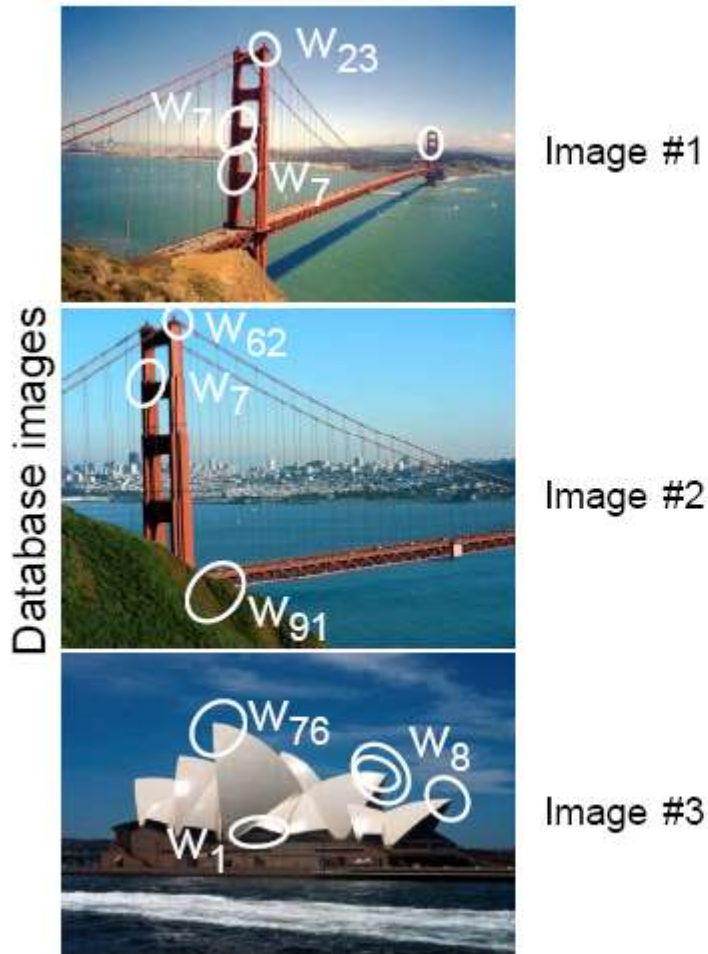


# Visual vocabulary formation

Issues:

- Vocabulary size, number of words
- Sampling strategy: where to extract features?
- Clustering / quantization algorithm
- Unsupervised vs. supervised
- What corpus provides features (universal vocabulary?)

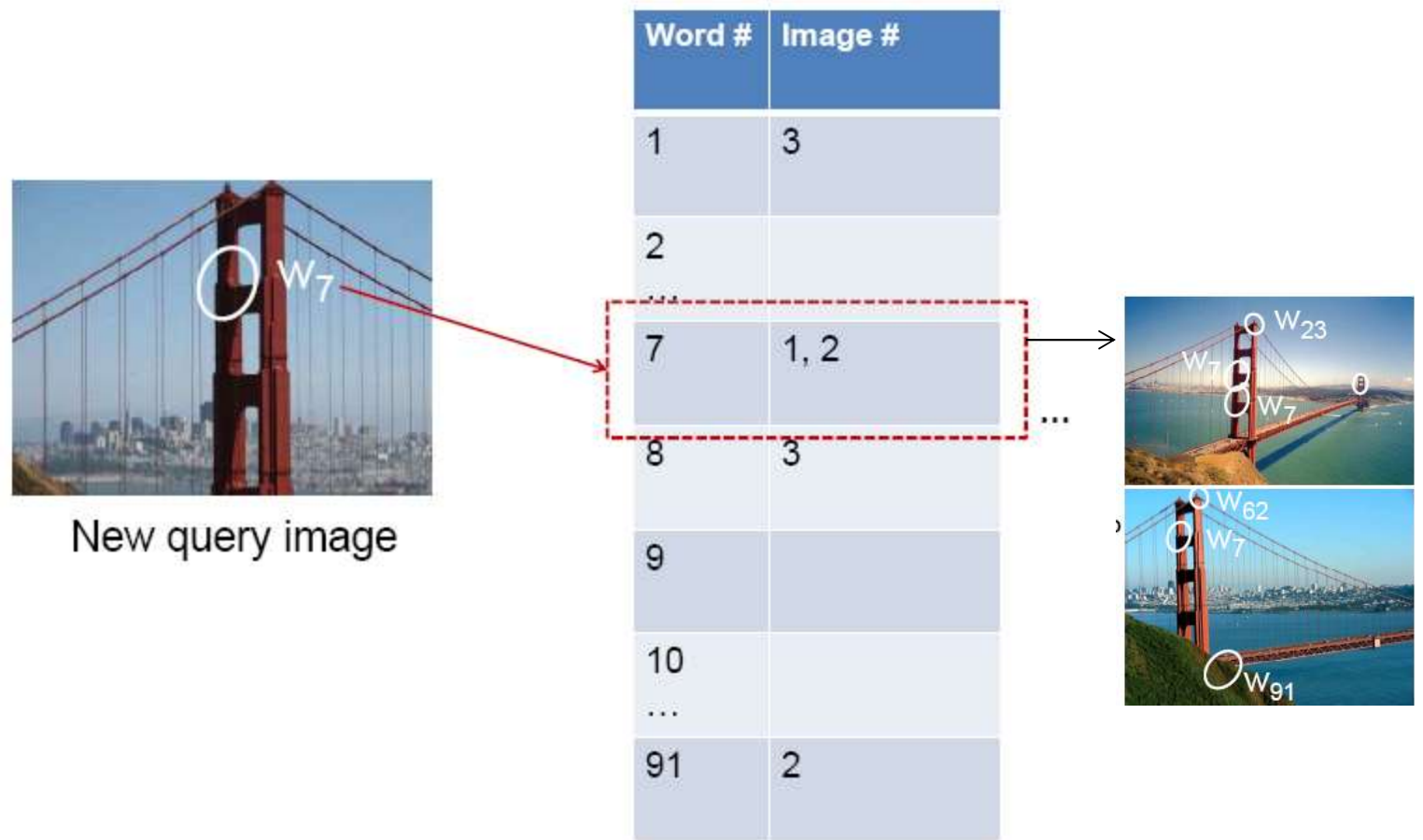
# Inverted file index



Word #	Image #
1	3
2	
...	
7	1, 2
8	3
9	
10	
...	
91	2

- 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.

# Instance recognition: remaining issues

- How to summarize the content of an entire image? And gauge overall similarity?
- How large should the vocabulary be? How to perform quantization efficiently?
- Is having the same set of visual words enough to identify the object/scene? How to verify spatial agreement?
- How to score the retrieval results?

# 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 messages that reach our eyes.

For a long time, the visual image was considered as a

centers in the visual

movie scene. The image

discoveries. We know that

perception is more complex

following the path to the various centers of the cortex,

Hubel and Wiesel have demonstrated that the *message about the image falling on the retina undergoes a*

*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 created by a predicted 30% increase in exports to \$750bn, compared with \$560bn in 2004.

The increase will annoy the US, which has long complained about China's trade policies.

China's government has agreed to a deal with the US, but the yuan is still not free to rise.

The government also needs to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

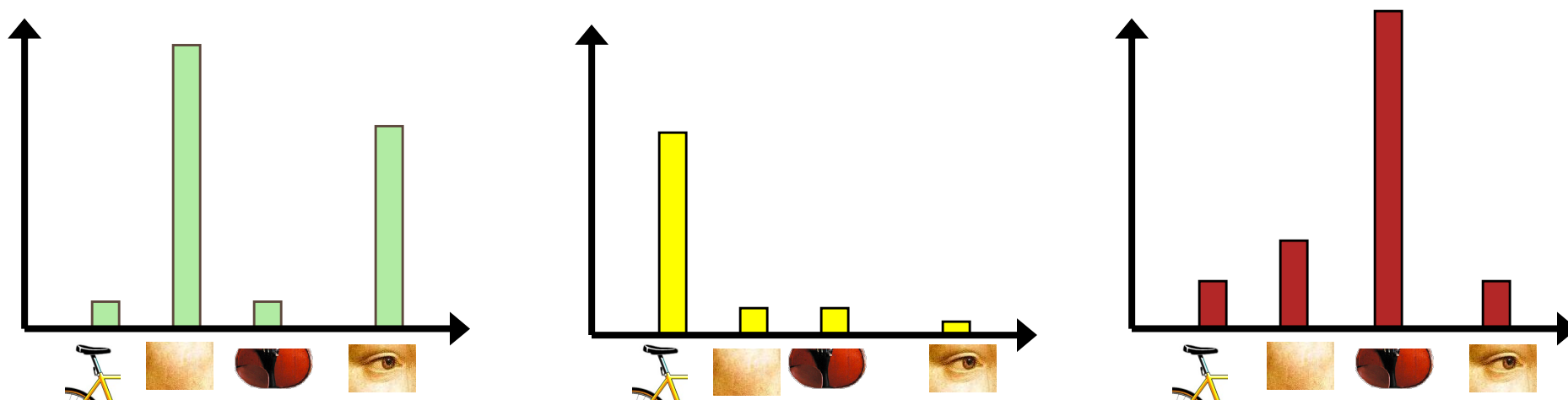
China's government has also needed to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

China's government has also needed to control the demand for the yuan in the country.

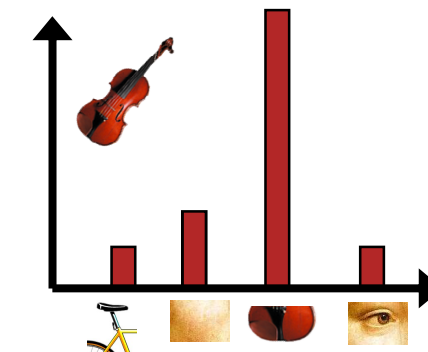
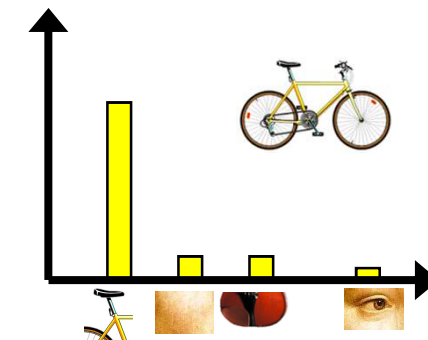
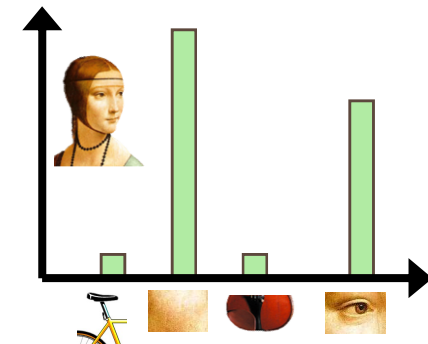
China's government has also needed to control the demand for the yuan in the country.

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



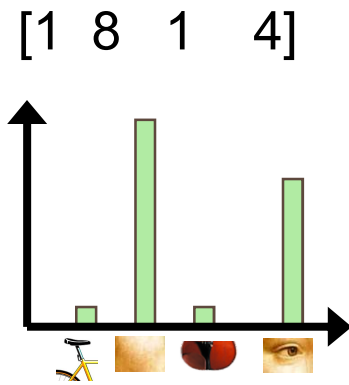
# Bags of visual words

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

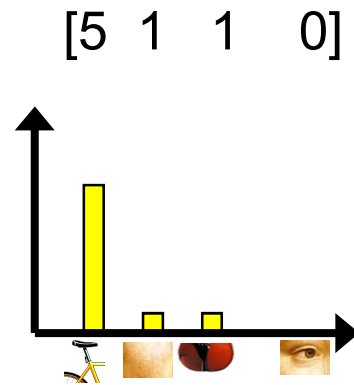


# Comparing bags of words

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



$\vec{d}_j$



$\vec{q}$

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

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

for vocabulary of  $V$  words

# *tf-idf* weighting

- Term frequency – inverse document 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

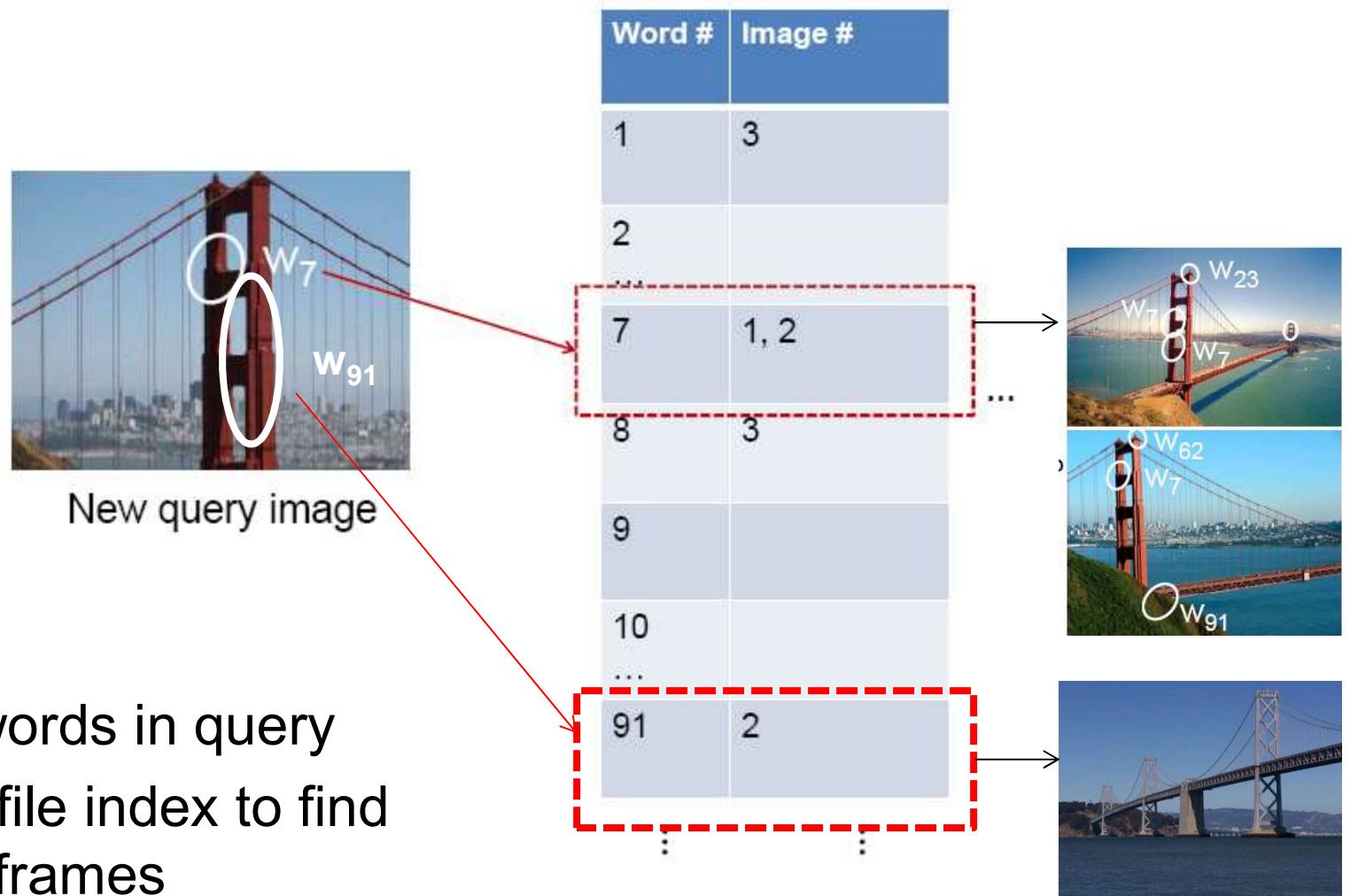
Number of words in  
document d

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

Total number of  
documents in  
database

Number of documents  
word i occurs in, in  
whole database

# Inverted file index and bags of words similarity

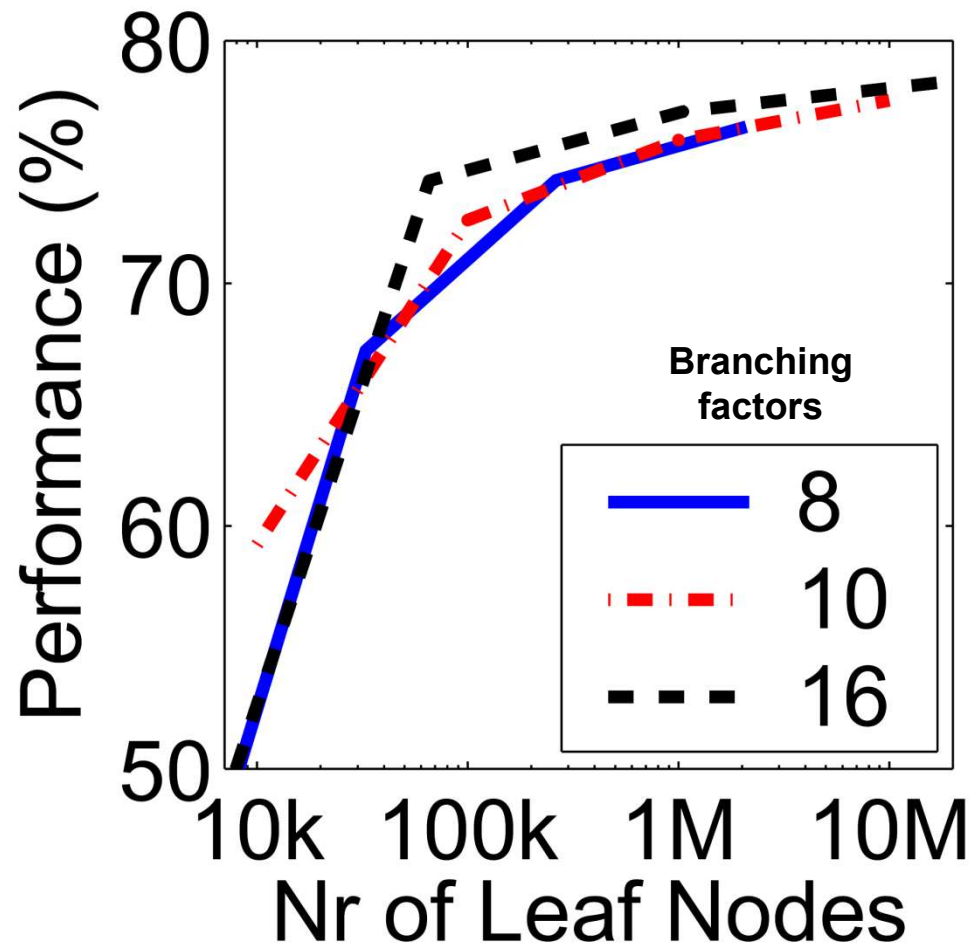


1. Extract words in query
2. Inverted file index to find relevant frames
3. Compare word counts

# Instance recognition: remaining issues

- How to summarize the content of an entire image? And gauge overall similarity?
- How large should the vocabulary be? How to perform quantization efficiently?
- Is having the same set of visual words enough to identify the object/scene? How to verify spatial agreement?
- How to score the retrieval results?

# Vocabulary size



Results for recognition task with 6347 images

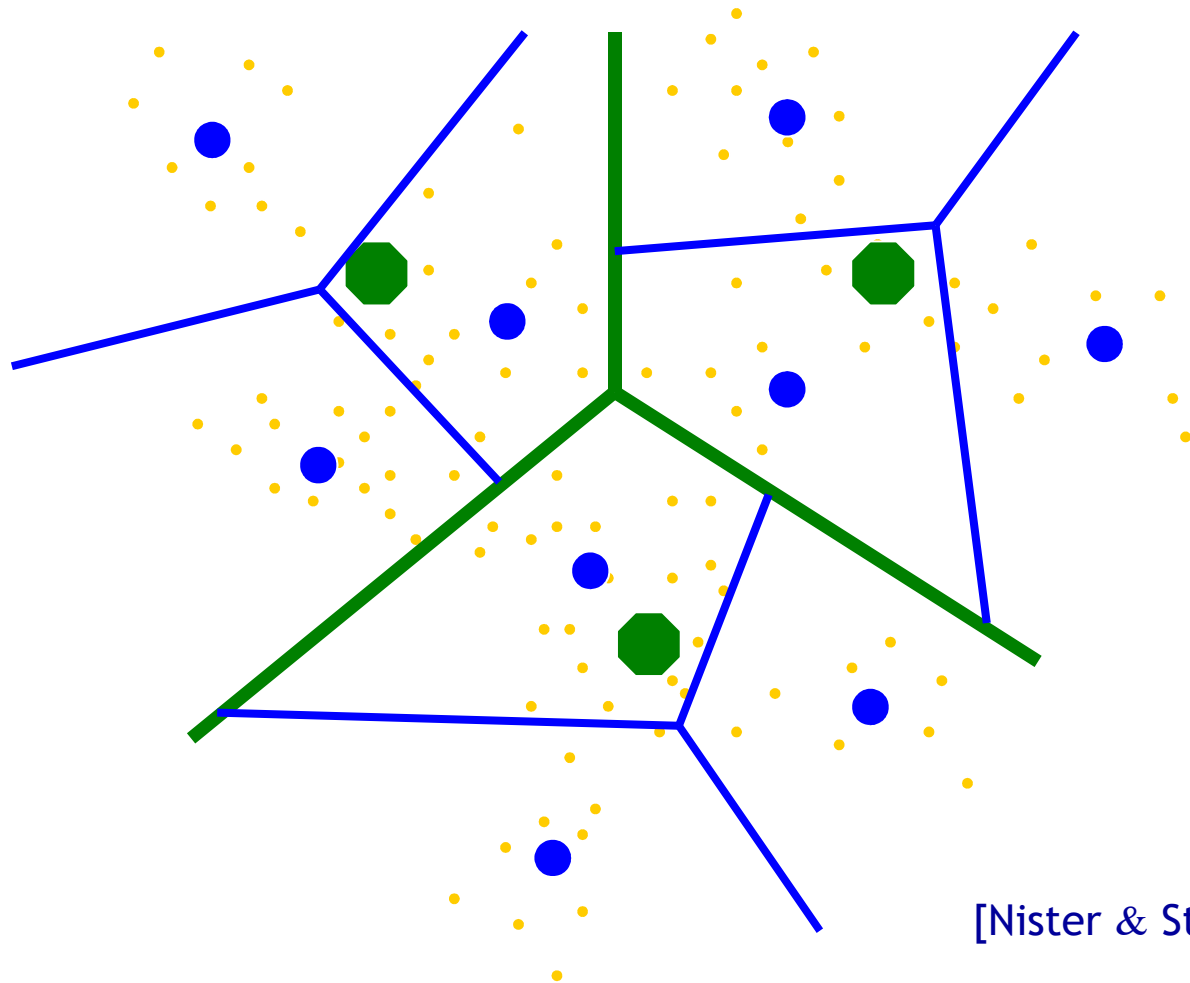


*Influence on performance, sparsity?*

Nister & Stewenius, CVPR 2006

# Vocabulary Trees: hierarchical clustering for large vocabularies

- Tree construction:

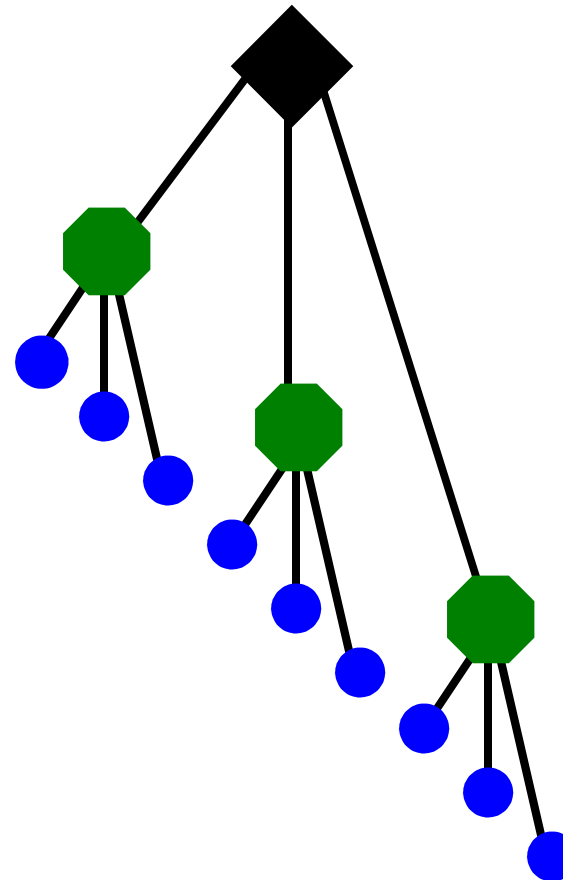


[Nister & Stewenius, CVPR'06]

K. Grauman, B. Leibe

Slide credit: David Nister

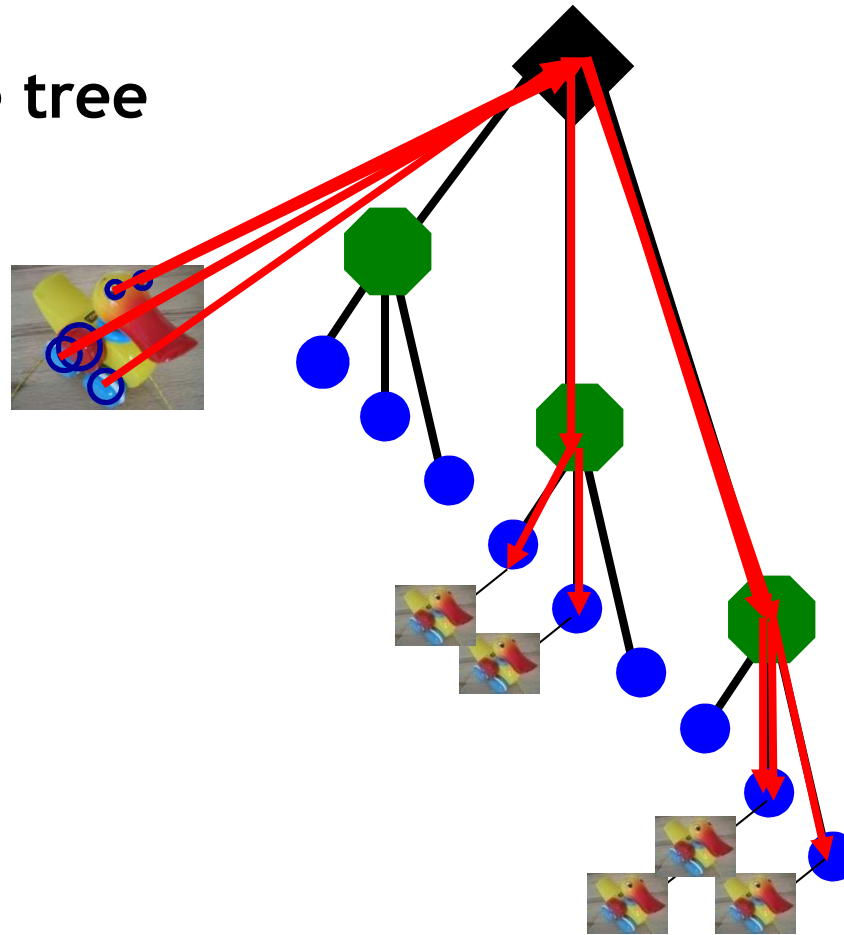
# Vocabulary Tree



[Nister & Stewenius, CVPR'06]

# Vocabulary Tree

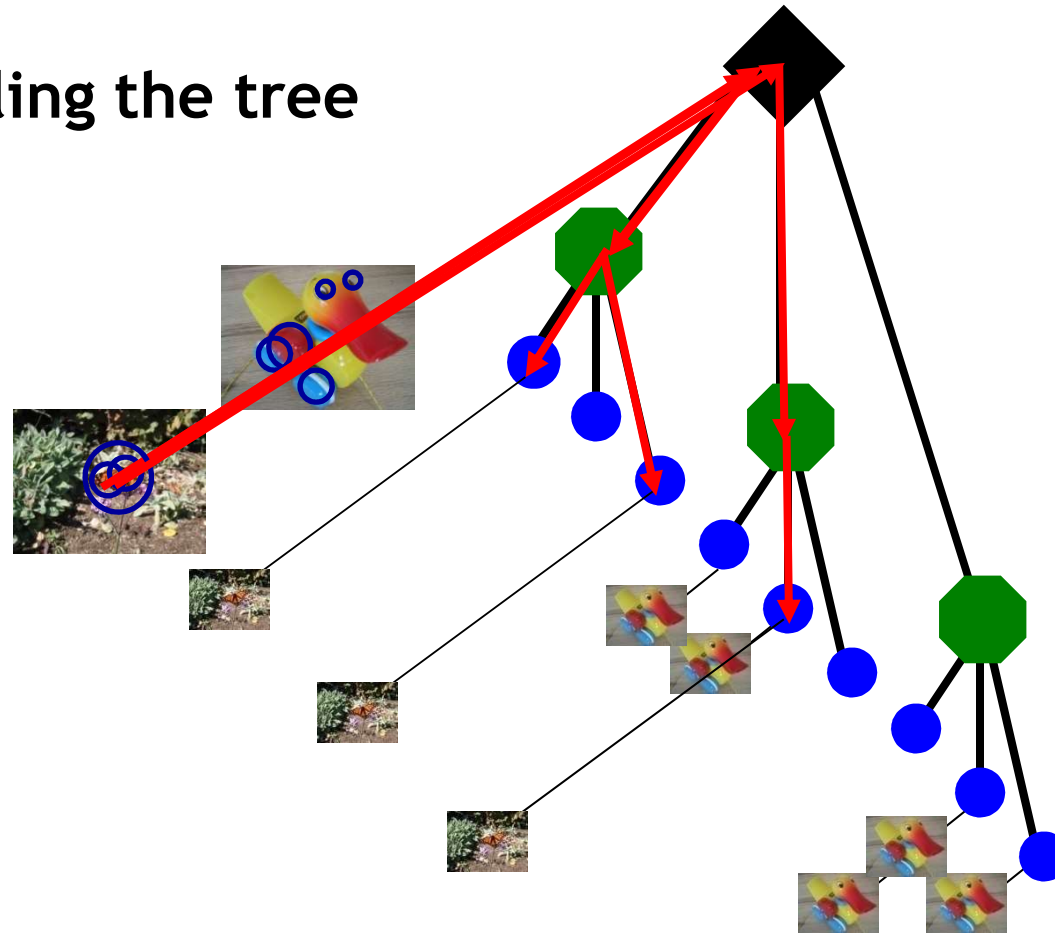
- Training: Filling the tree



[Nister & Stewenius, CVPR'06]

# Vocabulary Tree

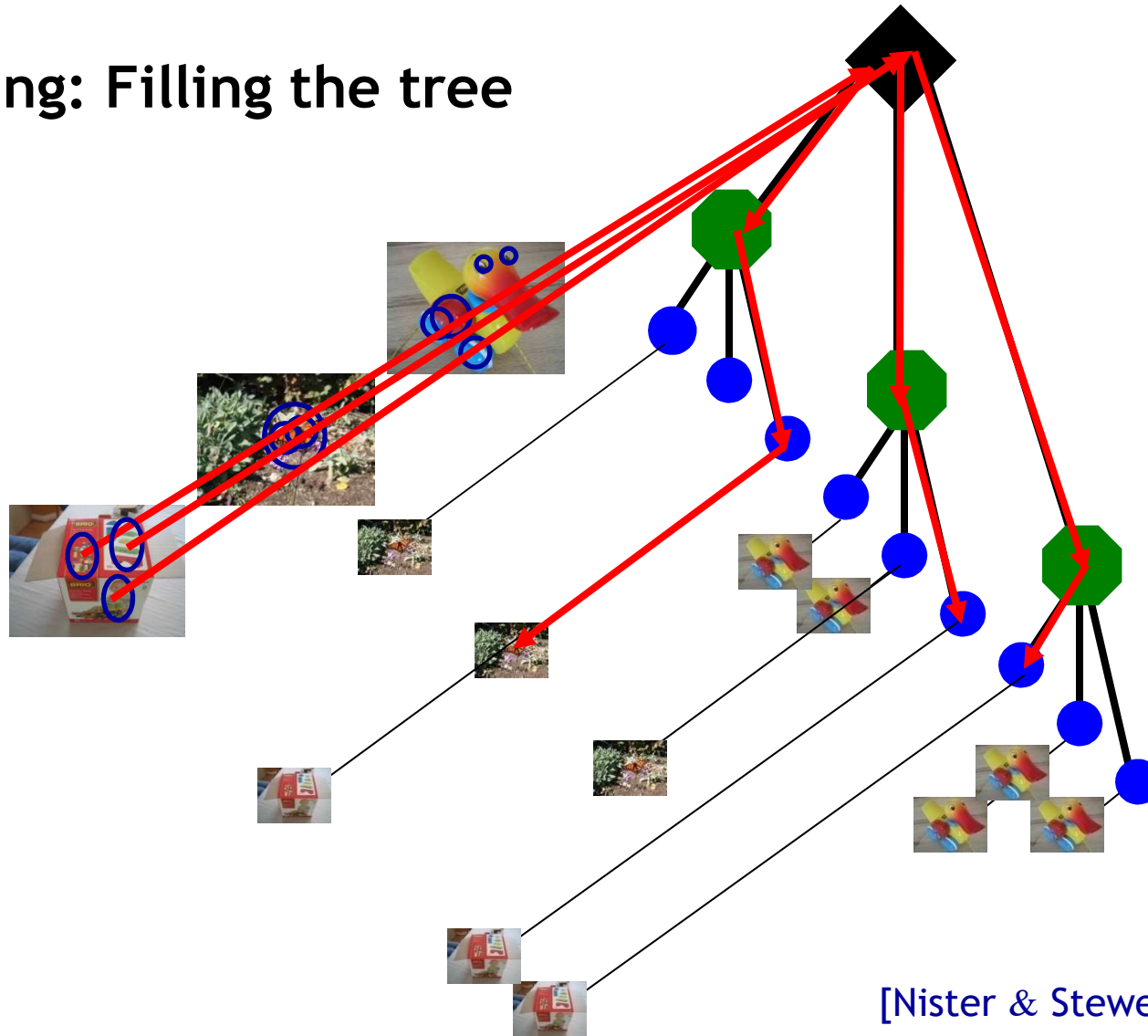
- Training: Filling the tree



[Nister & Stewenius, CVPR'06]

# Vocabulary Tree

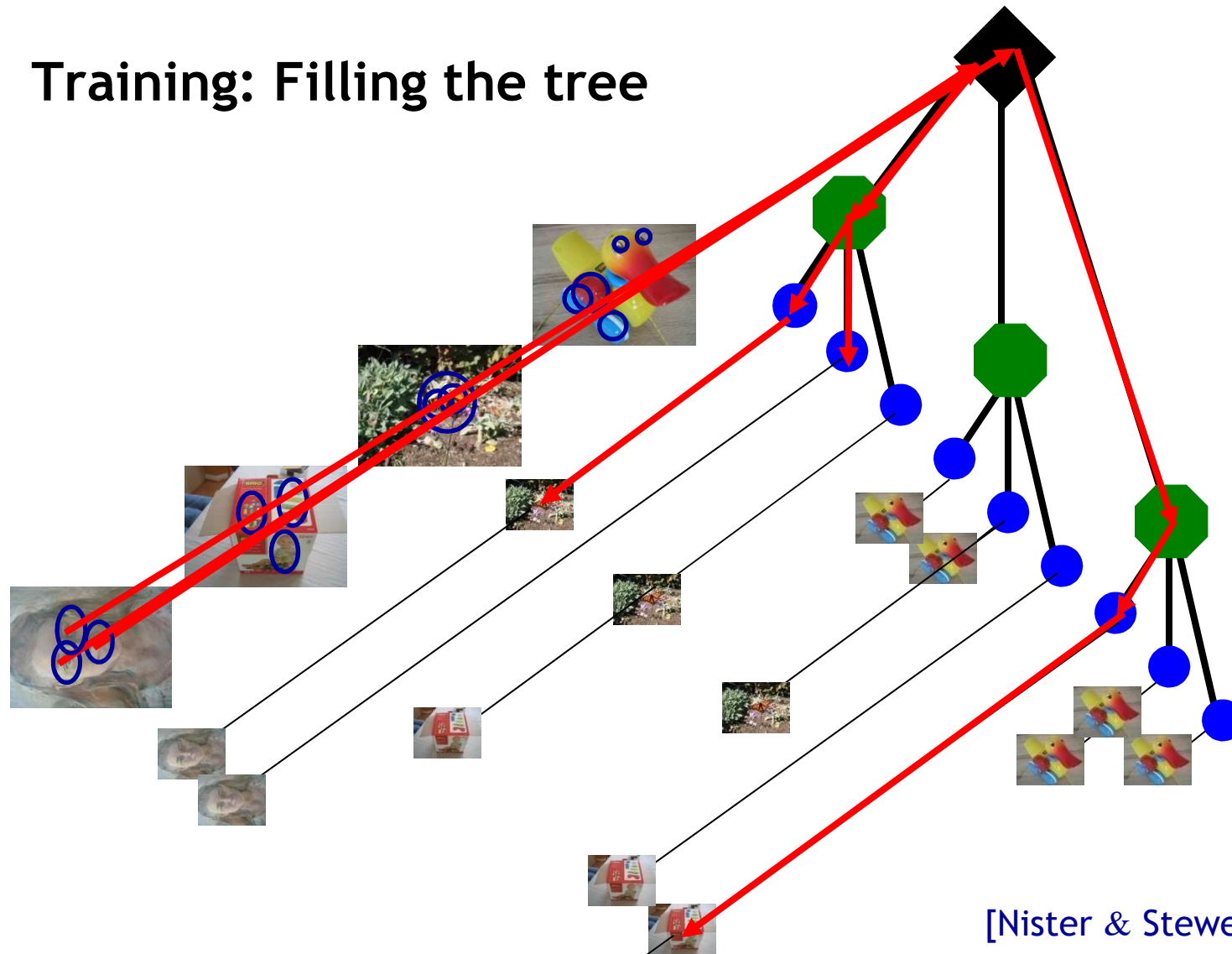
- Training: Filling the tree



[Nister & Stewenius, CVPR'06]

# Vocabulary Tree

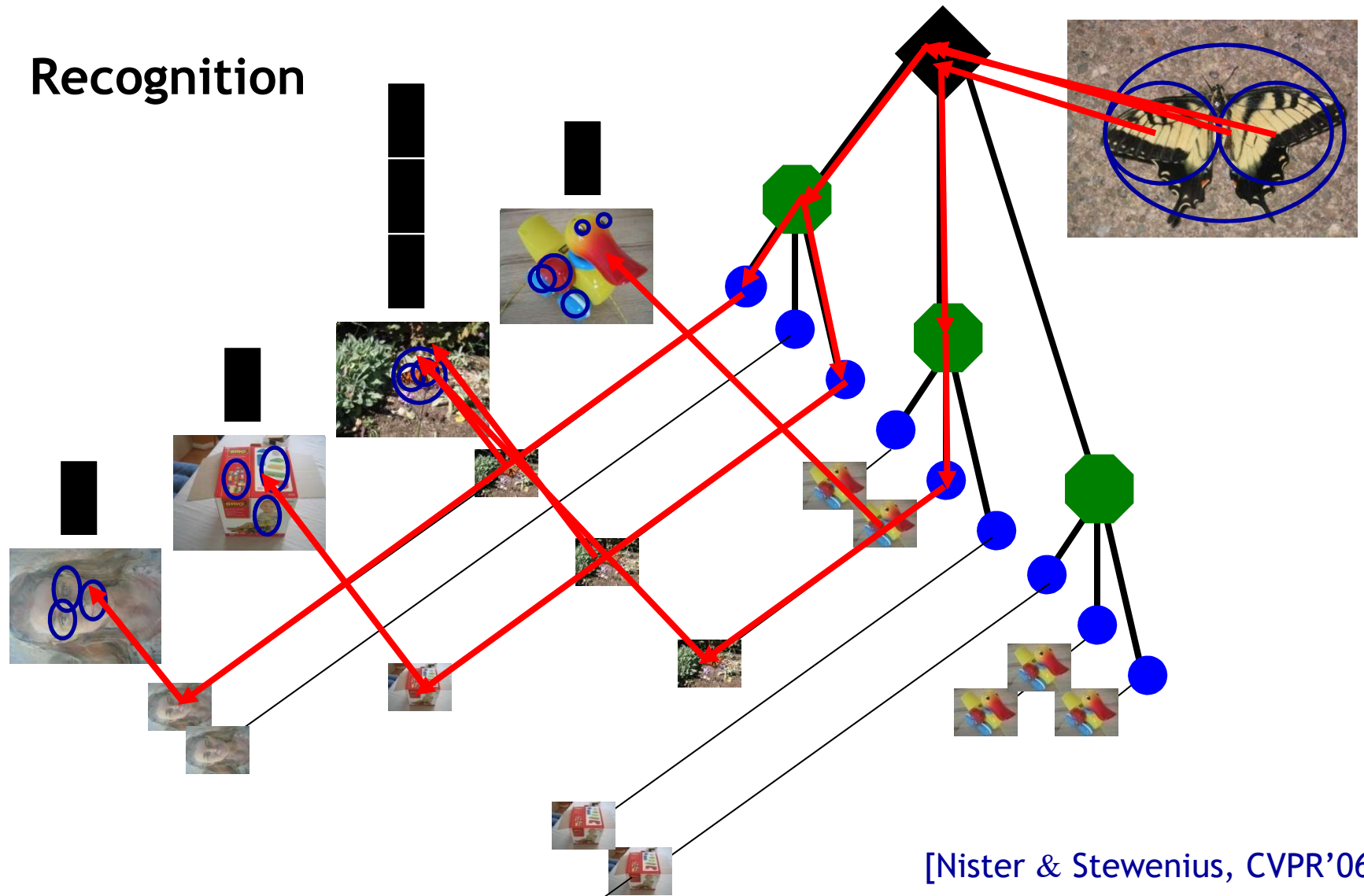
- Training: Filling the tree



[Nister & Stewenius, CVPR'06]

# Vocabulary Tree

- Recognition



[Nister & Stewenius, CVPR'06]

# Vocabulary trees: complexity

Number of words given tree parameters:  
branching factor and number of levels

Word assignment cost vs. flat vocabulary

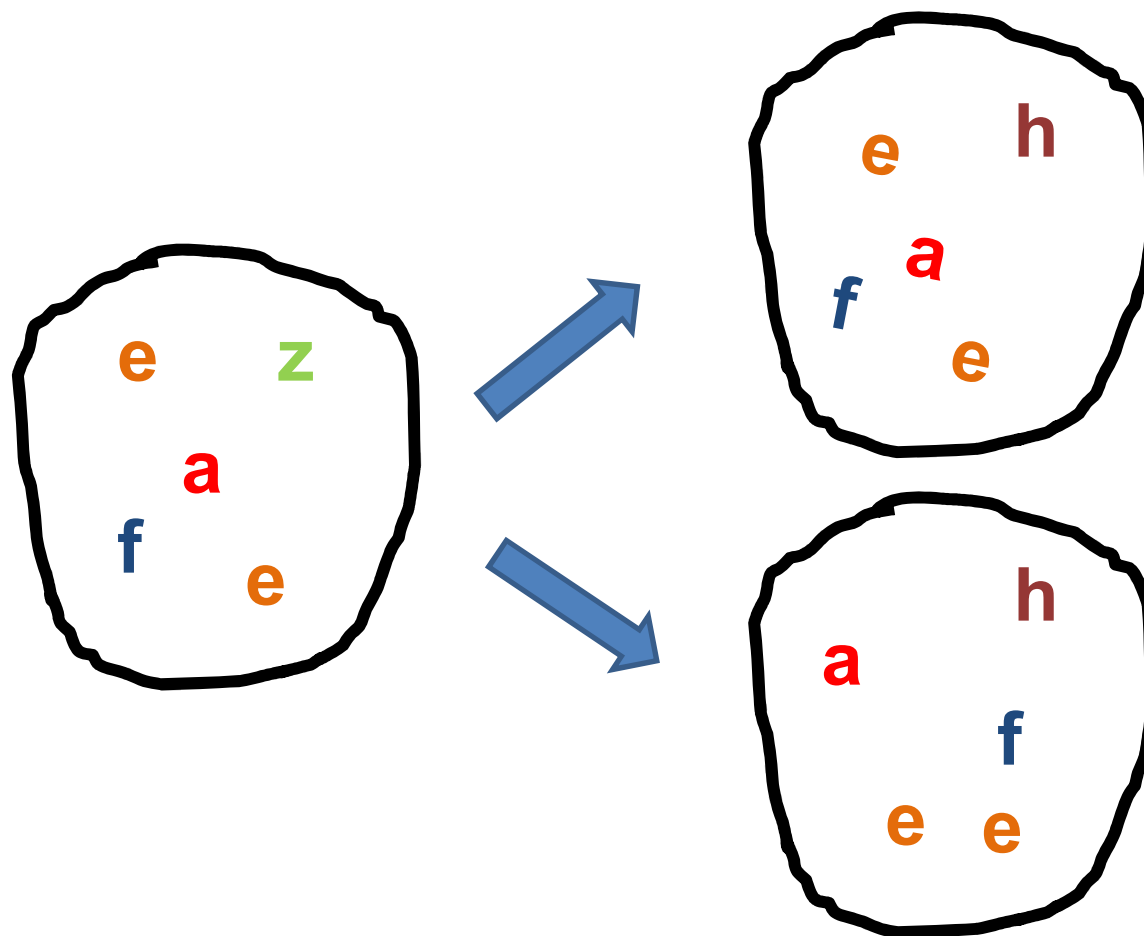
# Visual words/bags of words

- + flexible to geometry / deformations / viewpoint
- + compact summary of image content
- + provides vector representation for sets
- + very good results in practice
- background and foreground mixed when bag covers whole image
- optimal vocabulary formation remains unclear
- basic model ignores geometry – must verify afterwards, or encode via features

# Instance recognition: remaining issues

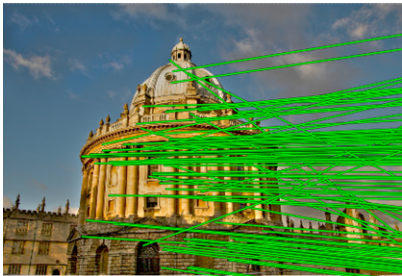
- How to summarize the content of an entire image? And gauge overall similarity?
- How large should the vocabulary be? How to perform quantization efficiently?
- Is having the same set of visual words enough to identify the object/scene? How to verify spatial agreement?
- How to score the retrieval results?

*Which matches better?*



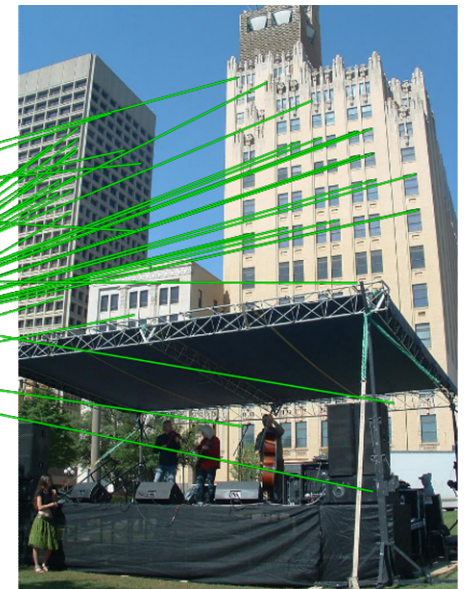
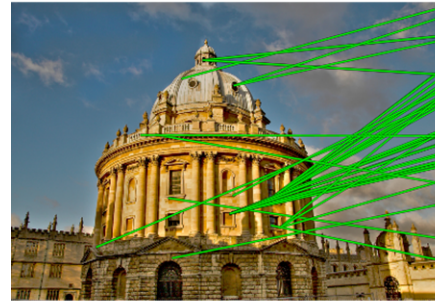
# Spatial Verification

Query



DB image with high BoW  
similarity

Query

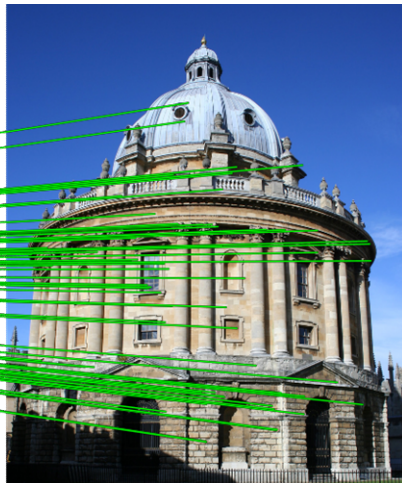
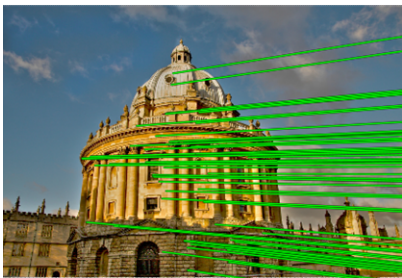


DB image with high BoW  
similarity

Both image pairs have many visual words in common.

# Spatial Verification

Query



DB image with high BoW similarity

Query



DB image with high BoW similarity

Only some of the matches are mutually consistent

# Spatial Verification: two basic strategies

- RANSAC
  - Typically sort by BoW similarity as initial filter
  - Verify by checking support (inliers) for possible affine transformations
    - e.g., “success” if find an affine transformation with  $> N$  inlier correspondences
- Generalized Hough Transform
  - Let each matched feature cast a vote on location, scale, orientation of the model object
  - Verify parameters with enough votes

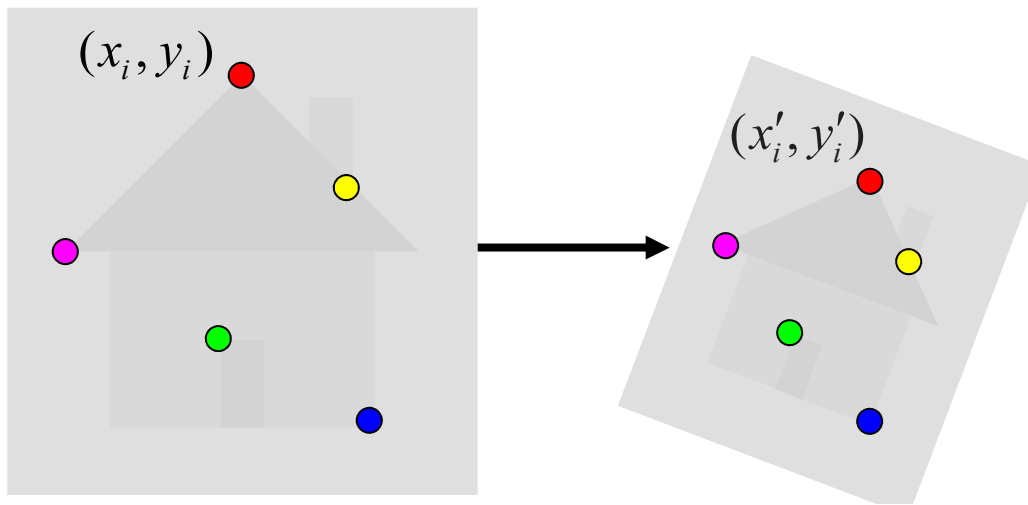
# RANSAC verification



For matching specific scenes/objects, common to use an **affine transformation** for spatial verification

# Fitting an affine transformation

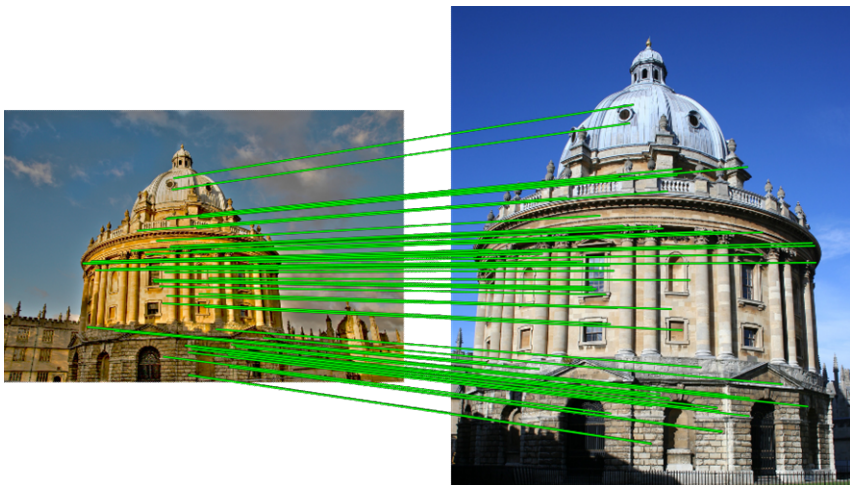
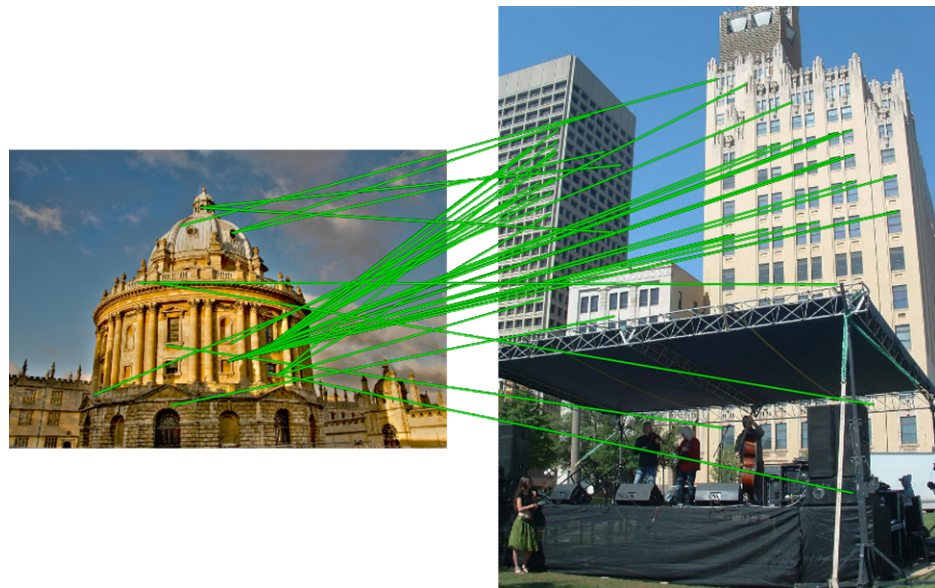
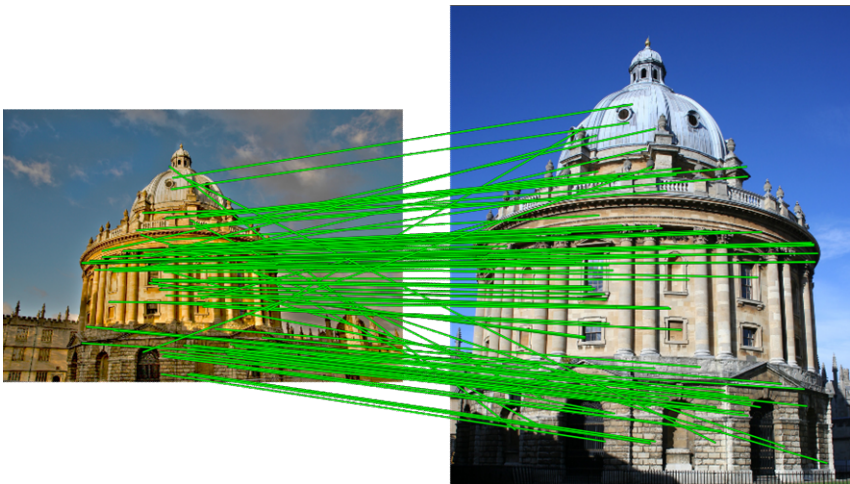
---



Approximates viewpoint changes for roughly planar objects and roughly orthographic cameras.

$$\begin{bmatrix} x'_i \\ y'_i \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

# RANSAC verification

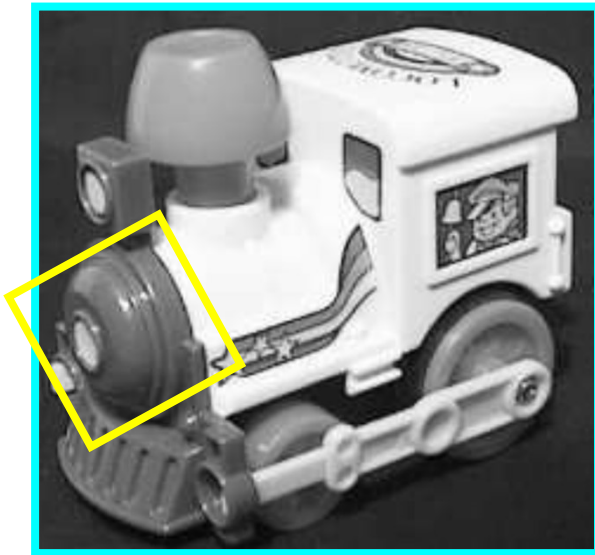


# Spatial Verification: two basic strategies

- RANSAC
  - Typically sort by BoW similarity as initial filter
  - Verify by checking support (inliers) for possible affine transformations
    - e.g., “success” if find an affine transformation with  $> N$  inlier correspondences
- Generalized Hough Transform
  - Let each matched feature cast a vote on location, scale, orientation of the model object
  - Verify parameters with enough votes

# Voting: Generalized Hough Transform

- If we use scale, rotation, and translation invariant local features, then each feature match gives an alignment hypothesis (for scale, translation, and orientation of model in image).



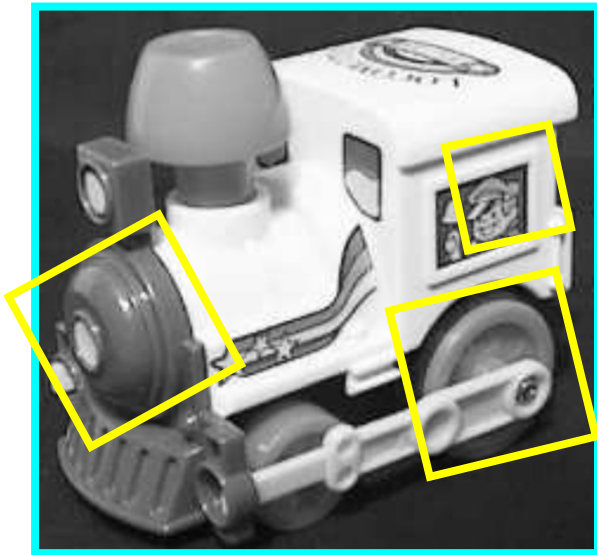
Model



Novel image

# Voting: Generalized Hough Transform

- A hypothesis generated by a single match may be unreliable,
- So let each match **vote** for a hypothesis in Hough space



Model



Novel image

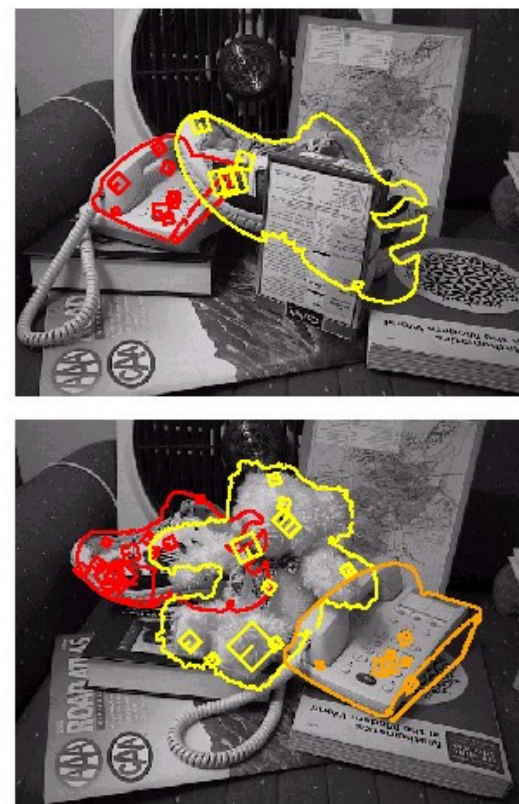
# Example result



Background subtract  
for model boundaries



Objects recognized,



Recognition in  
spite of occlusion

# Video Google System

1. Collect all words within query region
2. Inverted file index to find relevant frames
3. Compare word counts
4. Spatial verification

Sivic & Zisserman, ICCV 2003

- Demo online at :  
<http://www.robots.ox.ac.uk/~vgg/research/vgoogle/index.html>

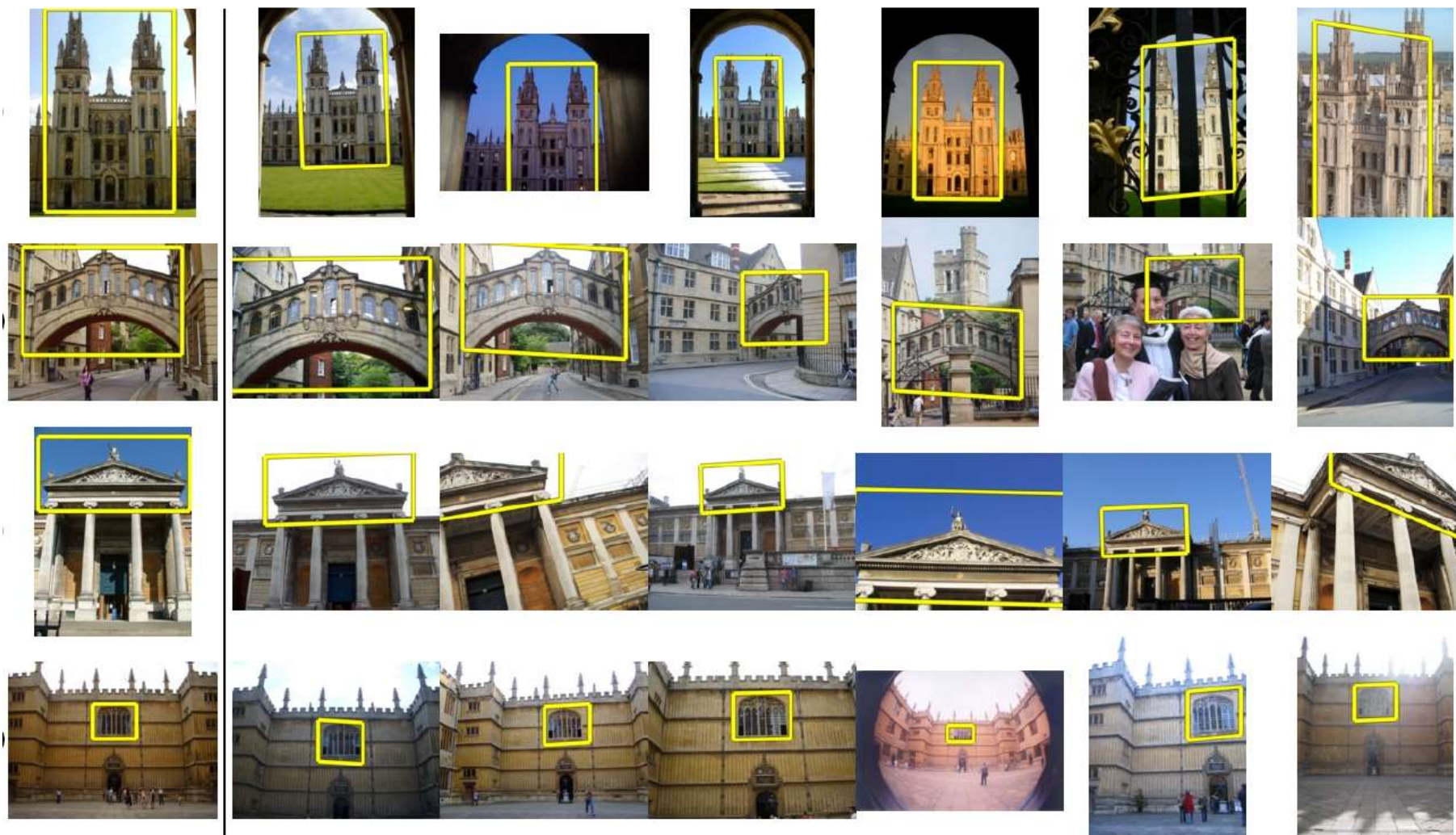


Query region



Retrieved frames

# Object retrieval with large vocabularies and fast spatial matching, Philbin et al., CVPR 2007



Query

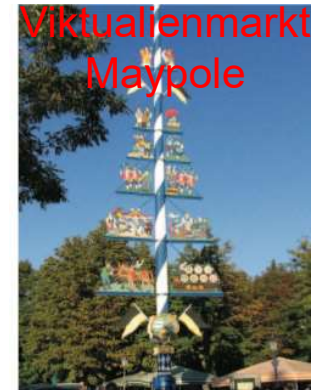
Results from 5k Flickr images (demo available for 100k set)

[Philbin CVPR'07]

# World-scale mining of objects and events from community photo collections, Quack et al., CIVR 2008



Auto-annotate by connecting to content on Wikipedia!



# Example Applications



## Mobile tourist guide

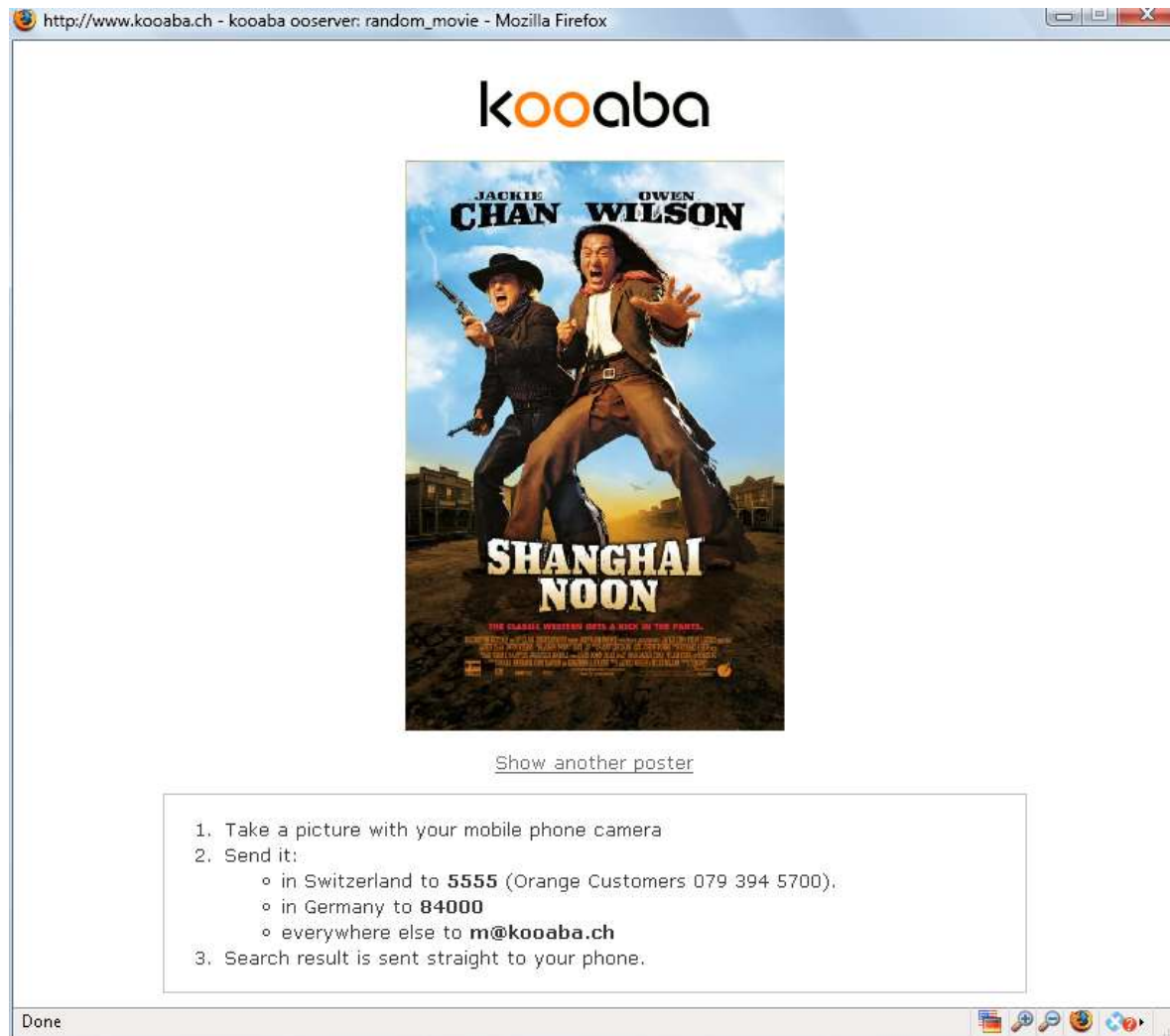
- Self-localization
- Object/building recognition
- Photo/video augmentation



# Web Demo: Movie Poster Recognition

50'000 movie posters indexed

Query-by-image from mobile phone available in Switzerland



[http://www.kooaba.com/en/products\\_engine.html#](http://www.kooaba.com/en/products_engine.html#)



# Google Goggles

Use pictures to search the web.

[▶ Watch a video](#)



## Get Google Goggles

**Android (1.6+ required)**

Download from Android Market.

[Send Goggles to Android phone](#)

**New! iPhone (iOS 4.0 required)**

Download [from the App Store](#).

[Send Goggles to iPhone](#)

New!



[Text](#)



[Landmarks](#)



[Books](#)



[Contact Info](#)



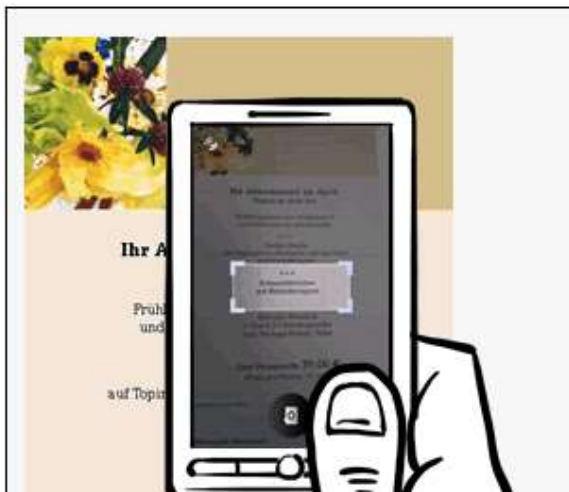
[Artwork](#)



[Wine](#)



[Logos](#)



# Recognition via feature matching+spatial verification

## **Pros:**

- Effective when we are able to find reliable features within clutter
- Great results for matching specific instances

## **Cons:**

- Scaling with number of models
- Spatial verification as post-processing – not seamless, expensive for large-scale problems
- Not suited for category recognition.

# What else can we borrow from text retrieval?

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 Natl 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	Choctawhatchee, 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	Ft 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; 95
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
	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
	Esterio; 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

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% increase in exports to \$750bn, compared with \$575bn in 2004, and a 10% increase in imports to \$660bn. The US Trade Representative said the surplus would annoy the US because it would mean China's trade policy was deliberately designed to create a surplus. China's government agrees that the surplus is a problem, but the government also needs to increase demand so that the surplus can be absorbed by the country. China has been increasing the value of the yuan against the dollar, but the US has not permitted it to trade within a narrow band, but the US wants the yuan to be allowed to move more freely. However, Beijing has made it clear that it will take its 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**

# Query expansion

Query: ***golf green***

Results:

- How can the grass on the ***greens*** at a ***golf*** course be so perfect?
- For example, a skilled ***golfer*** expects to reach the ***green*** on a par-four hole in ...
- Manufactures and sells synthetic ***golf*** putting ***greens*** and mats.

Irrelevant result can cause a `topic drift`:

- Volkswagen ***Golf***, 1999, ***Green***, 2000cc, petrol, manual, , hatchback, 94000miles, 2.0 GTi, 2 Registered Keepers, HPI Checked, Air-Conditioning, Front and Rear Parking Sensors, ABS, Alarm, Alloy

# Query Expansion

Results

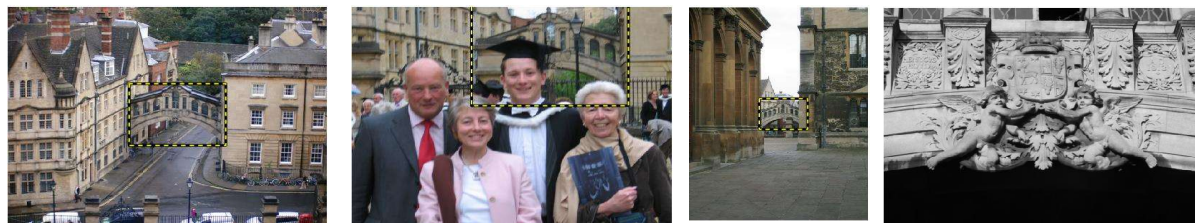


Query image

↓ Spatial verification



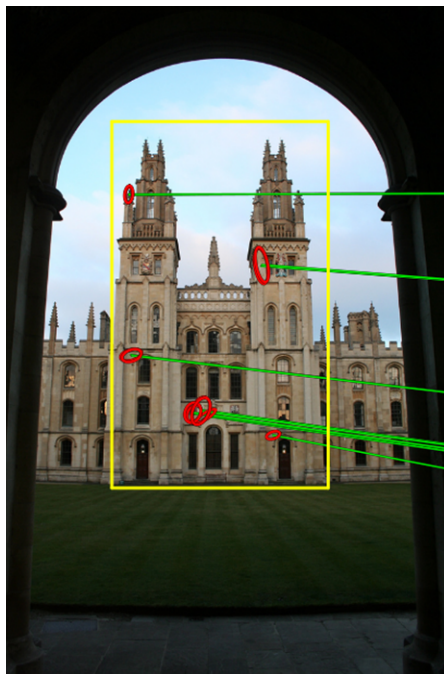
New results



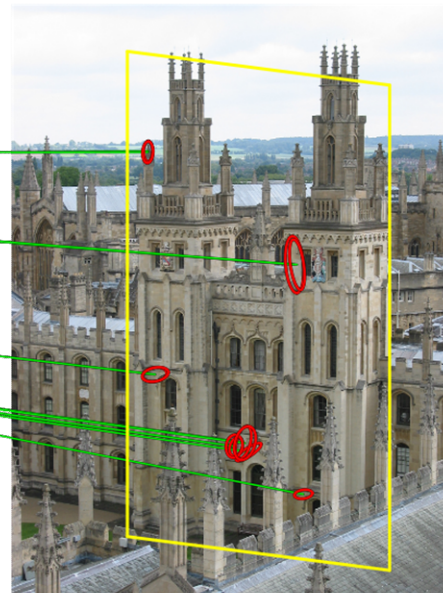
New query

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

# Query Expansion Step by Step



Query Image

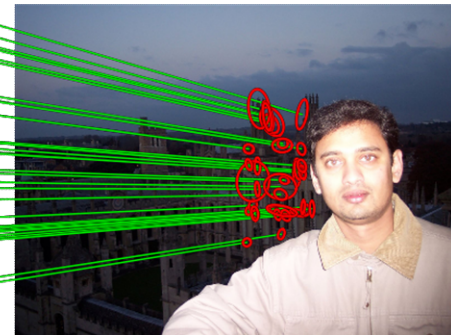
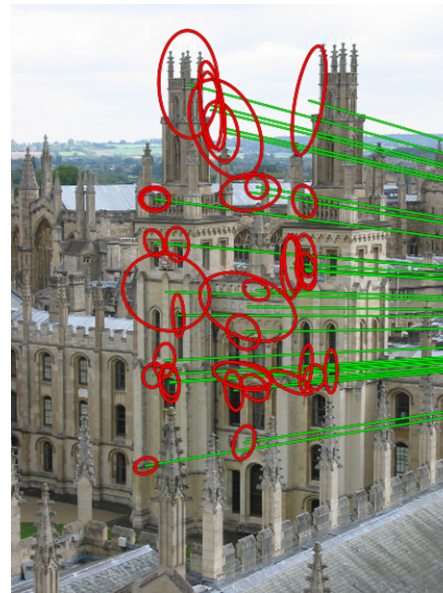
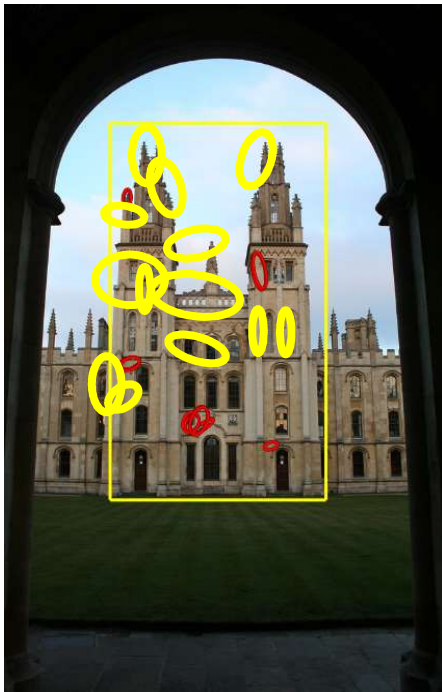


Retrieved image



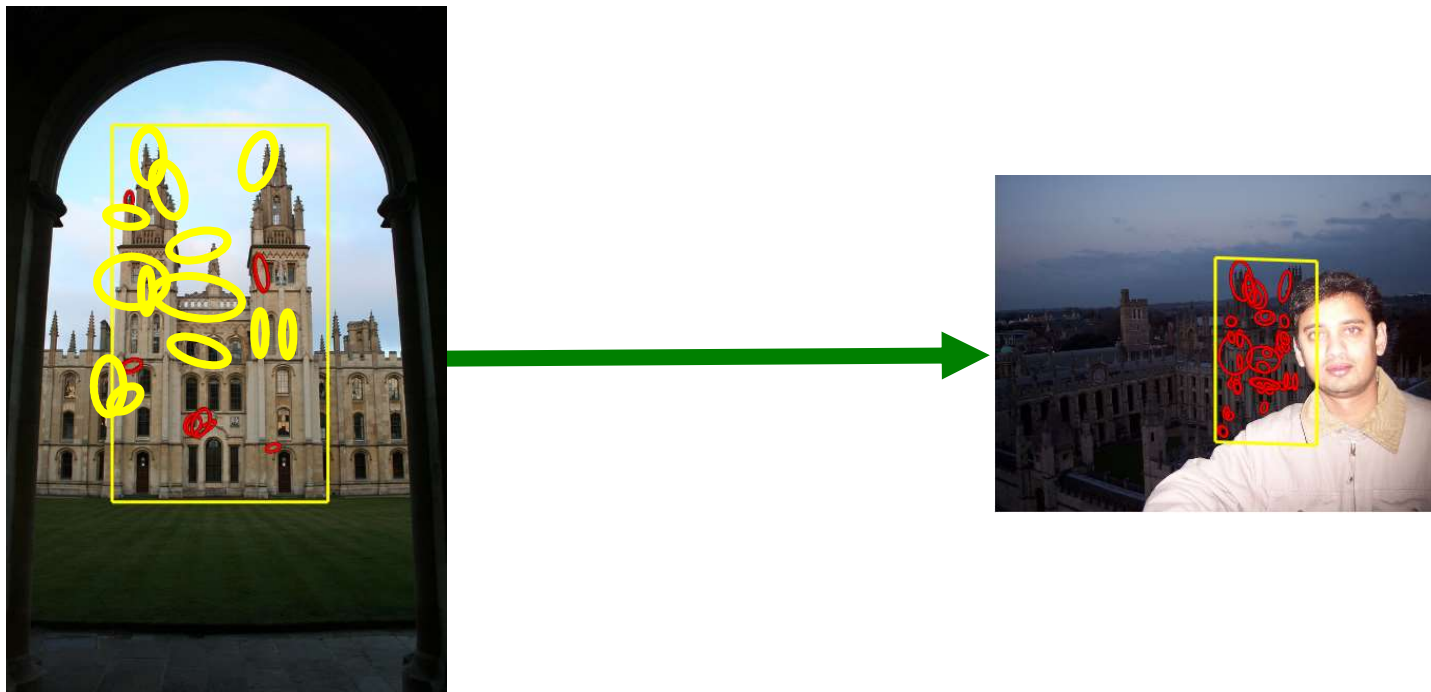
Originally not retrieved

# Query Expansion Step by Step



Slide credit: Ondrej Chum

# Query Expansion Step by Step



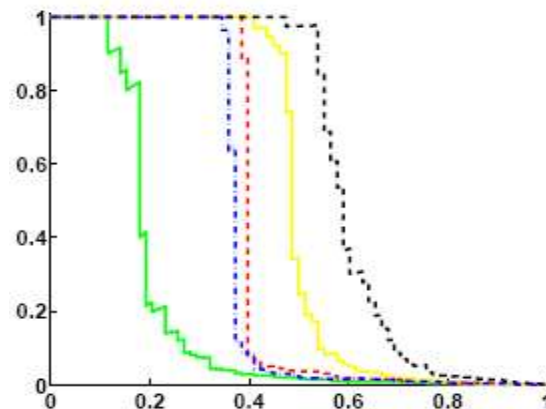
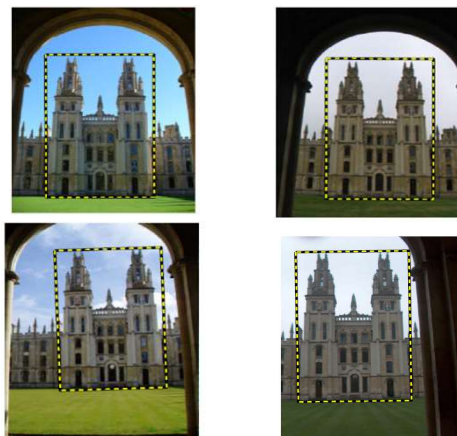
Slide credit: Ondrej Chum

# Query Expansion Results

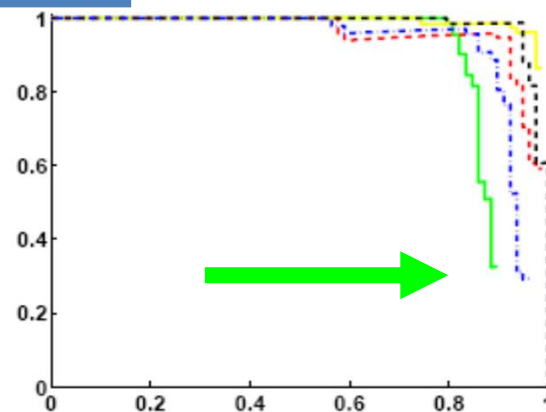


Query image

Original results (good)



Expanded results (better)



# Summary

- **Matching local invariant features**
  - Useful not only to provide matches for multi-view geometry, but also to find objects and scenes.
- **Bag of words** representation: quantize feature space to make discrete set of visual words
  - Summarize image by distribution of words
  - Index individual words
- **Inverted index**: pre-compute index to enable faster search at query time
- **Recognition of instances via alignment**: matching local features followed by spatial verification
  - Robust fitting : RANSAC, GHT