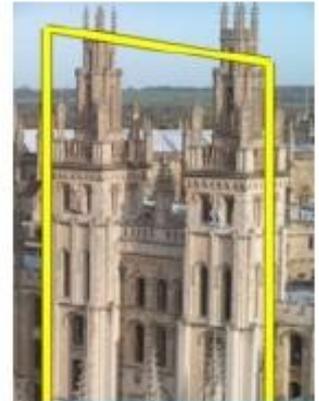
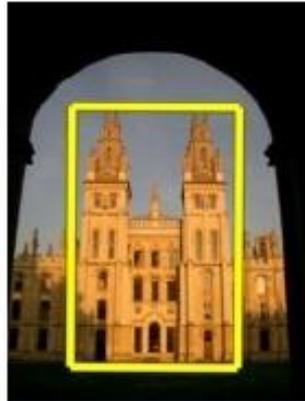
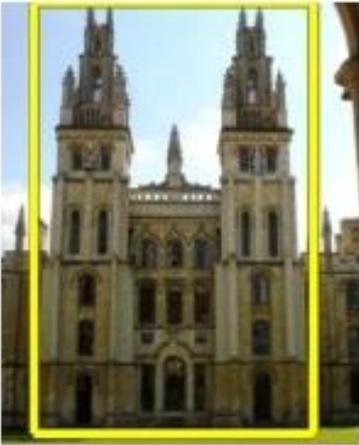
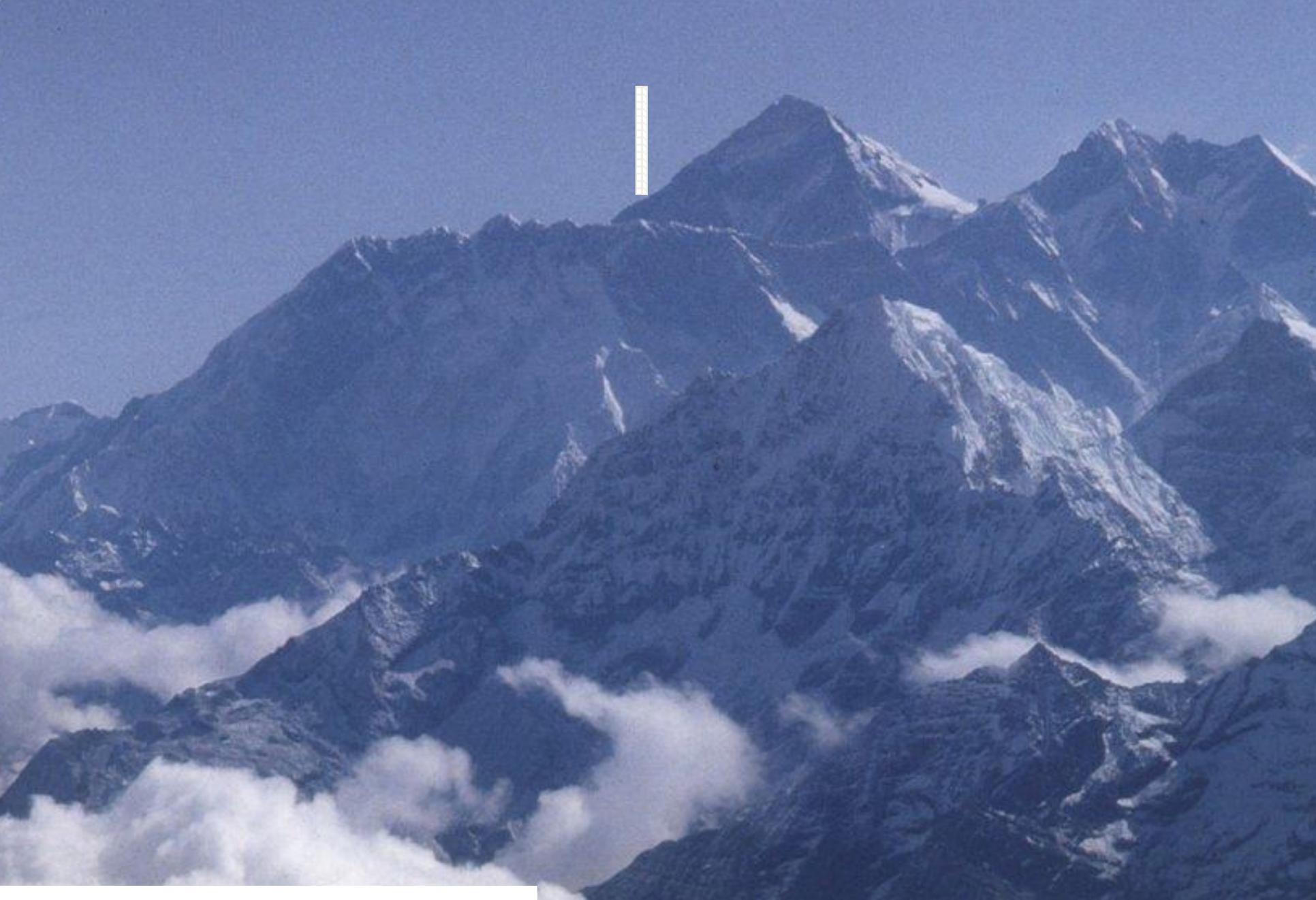


***note:
black & white***



Recap: Bag of Words for Large Scale Retrieval





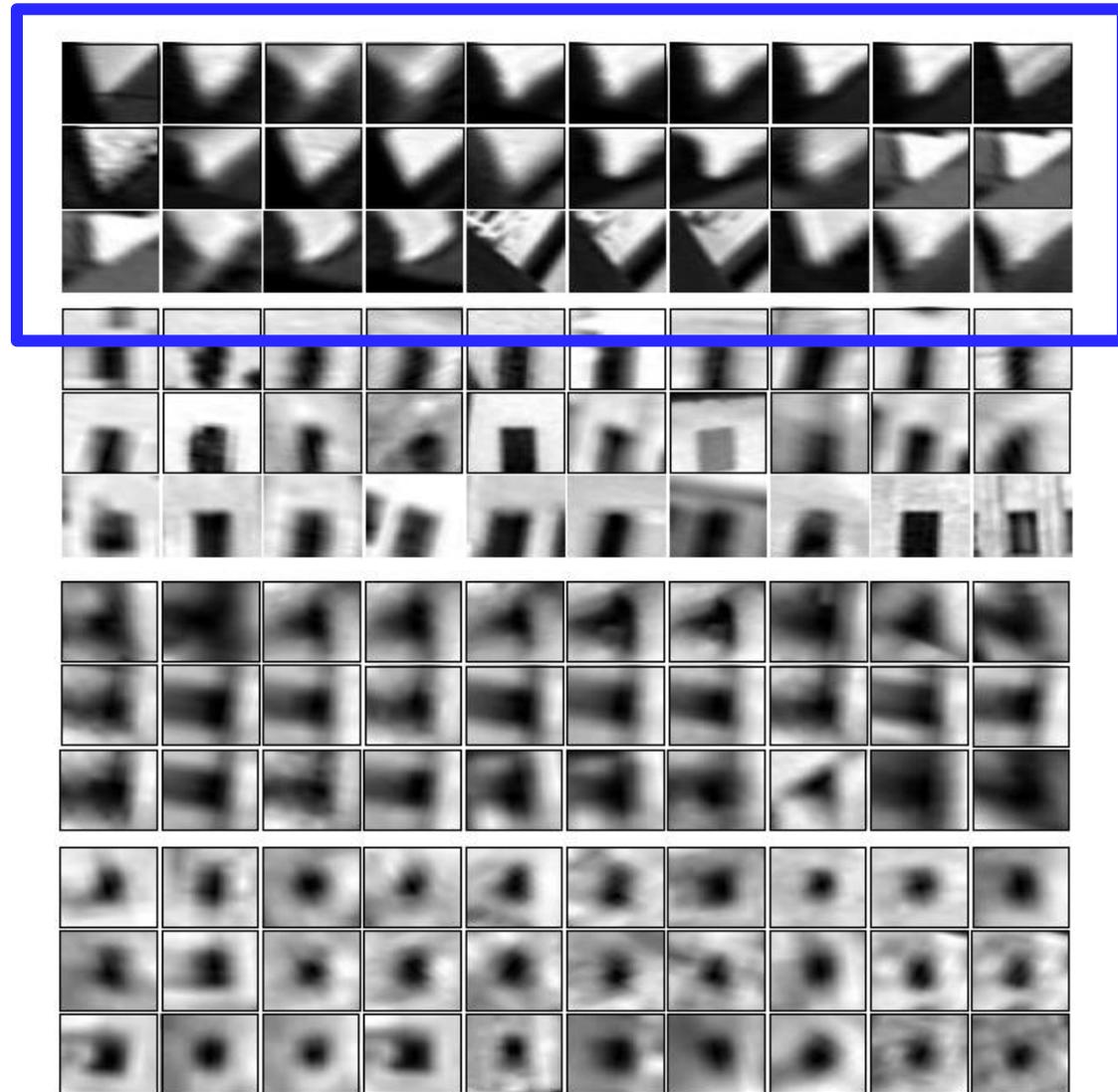
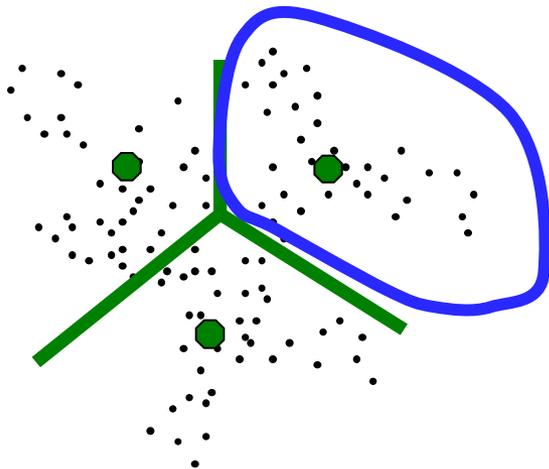
Slide Credit: Nister

Summary – large scale retrieval

- We want to do feature matching (project 2) with a billion images
- Problem: the all-pairs local feature matching is slow!
 - Solution: quantize features and build bag of feature representation. **Lossy!** But spatial verification can help.

Visual words

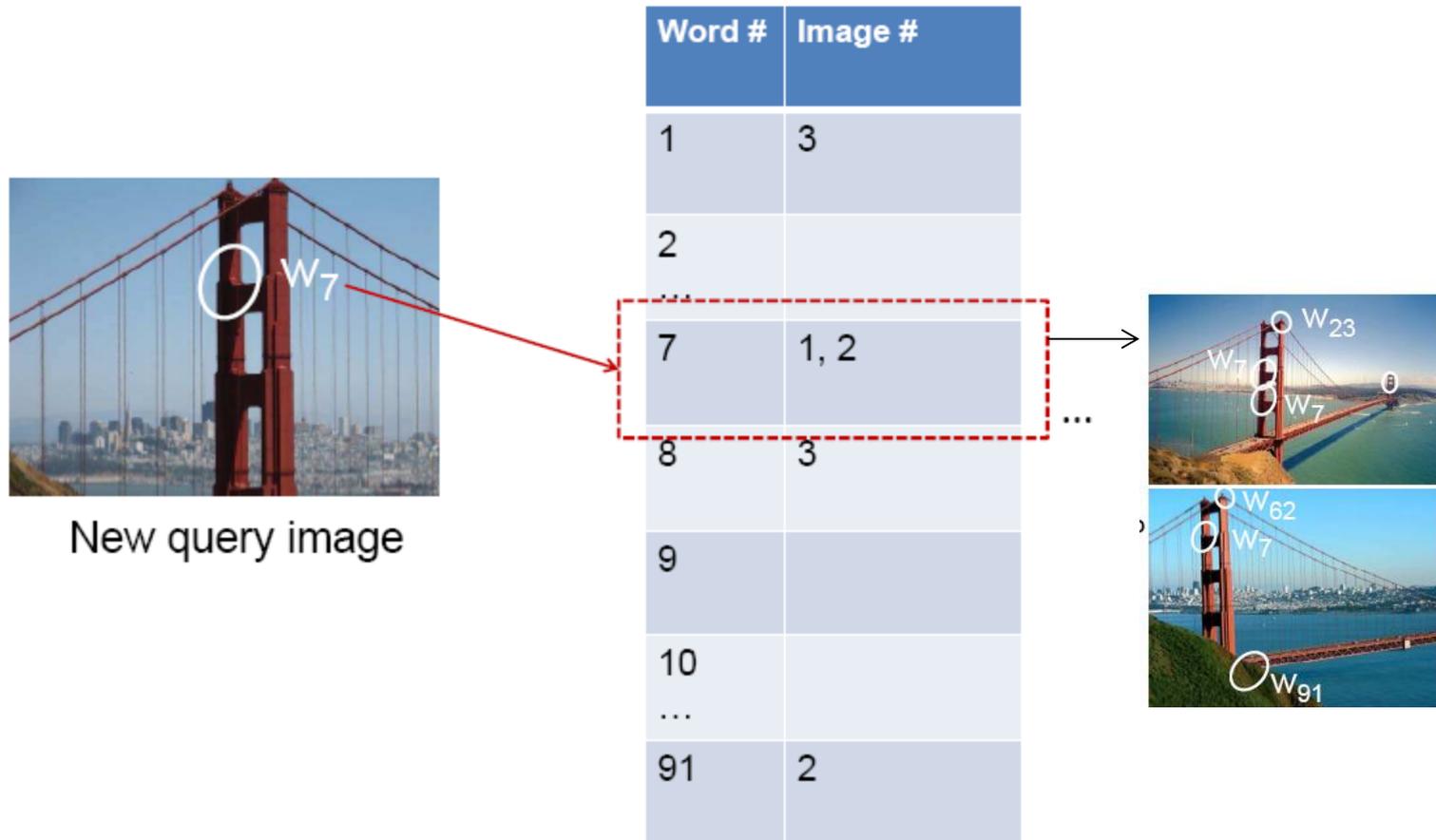
- Example: each group of patches belongs to the same visual word



Summary – large scale retrieval

- We want to do feature matching (project 2) with a billion images
- Problem: the all-pairs local feature matching is slow!
 - Solution: quantize features and build bag of feature representation. **Lossy!** But spatial verification can help.
- Problem: Finding the overlap in visual words based on the Bags of Features is still too slow!
 - Solution: inverted file index, one lookup per word.

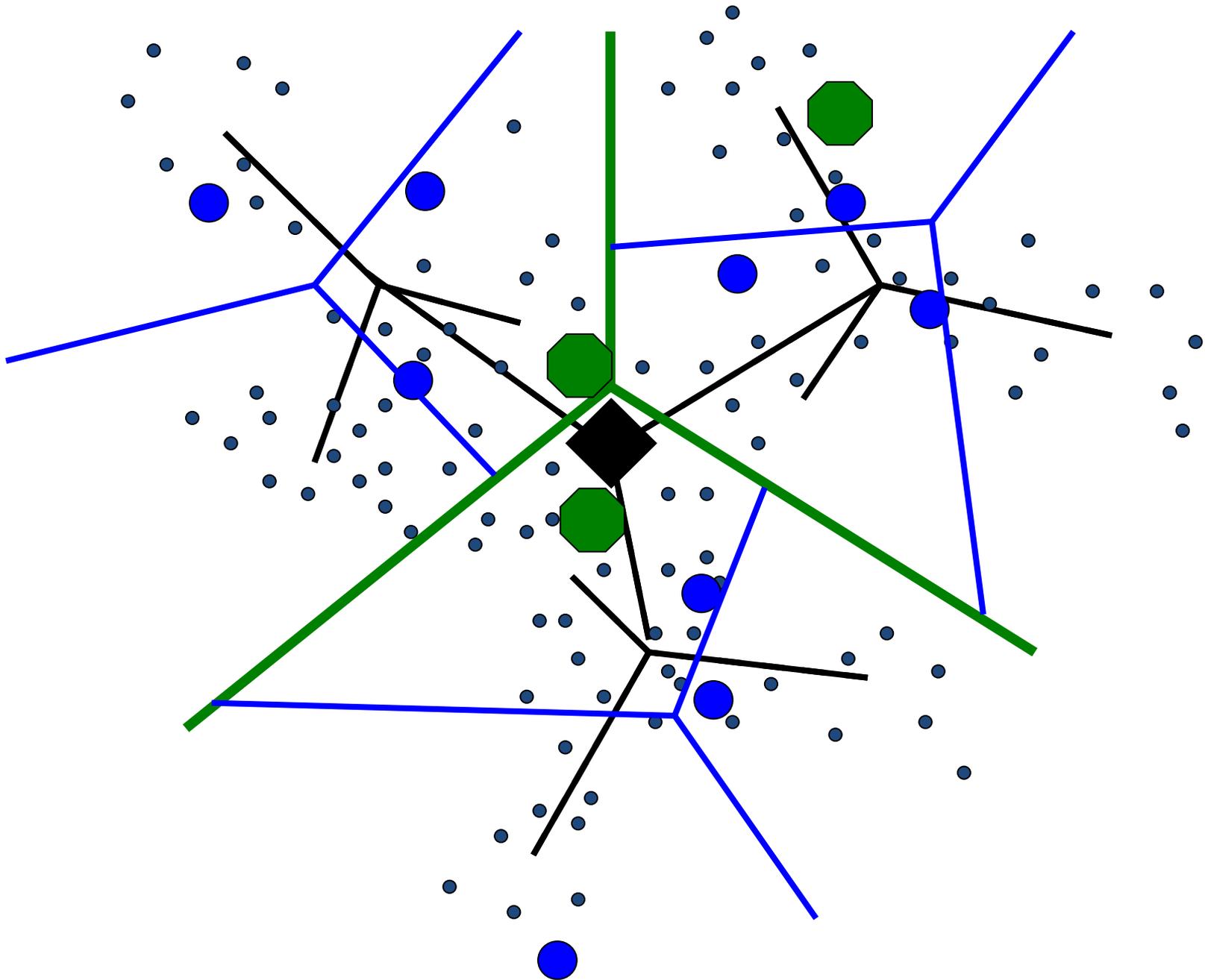
Inverted file index



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

Summary – large scale retrieval

- We want to do feature matching (project 2) with a billion images
- Problem: the all-pairs local feature matching is slow!
 - Solution: quantize features and build bag of feature representation. **Lossy!** But spatial verification can help.
- Problem: Finding the overlap in visual words based on the Bags of Features is still too slow!
 - Solution: inverted file index, one lookup per word.
- Problem: Even quantizing the local features into a visual word is too slow!
 - Solution: vocabulary tree. **Lossy!**



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 Is) - 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 Mailman; 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
	Epiphyles; 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

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 \$566bn. The surplus will annoy the US because China's deliberate policy is to agree to a yuan is also needed to demand so much country. China's yuan against the dollar and permitted it to trade within a narrow band but the US wants the yuan to be allowed to trade freely. However, Beijing has made it clear that it will take its time and tread carefully before allowing the yuan to rise further in value.



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

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

Scoring retrieval quality



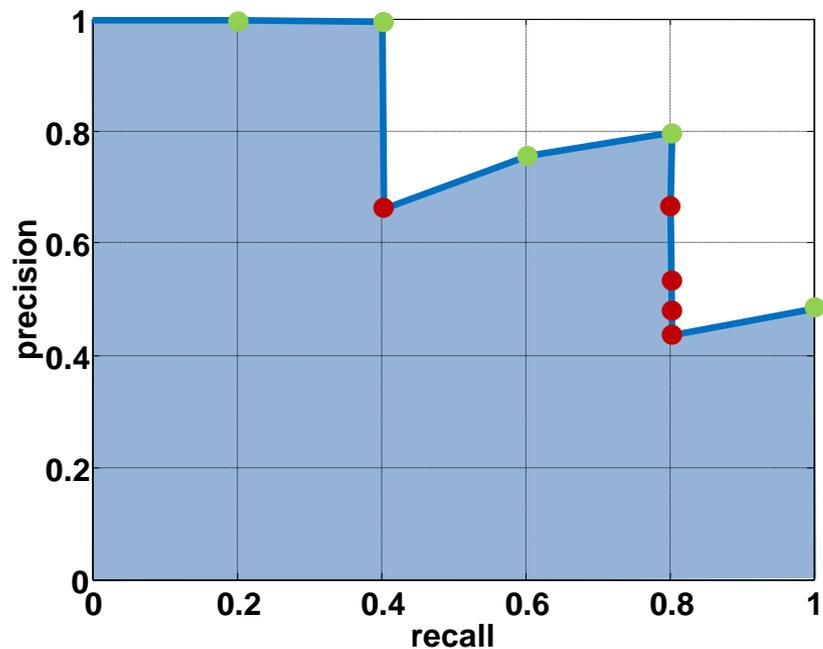
Query

Database size: 10 images

Relevant (total): 5 images

precision = $\frac{\text{\#relevant}}{\text{\#returned}}$

recall = $\frac{\text{\#relevant}}{\text{\#total relevant}}$



Results (ordered):



- Sliding window detector must evaluate tens of thousands of location/scale combinations
- Faces are rare: 0–10 per image
 - For computational efficiency, we should try to spend as little time as possible on the non-face windows
 - A megapixel image has $\sim 10^6$ pixels and a comparable number of candidate face locations
 - To avoid having a false positive in every image, our false positive rate has to be less than 10^{-6}

Sliding Window Face Detection

with Viola-Jones

Computer Vision

CS 143, Brown

James Hays

Face detection and recognition



Detection



Recognition

"Sally"

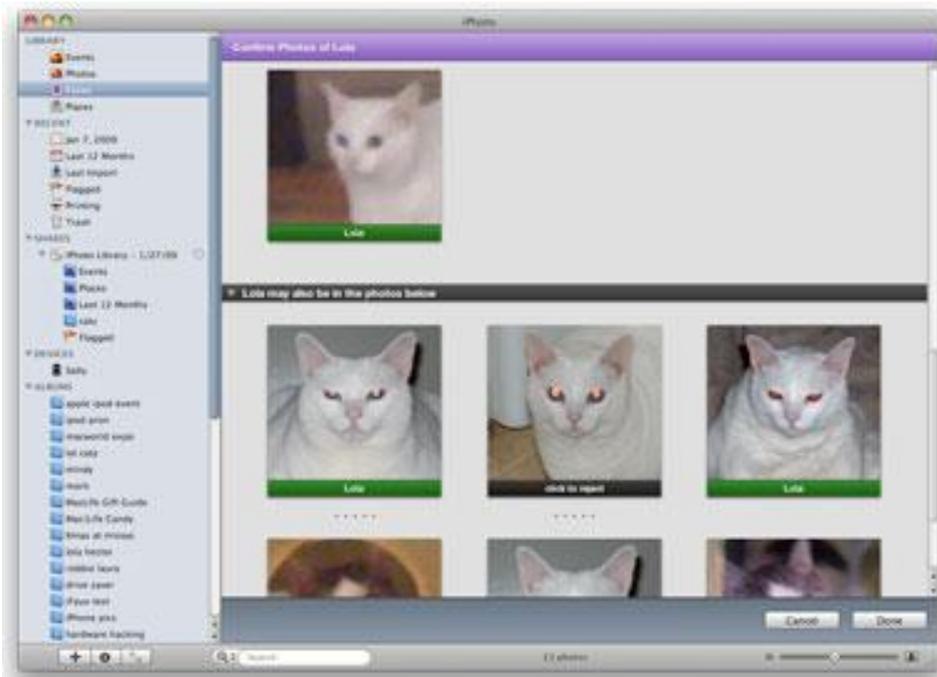
Consumer application: Apple iPhoto



<http://www.apple.com/ilife/iphoto/>

Consumer application: Apple iPhoto

Can be trained to recognize pets!



http://www.maclife.com/article/news/iphotos_faces_recognizes_cats

Consumer application: Apple iPhoto

Things iPhoto thinks are faces



Funny Nikon ads

"The Nikon S60 detects up to 12 faces."



Funny Nikon ads

"The Nikon S60 detects up to 12 faces."



Challenges of face detection

- Sliding window detector must evaluate tens of thousands of location/scale combinations
- Faces are rare: 0–10 per image
 - For computational efficiency, we should try to spend as little time as possible on the non-face windows
 - A megapixel image has $\sim 10^6$ pixels and a comparable number of candidate face locations
 - To avoid having a false positive in every image image, our false positive rate has to be less than 10^{-6}

The Viola/Jones Face Detector

- A seminal approach to real-time object detection
- Training is slow, but detection is very fast
- Key ideas
 - *Integral images* for fast feature evaluation
 - *Boosting* for feature selection
 - *Attentional cascade* for fast rejection of non-face windows

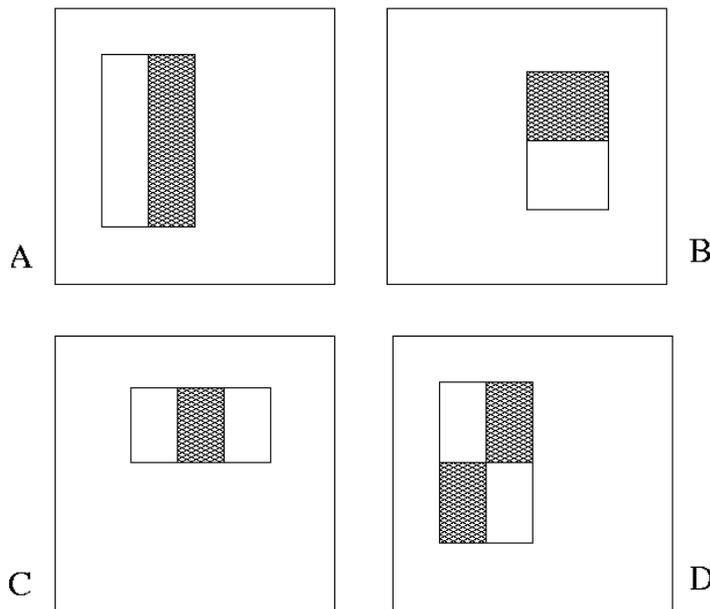
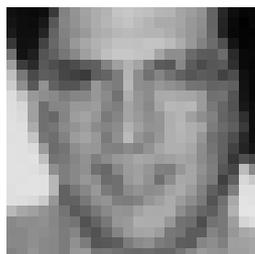
P. Viola and M. Jones. [*Rapid object detection using a boosted cascade of simple features.*](#) CVPR 2001.

P. Viola and M. Jones. [*Robust real-time face detection.*](#) IJCV 57(2), 2004.

~8000 citations!

Image Features

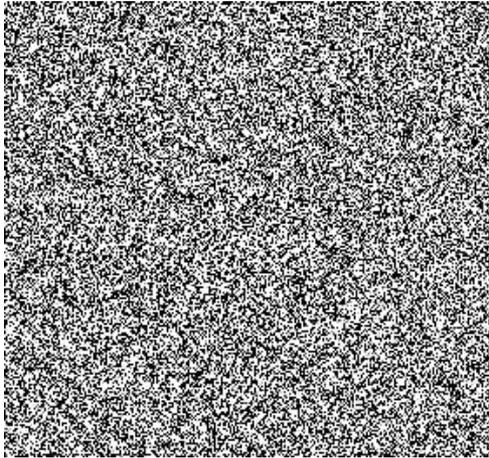
“Rectangle filters”



Value =

$$\sum (\text{pixels in white area}) - \sum (\text{pixels in black area})$$

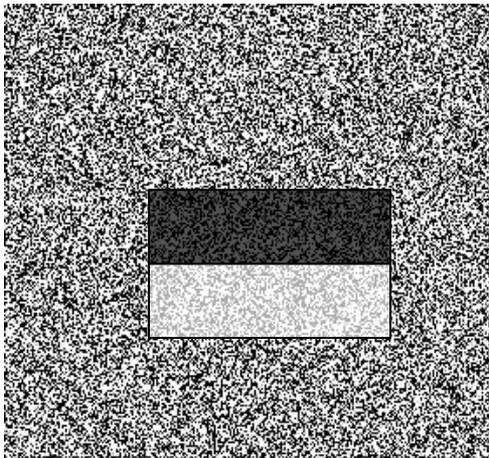
Example



Source

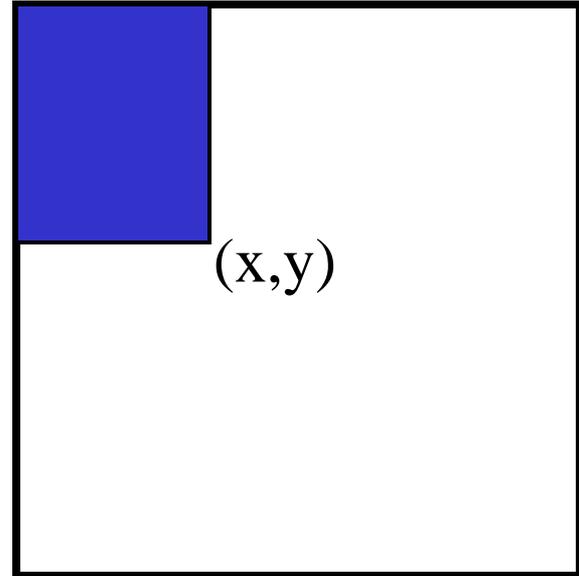


Result

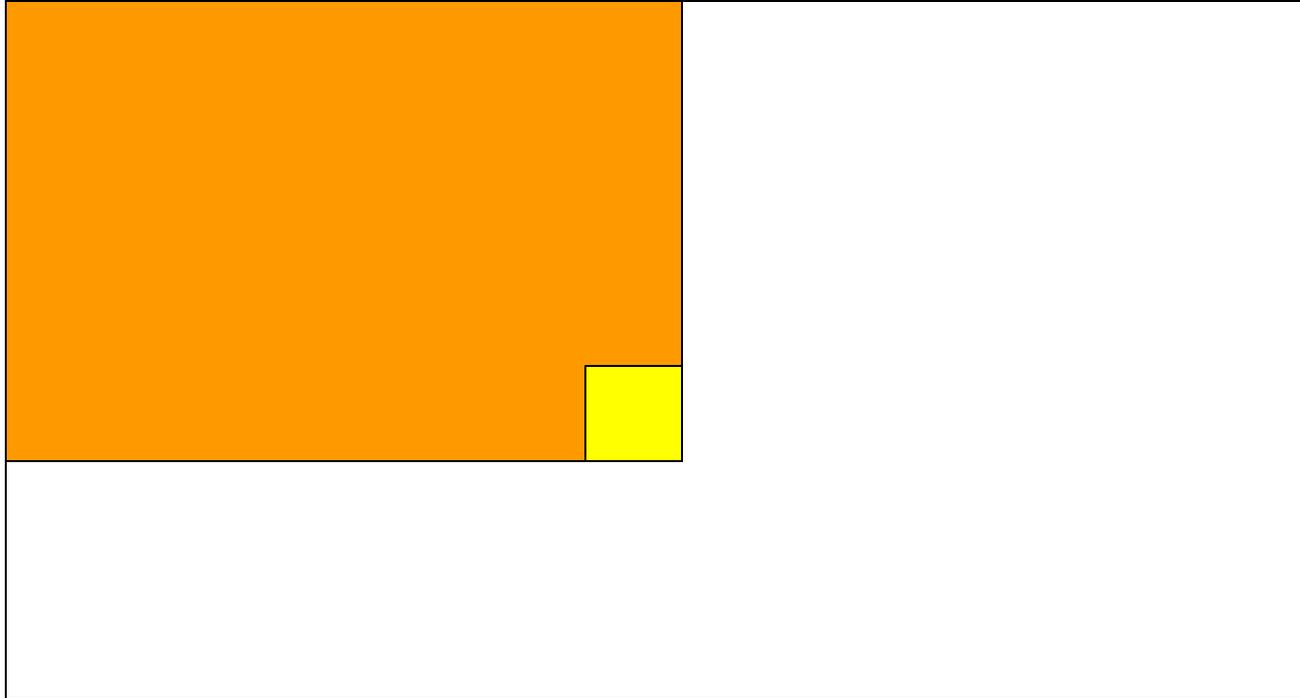


Fast computation with integral images

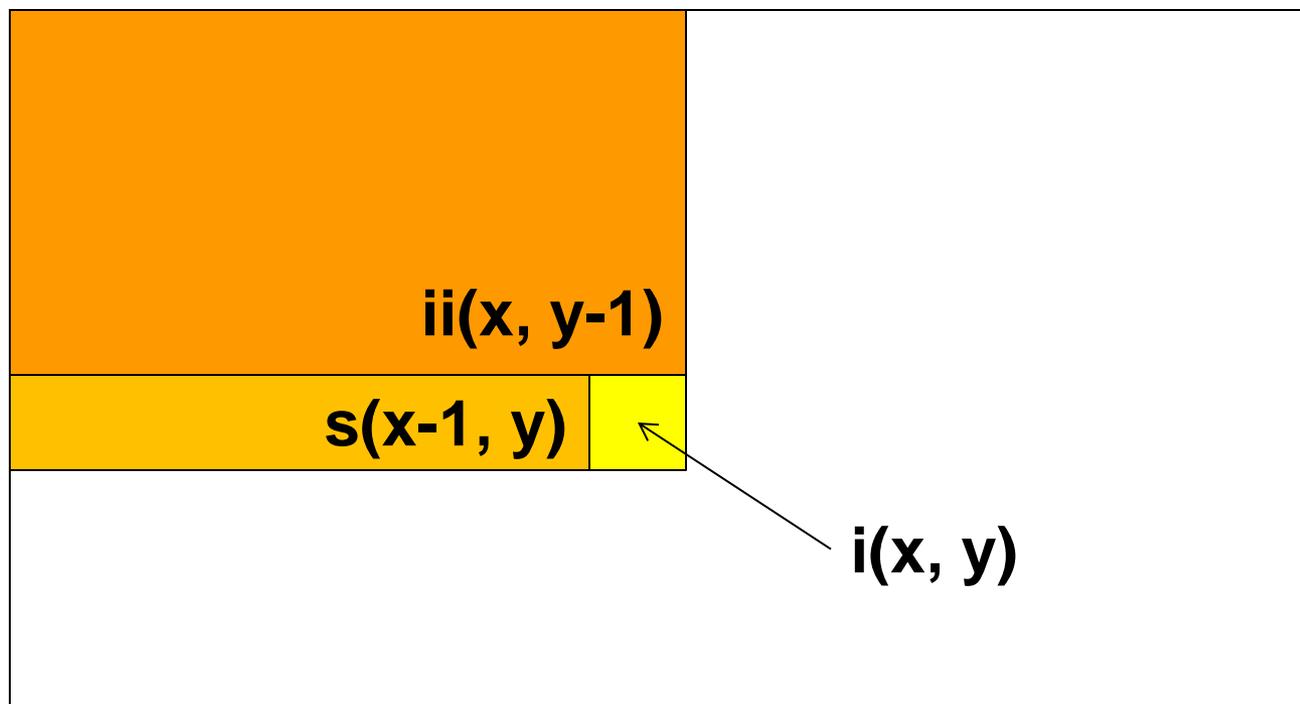
- The *integral image* computes a value at each pixel (x,y) that is the sum of the pixel values above and to the left of (x,y) , inclusive
- This can quickly be computed in one pass through the image



Computing the integral image



Computing the integral image



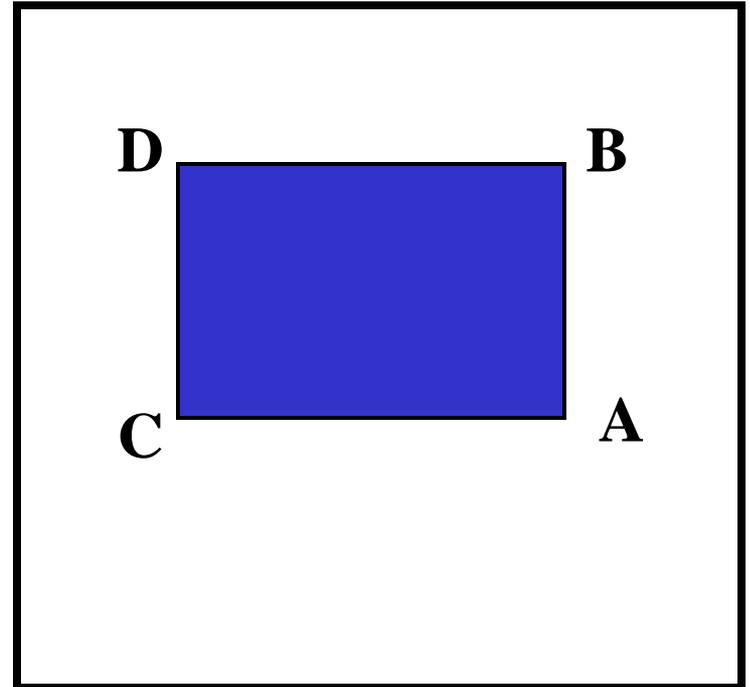
Cumulative row sum: $s(x, y) = s(x-1, y) + i(x, y)$

Integral image: $ii(x, y) = ii(x, y-1) + s(x, y)$

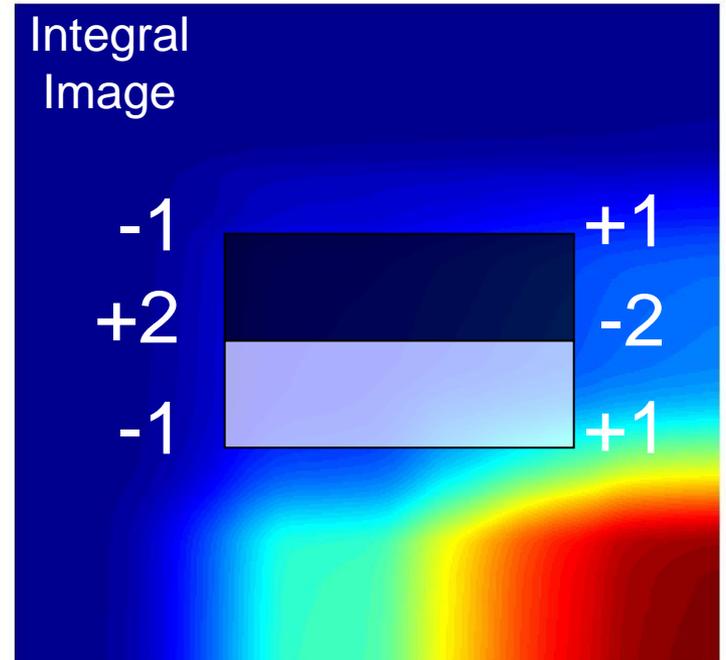
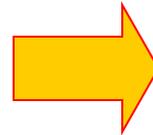
MATLAB: `ii = cumsum(cumsum(double(i)), 2);`

Computing sum within a rectangle

- Let A,B,C,D be the values of the integral image at the corners of a rectangle
- Then the sum of original image values within the rectangle can be computed as:
$$\text{sum} = A - B - C + D$$
- Only 3 additions are required for any size of rectangle!

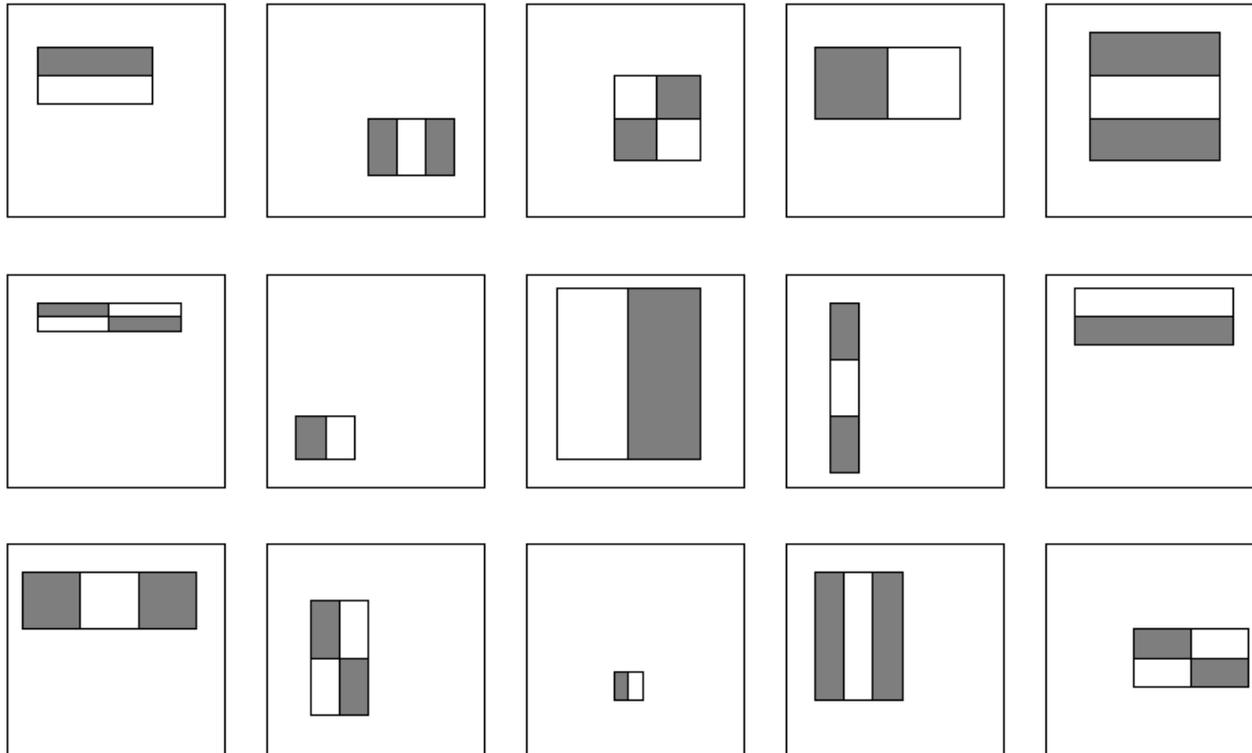


Computing a rectangle feature



Feature selection

- For a 24x24 detection region, the number of possible rectangle features is ~160,000!



Feature selection

- For a 24x24 detection region, the number of possible rectangle features is ~160,000!
- At test time, it is impractical to evaluate the entire feature set
- Can we create a good classifier using just a small subset of all possible features?
- How to select such a subset?

Boosting

- *Boosting* is a classification scheme that combines *weak learners* into a more accurate *ensemble classifier*
- Weak learners based on rectangle filters:

$$h_t(x) = \begin{cases} 1 & \text{if } p_t f_t(x) > p_t \theta_t \\ 0 & \text{otherwise} \end{cases}$$

Diagram annotations for the weak classifier equation:

- value of rectangle feature (points to $f_t(x)$)
- parity (points to p_t)
- threshold (points to θ_t)
- window (points to $h_t(x)$)

- Ensemble classification function:

$$C(x) = \begin{cases} 1 & \text{if } \sum_{t=1}^T \alpha_t h_t(x) > \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{otherwise} \end{cases}$$

Diagram annotation for the ensemble classification function:

- learned weights (points to α_t)

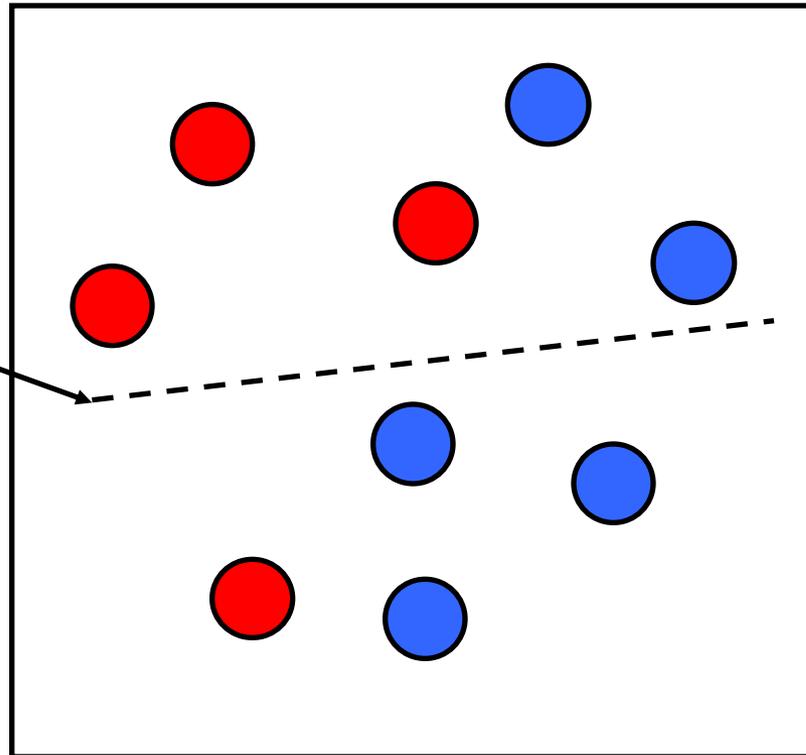
Training procedure

- Initially, weight each training example equally
- In each boosting round:
 - Find the weak learner that achieves the lowest *weighted* training error
 - Raise the weights of training examples misclassified by current weak learner
- Compute final classifier as linear combination of all weak learners (weight of each learner is directly proportional to its accuracy)
 - Exact formulas for re-weighting and combining weak learners depend on the particular boosting scheme (e.g., AdaBoost)

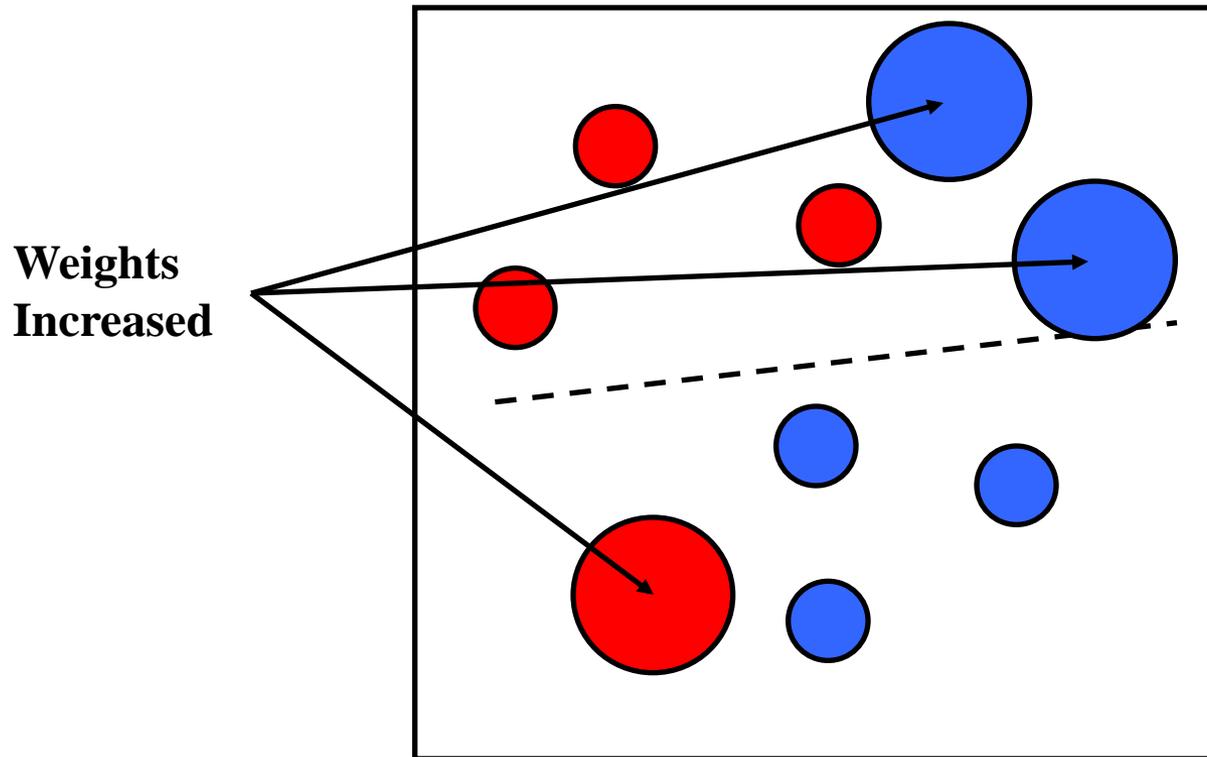
Y. Freund and R. Schapire, [A short introduction to boosting](#), *Journal of Japanese Society for Artificial Intelligence*, 14(5):771-780, September, 1999.

Boosting intuition

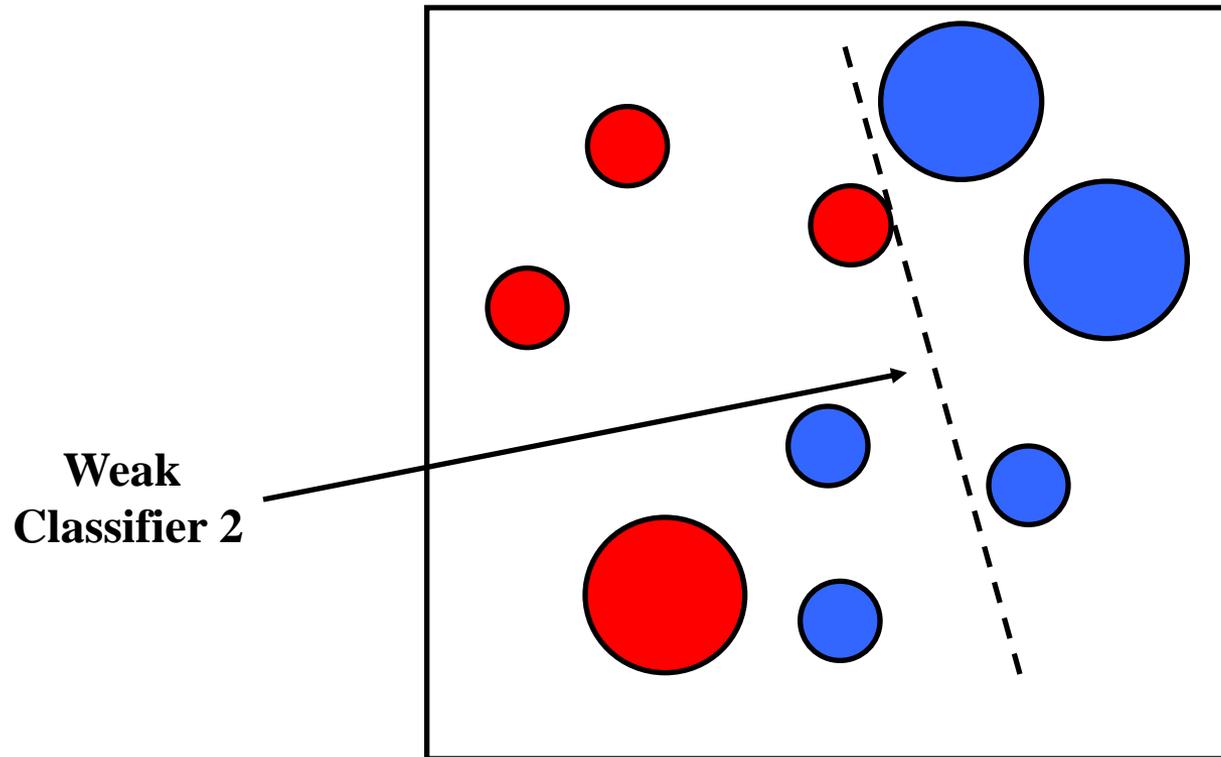
**Weak
Classifier 1**



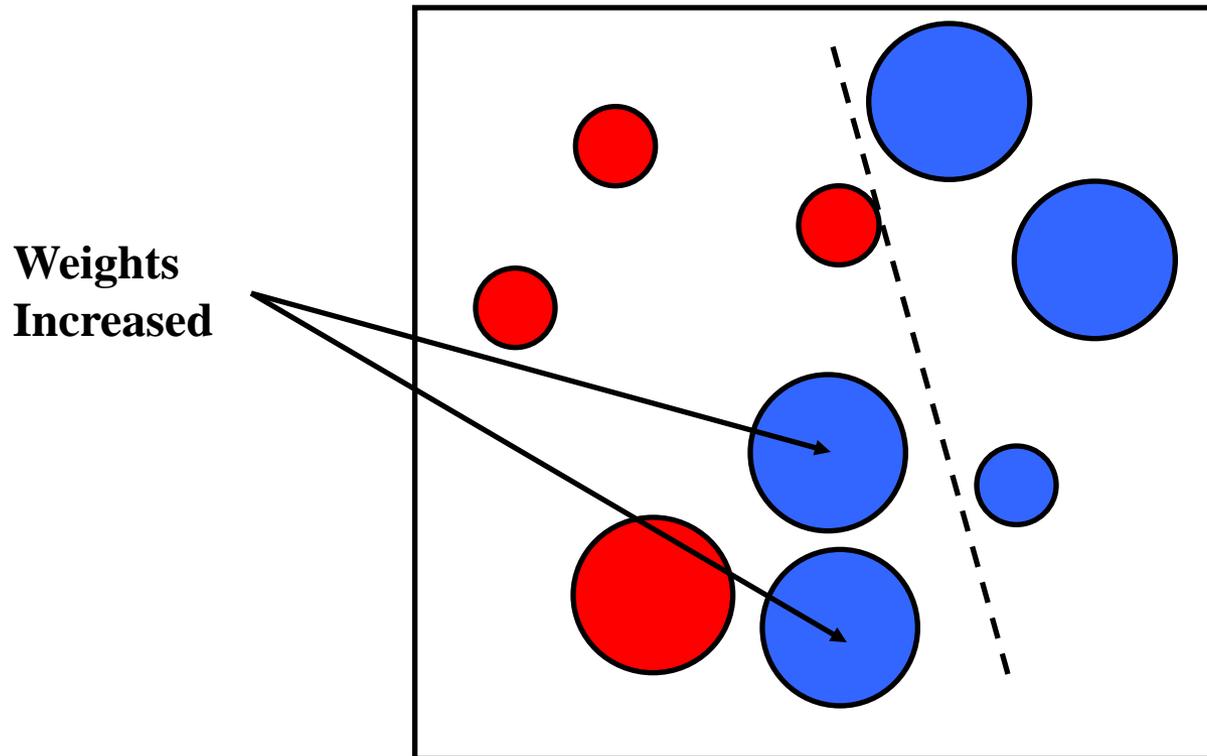
Boosting illustration



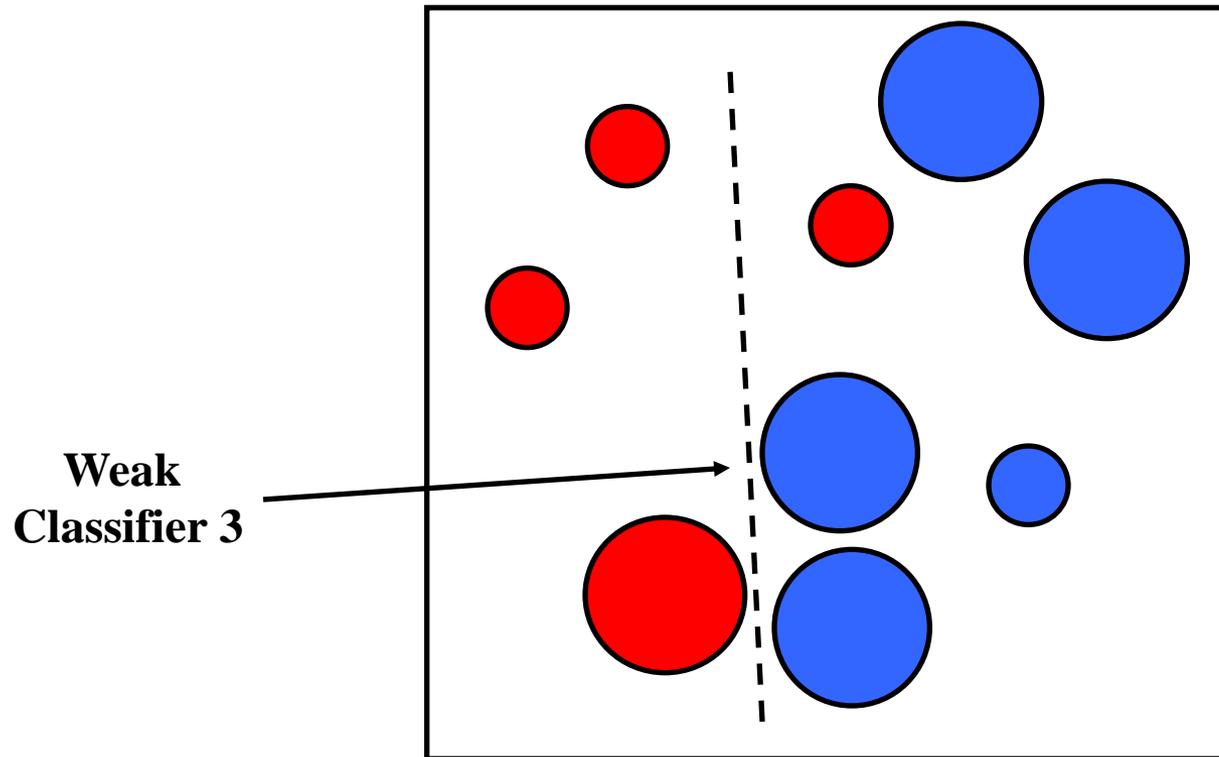
Boosting illustration



Boosting illustration

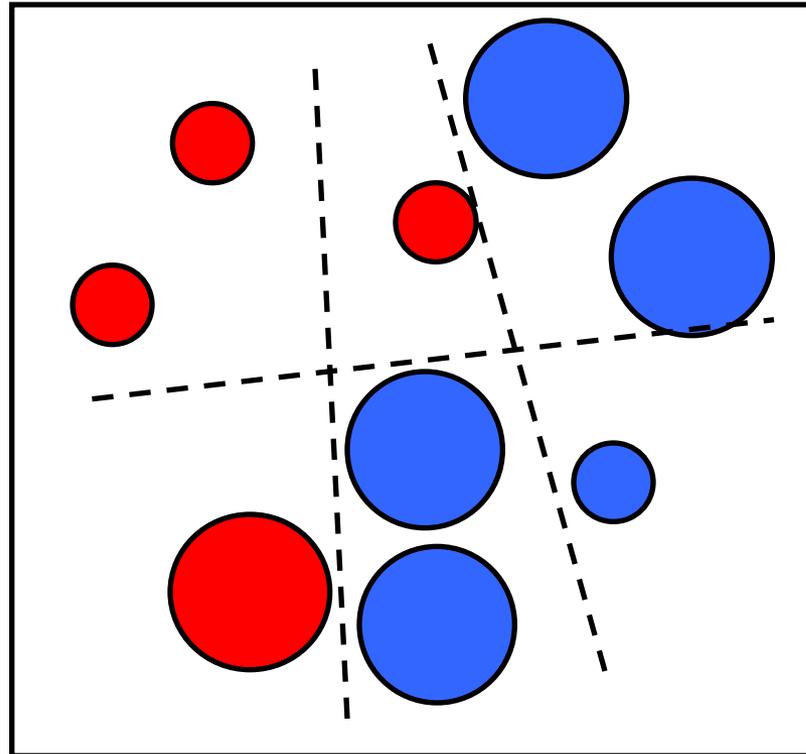


Boosting illustration



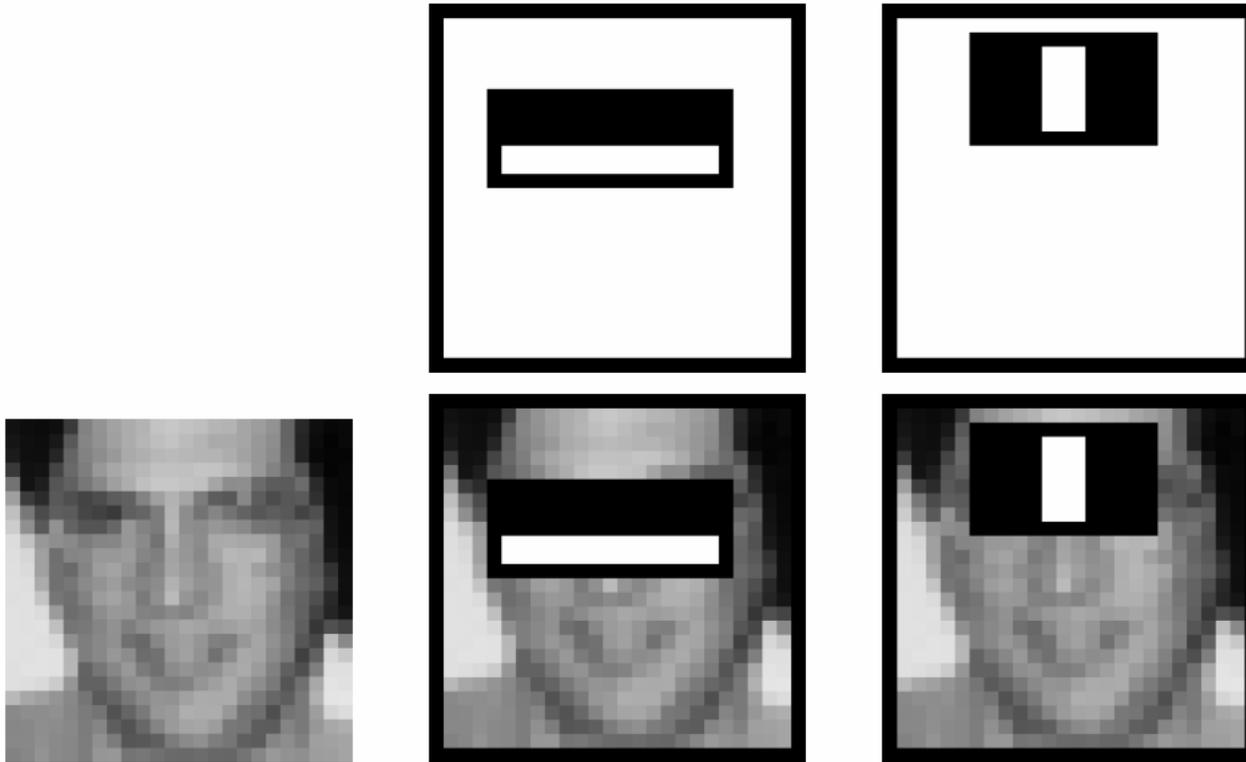
Boosting illustration

**Final classifier is
a combination of weak
classifiers**



Boosting for face detection

- First two features selected by boosting:



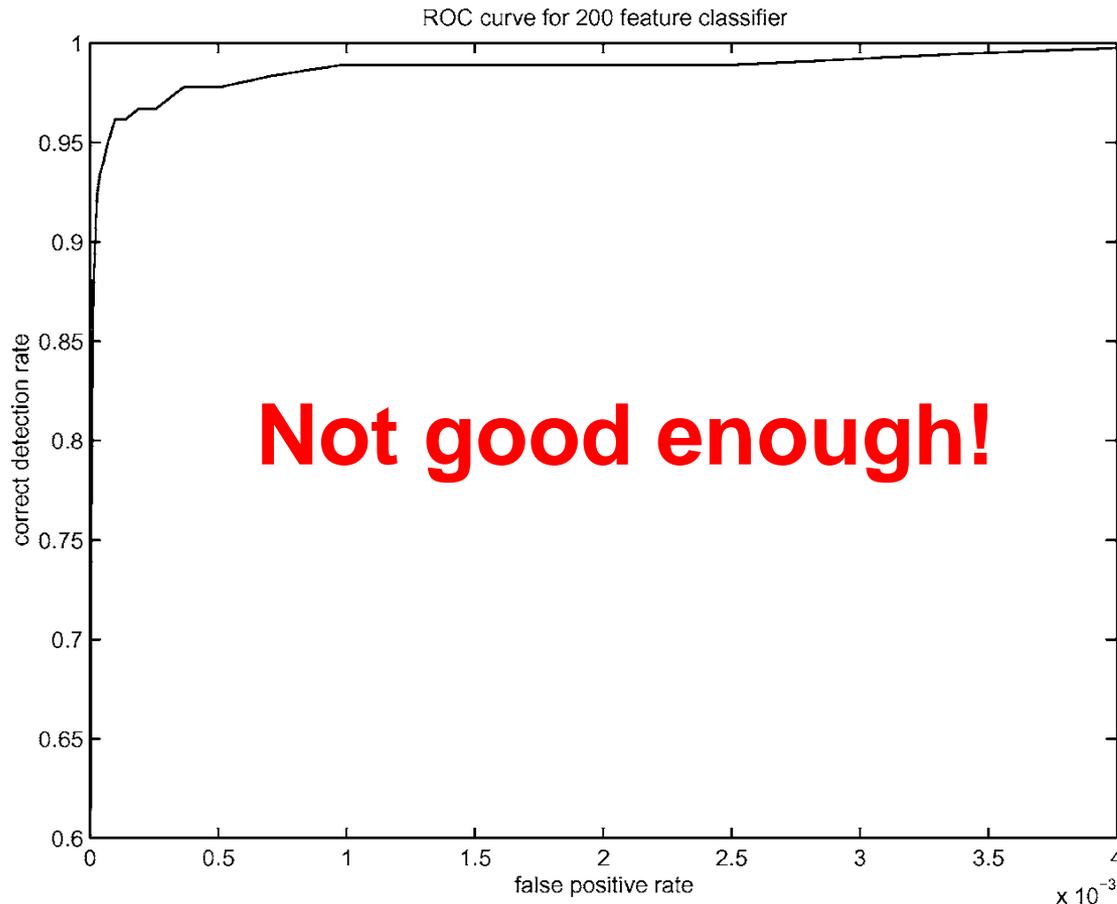
This feature combination can yield 100% detection rate and 50% false positive rate

Boosting vs. SVM

- **Advantages of boosting**
 - Integrates classifier training with feature selection
 - Complexity of training is linear instead of quadratic in the number of training examples
 - Flexibility in the choice of weak learners, boosting scheme
 - Testing is fast
 - Easy to implement
- **Disadvantages**
 - Needs many training examples
 - Training is slow
 - Often doesn't work as well as SVM (especially for many-class problems)

Boosting for face detection

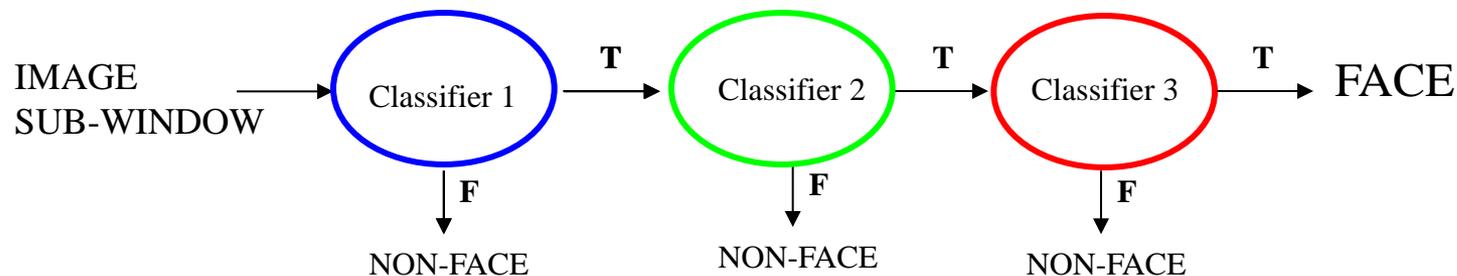
- A 200-feature classifier can yield 95% detection rate and a false positive rate of 1 in 14084



Receiver operating characteristic (ROC) curve

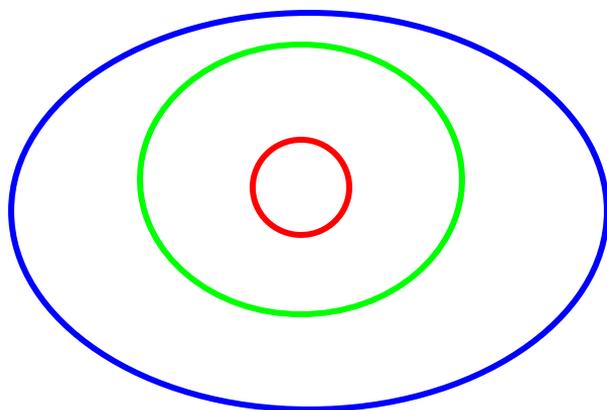
Attentional cascade

- We start with simple classifiers which reject many of the negative sub-windows while detecting almost all positive sub-windows
- Positive response from the first classifier triggers the evaluation of a second (more complex) classifier, and so on
- A negative outcome at any point leads to the immediate rejection of the sub-window

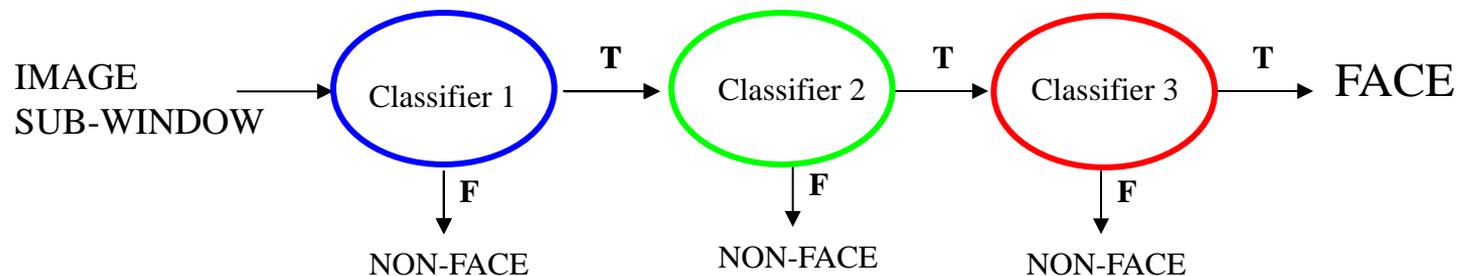
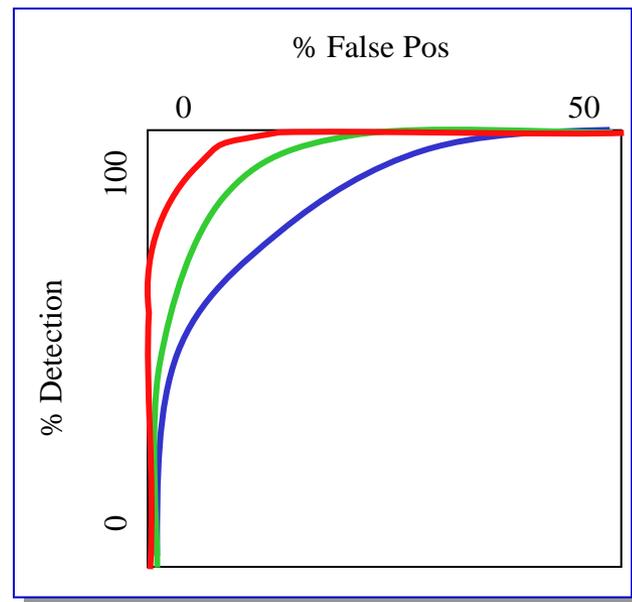


Attentional cascade

- Chain classifiers that are progressively more complex and have lower false positive rates:

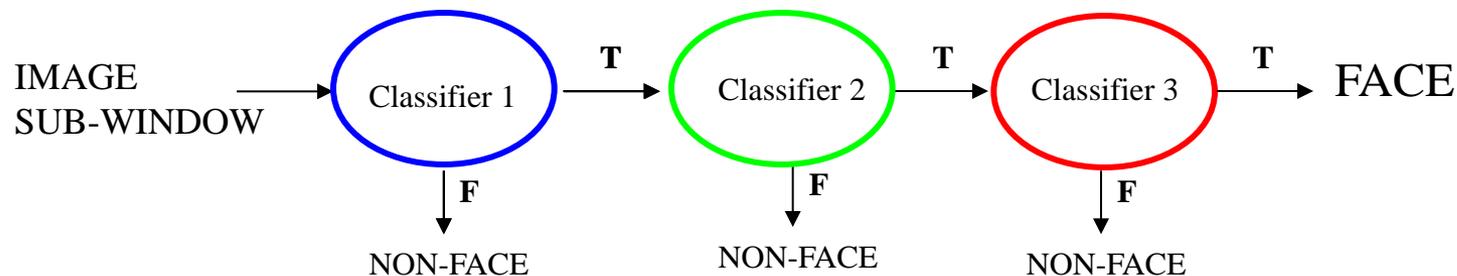


Receiver operating characteristic



Attentional cascade

- The detection rate and the false positive rate of the cascade are found by multiplying the respective rates of the individual stages
- A detection rate of 0.9 and a false positive rate on the order of 10^{-6} can be achieved by a 10-stage cascade if each stage has a detection rate of 0.99 ($0.99^{10} \approx 0.9$) and a false positive rate of about 0.30 ($0.3^{10} \approx 6 \times 10^{-6}$)



Training the cascade

- Set target detection and false positive rates for each stage
- Keep adding features to the current stage until its target rates have been met
 - Need to lower AdaBoost threshold to maximize detection (as opposed to minimizing total classification error)
 - Test on a *validation set*
- If the overall false positive rate is not low enough, then add another stage
- Use false positives from current stage as the negative training examples for the next stage

The implemented system

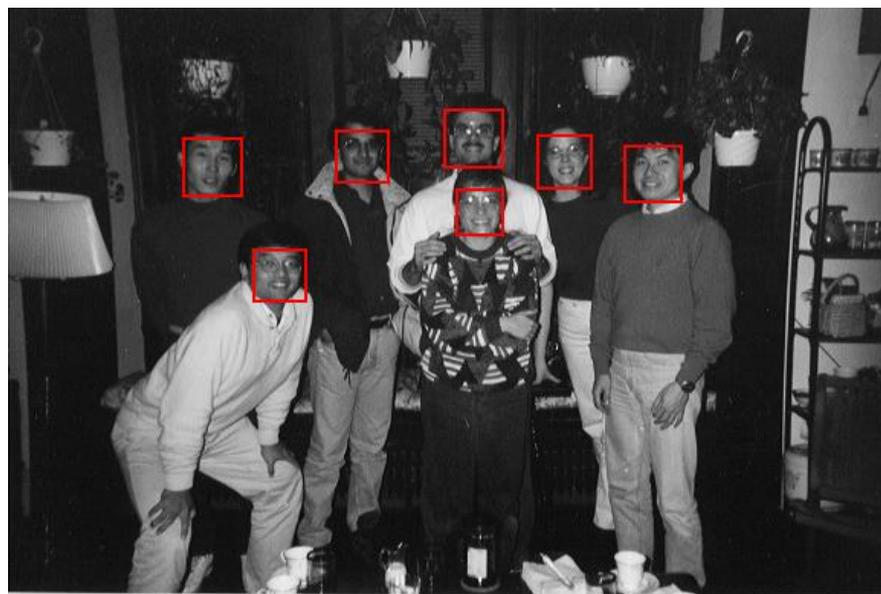
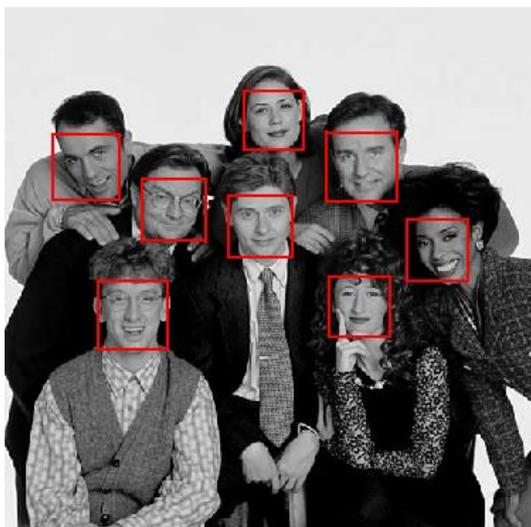
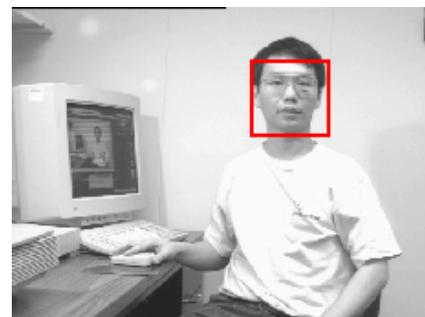
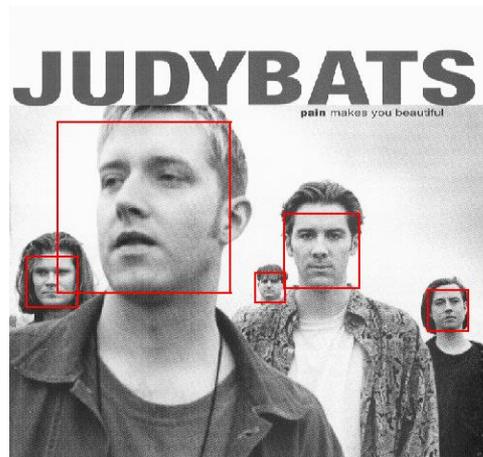
- Training Data
 - 5000 faces
 - All frontal, rescaled to 24x24 pixels
 - 300 million non-faces
 - 9500 non-face images
 - Faces are normalized
 - Scale, translation
- Many variations
 - Across individuals
 - Illumination
 - Pose



System performance

- Training time: “weeks” on 466 MHz Sun workstation
- 38 layers, total of 6061 features
- Average of 10 features evaluated per window on test set
- “On a 700 Mhz Pentium III processor, the face detector can process a 384 by 288 pixel image in about .067 seconds”
 - 15 Hz
 - 15 times faster than previous detector of comparable accuracy (Rowley et al., 1998)

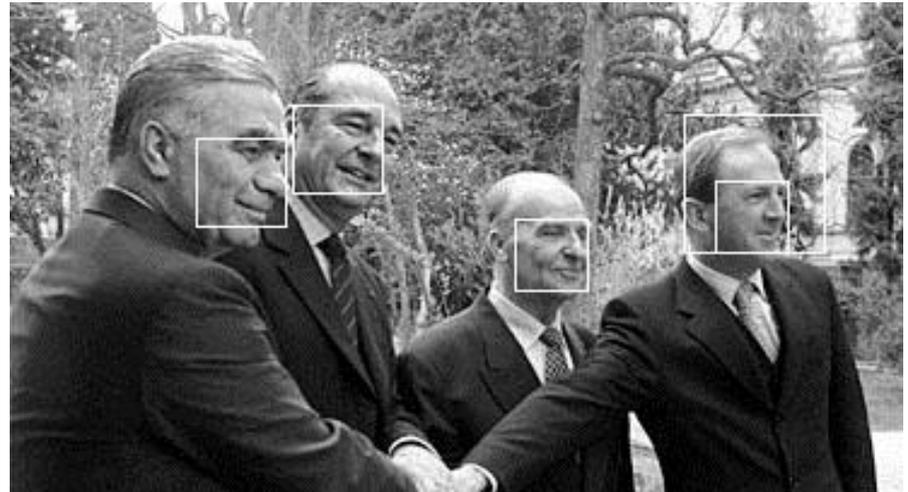
Output of Face Detector on Test Images



Other detection tasks

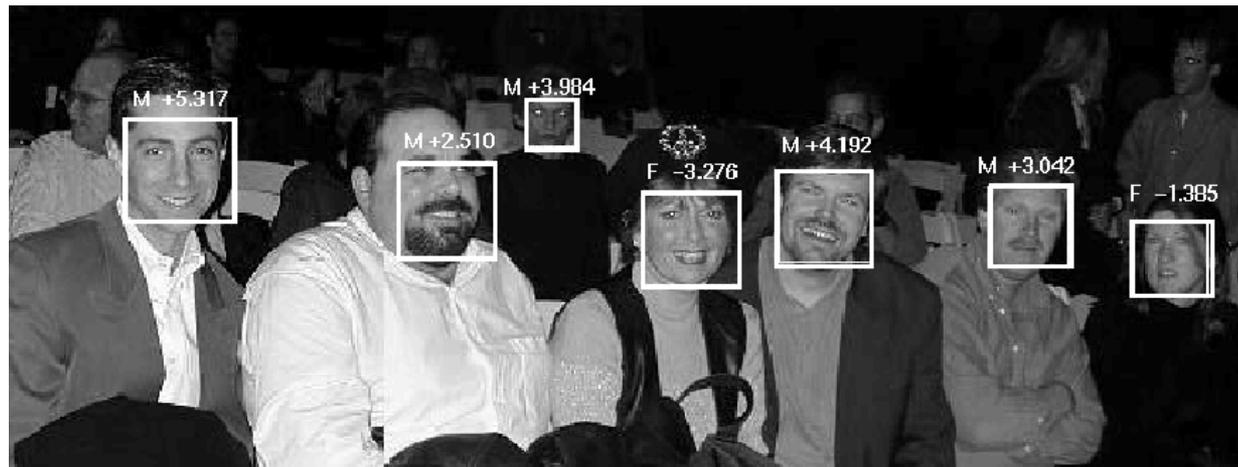


Facial Feature Localization

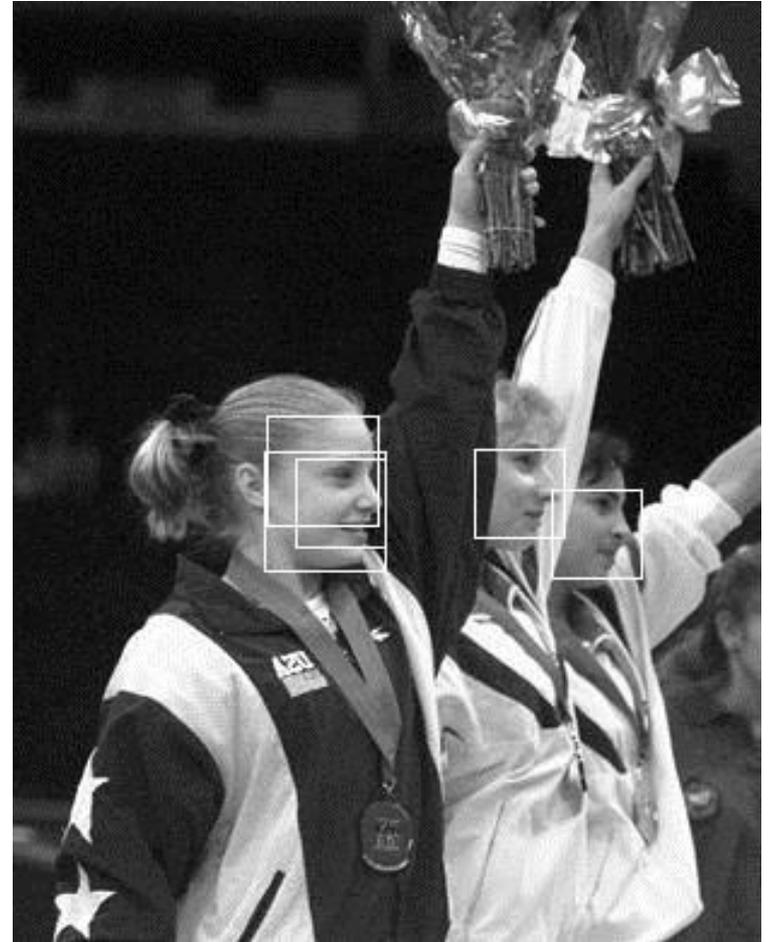


Profile Detection

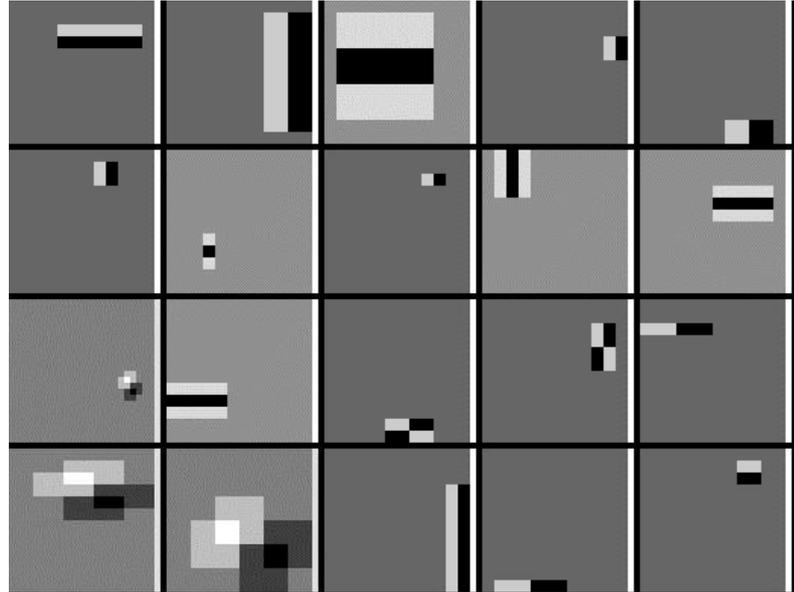
Male vs.
female



Profile Detection



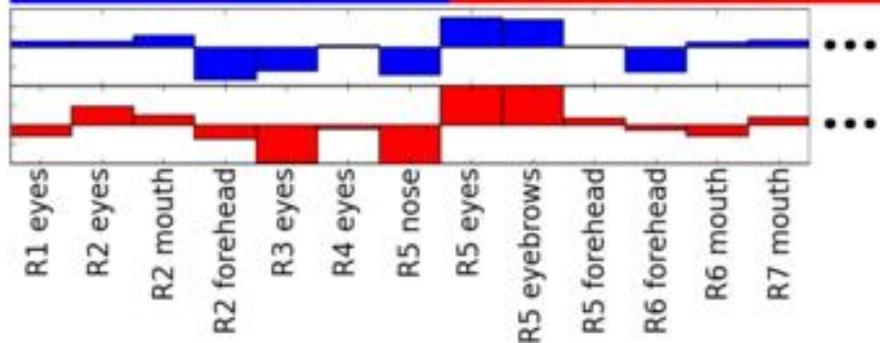
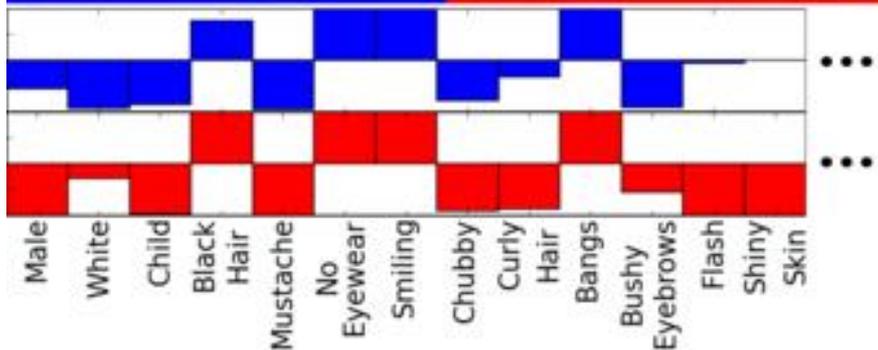
Profile Features



Summary: Viola/Jones detector

- Rectangle features
- Integral images for fast computation
- Boosting for feature selection
- Attentional cascade for fast rejection of negative windows

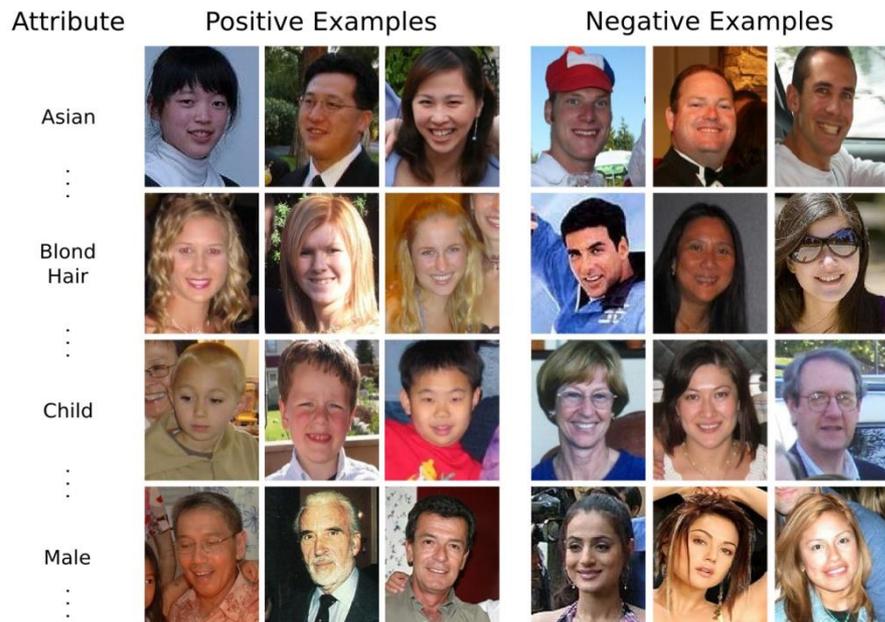
Face Recognition



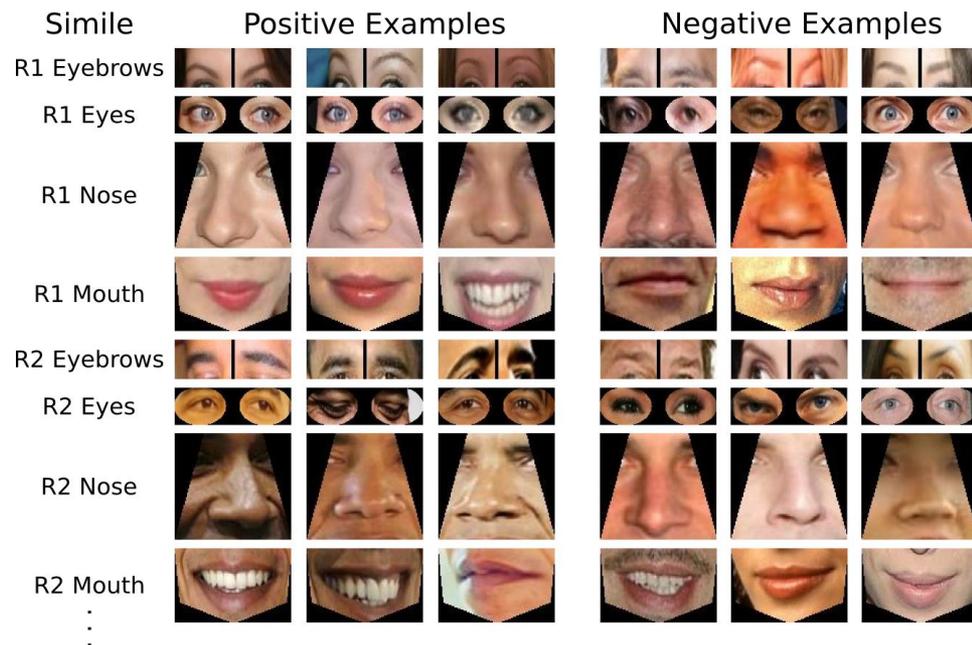
N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "[Attribute and Simile Classifiers for Face Verification](#)," ICCV 2009.

Face Recognition

Attributes for training



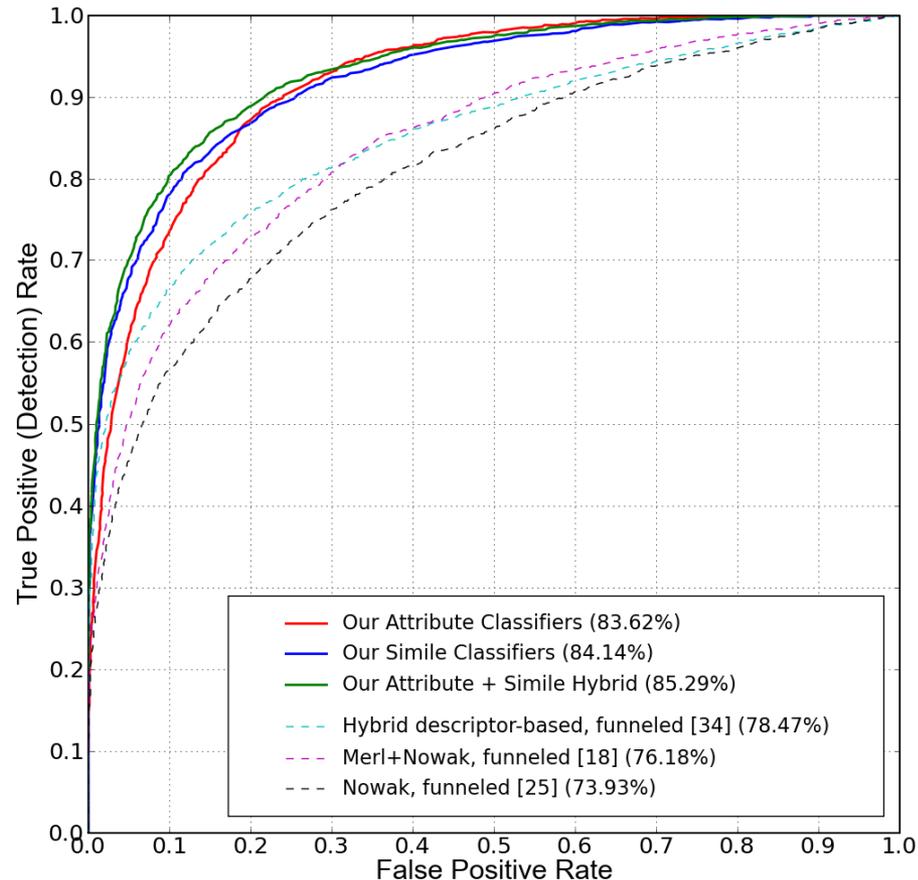
Similes for training



N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "[Attribute and Simile Classifiers for Face Verification](#)," ICCV 2009.

Face Recognition

Results on Labeled Faces in the Wild Dataset



N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and Simile Classifiers for Face Verification," ICCV 2009.