

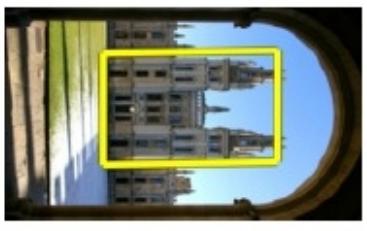
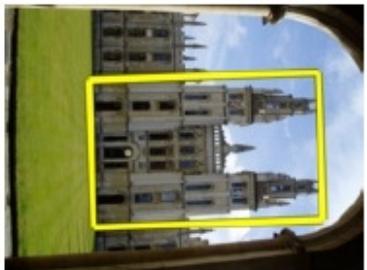
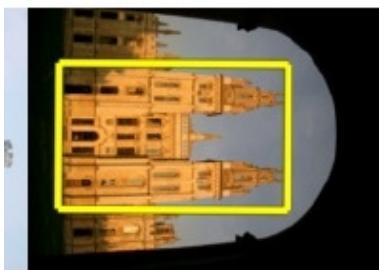
Large Scale Image Instance Recognition

Chetan Arora

Disclaimer: The contents of these slides are taken from various publicly available resources such as research papers, talks and lectures. To be used for the purpose of classroom teaching, and academic dissemination only.

Problem Statement

- How to quickly find images in a large database that match a given image region?

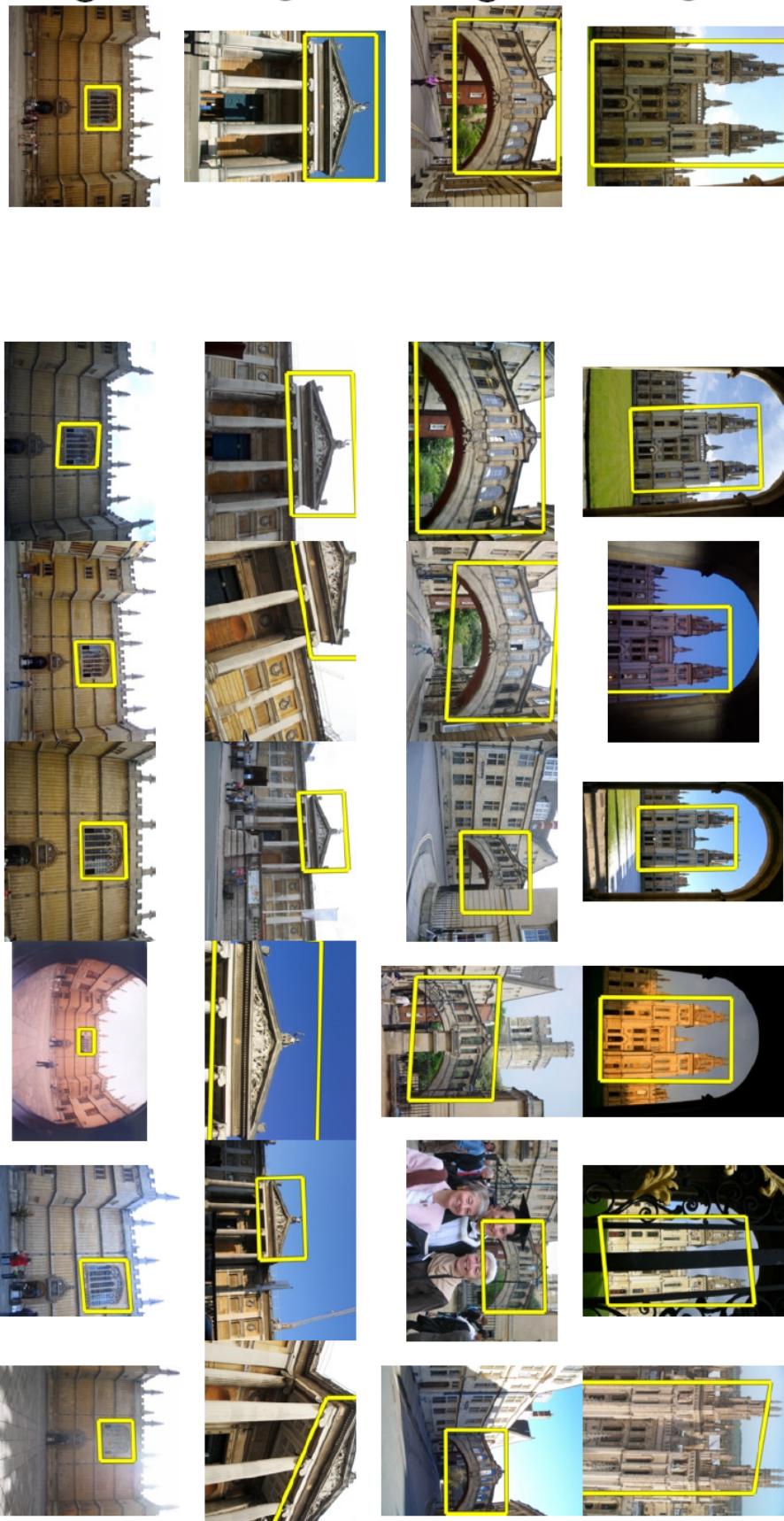




Applications: Large Scale Retrieval

Query

Results from 5k Flickr images





Applications: Image Auto Annotation



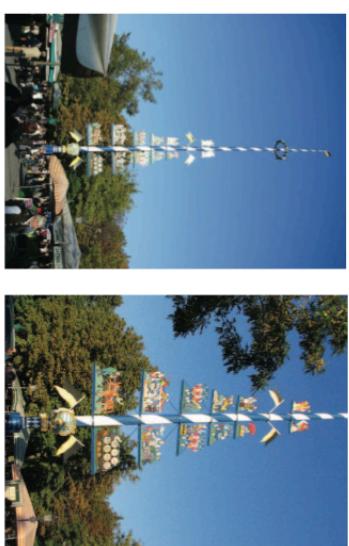
Moulin Rouge



Tour Montparnasse



Colosseum



Viktualienmarkt
Maypole



Old Town Square (Prague)

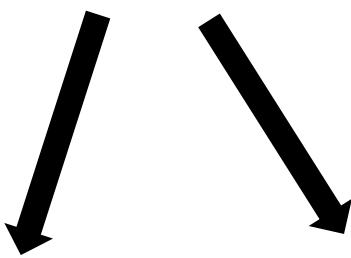
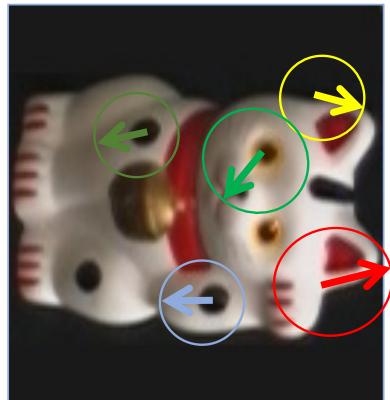


Left: Wikipedia image
Right: closest match from Flickr

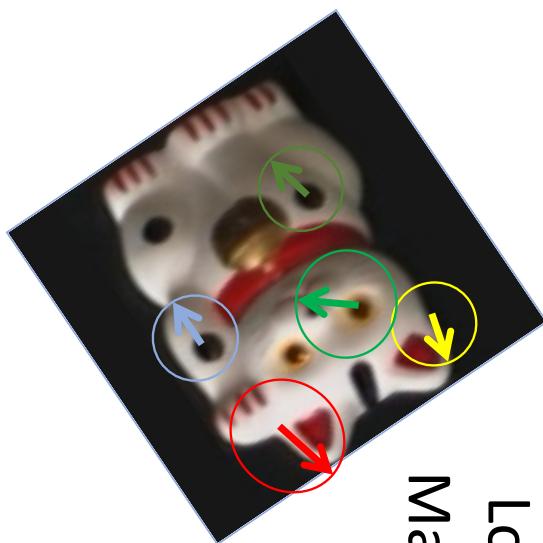


Image Retrieval: Naive Idea

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



Lots of
Matches



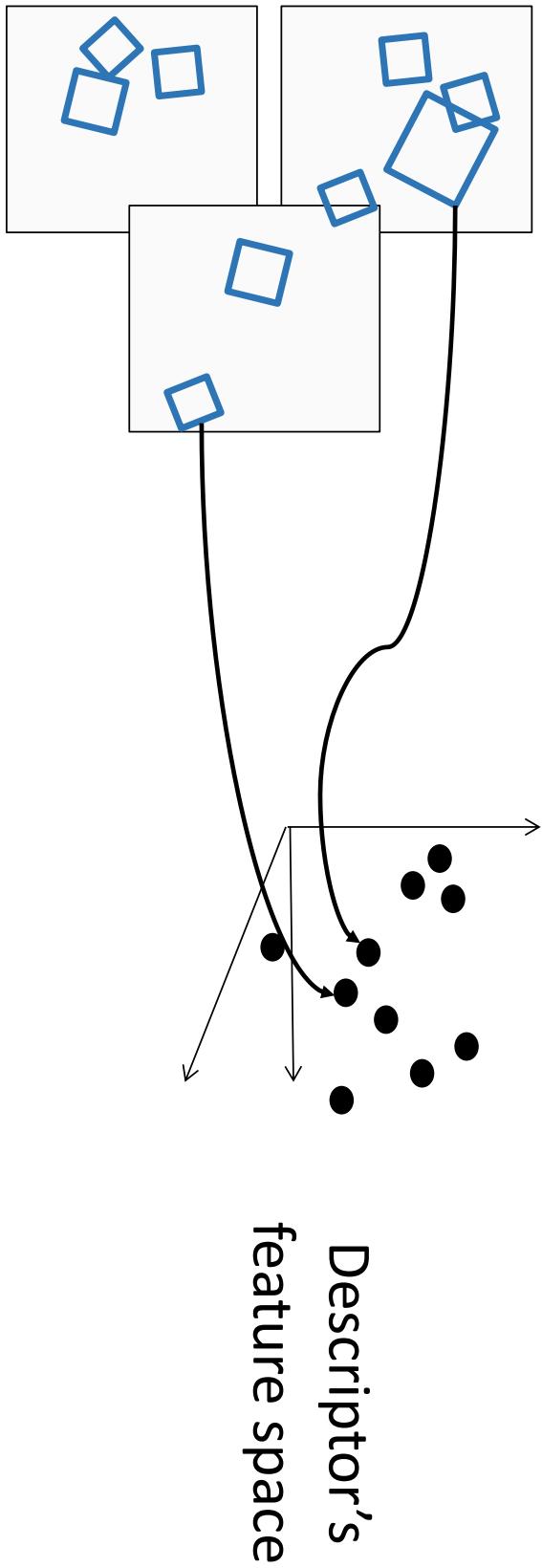
Few or No
Matches



Problem: This will be really, really slow!

Indexing Local Features

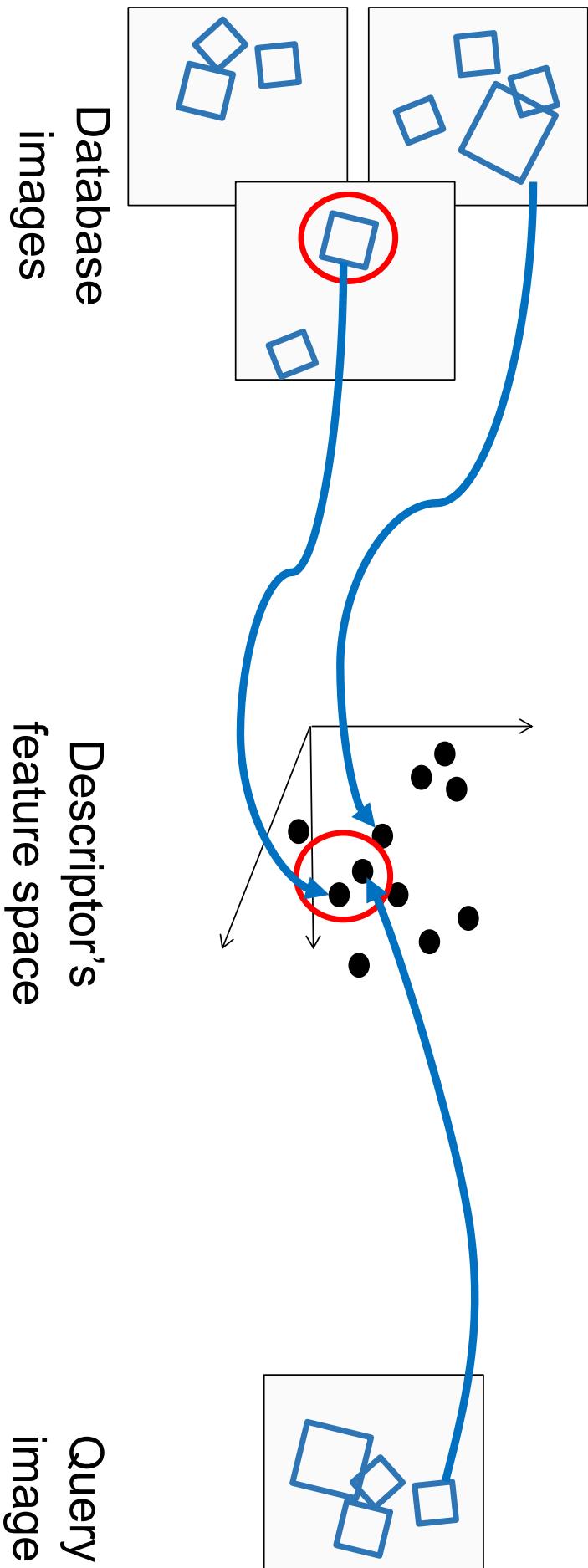
- Each patch / region has a descriptor, which is a point in some high-dimensional feature space (e.g., SIFT).





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

- 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”.

Index

"Along I-75," From Detroit to Florida; Inside Back cover

"Drive-I-95; From Boston to Florida; Inside Back cover

1920 Spanish Trail Roadway; 111-112; 104

5-11 Traffic Information; 89

AAA (and CAA); 83

National Office; 88

Colored 25-mile Maps; cover

Travel Services; 85

Aerospace Inspection Strs; 126

Automobile Inspection Strs; 126

Air Conditioning, First; 112

Alabama; 124

Alachua County; 132

Alaska; Fever; 143

Alaska, Salmon Gardens; 106

Aligator Alley; 155

Aligator Farm, St Augustine; 169

Aligator, Baby (definition); 157

Aligator, Buddy; 155

Alabamians; 130, 135, 138, 147-156

Alabamian Band; 170

Alamogordo; 146

Alameda, San Francisco; 142

Alpenhorn Mass Art; 136

Arabian Knights; 94

Art Museum, Ringling; 147

Arthur Beach, Cape; 183

Aucilla River Project; 106

Babcock-Wilcox Nuclear; 151

Baker Maritime Center; 194

Barefoot Mallin'; 182

Barge Canal; 137

Base Line Expy; 80

Bezel Outfitter Mall; 89

Bernard Castro; 136

BIG T; 165

BIG T, Inc. 158

BIG T, Inc., Master; 105

Big Swallow Sabot; 60

Blisswater River SP; 117

Blue Angels; 144

Bluebird Inn; 117

Autumn; 121

Blue Gas SP; 87

Blue Gas, Federal Highway; 125

Boca Grande; 150

Boca Raton; 182

Bonnie Blue Flag; 124

Bonnie Blue Trail; 188

Braydon's; 145-147

Breakers, The; Palm Beach; 181

Brenton Hotel, Miami; 185

Broadway; 107

Bromeliads (see Epiphytes)

Broward County; 159, 181

Brownie; 155

Bush, Plantation Farms; 171

Bush, Gov.; Job; 100

Butterfly Center, McGuire; 134

CAA (see AAA)

C-C-C, The; 111-113, 115, 135, 142

Ca. d'Zan; 147

Choccolocco River; 152

Namee; 110

Cameron New Seashore; 173

Carney Road; 106-109

Cape Canaveral; 174

Cassidy San Marcos; 169

Cave Diving; 131

Cayo Coasta; 98

Cebuano, Namee; 150

Chamberlain, Horace; 149

Charlevoix; 116

Chippewa; 114

Namee; 115

Choctawhatchee, Namee; 115

Circus Museum; Ringling; 147

Church; 88, 97, 130, 136, 140, 180

City Beach; 180

Florida Expressways; 134-195

Jacksonville Expressways; 163

Kosummarie Expressways; 192-193

Miami Expressways; 194-195

Orlando Expressways; 192-193

Pensacola; 20

Tallahassee; 119

Timeless Sunshine; 63

St. Augustine; 181

Civil War; 100, 108-127, 138, 141

Cleaveland Marine Aquarium; 187

Collier County; 154

Collier, Barron; 152

Colonial Spanish Quartermars; 168

Comberbank, John; 101-128

Commons Building, Tallahassee; 165

Cookhouse Inn, Namee; 154

Cookways; 95

Cracker, Florida; 88, 95, 132

Crossover Express; 134, 135, 98, 143

Cuban Bread; 154

Dade Battlements; 149

Dade, Big, railroads; 139-140, 161

Daniel Boone, Florida Walk; 117

Daytona Beach; 172-173

De Land; 87

De Soto, Hernando; 189

Amber; 108-109, 146

County; 149

Laporte; 146

Native Americans; 163

National Park; 147

Tallahassee; 147

DeFunak Springs; 116

Namee; 115

DeWitt Wrights Gas SP; 155

Dennies, Matt; St. Augustine; 109

Dickie, Azalea Park; 69

Dimocrot World; 88

Discovery Cove; 90

Dixie Highway; 96

Don Garita, Drag Slacking Mtns; 138

Douglas, Marjorie St. John; 159

Driving Lanes; 85

Duval County; 163

Eau Gallie; 175

Edison, Thomas; 182

Eglin AB; 116-118

Elliott, Franklin; 176

Elmer, Point Wreck; 20

Emergency Callboxes; 83

Epiphyses; 142, 148, 157, 159

Escambia Bay; 119

Bridge (I-10); 119

County; 120

Ester; 153

Evans; 90, 95, 139-140, 154-160

Welles MA; 160

Wonder Gardens; 154

Florida; 12,000 years ago; 187

Caverns SP; 114

Fantasy of Flight; 95

Fayev Dikov; SP; 171

Fire; 168

Forest, Piney Woods; 148

Fees; 155

Flagler County; 171

Flagler, Henry; 171, 165, 171

Florida Aqualure; 186

Wright MA; 160

Falling Waters SP; 115

Fantasy of Flight; 95

Fayev Dikov; SP; 171

Fire; 168

Forest, Piney Woods; 148

Fees; 155

Flagler County; 171

Flagger, Henry; 171, 165, 171

Florida, 12,000 years ago; 187

Caveins SP; 114

Map of all Expressways; 2-3

Map of National Parks; 134

National Committee; 141

Part of Africa; 177

Platorm; 187

Sheriff's Boys Camp; 126

Sports Hall of Farmer; 130

Sun n Fun Museum; 97

Supreme Court; 102

Florida Strip; 171, 172, 189

Administration; 169

Florida Strip; 171, 172, 189

Administrative; 169

Corn System; 190

Exit System; 190

Exit; 189

HETF; 76, 161, 190

History; 189

Names; 189

Service Plan; 190

Service Plan; 190

Service Plan; 190

Toll Plazas; 180

Ford, Harry; 152

Ford Bartrams; 122

Burned Aves; 123

Fort Caroline; 164

Fort Davis; 161

Fort De Soto; 180

Fort Lauderdale; 161-184

Fort Myers; 152-153

Fort Pierce; 177-178

Farmers Market; 180

Fountain of Youth; 170

Frank Lloyd Wright Center; 97

Garrison, County; 11

Garrison, County; 11-135, 146

Garrison, Daniel; 145

Gardner, of Eden; 112

Gasparsen, Prairie; 150

Gatorland; 134

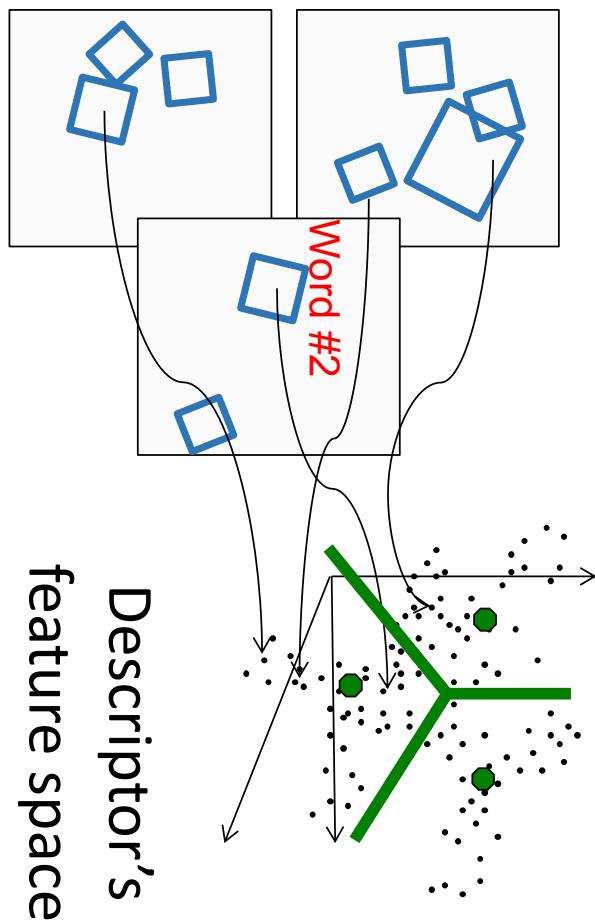
Gaylord Palms; 90

Geology; 102-103, 110, 111, 131-132



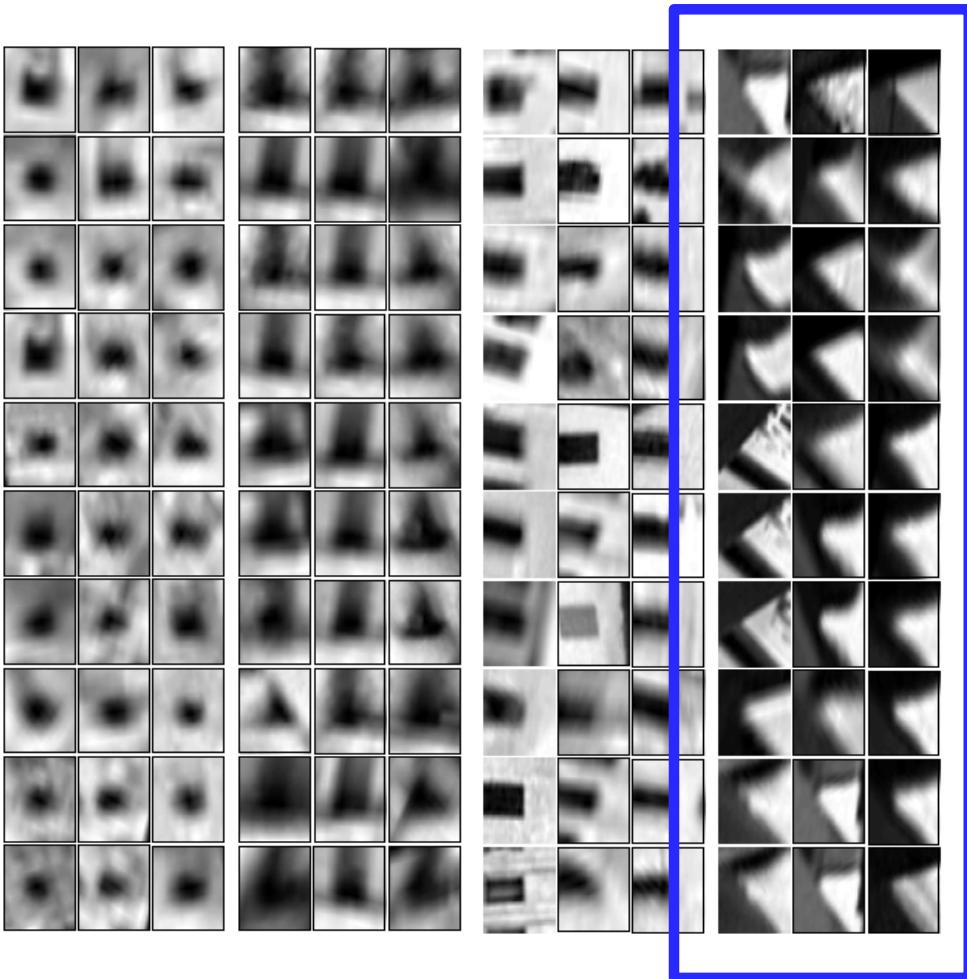
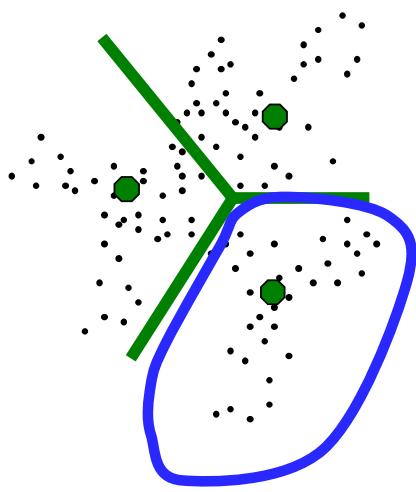
Visual Words

- 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





Inverted file index

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

Database images

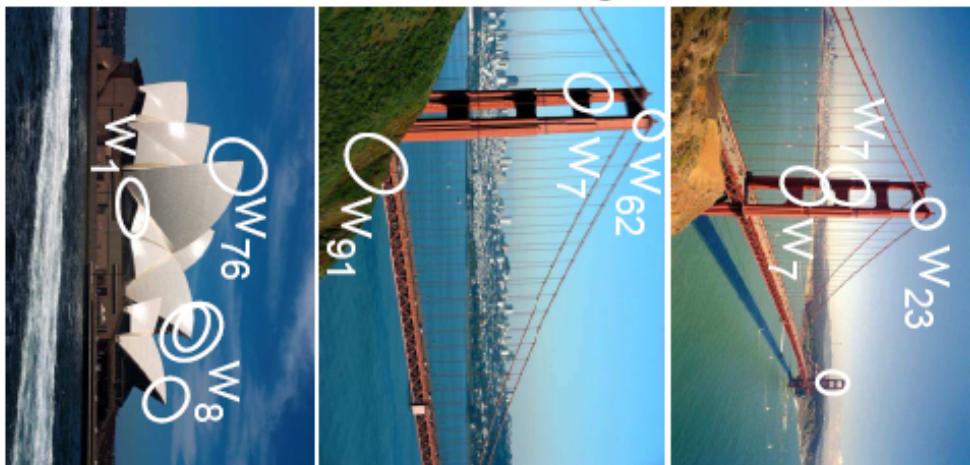


Image #1

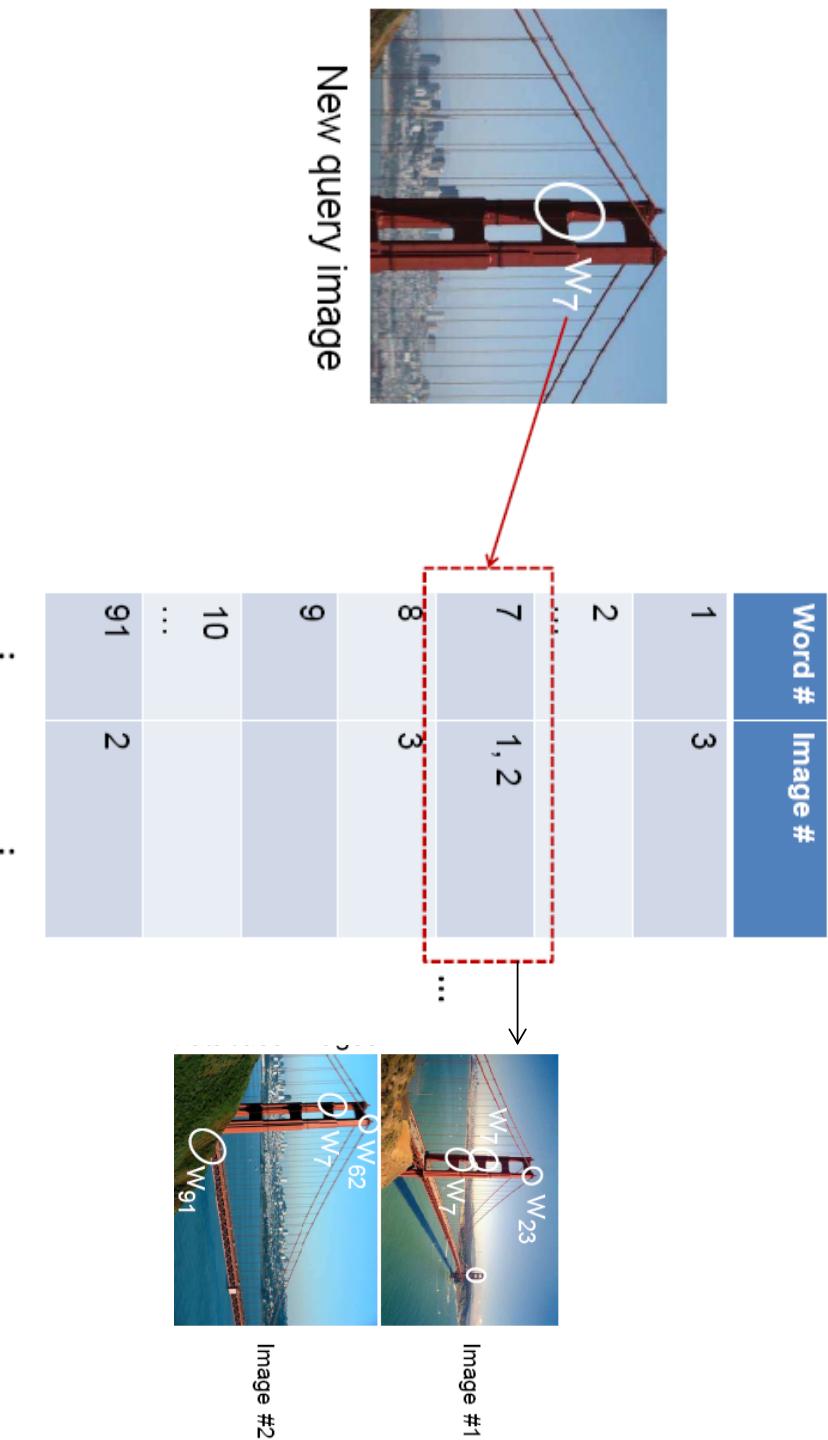
Image #2

Image #3

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

Inverted file index

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



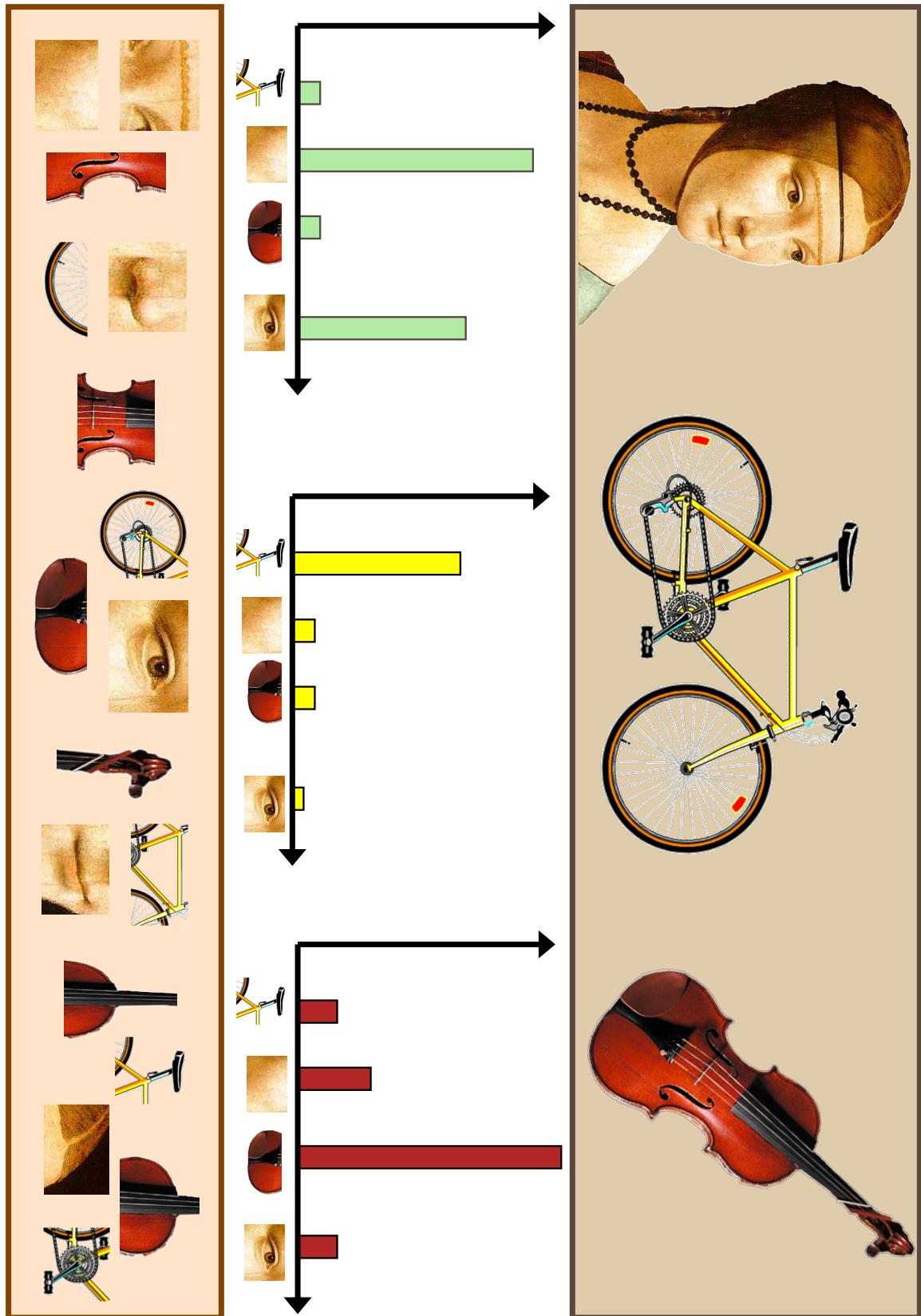
Instance Recognition: Other Issues

- How to summarize the content of an entire image? And gauge overall similarity?
- 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. The world around us is experienced through our eyes. For our eyes, the retina is the receptor organ. The visual information is transmitted to the brain via the optic nerve, where it is processed by the cerebral cortex. The process of perception begins in the brain, where the visual information is interpreted and integrated with other sensory inputs. By understanding the visual system, we can appreciate how the brain processes and interprets the world around us.

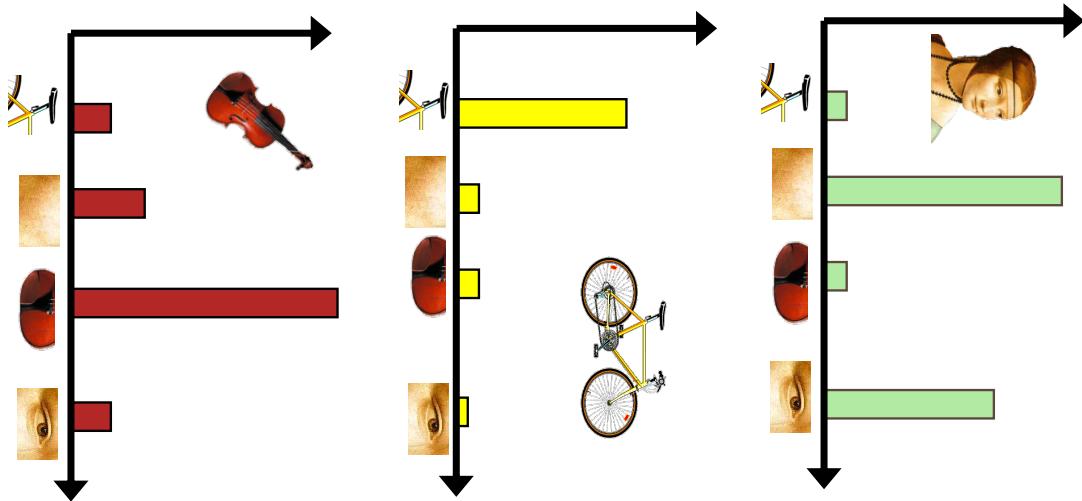




Bag of Visual Words

- Summarize entire image based on its distribution (histogram) of word occurrences.

- Analogous to bag of words representation commonly used for documents.





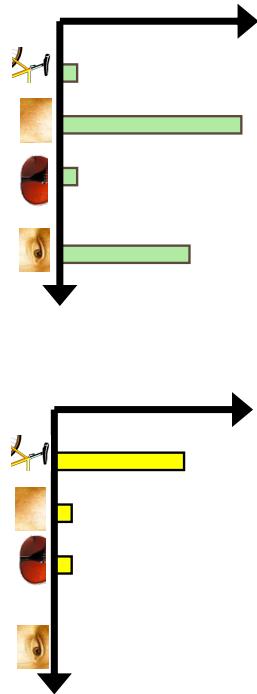
Comparing Bag of Words

- Rank frames by normalized scalar product between occurrence counts:
nearest neighbor search for similar images.

$$[1 \ 8 \ 1 \ 4]$$

$$[5 \ 1 \ 1 \ 0]$$

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



$$\vec{d}_j$$

$$\vec{q}$$

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

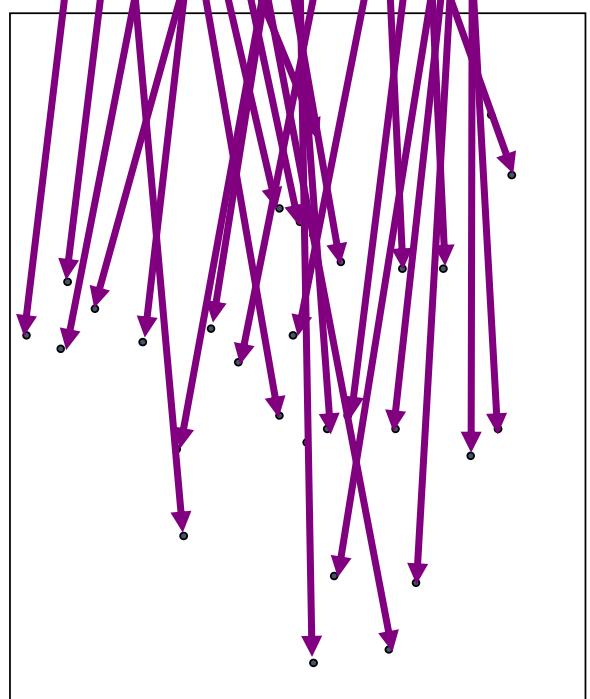
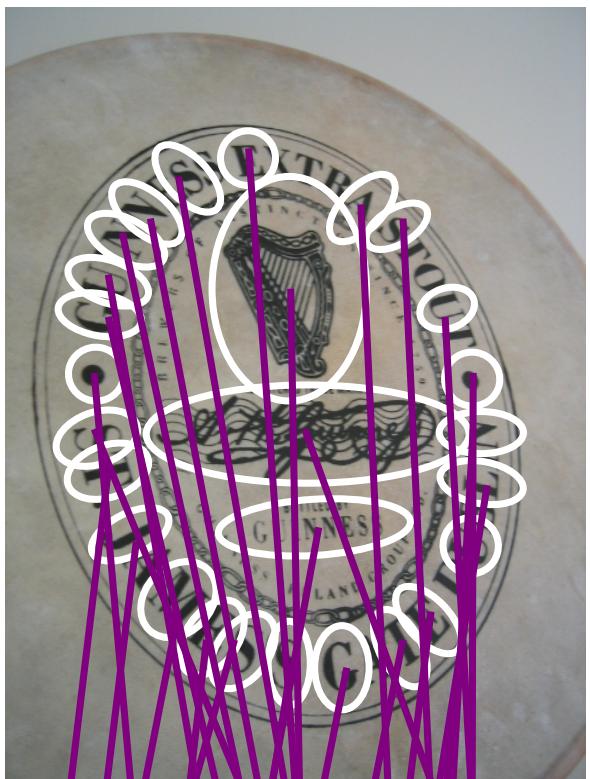
for vocabulary of V words

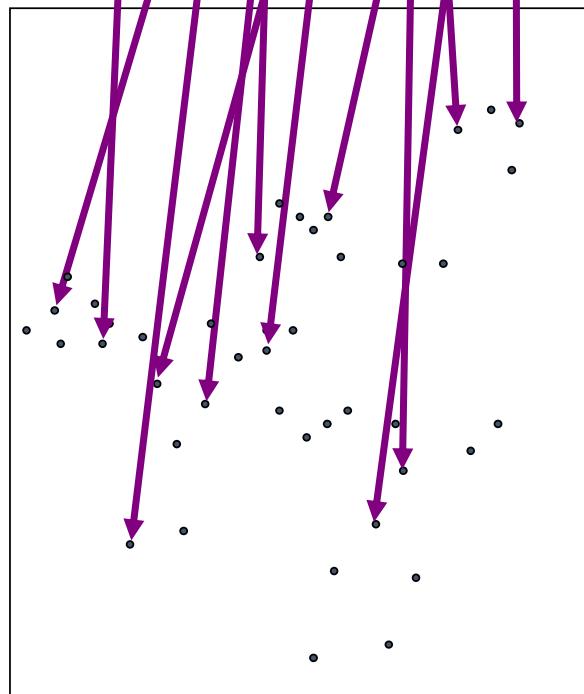
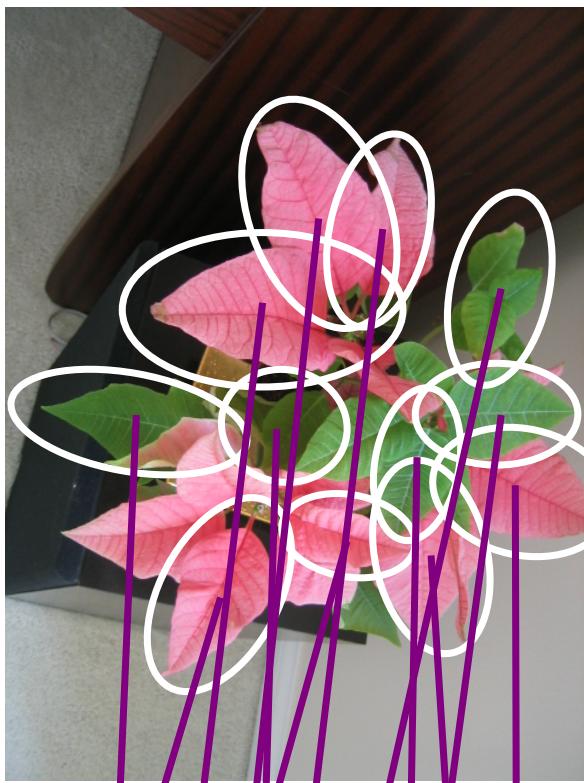


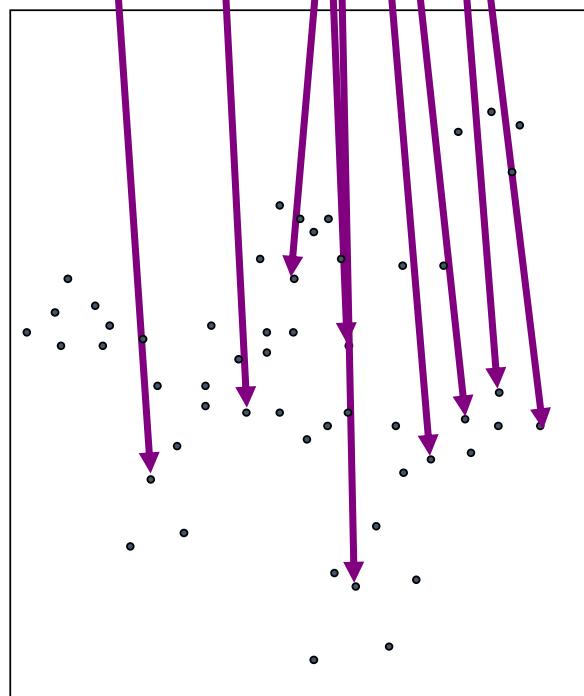
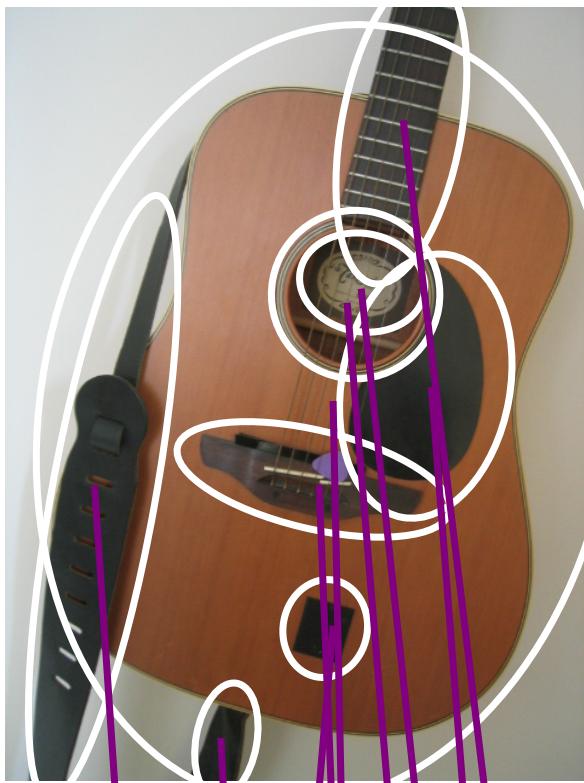
Instance Recognition: Other Issues

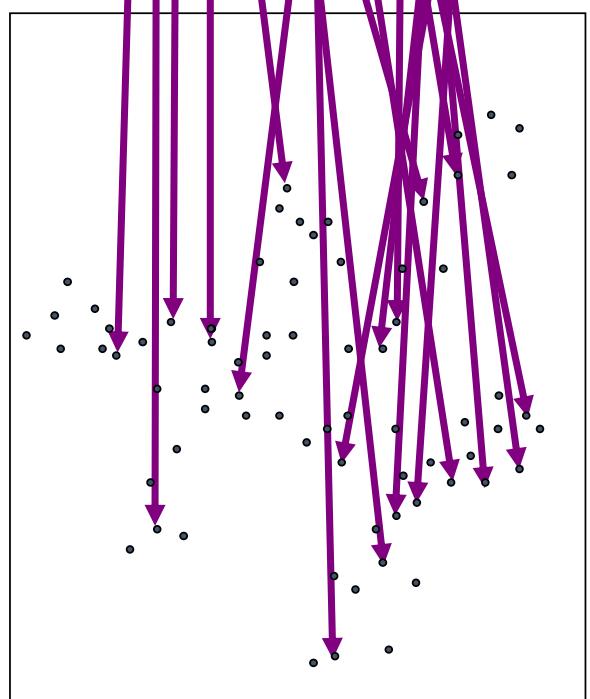
- How to summarize the content of an entire image? And gauge overall similarity?
- How to perform quantization efficiently?
- Is having the same set of visual words enough to identify the object/scene?
How to verify spatial agreement?
- How to score the retrieval results?

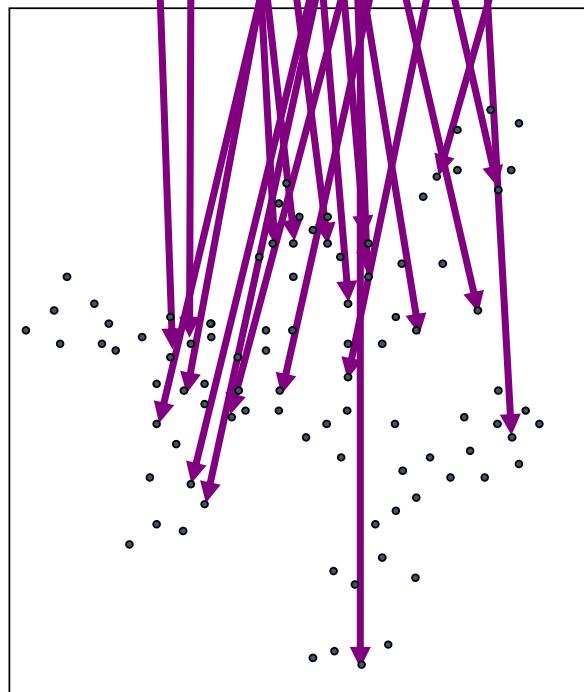
Recognition with K-Tree

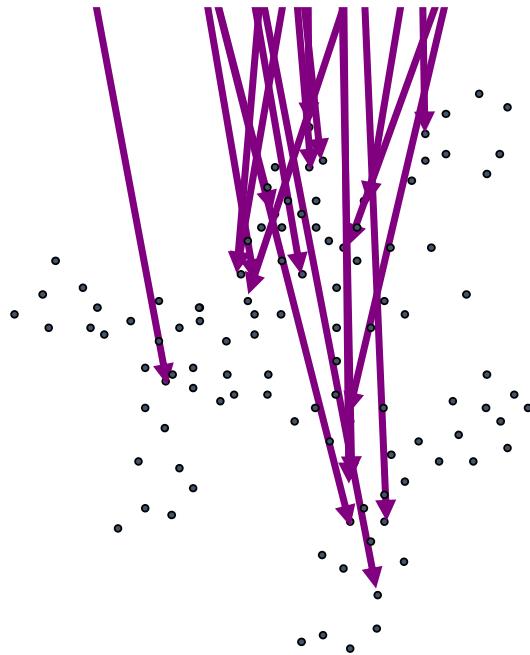


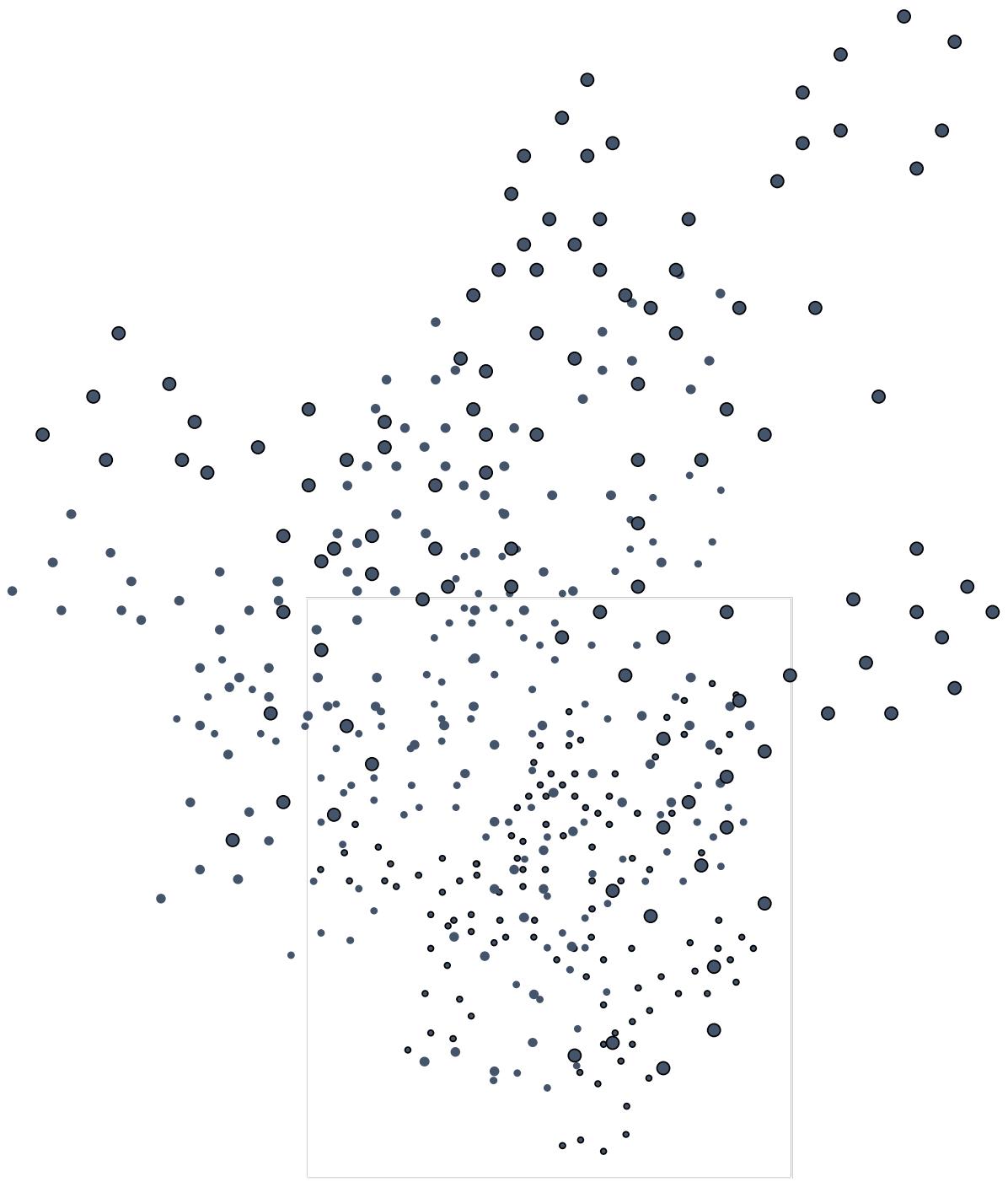


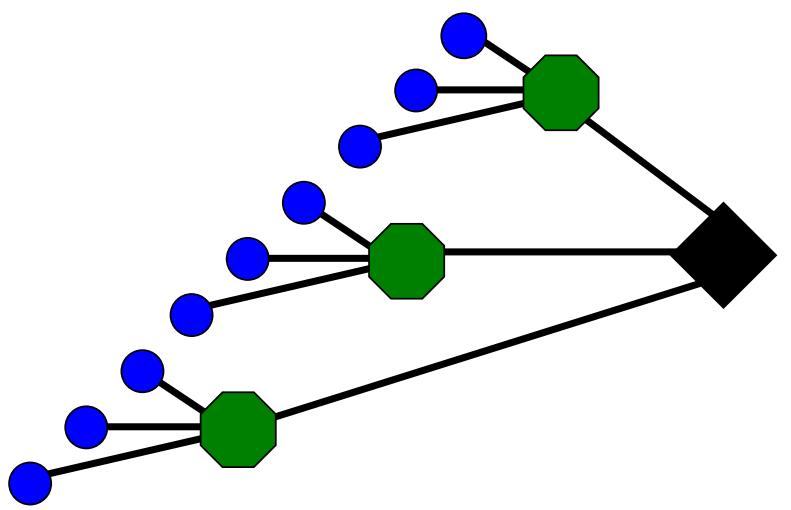


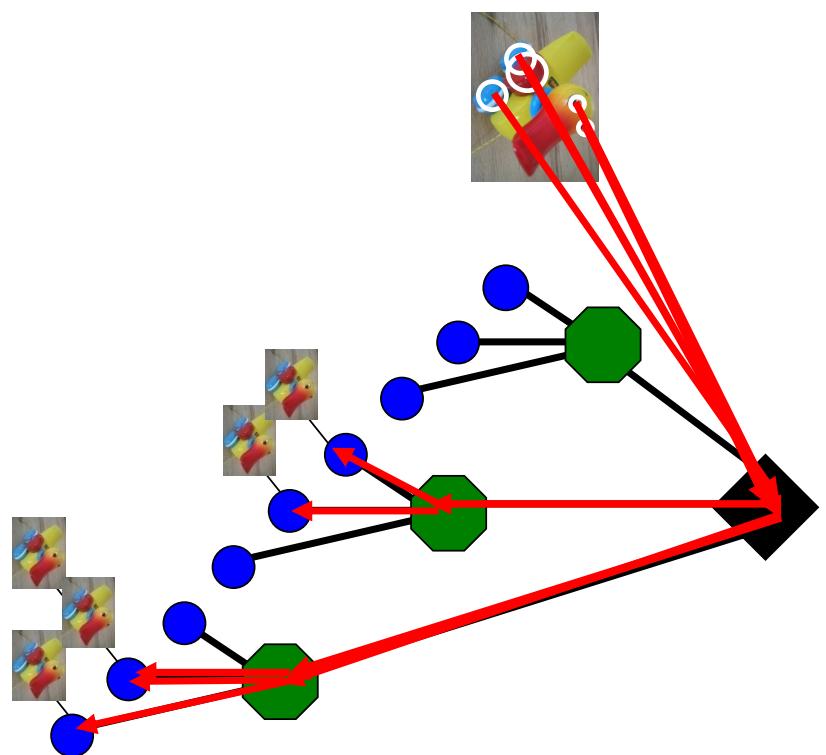


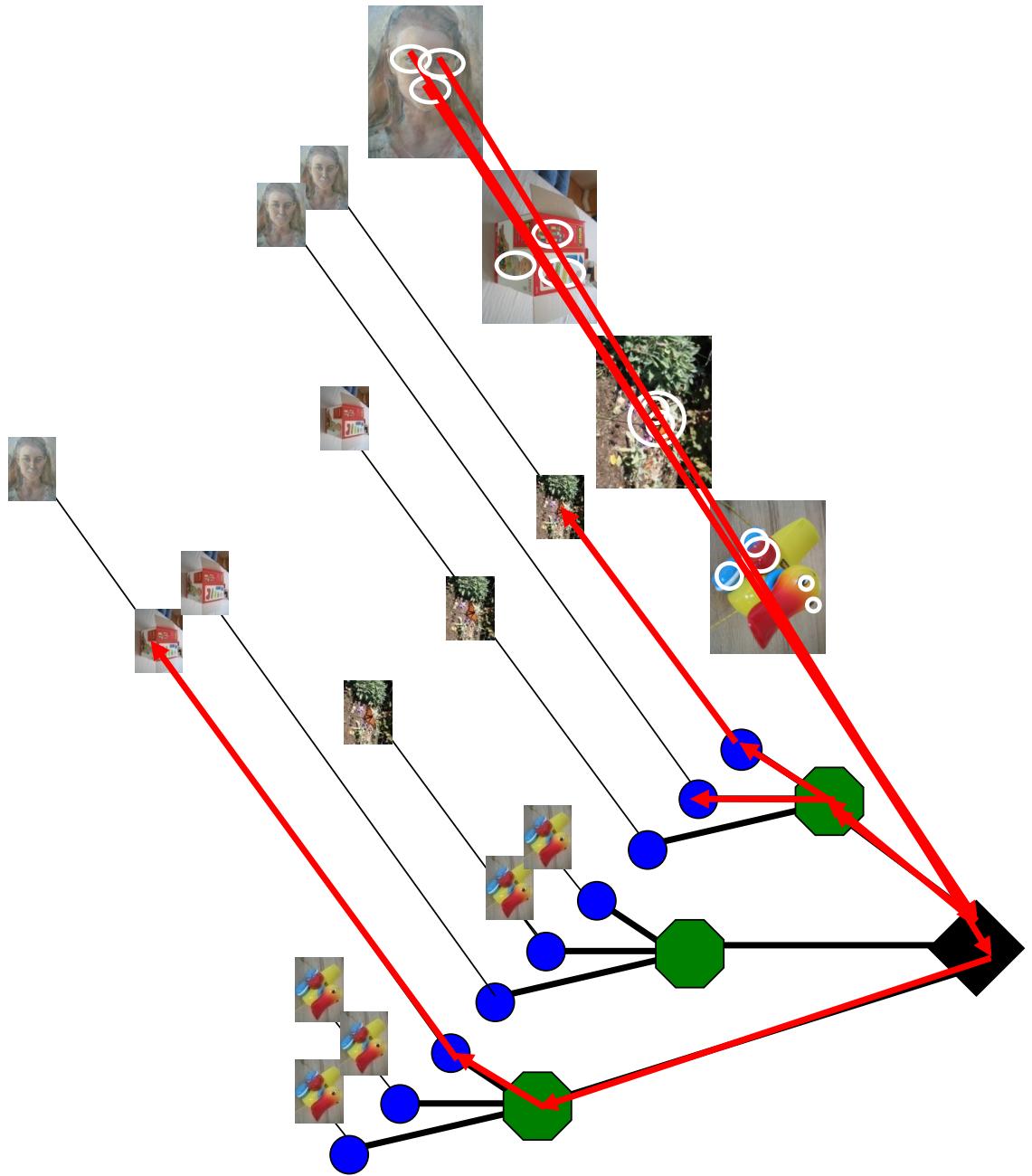


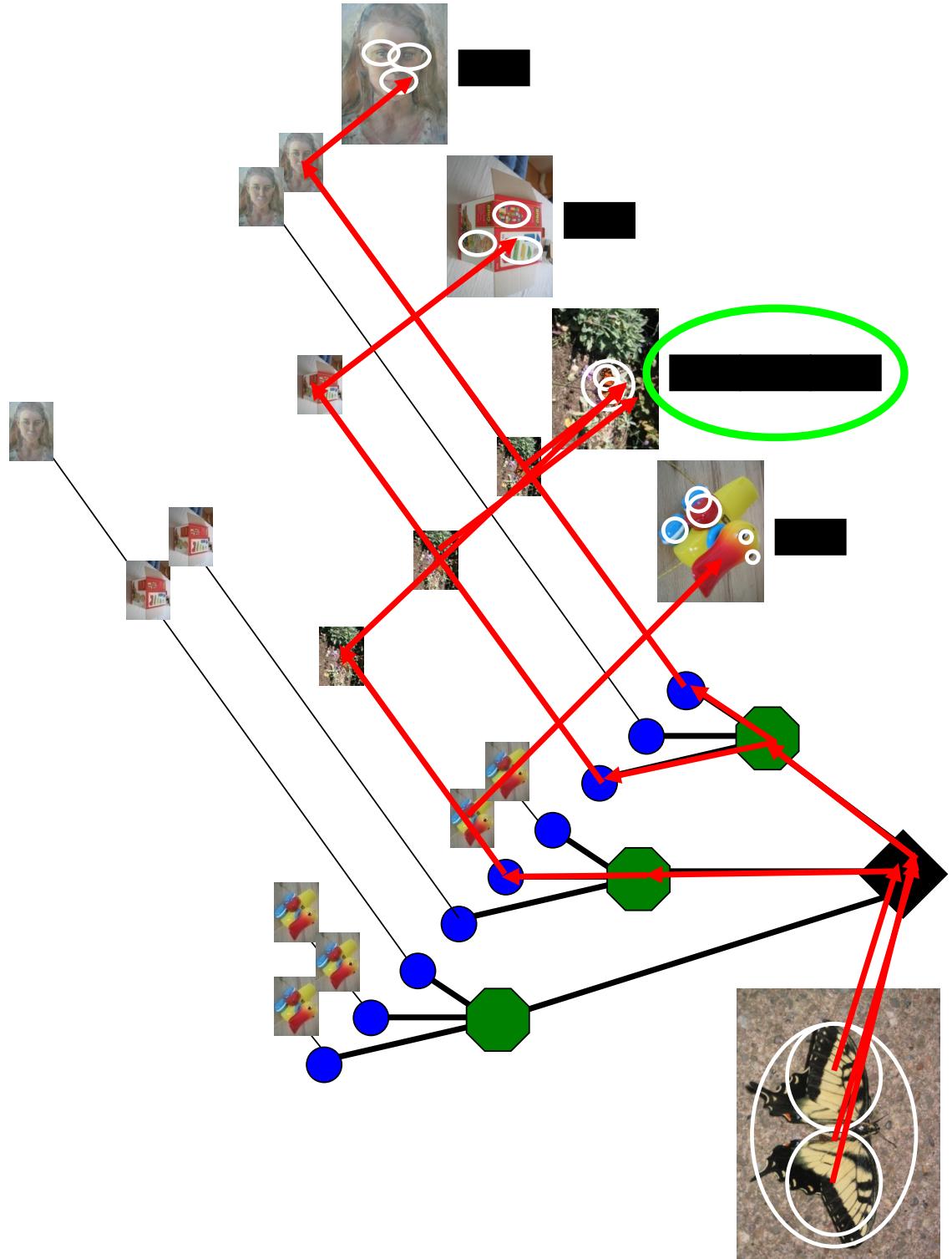










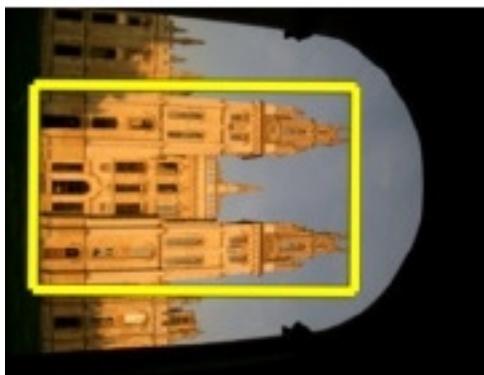


Instance Recognition: Other Issues

- How to summarize the content of an entire image? And gauge overall similarity?
- 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?

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

Query

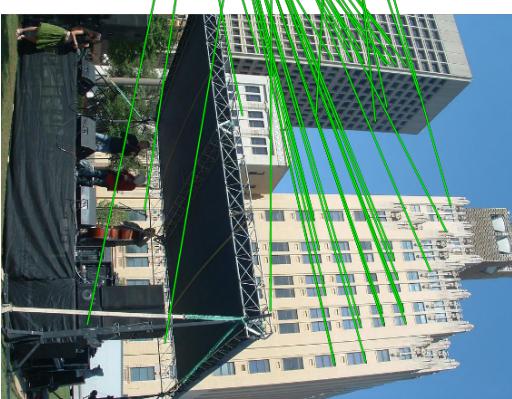


DB image with high Bow similarity

Query



DB image with high Bow similarity



Both image pairs have many visual words in common.

But only some of the matches are mutually consistent

Instance Recognition: Other Issues

- How to summarize the content of an entire image? And gauge overall similarity?
- 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?**



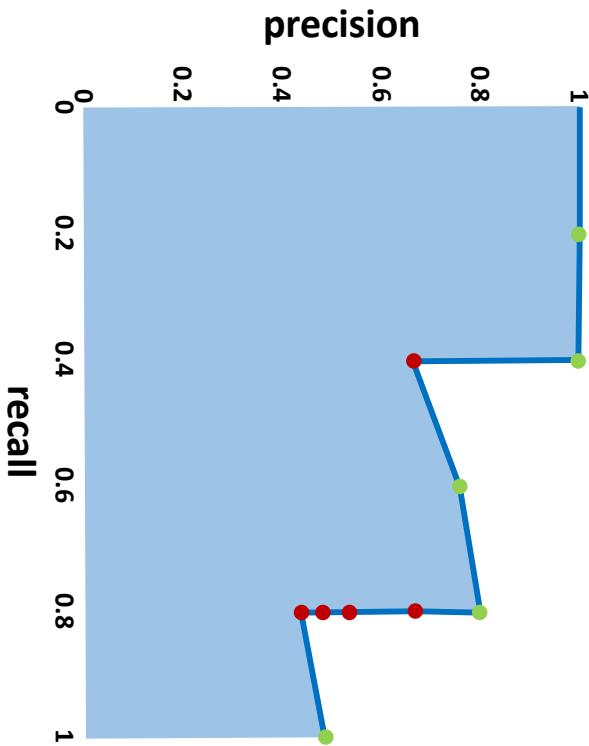
Scoring Quality of Image Retrieval



Database size: 10 images
Relevant (total): 5 images

precision = #correct-returned / #returned
recall = #correct-returned / #total relevant

Query



Results (ordered):





What Else Can We Borrow from Text Retrieval?

- TF-IDF: Term Frequency – Inverse Document Frequency
 - Describe frame by frequency of each word within it, down-weight words that appear often in the database

Number of occurrences
of word i in document d

Total number of
documents in database

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

Number of words in
document d

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