Big ideas

- Straightforward discriminative methods are powerful
- Spot locally discriminative properties
  - skin
  - faces
  - motion
- Pool these for detection
- Reason about relations

Finding skin

- Skin has a very small range of (intensity independent) colours, and little texture
  - Compute an intensity-independent colour measure, check if colour is in this range, check if there is little texture (median filter)
  - See this as a classifier - we can set up the tests by hand, or learn them.

Histogram based classifiers

- Represent class-conditional densities with histogram
- Advantage:
  - estimates become quite good
    - (with enough data!)
- Disadvantage:
  - Histogram becomes big with high dimension
    - but maybe we can assume feature independence?

Histogram classifier for skin

$$\frac{P(rgb \mid skin)}{P(rgb \mid \neg skin)} \geq \Theta$$

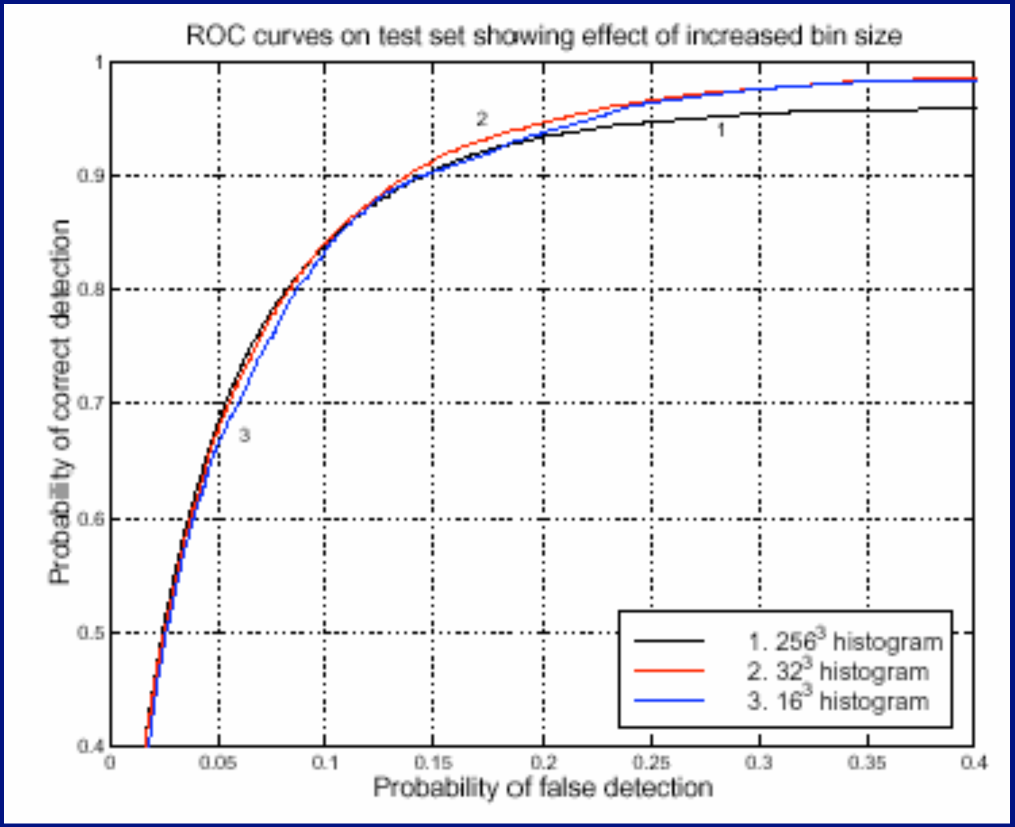


Figure from Jones+Rehg, 2002

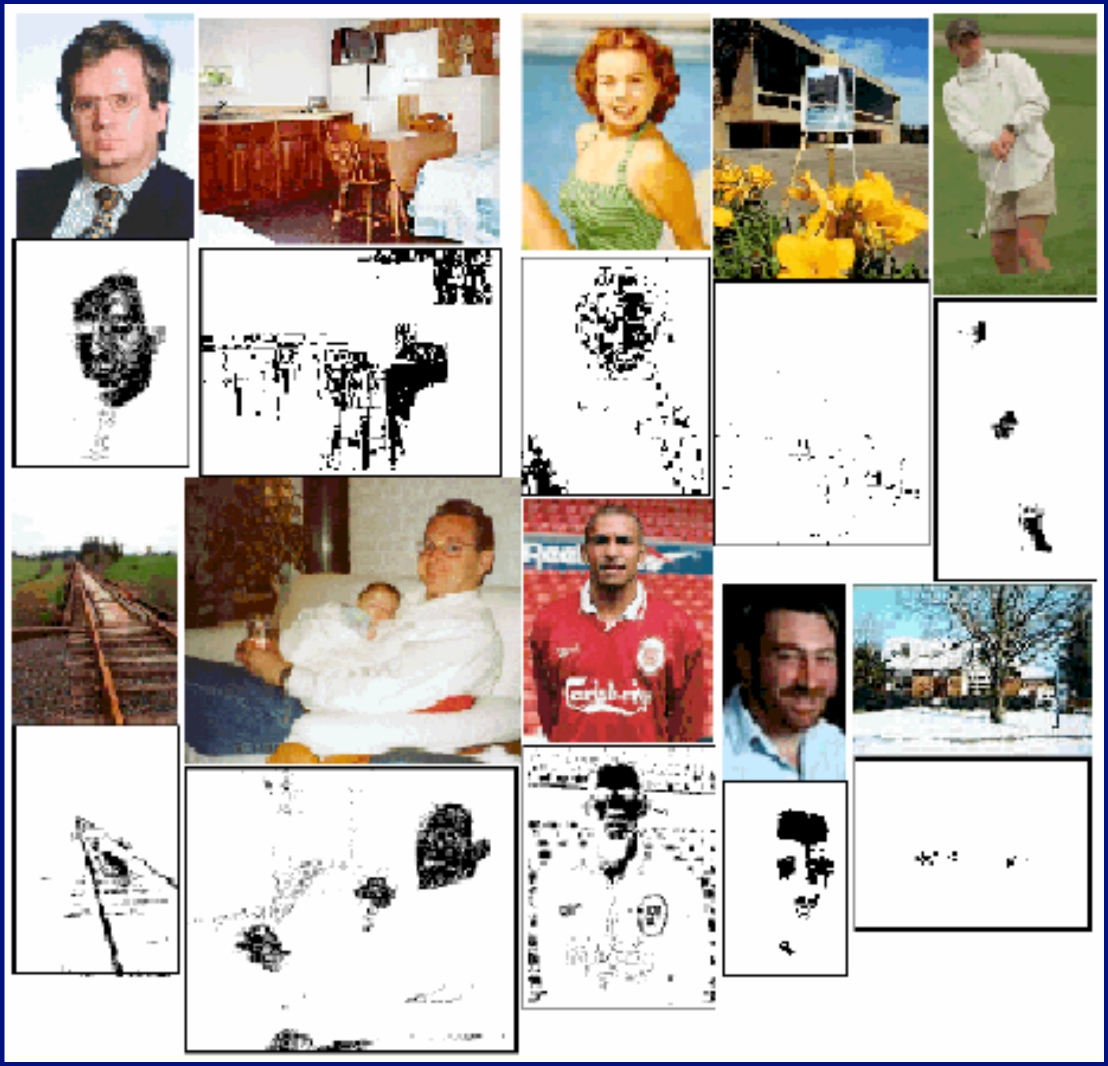


Figure from Jones+Rehg, 2002

Curse of dimension - I

- This won’t work for many features
  - try R, G, B, and some texture features
  - too many histogram buckets

Finding faces

- Faces “look like” templates (at least when they’re frontal).
- General strategy:
  - search image windows at a range of scales
  - Correct for illumination
  - Present corrected window to classifier
- Issues
  - How corrected?
  - What features?
  - What classifier?
  - what about lateral views?

The Thatcher Illusion
Figures by Henry Rowley,
http://www.cs.cmu.edu/~har/puzzle.html

The Thatcher Illusion
Figures by Henry Rowley,
http://www.cs.cmu.edu/~har/puzzle.html

Naive Bayes

- Previously, we detected with a likelihood ratio test

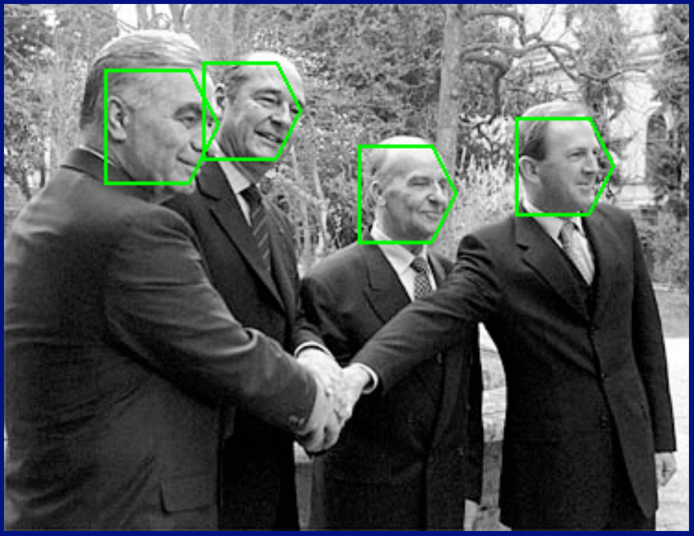
$$\frac{P(\text{features}|\text{event})}{P(\text{features}|\text{not event})} > \text{threshold}$$

- Now assume that features are conditionally independent given event

$$P(f_0, f_1, f_2, \ldots, f_n|\text{event}) = P(f_0|\text{event})P(f_1|\text{event})P(f_2|\text{event})\ldots P(f_n|\text{event})$$

Naive Bayes

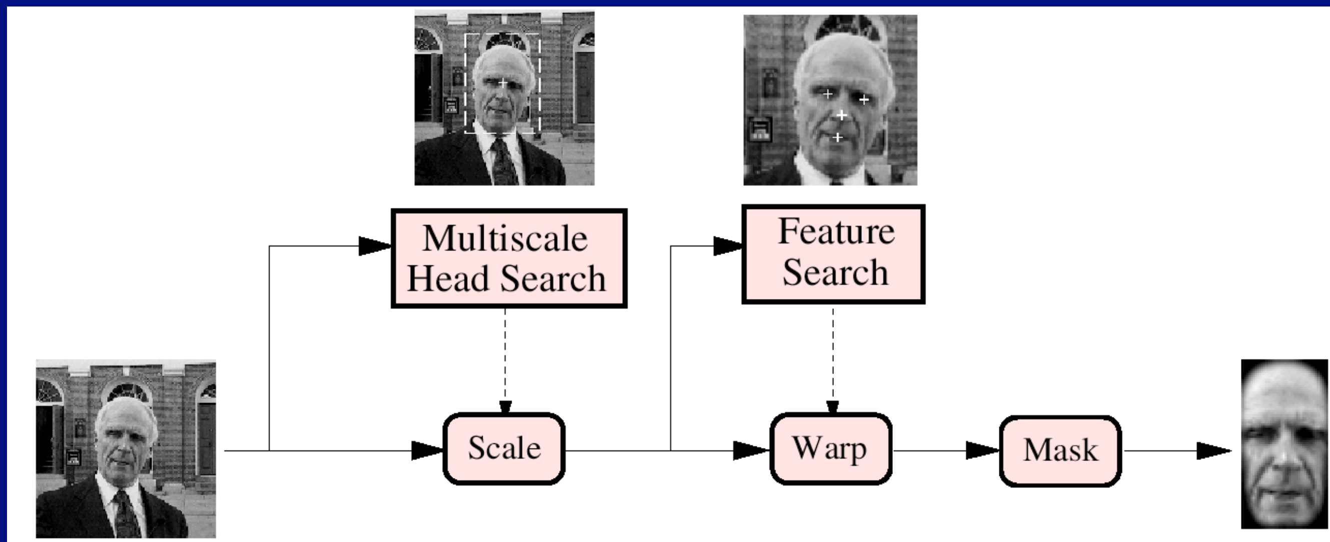
- (not necessarily perjorative)
- Histogram doesn’t work when there are too many features
  - the curse of dimension, first version
  - assume they’re independent conditioned on the class, cross fingers
  - reduction in degrees of freedom
  - very effective for face finders
    - relations may not be all that important
  - very effective for high dimensional problems
    - bias vs. variance



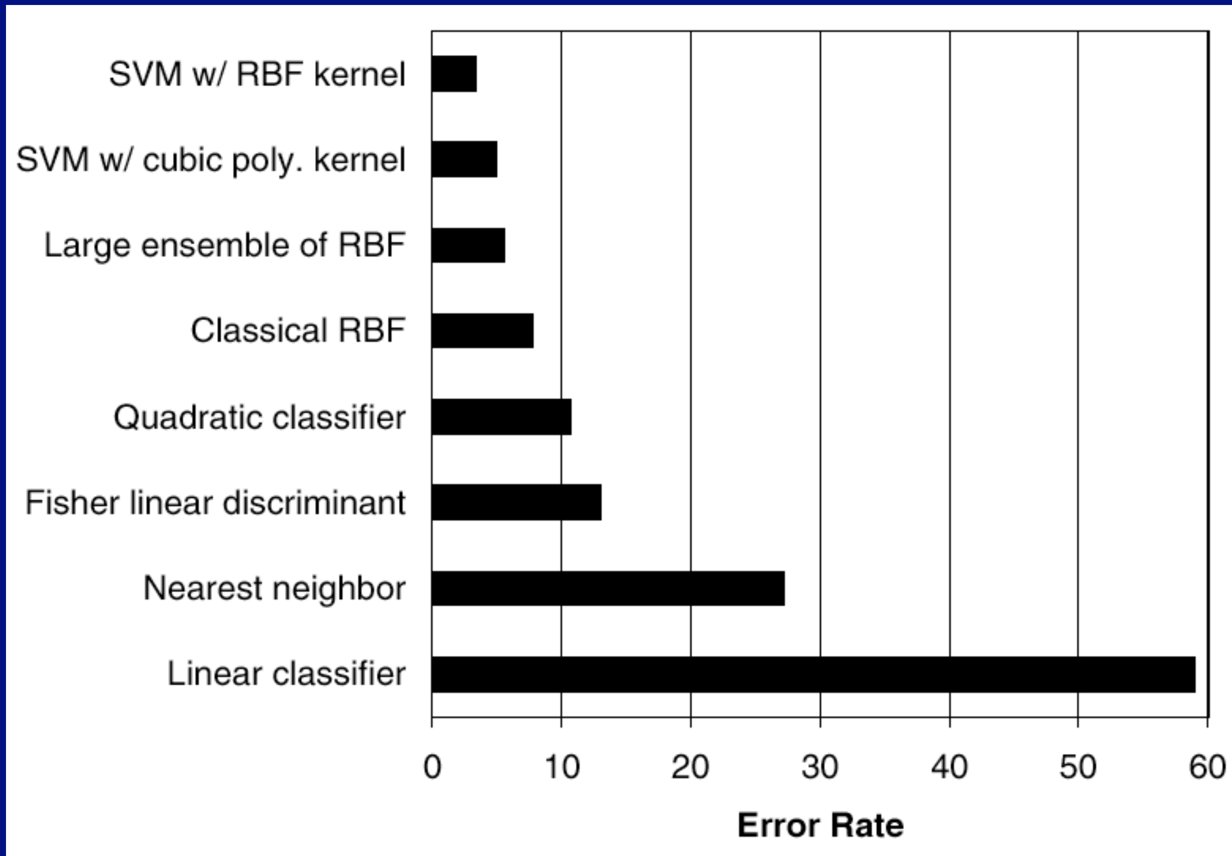
Work by Schneiderman and Kanade,
http://www.cs.cmu.edu/afs/cs.cmu.edu/user/hws/www/hws.html

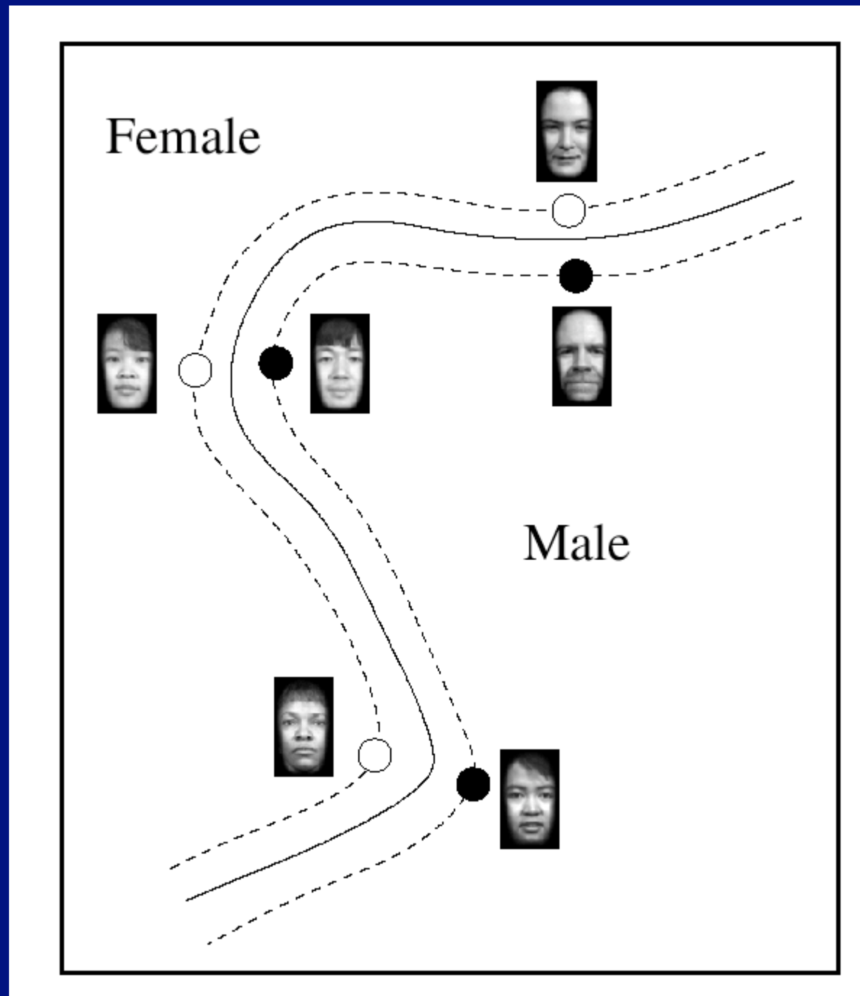

Work by Schneiderman and Kanade,
http://www.cs.cmu.edu/afs/cs.cmu.edu/user/hws/www/hws.html

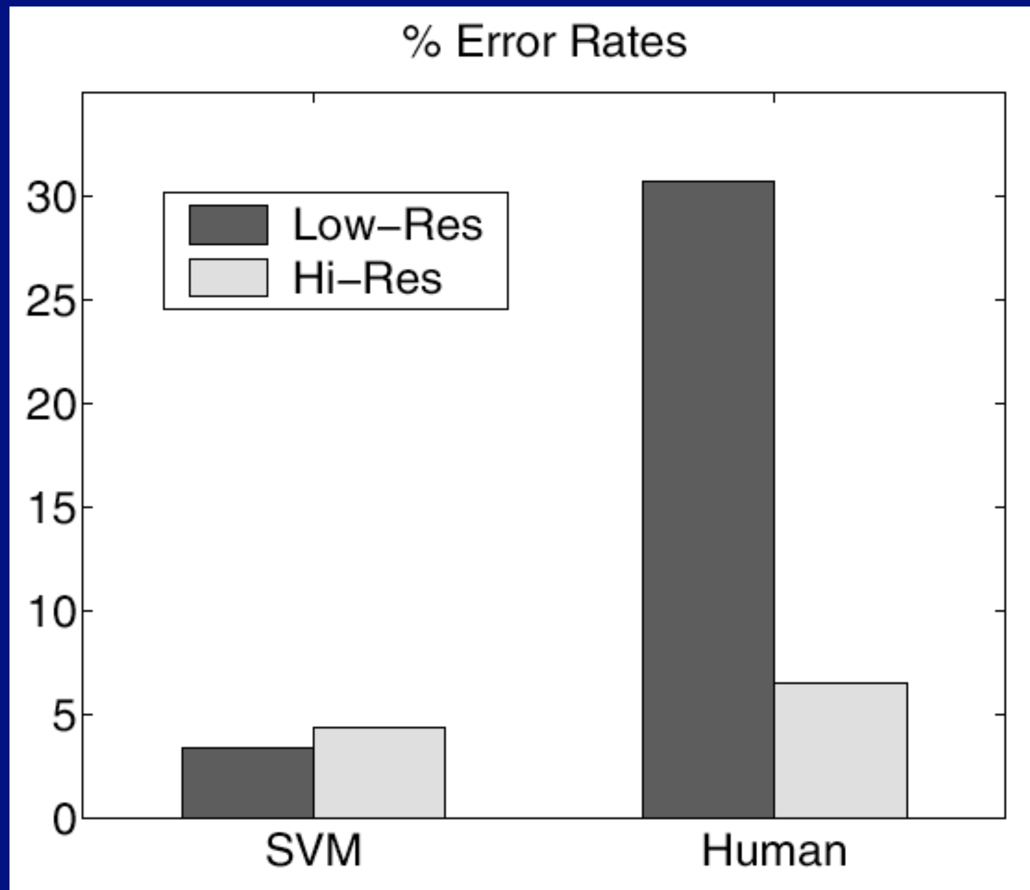
Many more face finders on the face detection home page
http://home.t-online.de/home/Robert.Frischholz/face.htm



From “Gender Classication with
Support Vector Machines”
Baback Moghaddam
Ming-Hsuan Yang, MERL TR


SVM w/ RBF kernel
SVM w/ cubic poly. kernel
Large ensemble of RBF
Classical RBF
Quadratic classifier
Fisher linear discriminant
Nearest neighbor
Linear classifier
0
10
20
30
40
50
60
Error Rate


Female
Male


% Error Rates
Low–Res
Hi–Res
30
25
20
15
10
5
0
SVM
Human

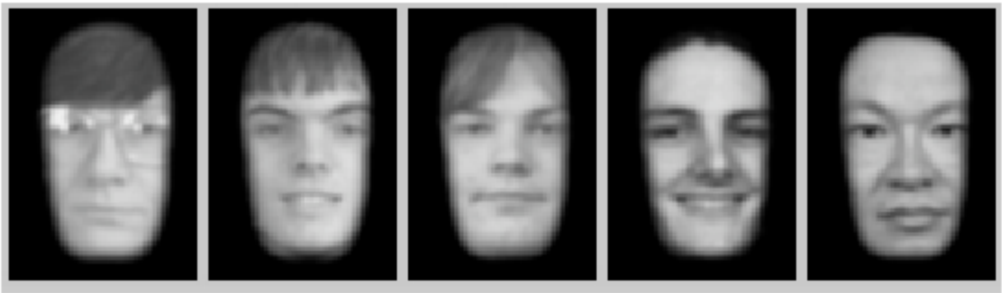
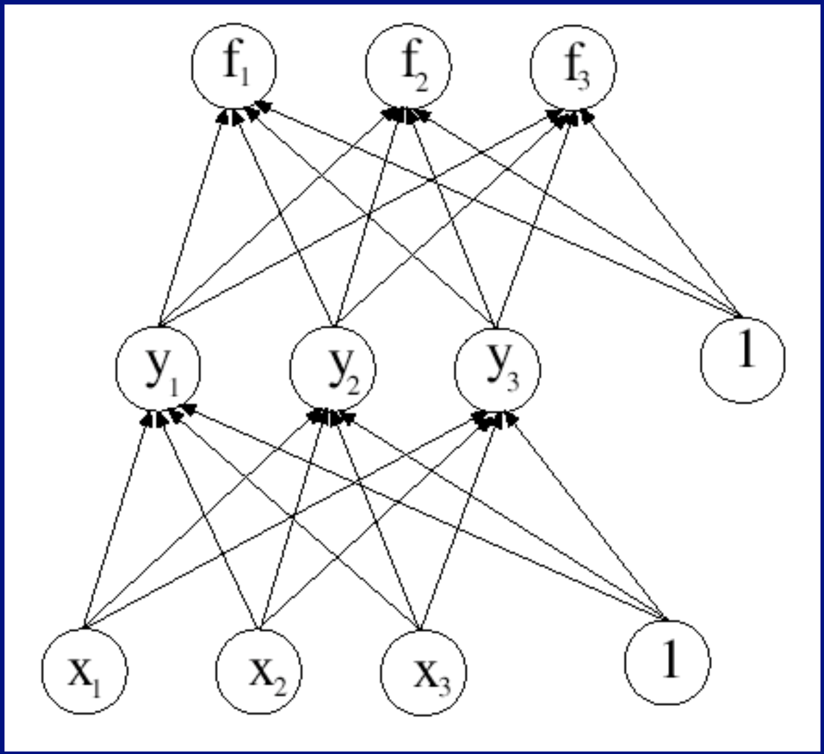


Figure 7. Top five human misclassifications

Neural networks

- Logistic regression heavily generalized



$$\boldsymbol{g}(\boldsymbol{x}) \approx \boldsymbol{f}(\boldsymbol{x}) = [\phi(\boldsymbol{w}_{21} \cdot \boldsymbol{y}), \phi(\boldsymbol{w}_{22} \cdot \boldsymbol{y}), \ldots \phi(\boldsymbol{w}_{2n} \cdot \boldsymbol{y})]$$

$$\boldsymbol{y}(\boldsymbol{z}) = [\phi(\boldsymbol{w}_{11} \cdot \boldsymbol{z}), \phi(\boldsymbol{w}_{12} \cdot \boldsymbol{z}), \ldots \phi(\boldsymbol{w}_{1m} \cdot \boldsymbol{z}), 1]$$

$$\boldsymbol{z}(\boldsymbol{x}) = [x_1, x_2, \ldots, x_p, 1]$$

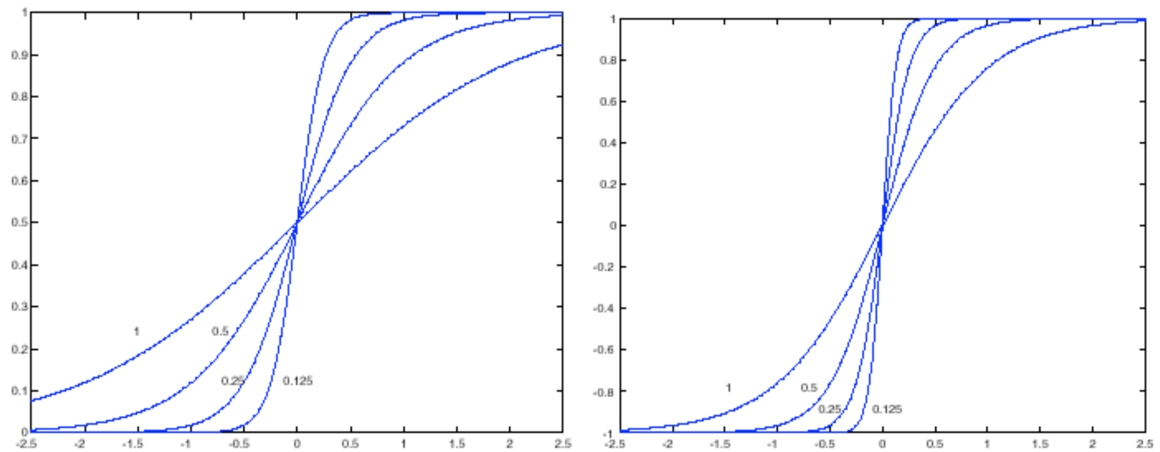


Figure 22.14. On the **left**, a series of squashing functions obtained using $\phi(x;\nu) = \frac{e^{x/\nu}}{1+e^{x/\nu}}$, for different values of ν indicated on the figure. On the **right**, a series of squashing functions obtained using $\phi(x;\nu,A) = A\tanh(x/\nu)$ for different values of ν indicated on the figure. Generally, for x close to the center of the range, the squashing function is linear; for x small or large, it is strongly non-linear.

Training

- Choose parameters to minimize error on training set

$$Error(\boldsymbol{p}) = \left(\frac{1}{2}\right)\sum_{e} |\boldsymbol{n}(\boldsymbol{x}^e;\boldsymbol{p}) - \boldsymbol{o}^e|^2$$

- Stochastic gradient descent, computing gradient using trick (backpropagation, aka the chain rule)
- Stop when error is low, and hasn't changed much

Rowley-Baluja-Kanade face finder (1)

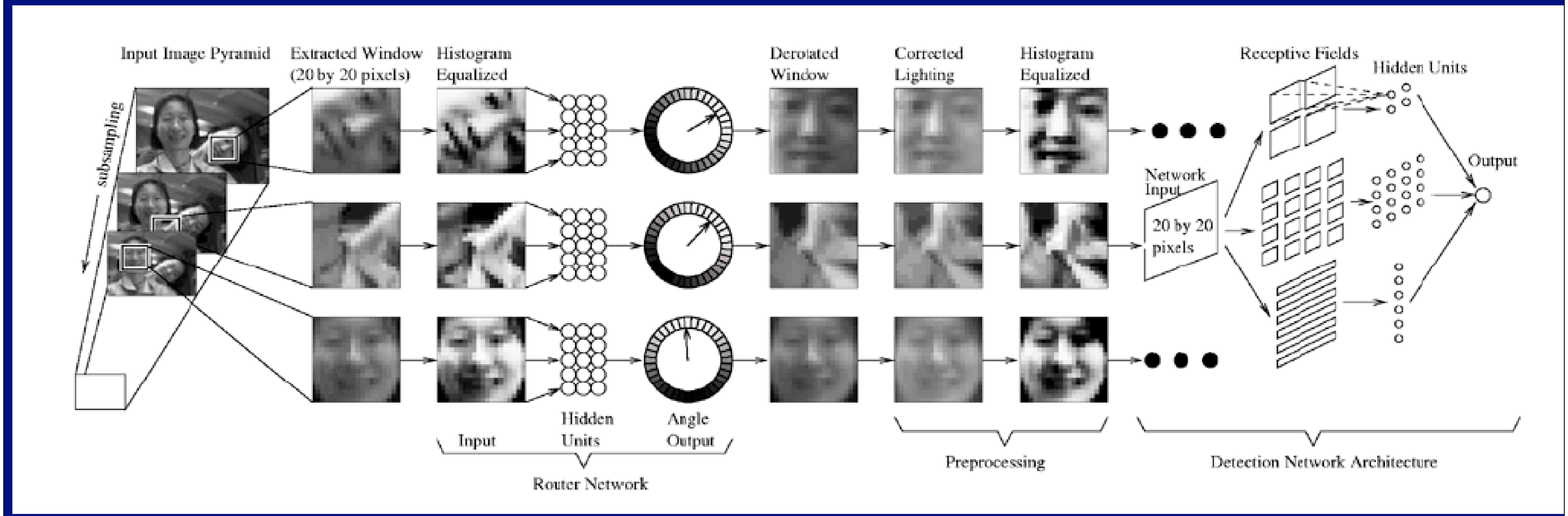


Figure from "Rotation invariant neural-network based face detection,"
H.A. Rowley, S. Baluja and T. Kanade, Proc. Computer Vision and Pattern Recognition, 1998, c 1998, IEEE as shown in Forsyth and Ponce, p589

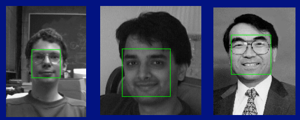
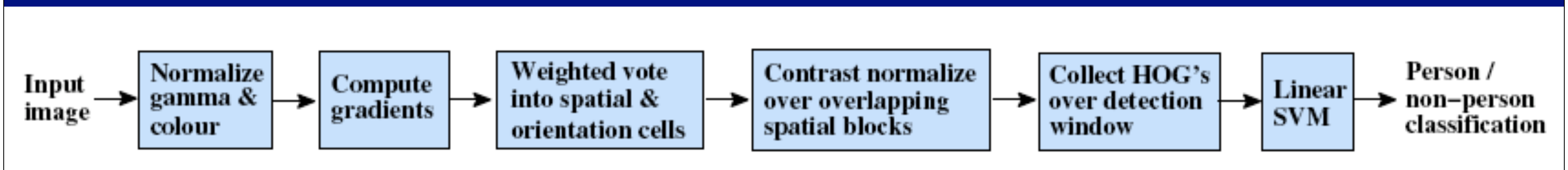


Figure from “Rotation invariant neural-network based face detection,”
H.A. Rowley, S. Baluja and T. Kanade, Proc. Computer Vision and Pattern Recognition,
1998, c 1998, IEEE as shown in Forsyth and Ponce, p589

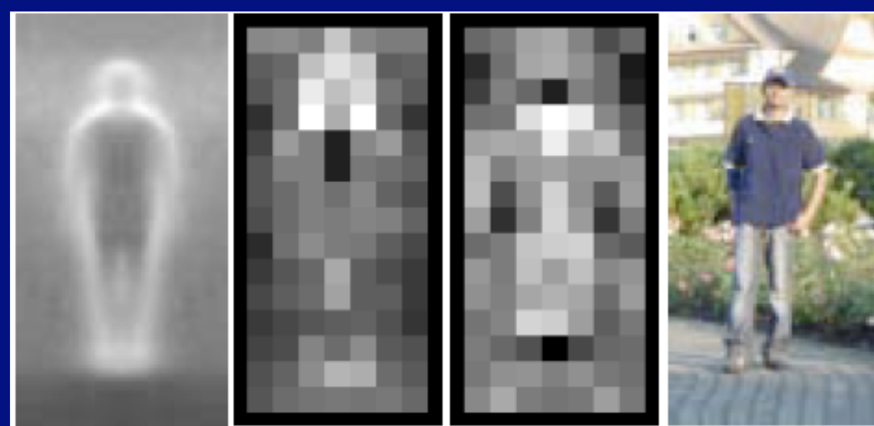
Pedestrian detection with an SVM

Dalal+Triggs 05

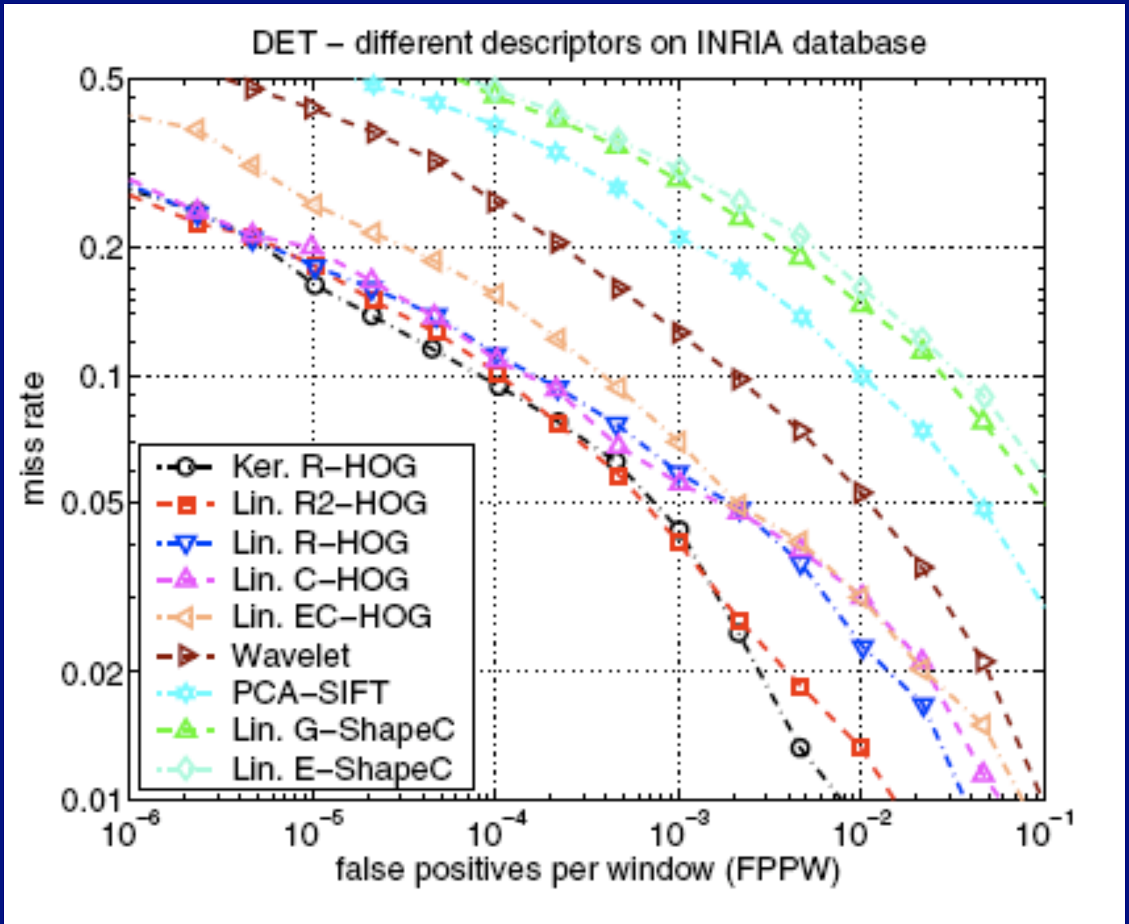
Features



Dalal+Triggs 05



Performance



Dalal+Triggs 05

Face Recognition

- Whose face is this? (perhaps in a mugshot)
- Issue:
  - What differences are important and what not?
  - Reduce the dimension of the images, while maintaining the “important” differences.
- One strategy:
  - Principal components analysis, then nearest neighbours
  - Many face recognition strategies at http://www.cs.rug.nl/users/peterkr/FACE/face.html

Curse of dimension-II

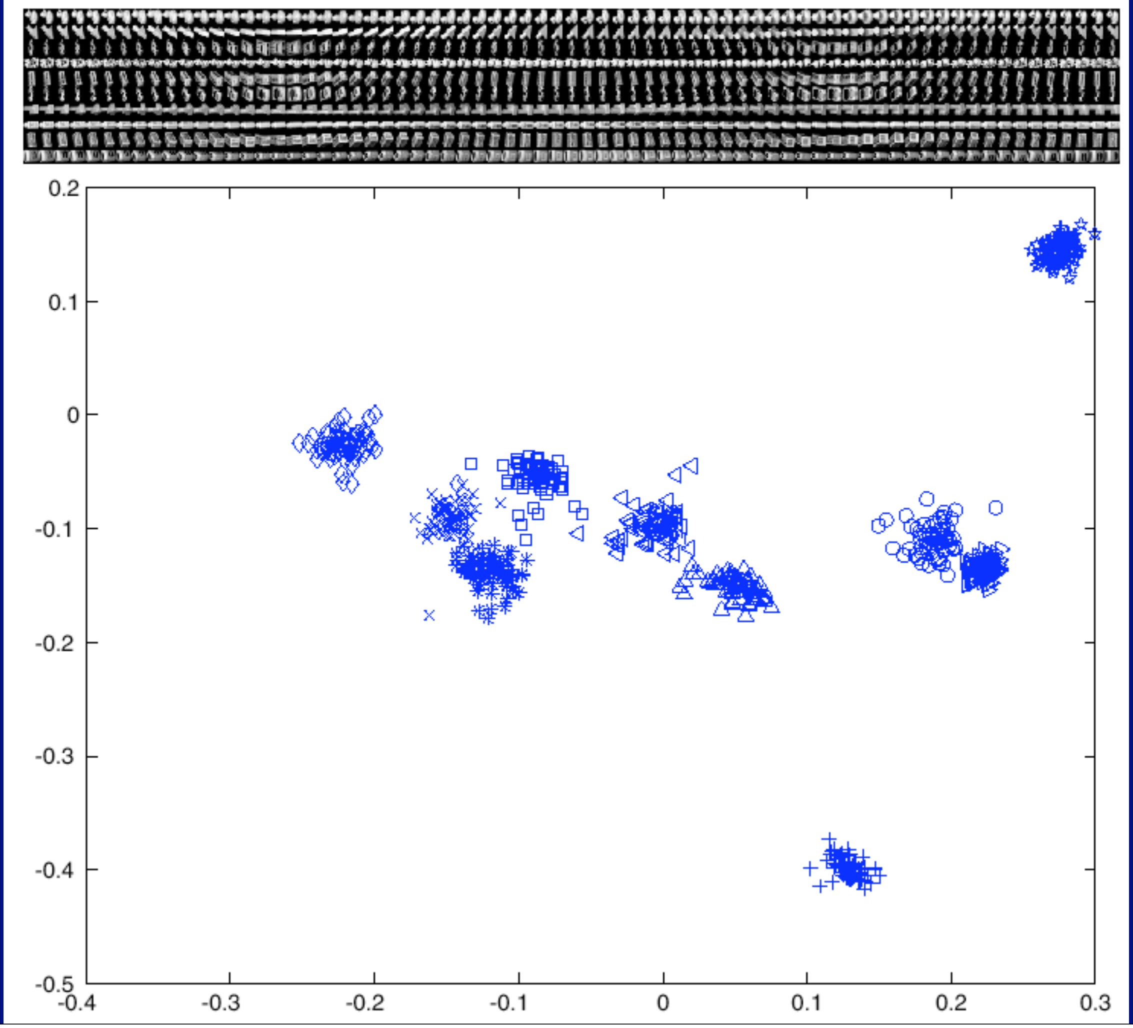
- General phenomenon of high dimensions
  - volume is concentrated at the boundary
- Parameter estimation is hard for high dimensional distributions
  - even Gaussians
    - where probability is concentrated further and further from the mean
    - and covariance has too many parameters
      - dodge: assume covariance is diagonal
- Idea: reduce the dimension of the feature set
  - Principal components
  - Linear discriminants

Principal components for face images, from
http://vismod.www.media.mit.edu/vismod/demos/facerec/basic.html

Linear discriminant analysis

- Principal components do not preserve discrimination
  - so we could have features that don't distinguish, see picture
- Assume (pretend) class conditional densities are normal, with the same covariance
  - Choose linear features so that
    - between class variation is big compared to within class variation
      - between class variation
        - covariance of class means
      - within class variation
        - class covariance

First two canonical variates for well known image collection

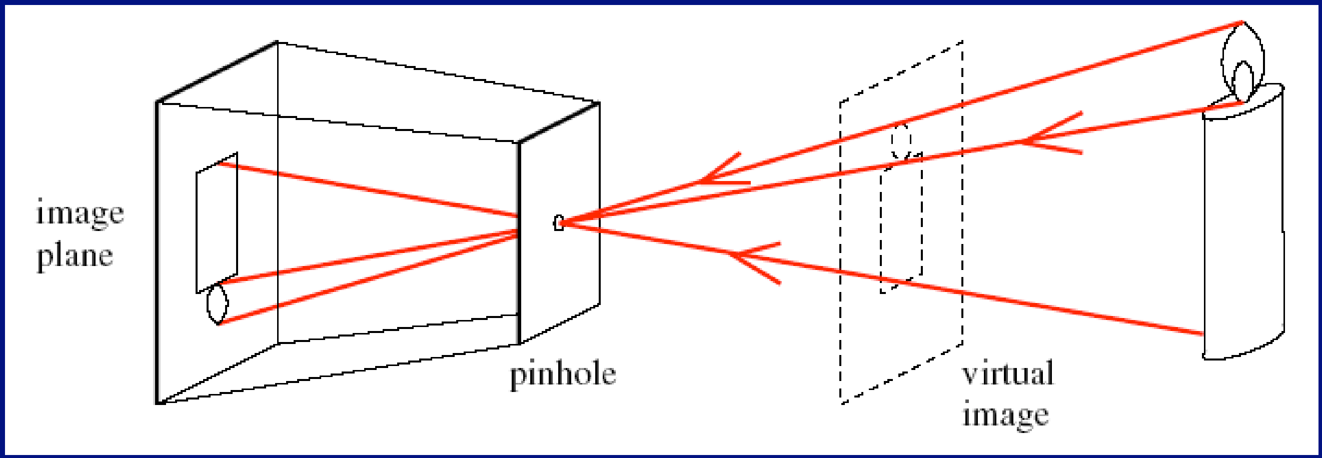


More complex template matching

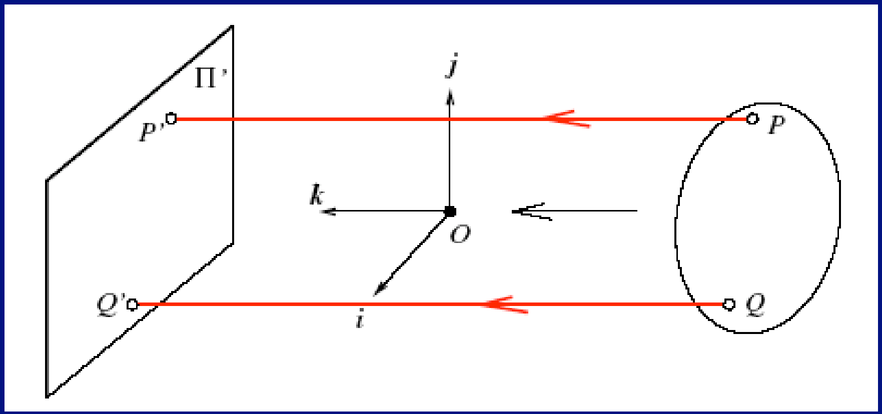
- Encode an object as a set of patches
  - centered on interest points
  - match by
    - voting
    - spatially censored voting
    - inference on a spatial model
- Patches are small
  - even if they’re on a curved surface, we can think of them as being plane

View variation for a plane patch

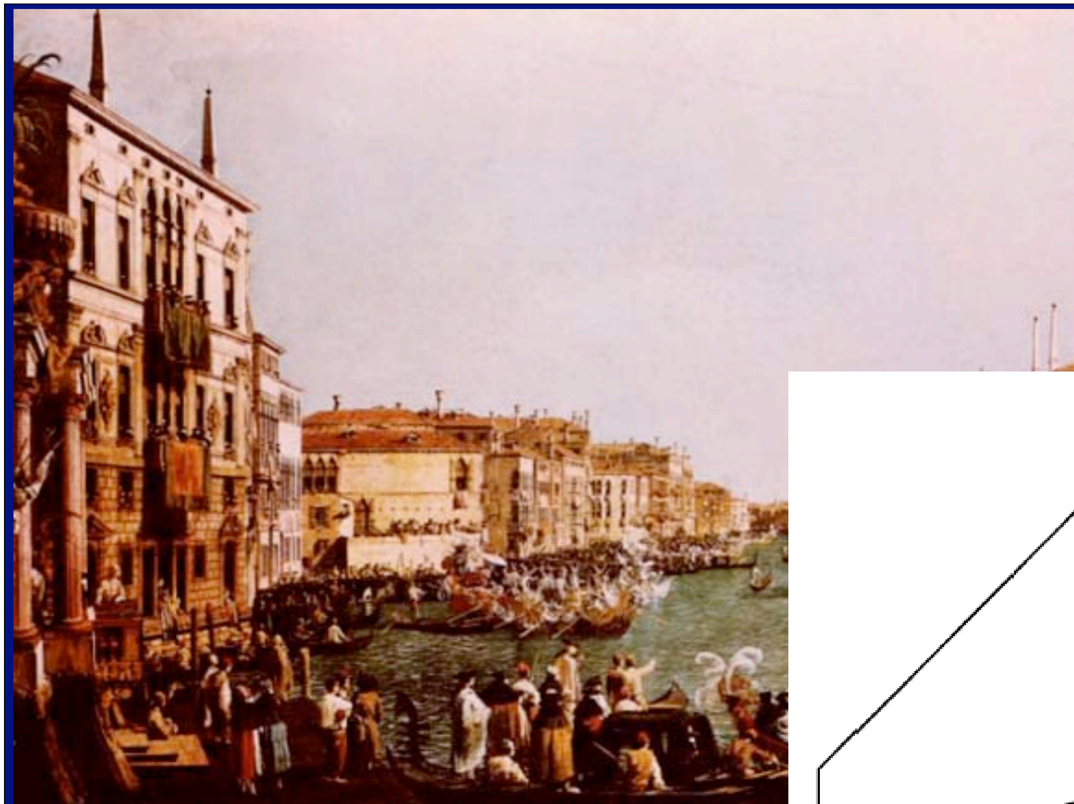
- Plane patches look different in different views

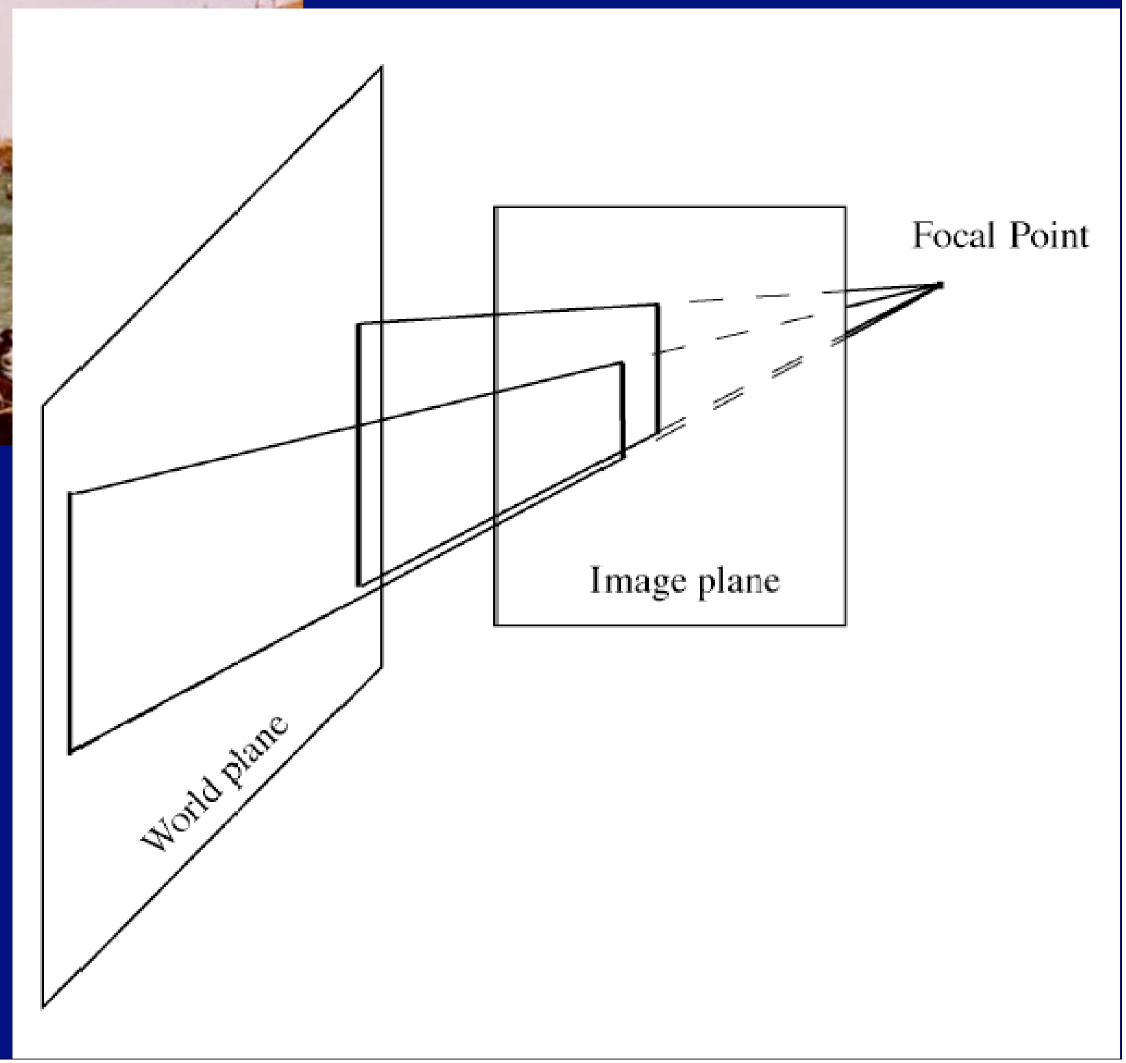


Pinhole camera (F+P, p31)

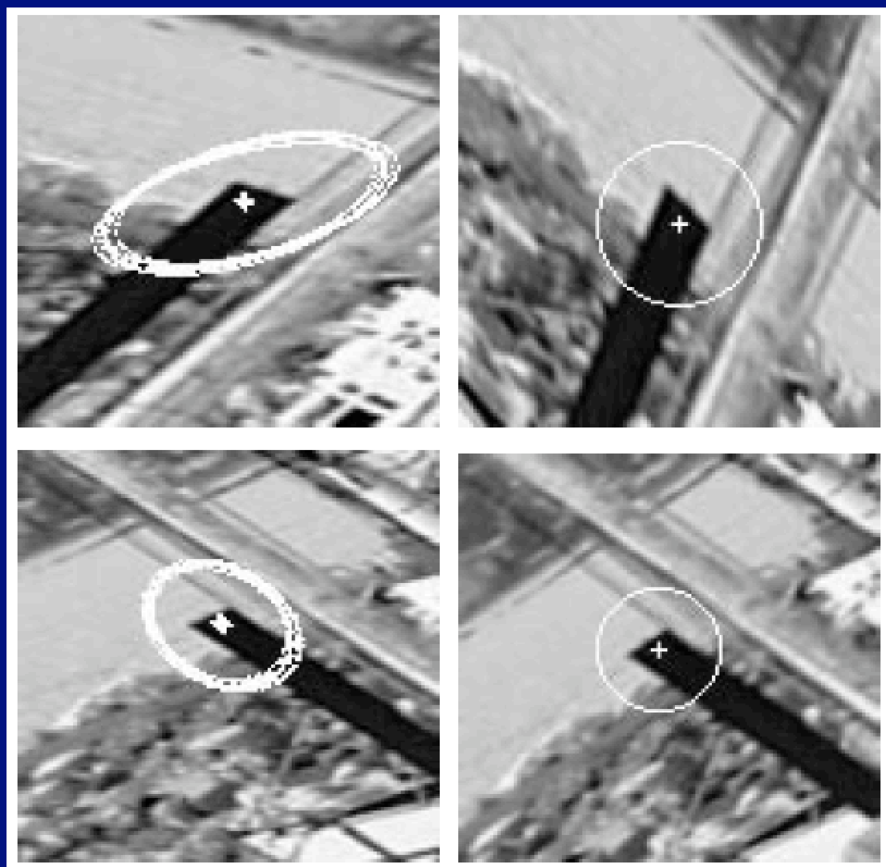


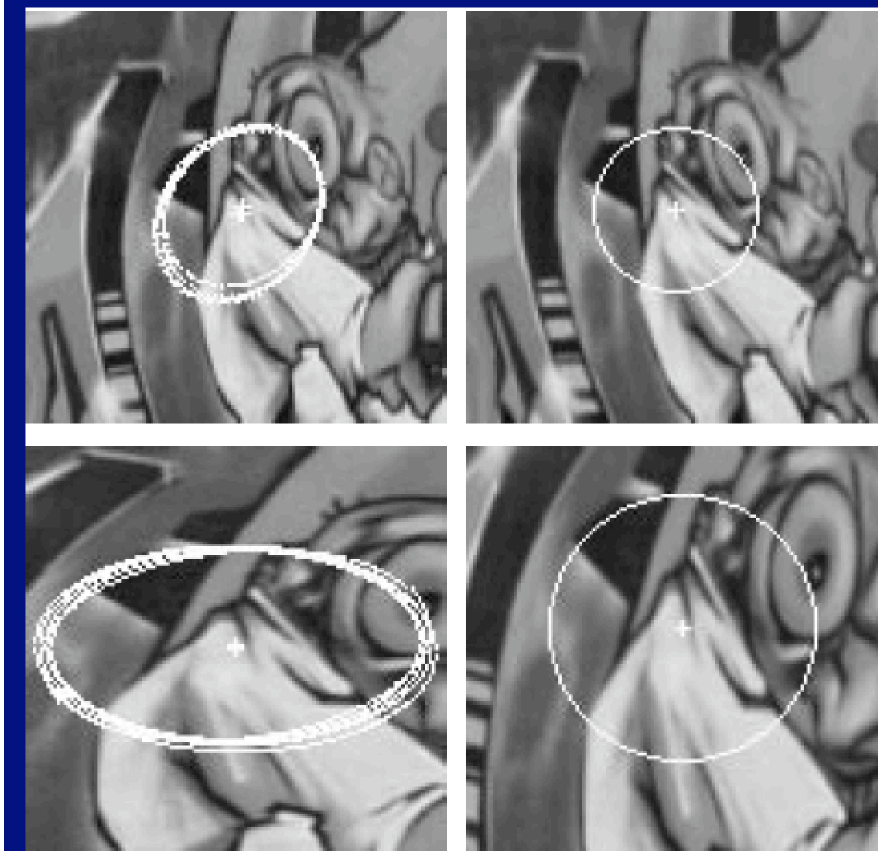
Orthographic camera (F+P, p33)

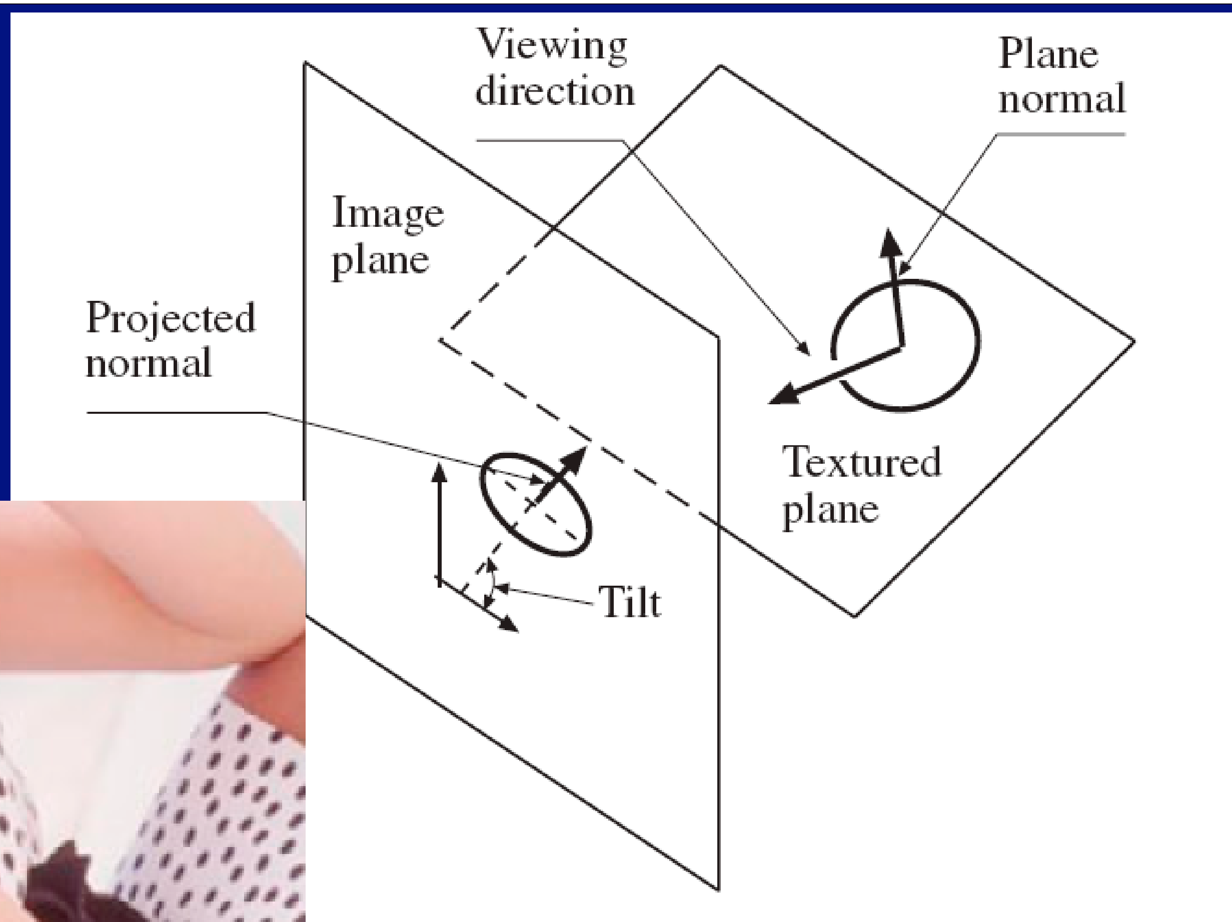




Views of plane patches in perspective cameras induce homographies







Views of plane patches in orthographic cameras induce a special subset of homographies

Interest points and local descriptions

- Find localizable points in the image
  - e.g. corners - established technology,
    - eg find image windows where there tend to be strong edges going in several different directions
- Build at each point
  - a local, canonical coordinate frame
    - Euclidean+scale
    - Affine
  - Do this by searching for a coordinate frame within which some predicate applies
    - E.g. Rotation frame from orientation of gradients
    - E.g. Rotation + scale orientation of gradients, maximum filter response
  - a representation of the image within that coordinate frame
    - this representation is invariant because frame is covariant

Example: Lowe, 99

- Find localizable points in the image
  - find maxima, minima of response to difference of gaussians
    - over space
    - over scale
    - using pyramid
- Build at each point
  - a Euclidean + scale coordinate frame
    - scale from scale of strongest response
    - rotation from peak of orientation histogram within window
  - Representation
    - SIFT features

Lowe’s SIFT features

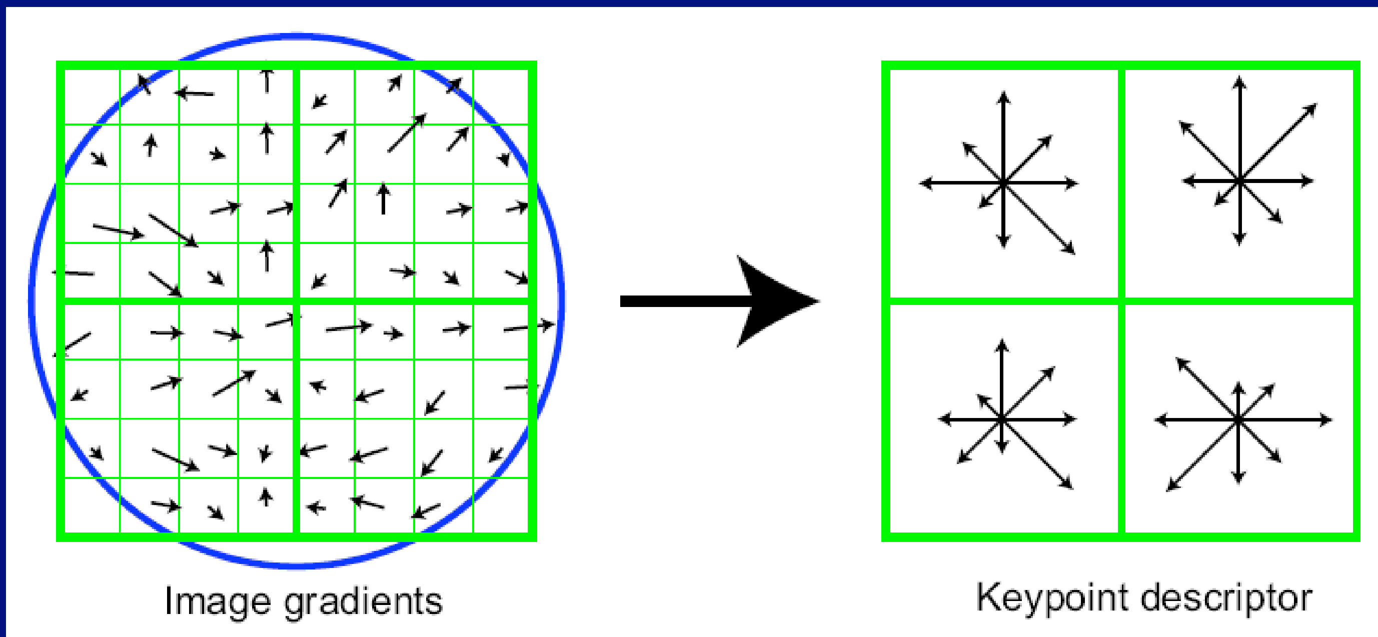
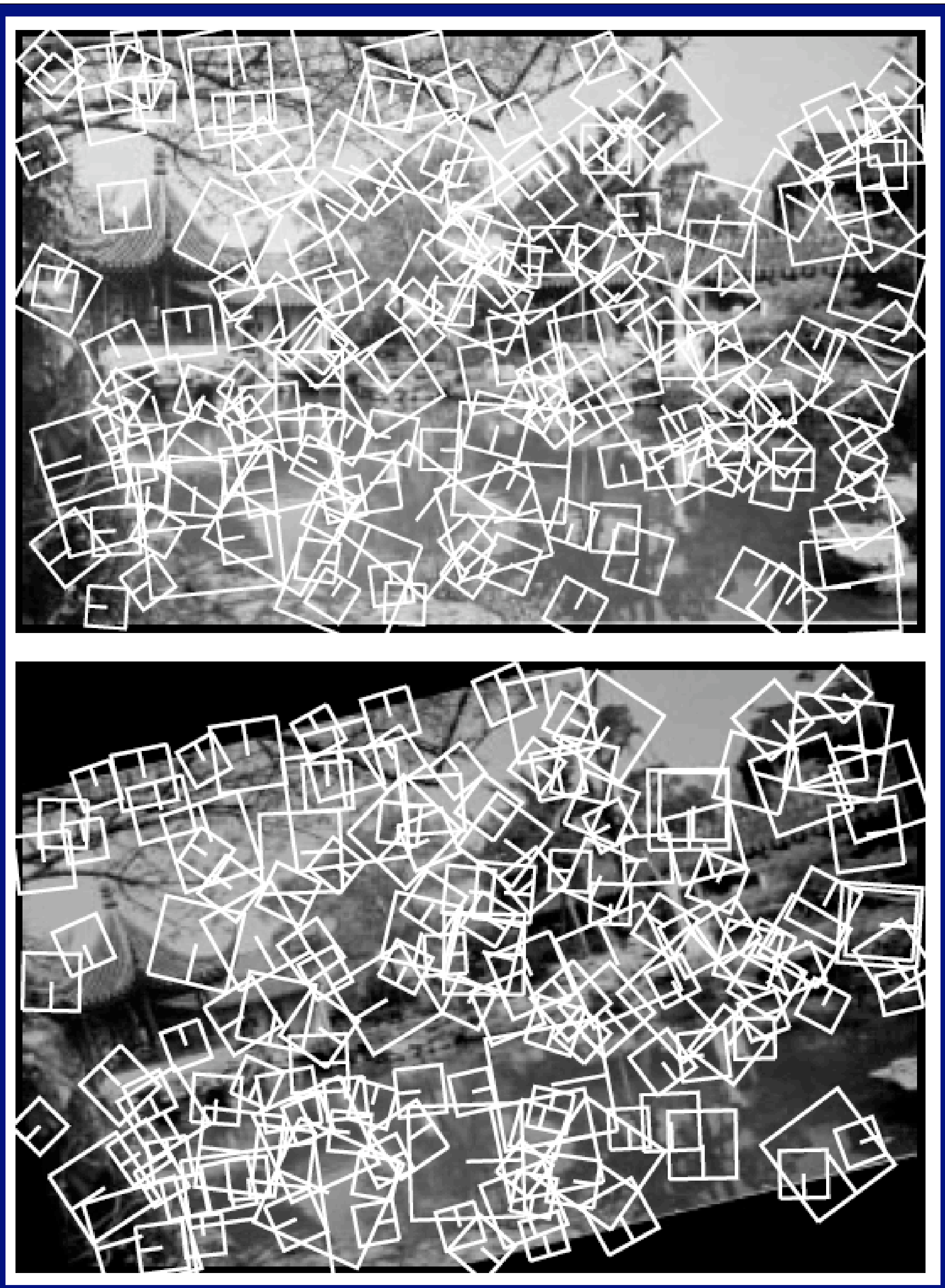
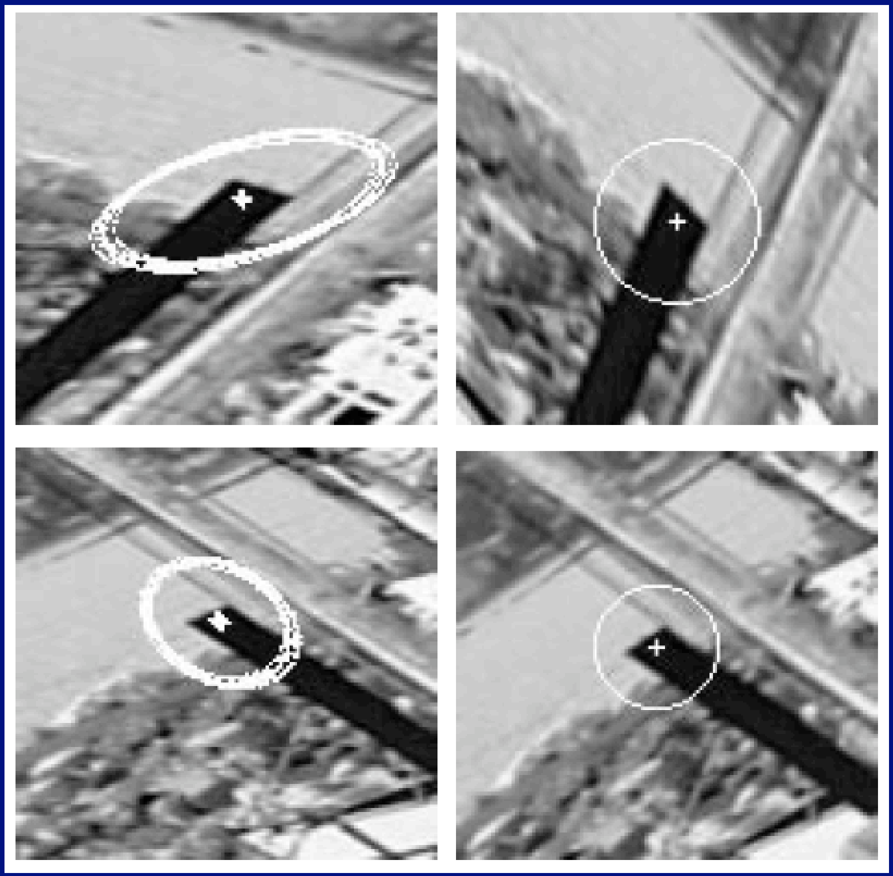


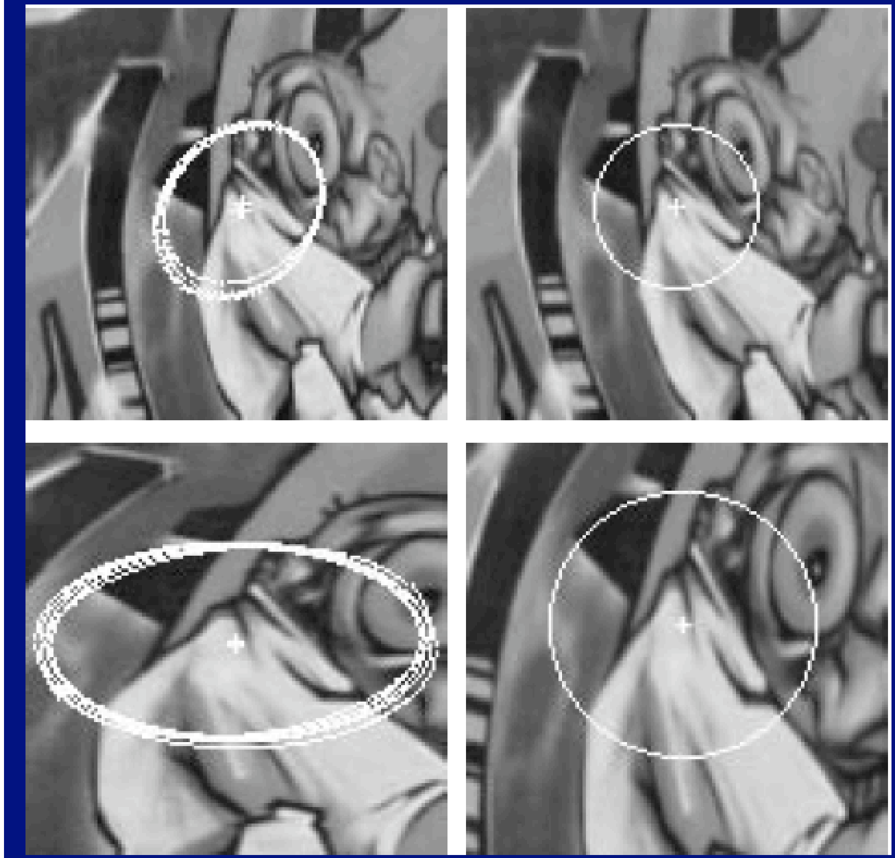
Fig 7 from:
Distinctive image features from scale-invariant keypoints
David G. Lowe, *International Journal of Computer Vision,* 60, 2 (2004), pp. 91-110.



From Lowe, 99, **Object Recognition from Local Scale-Invariant Features**

Mikolaczyk/Schmid coordinate frames





Matching objects with point features

- Voting
  - each point feature votes for every object that contains it
  - object with most votes wins
  - Startlingly effective (see figures)

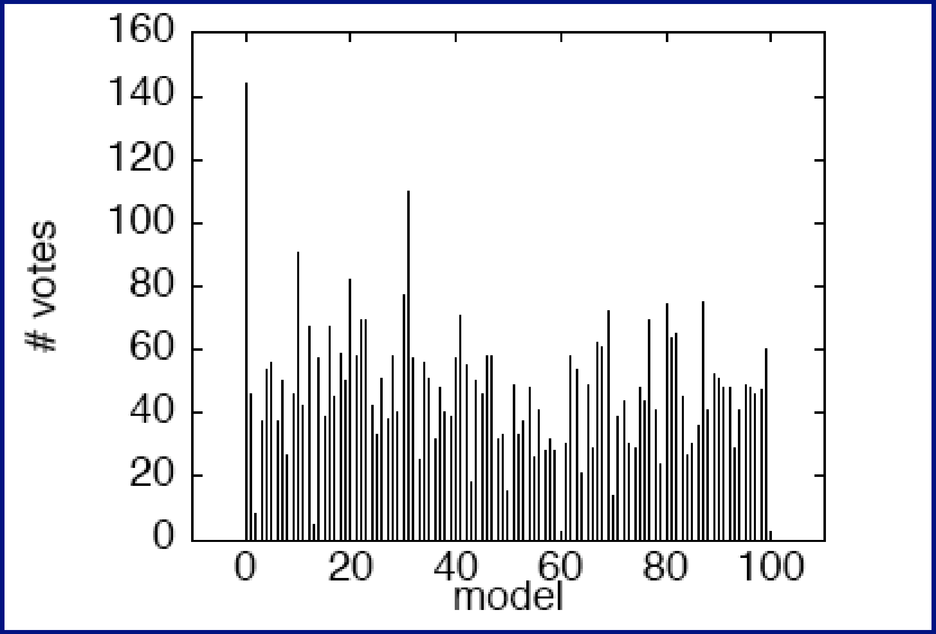
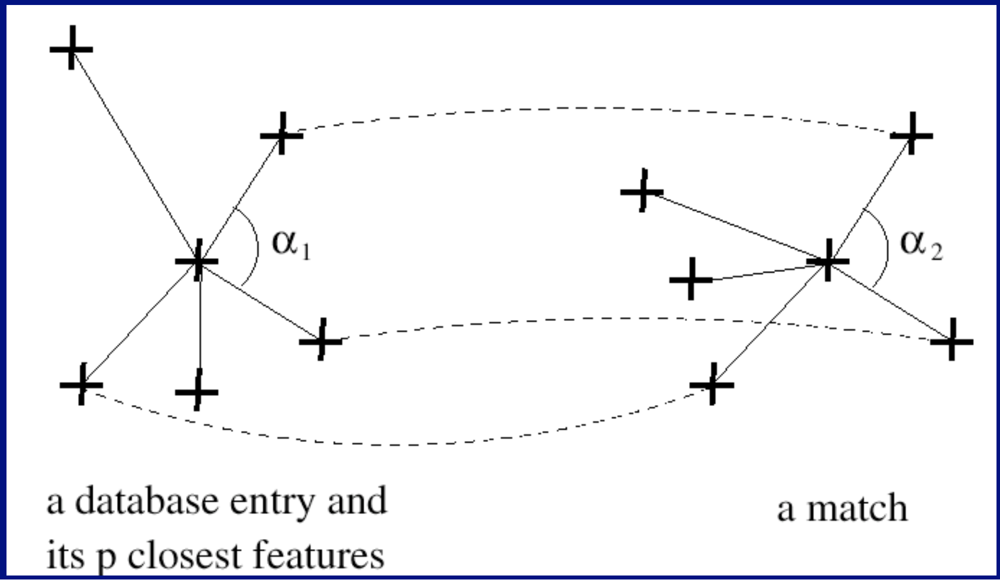
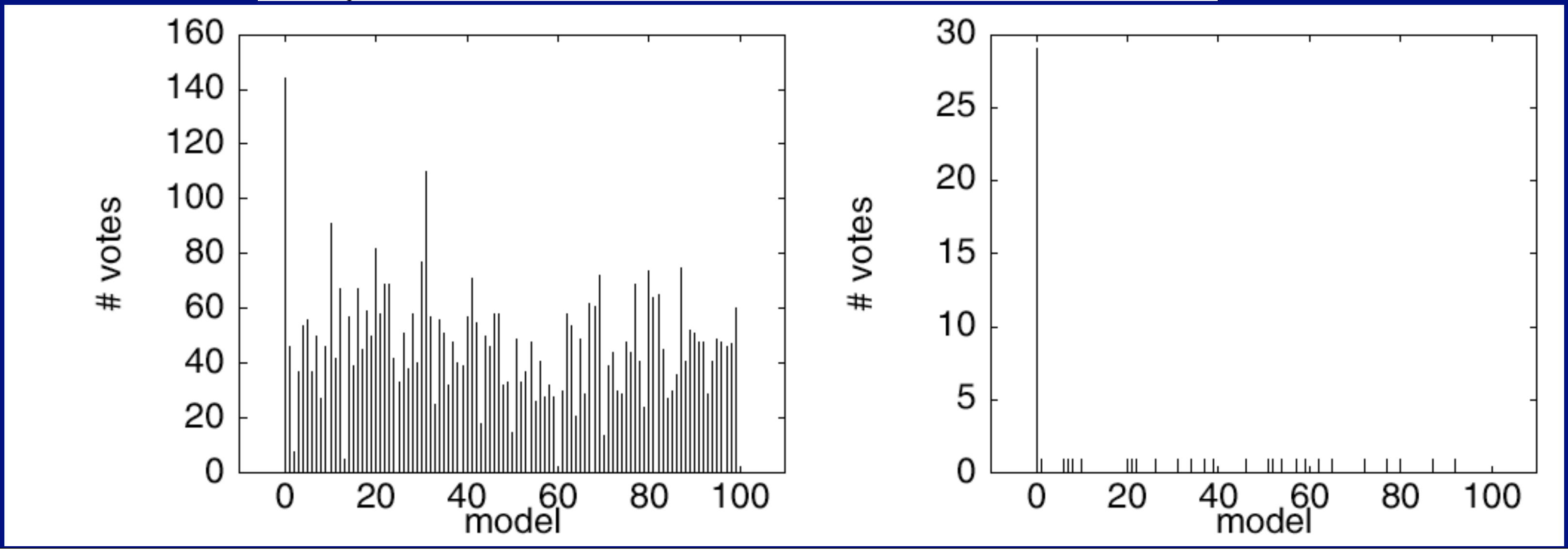


Figure from "Local grayvalue invariants for image retrieval," by C. Schmid and R. Mohr, IEEE Trans. Pattern Analysis and Machine Intelligence, 1997 c 1997, IEEE as used in Forsyth + Ponce, p 609

Employ spatial relations

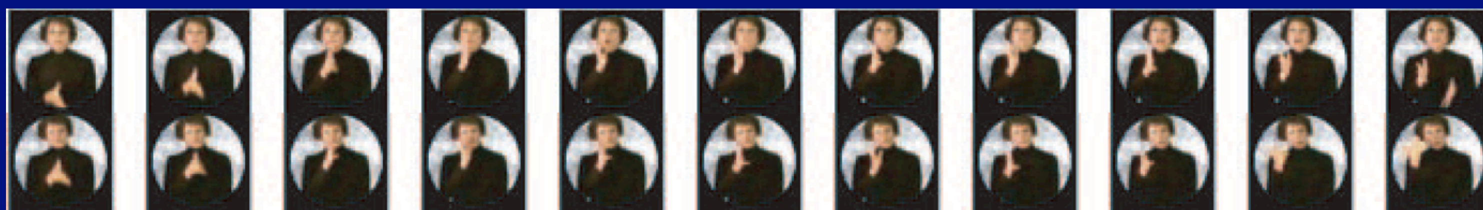
Figure from "Local grayvalue invariants for image retrieval," by C. Schmid and R. Mohr, IEEE Trans. Pattern Analysis and Machine Intelligence, 1997 c 1997, IEEE as used in Forsyth + Ponce, p 609



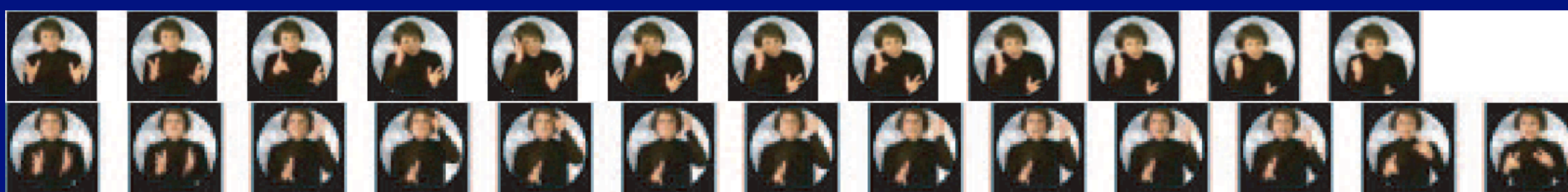


ASL translation

ASL



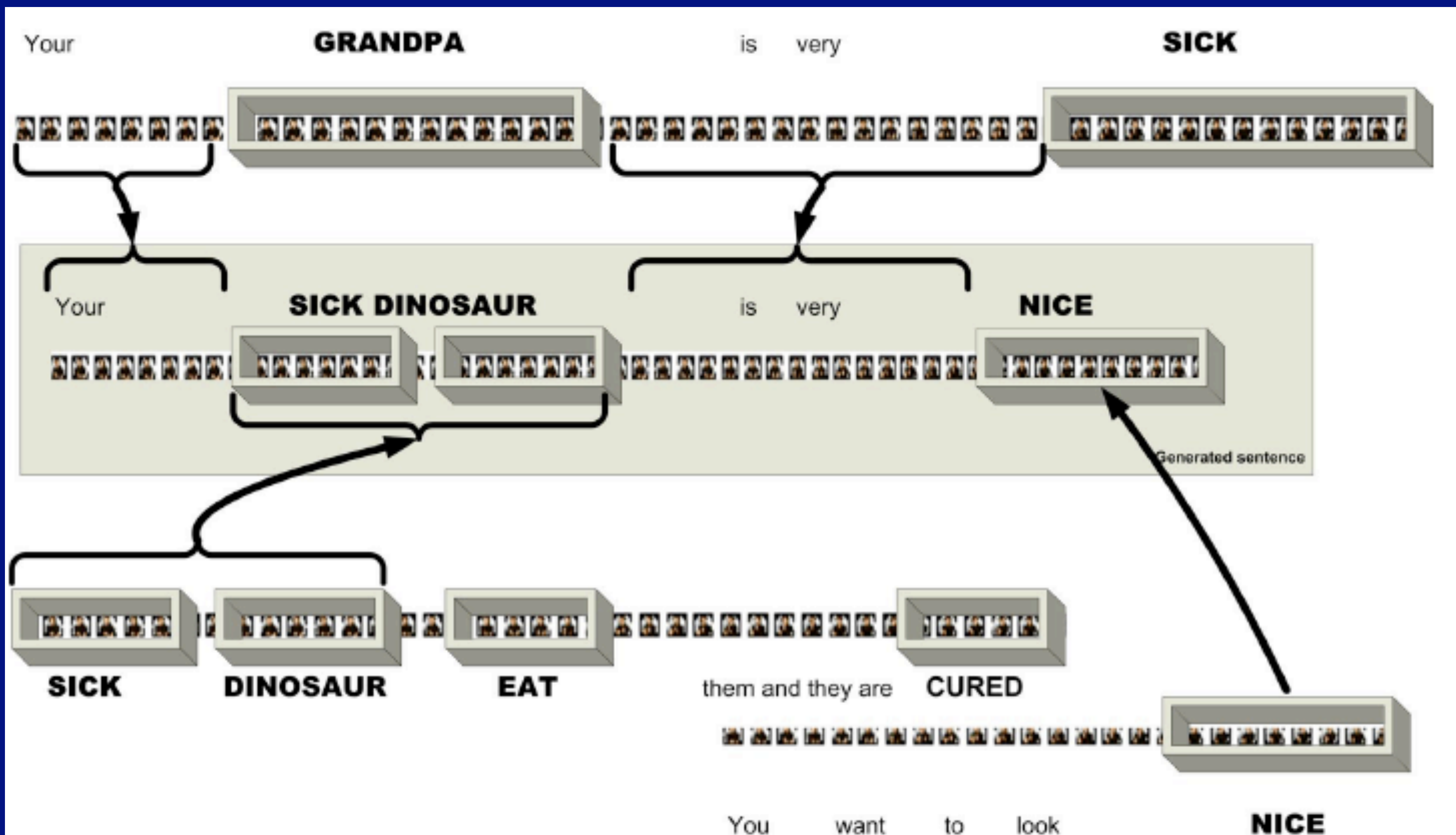
Grandma



Grandpa

- Notice:
  - intra-class variation (timing, role of hands)
  - carefully dressed narrator
  - low resolution
- Would like to spot words in text

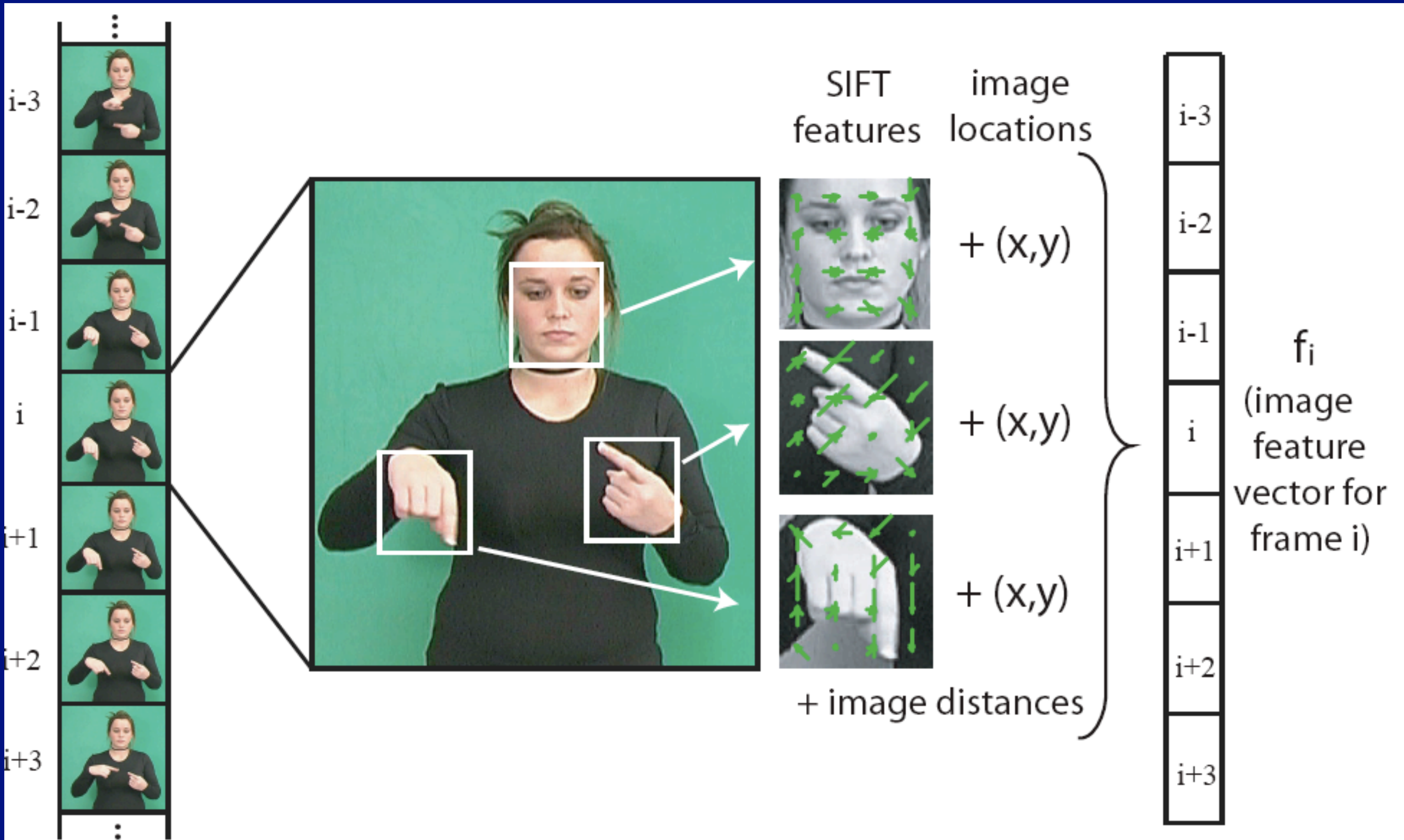
Wordspotting allows translations



Wordspotting

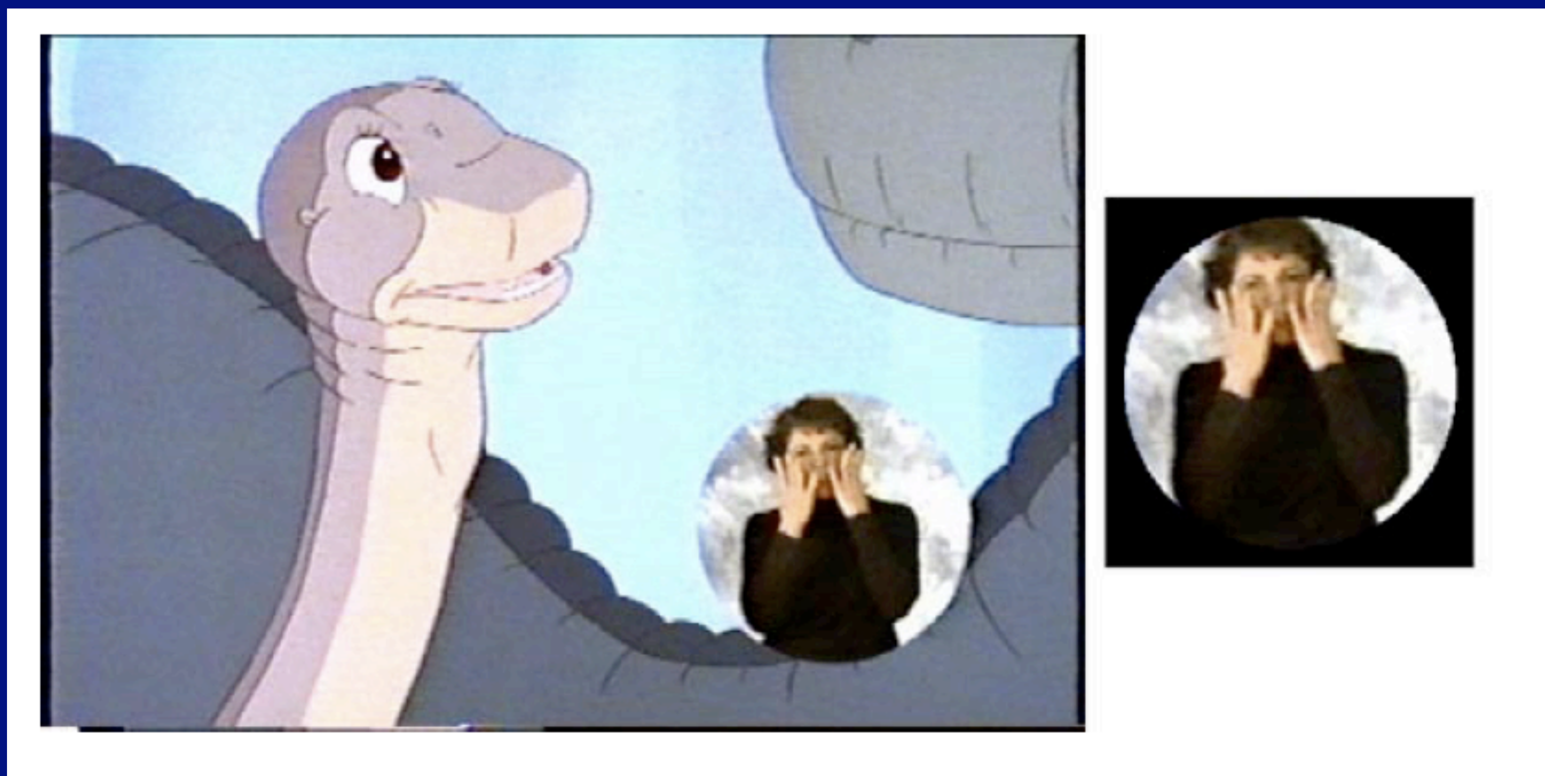
- Difficulties
  - building a discriminative word model for a large vocabulary is hard
    - need lots of examples of each word
  - building a generative word model is hard, too
    - no widely available pronunciation dictionary
  - very large number of features
    - next slide
- Strategy
  - build spotters for some words (base words)
  - now use the output of those spotters as features for other words
    - uses less training data
    - because features are discriminative

Features



i-3
i-2
i-1
i
i+1
i+2
i+3
SIFT features
image locations
+ (x,y)
+ (x,y)
+ (x,y)
+ image distances
i-3
i-2
i-1
i
i+1
i+2
i+3
f_i
(image feature vector for frame i)

ASL translation

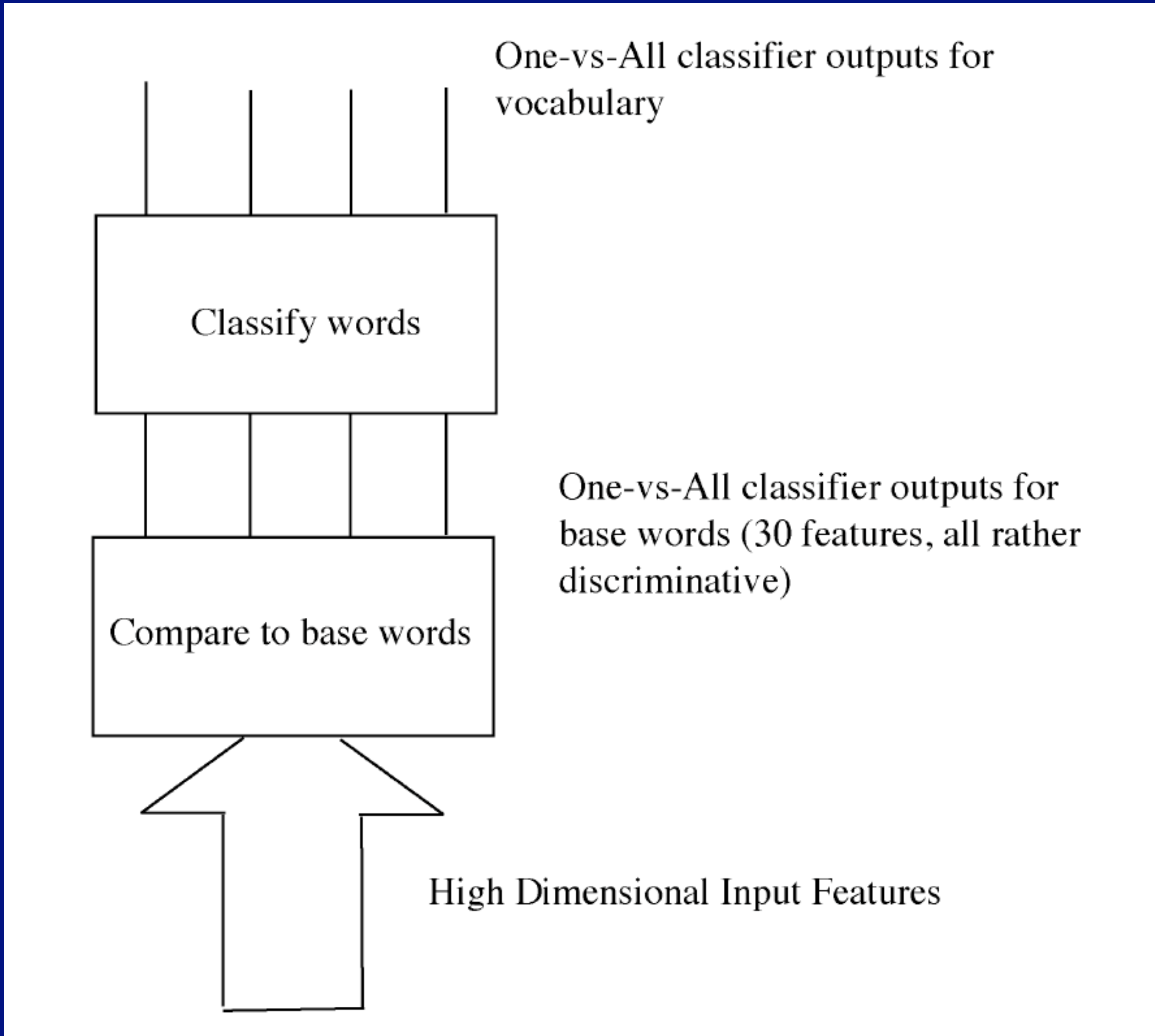


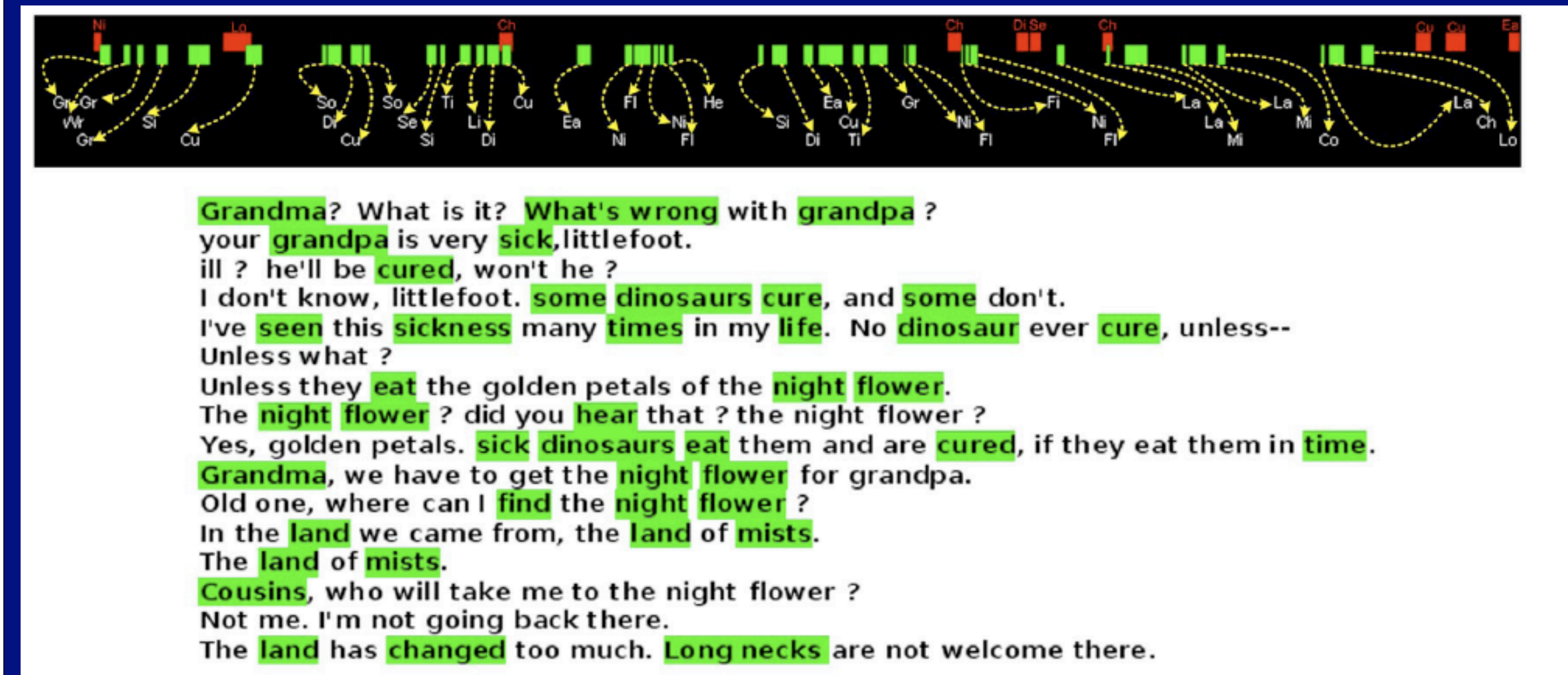
Logistic Regression

- Build a parametric model of the posterior,
  - p(class|information)
- For a 2-class problem, assume that
  - log(P(1|data))-log(P(0|data))=linear expression in data
- Training
  - maximum likelihood on examples
  - problem is convex
- Classifier boundary
  - linear

Multiclass classification

- Many strategies
  - 1-vs-all
    - for each class, construct a two class classifier comparing it to all other classes
    - take the class with best output
      - if output is greater than some value
  - Multiclass
    - log(P(i|features))-log(P(k|features))=(linear expression)
    - many more parameters
    - harder to train with maximum likelihood
    - still convex


One-vs-All classifier outputs for vocabulary
Classify words
One-vs-All classifier outputs for base words (30 features, all rather discriminative)
Compare to base words
High Dimensional Input Features



Forsyth Farhadi 06

Pose consistency

- A match between an image structure and an object structure implies a pose
  - we can vote on poses, objects

From Lowe, 99, **Object Recognition from Local Scale-Invariant Features**

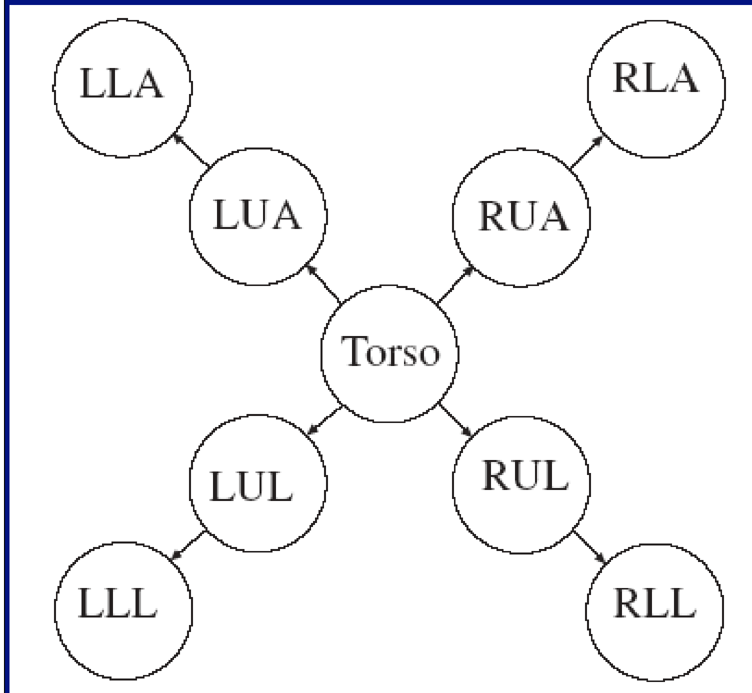
From Lowe, 99, **Object Recognition from Local Scale-Invariant Features**

Kinematic grouping

- Assemble a set of features to present to a classifier
  - which tests
    - appearance
    - configuration
    - whatever
- Classifier could be
  - handwritten rules (e.g. Fleck-Forsyth-Bregler 96)
  - learned classifier (e.g. Ioffe-Forsyth 99)
  - likelihood (e.g. Felzenszwalb-Huttenlocher 00)
  - likelihood ratio test (e.g. Leung-Burl-Perona 95; Fergus-Perona-Zisserman 03)

Pictorial structures

- For models with the right form, one can test "everything"
  - model is a set of cylindrical segments linked into a tree structure
    - model should be thought of as a 2D template
      - segments are cylinders, so no aspect issue there
      - 3D segment kinematics implicitly encoded in 2D relations
      - easy to build in occlusion
  - putative image segments are quantized
  - => dynamic programming to search all matches
  - What to add next? (DP deals with this)
  - Pruning? (Irrelevant)
  - Can one stop?
    - (Use a mixture of tree models, with missing segments marginalized out)
  - Known segment colour - Felzenszwalb-Huttenlocher 00
  - Learned models of colour, layout, texture - Ramanan Forsyth 03, 04



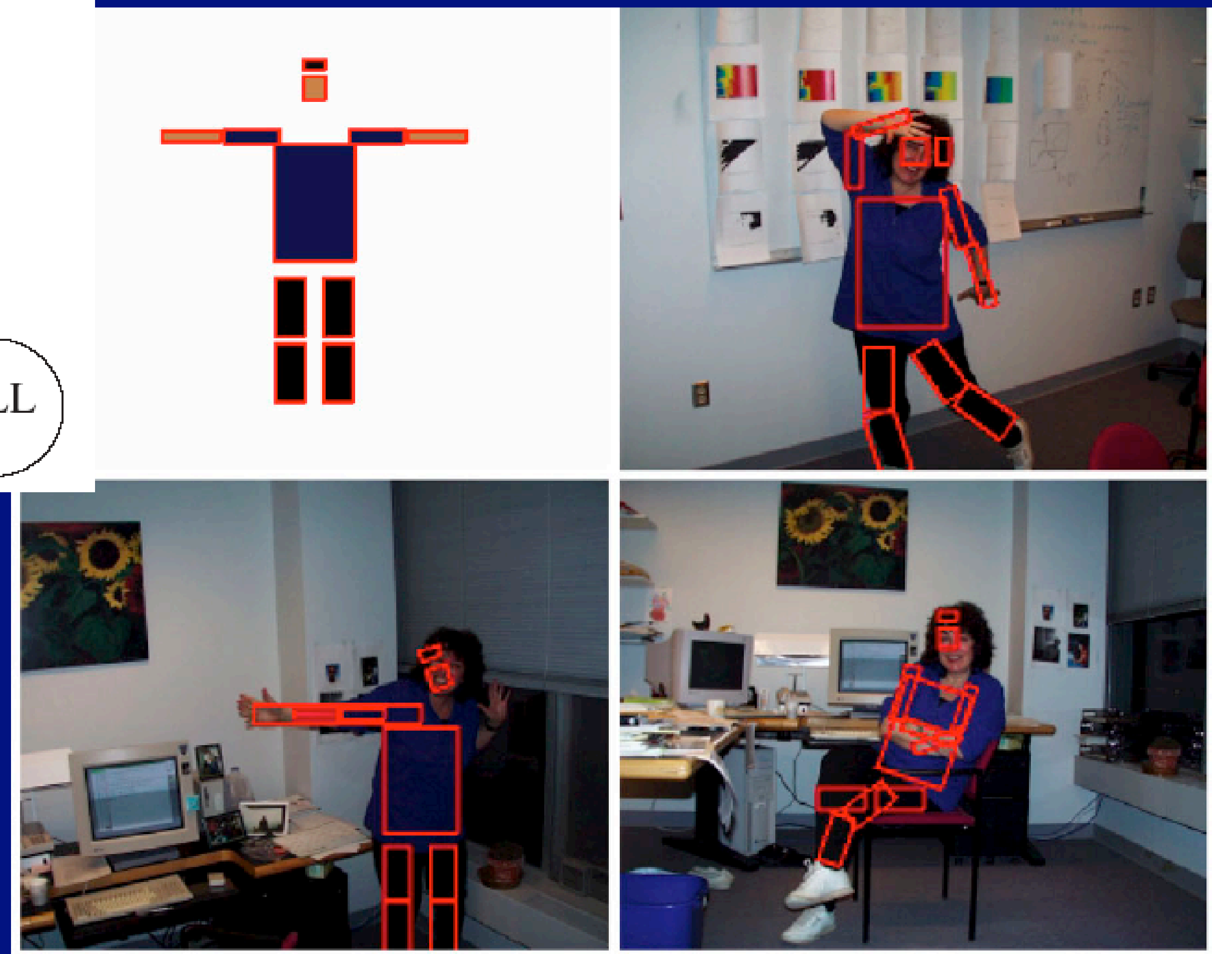
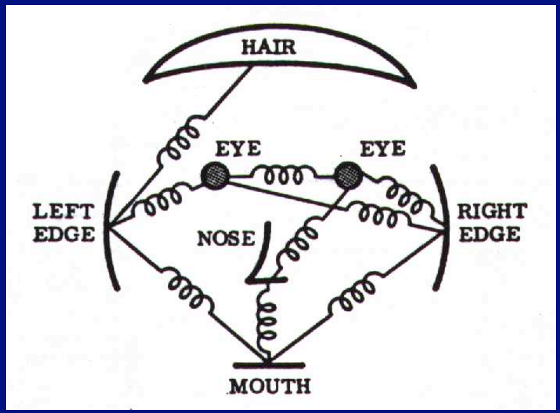
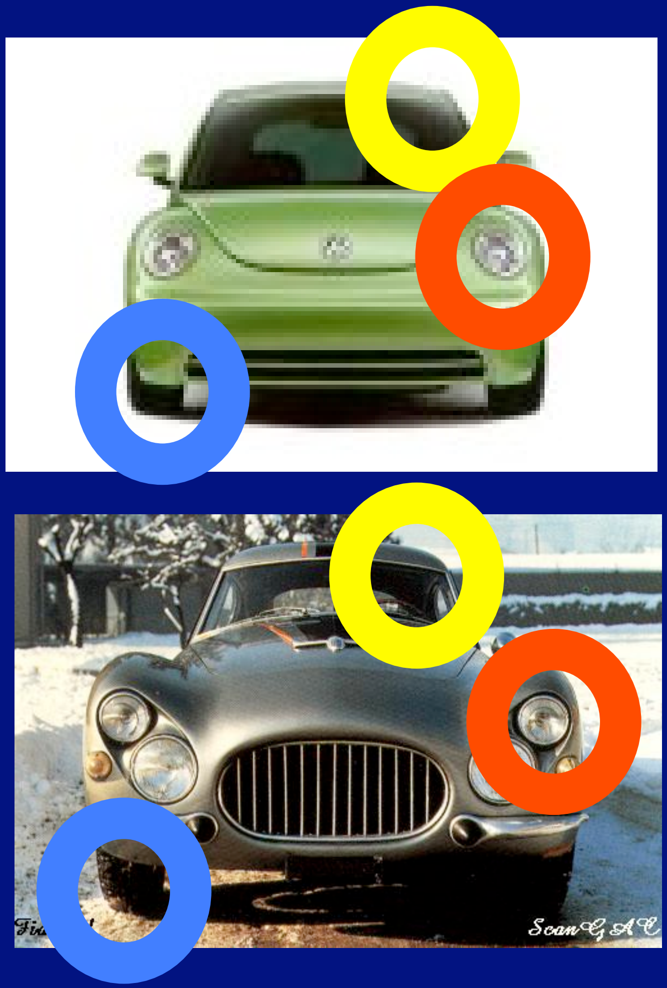


Figure from "Efficient Matching of Pictorial Structures,"
P. Felzenszwalb and D.P. Huttenlocher, Proc. Computer Vision and Pattern Recognition
2000, c 2000, IEEE as used in Forsyth+Ponce, pp 636, 640

Constellations of parts





Fischler & Elschlager 1973

Yuille '91
Brunelli & Poggio '93
Lades, v.d. Malsburg et al. '93
Cootes, Lanitis, Taylor et al. '95
Amit & Geman '95, '99
Perona et al. '95, '96, '98, '00
Agarwal & Roth '02

Generative model for plane templates (Constellation model)

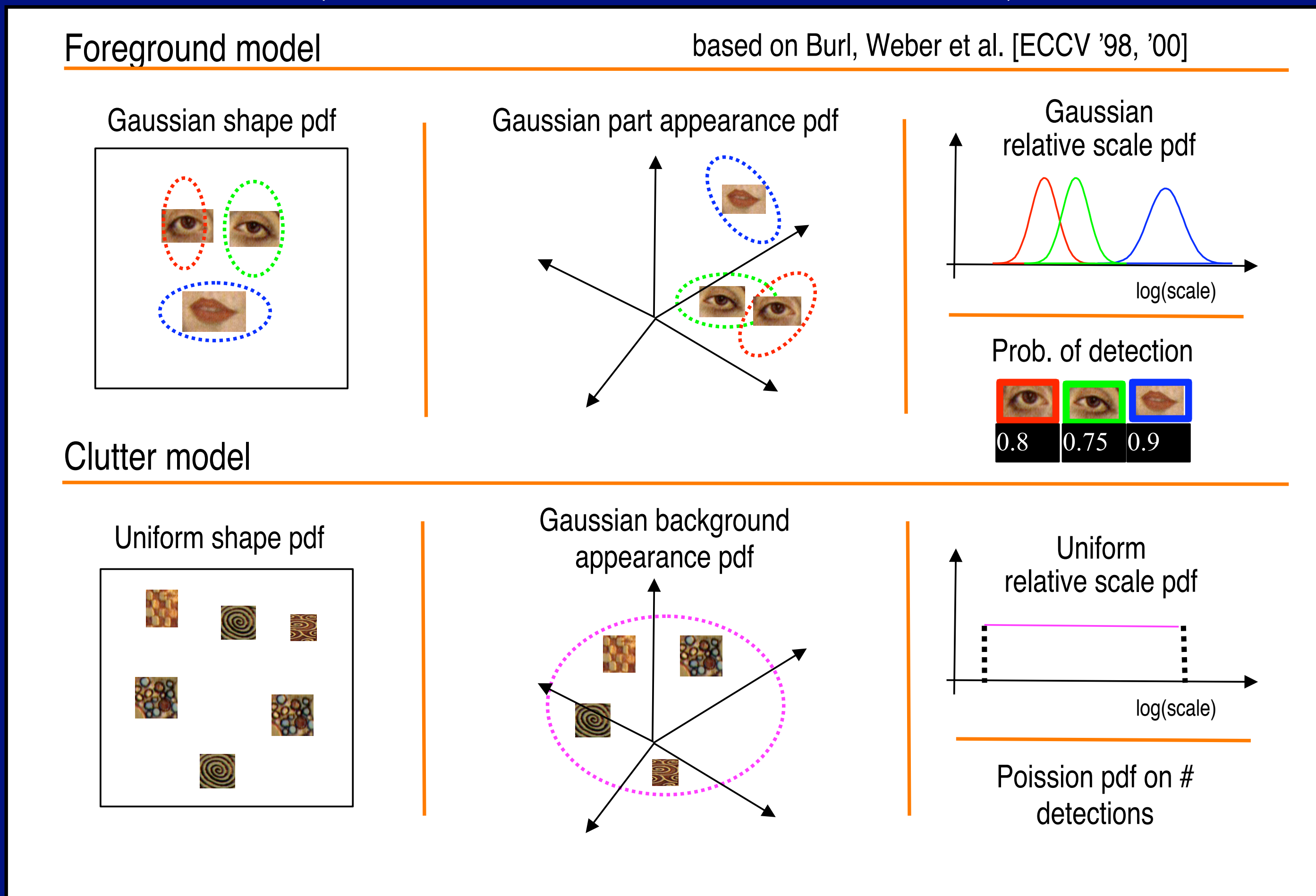
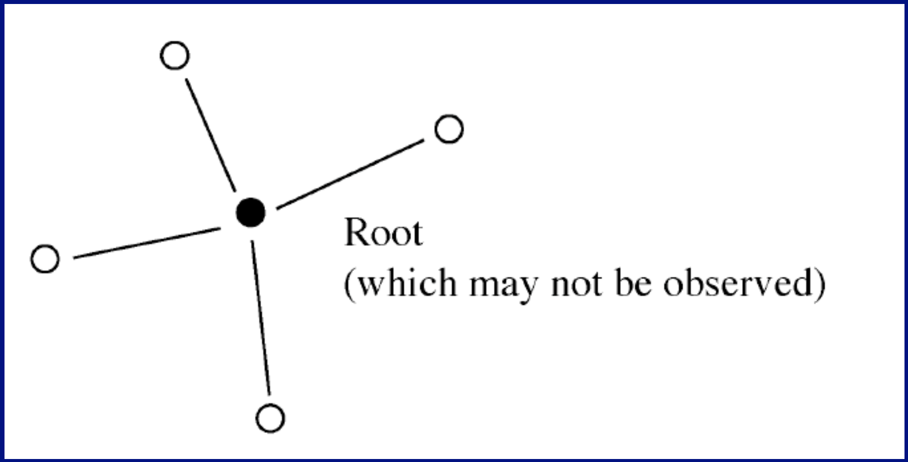


Figure after Fergus et al, 03; see also Fergus et al, 04

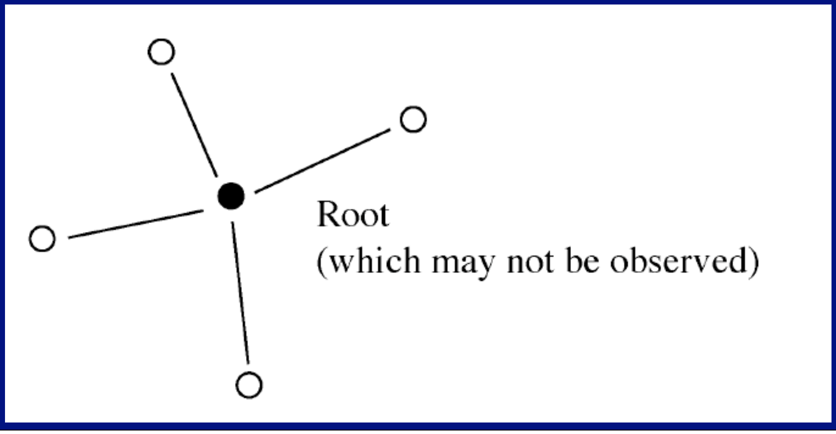
Star-shaped models

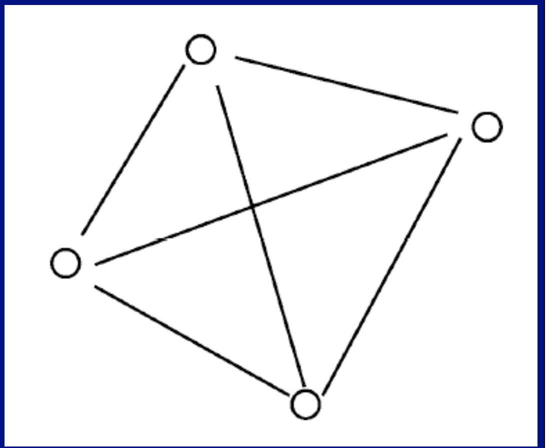
- Features generated at parts
  - at image points that are conditionally independent given part location
  - with appearance that is conditionally independent given part type
- Part locations are conditioned on root
- Easy to deal with
  - very like a pictorial structure
  - inference is dynamic programming
  - localization easy
  -



Other types of model

- We've already seen a tree-structured model!
  - (pictorial structure)
- Complete models are much more difficult to work with
  - because there is no conditional independence
  - means fewer features



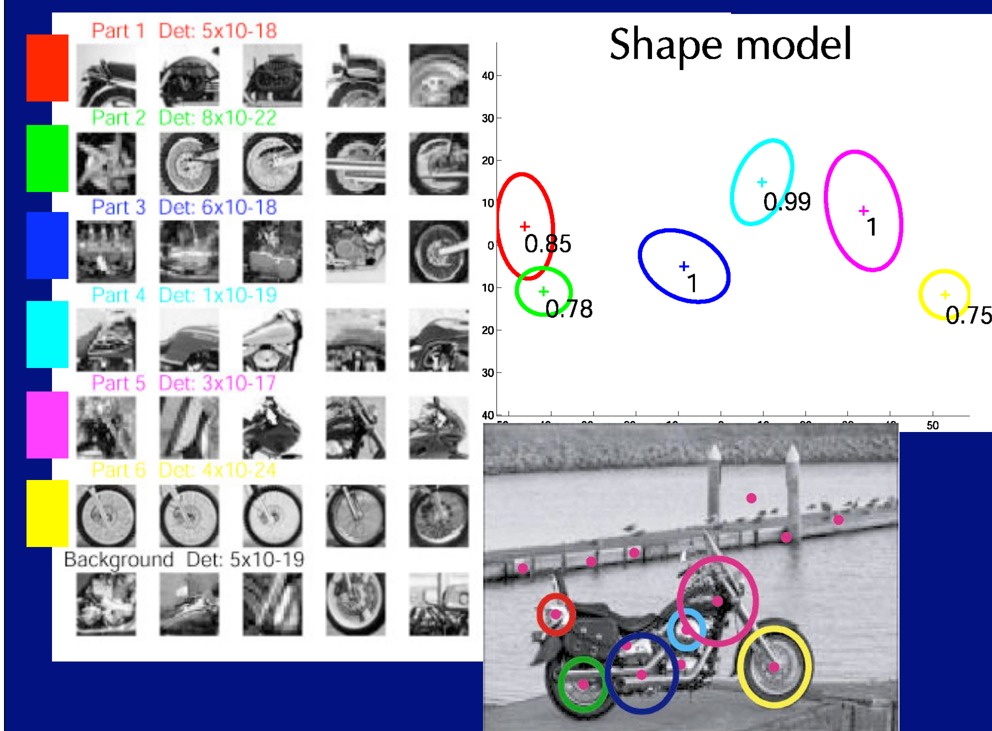


Constellation models

- Learning model
  - on data set consisting of instances, not manually segmented
  - choose number of features in model
  - run point feature detector
  - each response is from either one “slot” in the model, or bg
    - this known, easy to estimate parameters
    - parameters known, this is easy to estimate
  - missing variable problem -> EM
- Detecting instance
  - search for allocation of feature instances to slots that maximizes likelihood ratio
  - detect with likelihood ratio test

Typical models

Motorbikes



Spotted cats

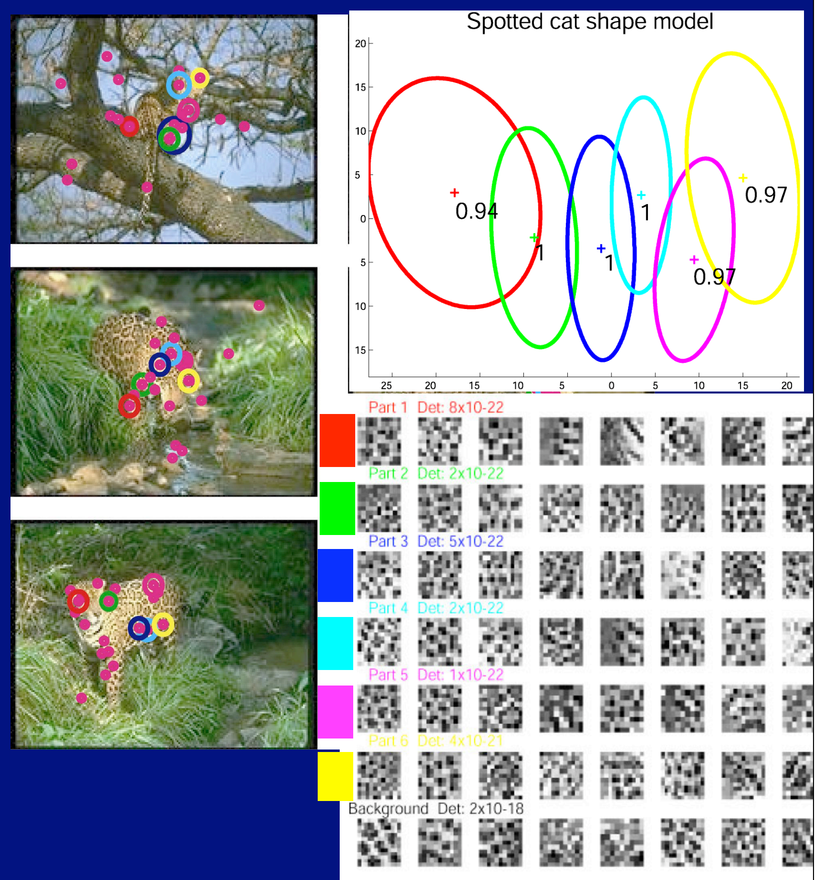


Figure after Fergus et al, 03; see also Fergus et al, 04

Summary of results

Dataset	Fixed scale experiment	Scale invariant experiment
Motorbikes	7.5	6.7
Faces	4.6	4.6
Airplanes	9.8	7.0
Cars (Rear)	15.2	9.7
Spotted cats	10.0	10.0

% equal error rate

Note: Within each series, same settings used for all datasets

Figure after Fergus et al, 03; see also Fergus et al, 04

Caution: dataset is known to have some quirky features

LOTS of BIG collections of images

Corel Image Data	40,000 images
Fine Arts Museum of San Francisco	83,000 images online
Cal-flora	20,000 images, species information
News photos with captions (yahoo.com)	1,500 images per day available from yahoo.com
Hulton Archive	40,000,000 images (only 230,000 online)
internet.archive.org	1,000 movies with no copyright
TV news archives (televisionarchive.org, informedia.cs.cmu.edu)	Several terabytes already available
Google Image Crawl	>330,000,000 images (with nearby text)
Satellite images (terrarserver.com, nasa.gov, usgs.gov)	(And associated demographic information)
Medial images	(And associated with clinical information)

* and the BBC is releasing its video archive, too;
and we collected 500,000 captioned news images;
and it’s easy to get scanned mediaeval manuscripts;
etc., etc.,

Imposing order

- Iconic matching
  - child abuse prosecution
  - managing copyright (BayTSP)
- Clustering
  - Browsing for:
    - web presence for museums (Barnard et al, 01)
    - home picture, video collections
    - selling pictures
- Searching
  - scanned writing (Manmatha, 02)
  - collections of insects
- Building world knowledge
  - a face gazetteer (Miller et al, 04)

Current, practical applications

Maybe applications

Maybe applications

Search is well studied

- Metadata indexing
  - keywords, date of photo, place, etc.
- Content based retrieval
  - query by example with
    - global features
      - (e.g. Flickner et al. 95, Carson et al. 99, Wang 00, various entire conferences)
    - local features
      - (e.g. Photobook - Pentland et al 96; Blobworld - Carson et al, 98)
    - relevance feedback
      - (e.g. Cox et al 00; Santini 00; Schettini 02; etc.)
  - query by class
    - naughty pictures
      - (eg Forsyth et al. 96, 99; Wang et al. 98; Chan et al 99)

What will users pay for?

- Work by Peter Enser and colleagues on the use of photø movie collections
  (Enser McGregor 92; Ornager 96; Armitage Enser 97; Markkula Sormunen 00; Frost et al 00;   Enser 00)
- Typical queries:

What is this about?

"… smoking of kippers…"

"The depiction of vanity in painting, the depiction of the female figure looking in the mirror, etc."

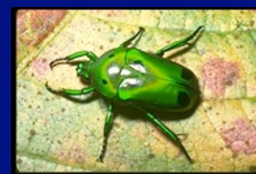
"Cheetahs running on a greyhound course in Haringey in 1932"

Annotation results in complementary words and pictures

Query on

“Rose”

Example from Berkeley
Blobworld system







Annotation results in complementary words and pictures

Query on

Example from Berkeley
Blobworld system

Annotation results in complementary words and pictures

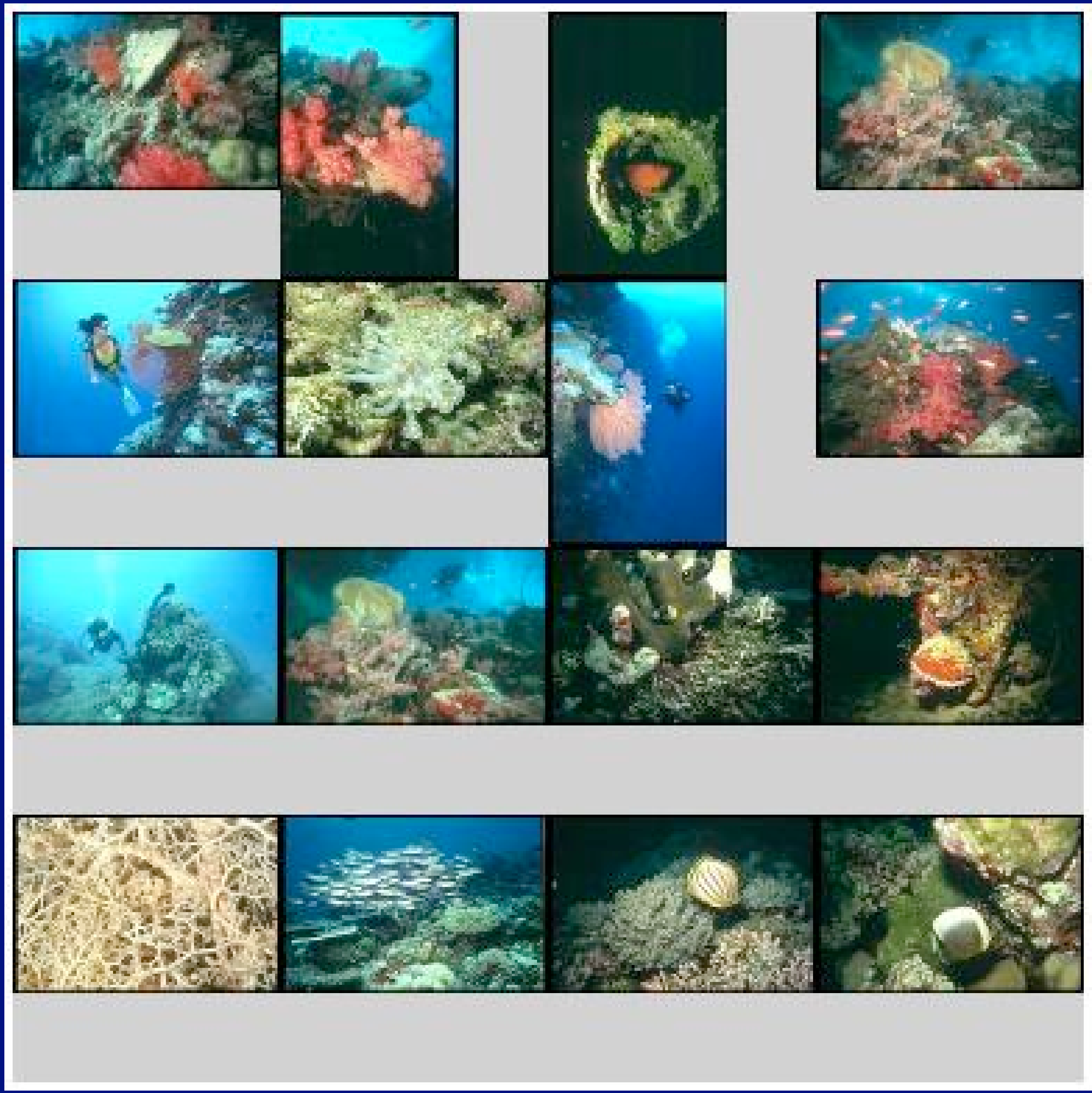
Query on

“Rose”

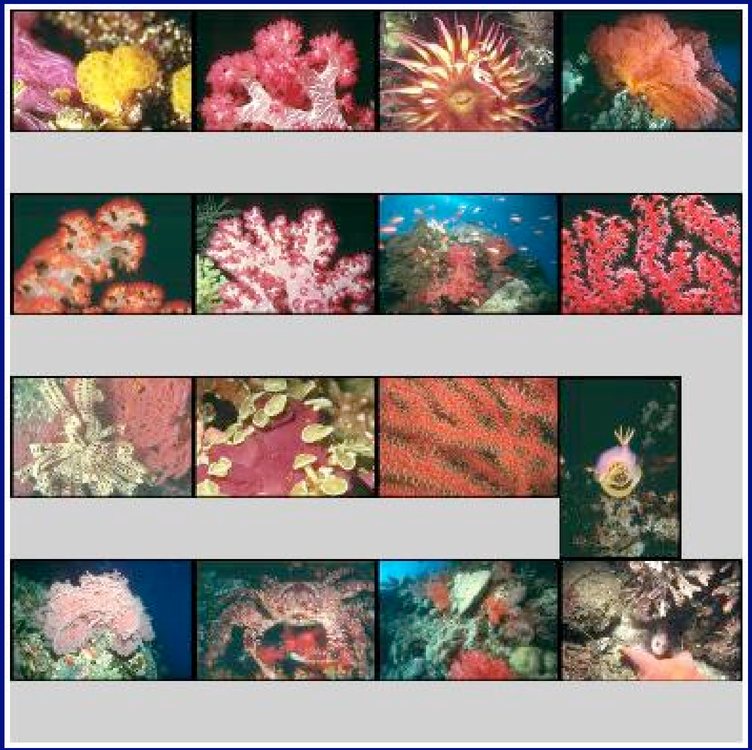
and

Example from Berkeley
Blobworld system

Cluster found using only text



Cluster found using only blob features



Adjacent clusters found using both text and blob features

Input

"This is a picture of the sun setting over the sea with waves in the foreground"

Image processing*

Each blob is a large vector of features

- Region size
- Position
- Colour
- Oriented energy (12 filters)
- Simple shape features

Language processing

sun sky waves sea

* Thanks to Blobworld team [Carson, Belongie, Greenspan, Malik], N-cuts team [Shi, Tal, Malik]

FAMSF Data

Web number: 4359202410830012

rec number: 2

Title: Le Matin

Primary class: Print

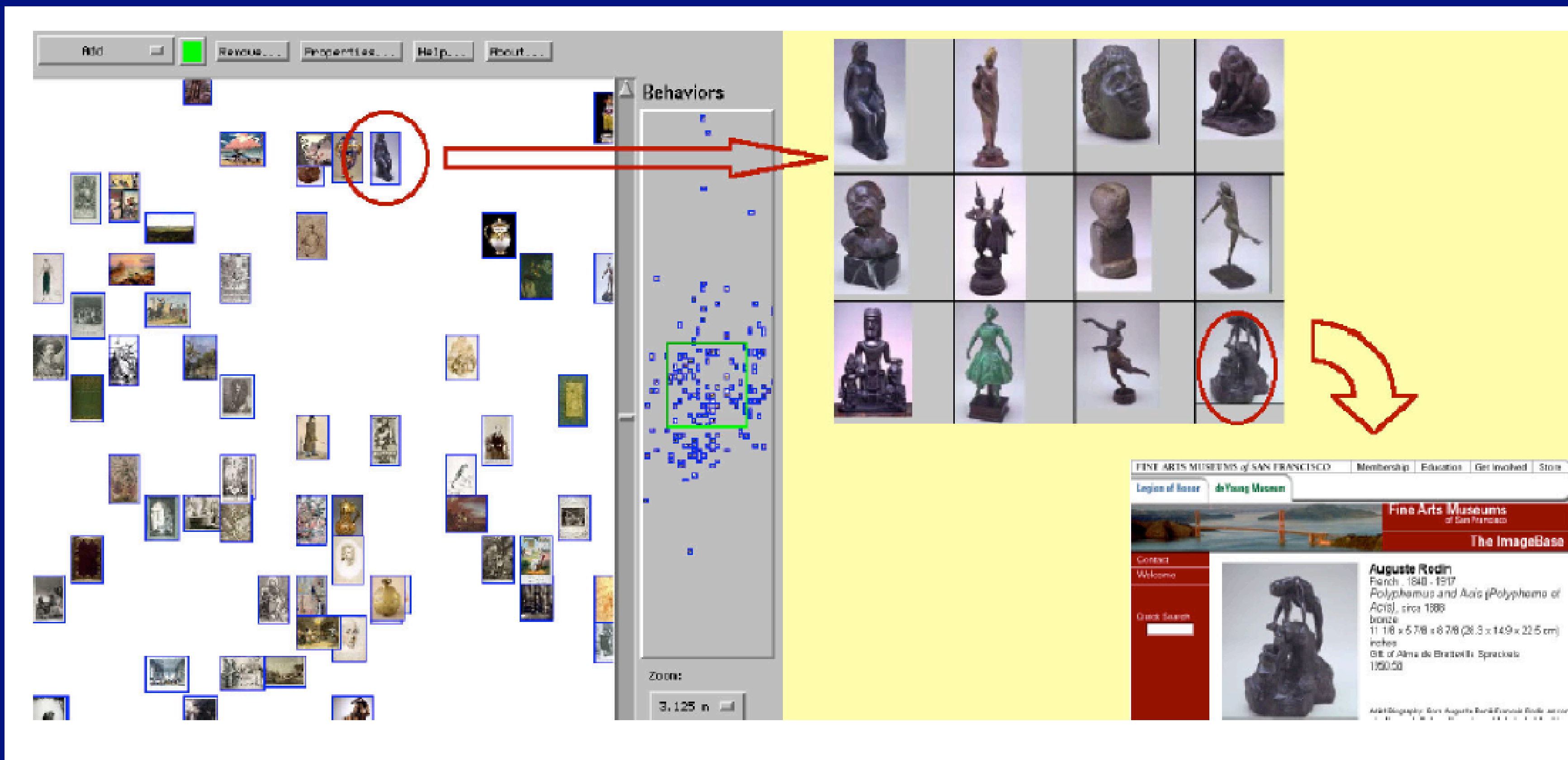
Artist: Tissot

Description:
serving woman stands in a dressing room, in front of vanity with chair, mirror and mantle, holding a tray with tea and toast

Display date: 1886

Country: France

83,000 images online, we clustered 8000


Behaviors
Zoom:
Fine Arts Museums
The ImageBase
Auguste Rodin

Searching

Compute P(document | query_items)

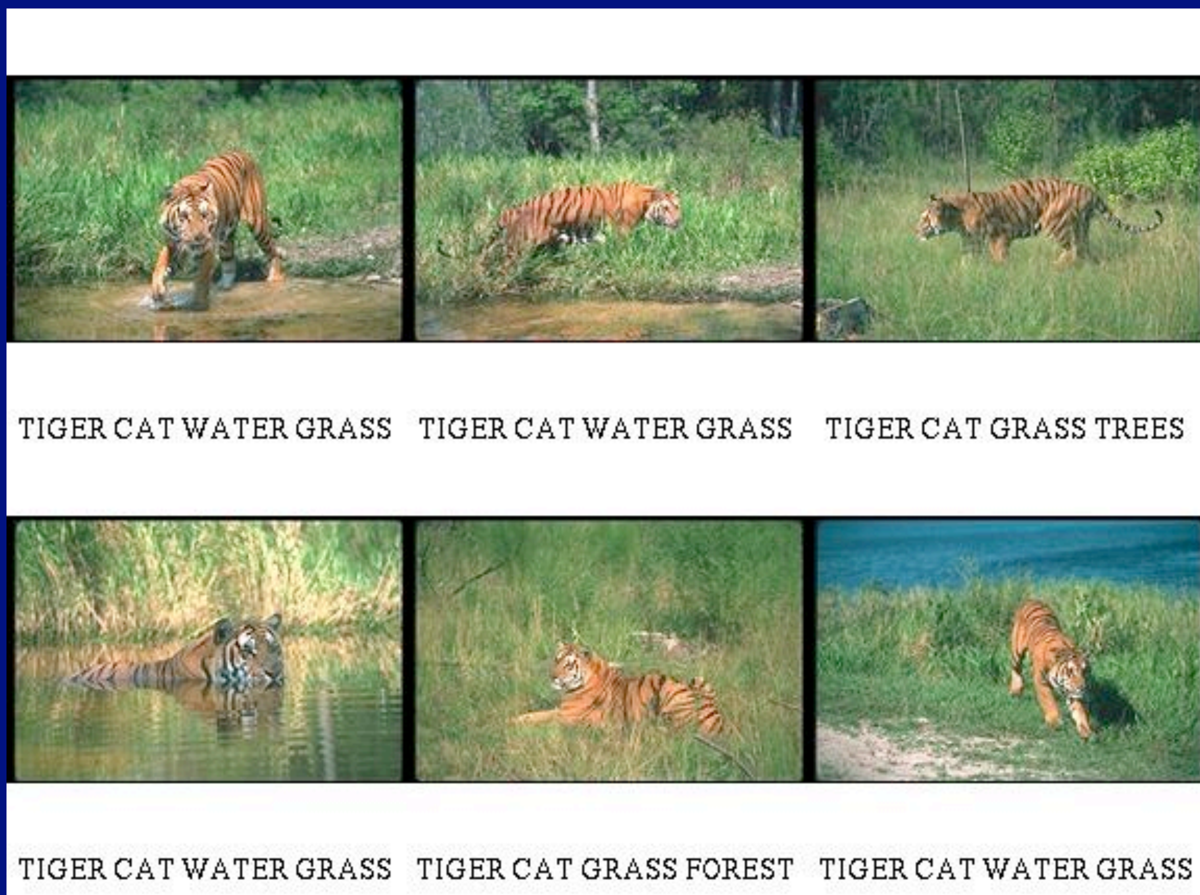
query_items can be words, features, or both

Natural way to express "soft queries"

Related retrieval work: Cascia, Sethi, and Sclaroff, 98; Berger and Lafferty, 98; Papadimitriou et al., 98

Query: “river tiger” from 5,000 Corel images
(The words never occur together.)

Retrieved items: rank order P(document | query)



Search

Compute P(document | query_items)

Related retrieval work: Cascia, Sethi, and Sclaroff, 98; Berger and Lafferty, 98; Papadimitriou et al., 98

Query: "water sky cloud  "

Retrieved items:

Pictures from Words (Auto-illustration)

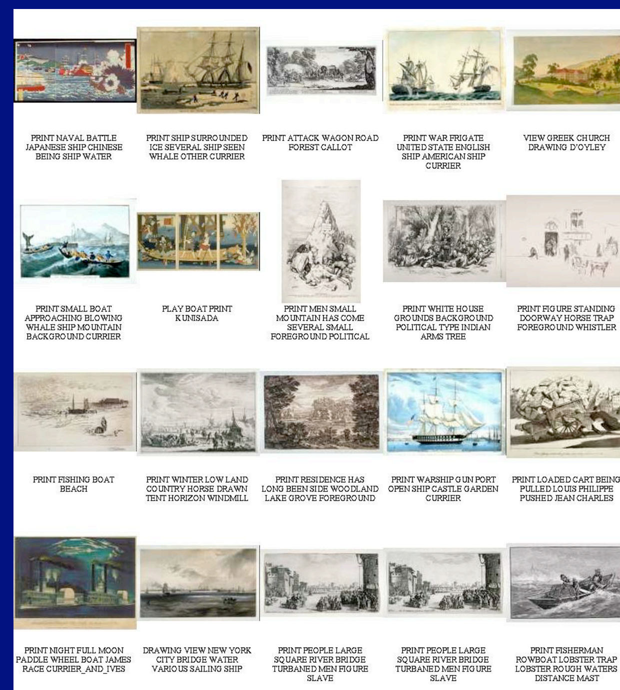
Text Passage (Moby Dick)

"The large importance attached to the harpooneer's vocation is evinced by the fact, that originally in the old Dutch Fishery, two centuries and more ago, the command of a whale-ship …"

Extracted Query

large importance attached fact old dutch century more command whale ship was person was divided officer word means fat cutter time made days was general vessel whale hunting concern british title old dutch ...

Retrieved Images



PRINT NAVAL BATTLE JAPANESE SHIP CHINESE BEING SHIP WATER

PRINT SHIP SURROUNDED ICE SEVERAL SHIP SEEN WHALE OTHER CURRIER

PRINT ATTACK WAGON ROAD FOREST CALLOT

PRINT WAR FRIGATE UNITED STATE ENGLISH SHIP AMERICAN SHIP CURRIER

PRINT SMALL BOAT APPROACHING BLOWING WHALE SHIP MOUNTAIN BACKGROUND CURRIER

PLAY BOAT PRINT KUNISADA

PRINT MEN SMALL MOUNTAIN HAS COME SEVERAL SMALL FOREGROUND POLITICAL

PRINT WHITE HOUSE GROUNDS BACKGROUND POLITICAL TYPE INDIAN ARMS TREE

Auto-annotation

- Predict words from pictures
  - Obstacle:
    - Hoffman's model uses document specific level probabilities
  - Dodge
    - smooth these empirically

- Attractions:
  - easy to score
  - large scale performance measures (how good is the segmenter?)
  - possibly simplify retrieval (Li+Wang, 03)


Keywords
GRASS TIGER CAT FOREST
Predicted Words (rank order)

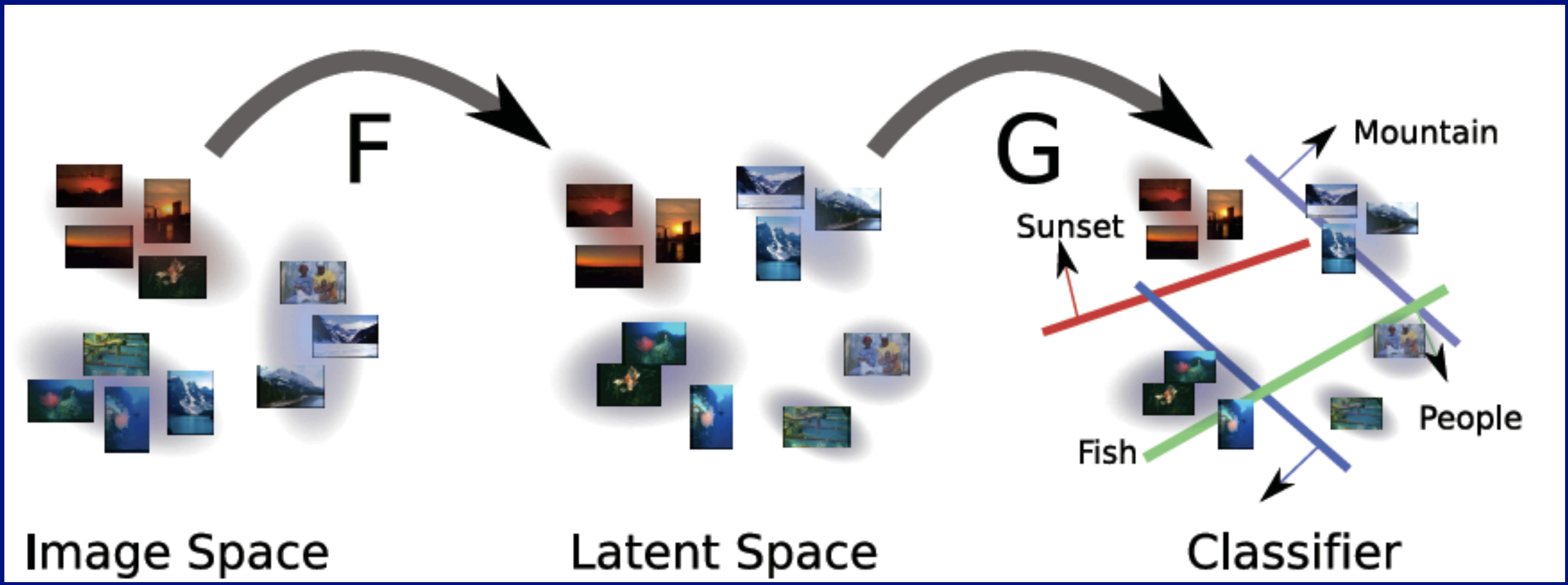
tiger cat grass people water bengal
buildings ocean forest reef


Keywords
HIPPO BULL mouth walk
Predicted Words (rank order)

water hippos rhino river grass
reflection one-horned head
plain sand


Keywords
FLOWER coralberry LEAVES
PLANT

Predicted Words (rank order)
fish reef church wall people water
landscape coral sand trees



Train a system of svm classifiers, one per word but penalize that matrix for rank, after Rennie+Srebro 05

The latent space reveals scenes because it is good at word prediction and takes appearance into account

Loeff Farhadi Forsyth??

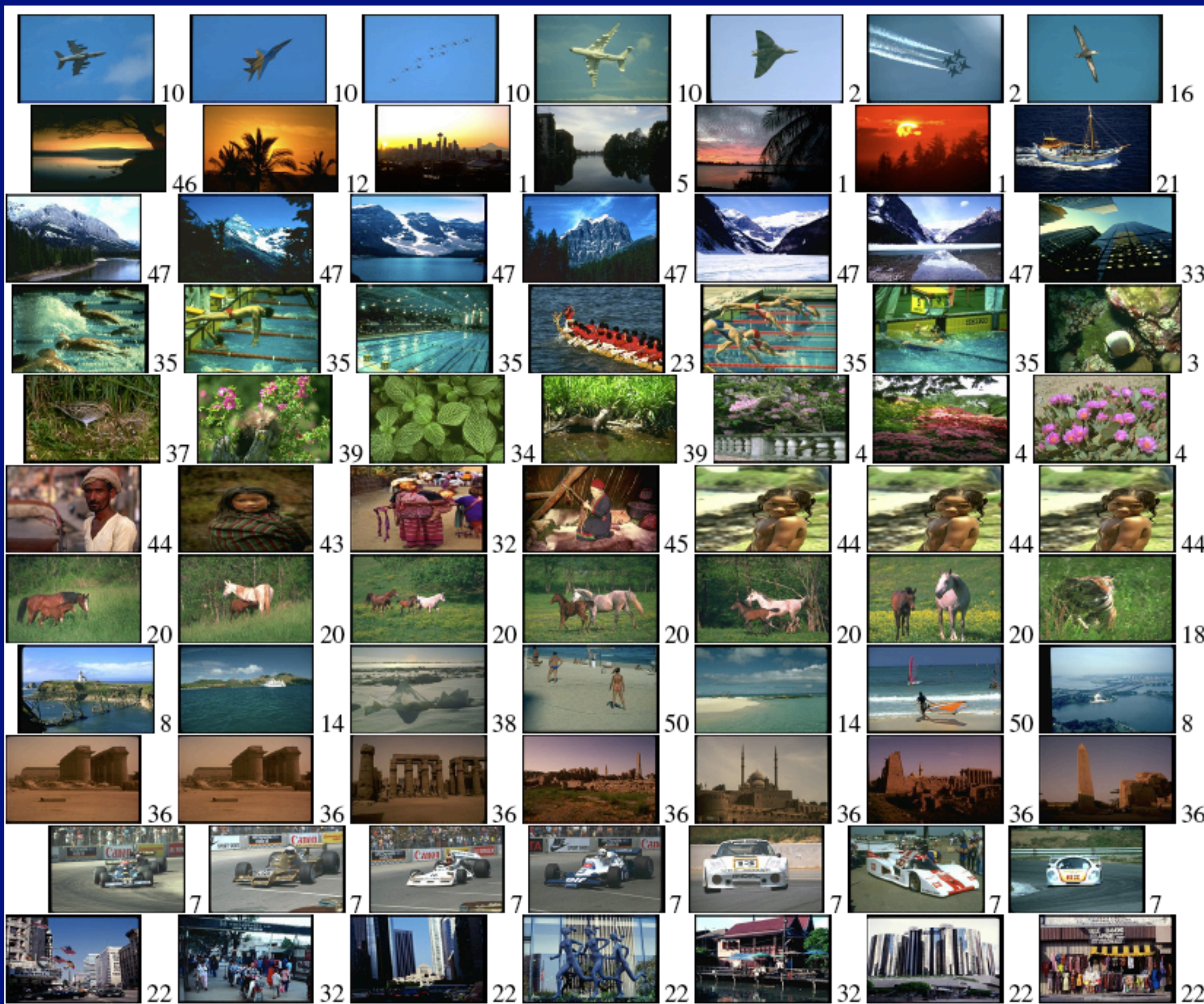
It was there and we didn't

It was there and we predicted it

It wasn't and we did

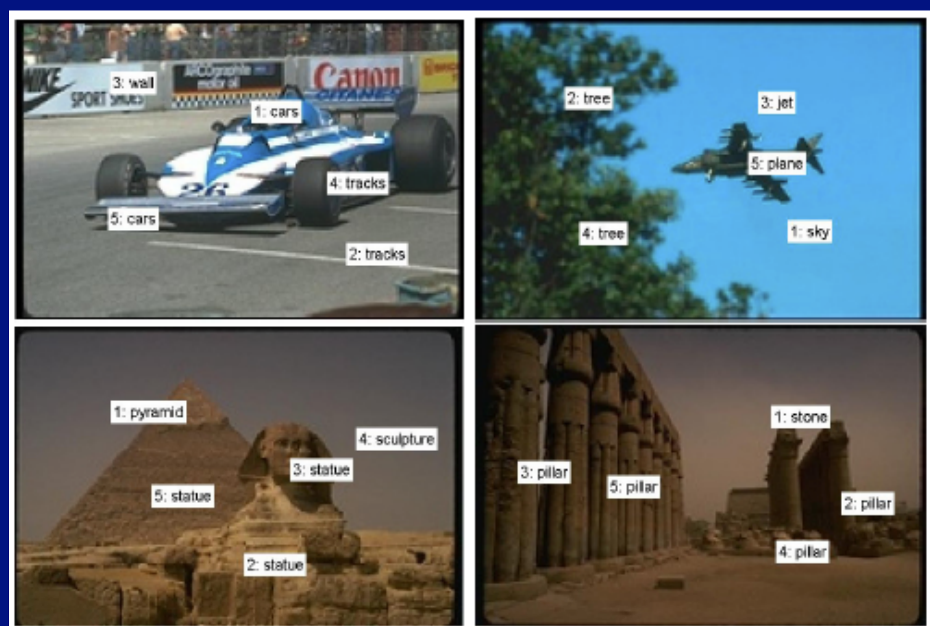
Loeff Farhadi Forsyth??

Scene



CD #
(rough proxy)

Loeff Farhadi Forsyth??



Loeff Farhadi Forsyth??

News dataset

- Approx 5e5 news images, with captions
  - Easily collected by script from Yahoo over the last 18 months or so
- Mainly people
  - politicians, actors, sportsplayers
  - long, long tails distribution
- Face pictures captured “in the wild”
- Correspondence problem
  - some images have many (resp. few) faces, few (resp. many) names (cf. Srihari 95)

President George W. Bush makes a statement in the Rose Garden while Secretary of Defense Donald Rumsfeld looks on, July 23, 2003. Rumsfeld said the United States would release graphic photographs of the dead sons of Saddam Hussein to prove they were killed by American troops. Photo by Larry Downing/Reuters

President George W. Bush makes a statement in the Rose Garden while Secretary of Defense Donald Rumsfeld looks on, July 23, 2003. Rumsfeld said the United States would release graphic photographs of the dead sons of Saddam Hussein to prove they were killed by American troops. Photo by Larry Downing/ Reuters

Data examples

Doctor Nikola shows a fork that was removed from an Israeli woman who swallowed it while trying to catch a bug that flew in to her mouth, in Poriah Hospital northern Israel July 10, 2003. Doctors performed emergency surgery and removed the fork. (Reuters)

President George W. Bush waves as he leaves the White House for a day trip to North Carolina, July 25, 2002. A White House spokesman said that Bush would be compelled to veto Senate legislation creating a new department of homeland security unless changes are made. (Kevin Lamarque/ Reuters)

Process

- Extract proper names
  - rather crudely, at present
- Detect faces
  - with Cordelia Schmid's face detector, (Vogelhuber Schmid 00)
- Rectify faces
  - by finding eye, nose, mouth patches, affine transformation
- Kernel PCA rectified faces
- Estimate linear discriminants
- Now have (face vector; name_1,...., name_k)

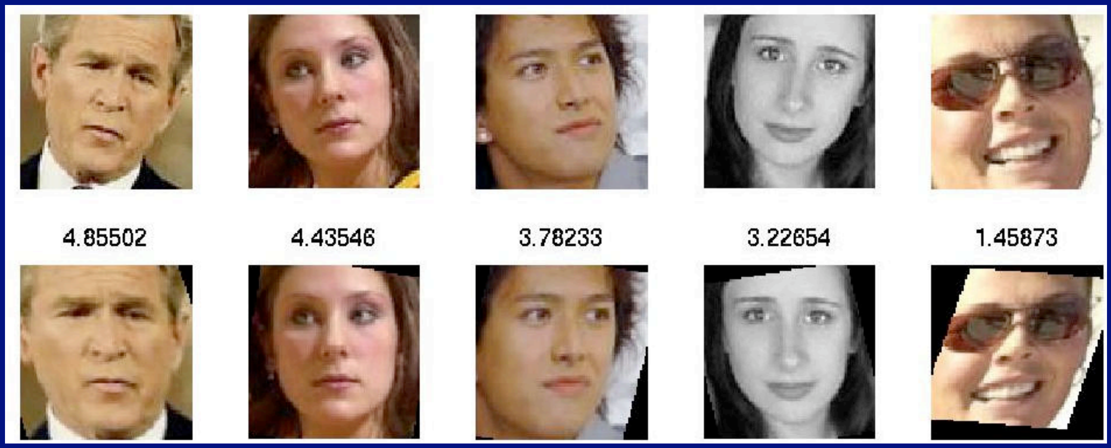
Scale

44773 big face responses

34623 properly rectified

27742 for k<=4

Rectification



- SVM based local feature detectors, affine transformation

Kernel PCA

- Find principal components - high variance features - in an abstract feature space by computing eigenvectors of the kernel matrix
- Problems:
  - kernel matrix dimension (number of data items, approx 40, 000)
  - can't compute all, many kernel matrix entries
  - incomplete cholesky no help
    - (column operations involve touching every face)
  - Nystrom approximation works, because matrix has very low rank

NxN

mxm

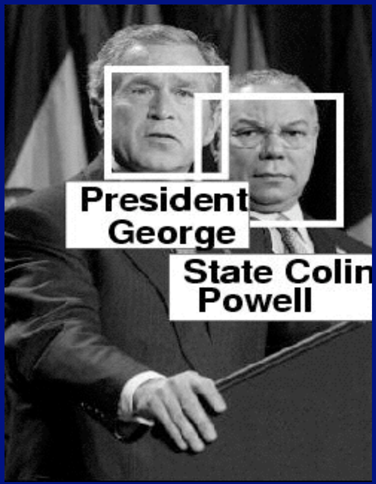
$$K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

Approximate C as:

$$C = B^T A^{-1} B$$

Building a face dictionary

- Compute linear discriminants
  - using single name, single face data items
  - we now have a set of clusters
- Now break correspondence with modified k-means
  - assign face to cluster with closest center,
    - chosen from associated names
  - recompute centers, iterate
  - using distance in LD space
- Now recompute discriminants, recluster with modified k-means



US President George W. Bush (L) makes remarks while Secretary of State Colin Powell (R) listens before signing the US Leadership Against HIV /AIDS , Tuberculosis and Malaria Act of 2003 at the Department of State in Washington, DC. The five-year plan is designed to help prevent and treat AIDS, especially in more than a dozen African and Caribbean nations(AFP/Luke Frazza)

German supermodel Claudia Schiffer gave birth to a baby boy by Caesarian section January 30, 2003, her spokeswoman said. The baby is the first child for both Schiffer, 32, and her husband, British film producer Matthew Vaughn, who was at her side for the birth. Schiffer is seen on the German television show 'Bet It...?!' ('Wetten Dass...?!') in Braunschweig, on January 26, 2002. (Alexandra Winkler/Reuters)

British director Sam Mendes and his partner actress Kate Winslet arrive at the London premiere of 'The Road to Perdition', September 18, 2002. The films stars Tom Hanks as a Chicago hit man who has a separate family life and co-stars Paul Newman and Jude Law. REUTERS/Dan Chung

and yields a useful little NLP module, too

IN Pete Sampras IN of the U.S. celebrates his victory over Denmark's **OUT Kristian Pless OUT** at the **OUT U.S. Open OUT** at Flushing Meadows August 30, 2002. Sampras won the match 6-3 7- 5 6-4. REUTERS/Kevin Lamarque
Germany's **IN Chancellor Gerhard Schroeder IN**, left, in discussion with France's **IN President Jacques Chirac IN** on the second day of the EU summit at the European Council headquarters in Brussels, Friday Oct. 25, 2002. EU leaders are to close a deal Friday on finalizing entry talks with 10 candidate countries after a surprise breakthrough agreement on Thursday between France and Germany regarding farm spending.(AP Photo/European Commission/HO)
'The Right Stuff' cast members **IN Pamela Reed IN**, (L) poses with fellow cast member **IN Veronica Cartwright IN** at the 20th anniversary of the film in Hollywood, June 9, 2003. The women played wives of astronauts in the film about early United States test pilots and the space program. The film directed by **OUT Philip Kaufman OUT**, is celebrating its 20th anniversary and is being released on DVD. REUTERS/Fred Prouser
Kraft Foods Inc., the largest U.S. food company, on July 1, 2003 said it would take steps, like capping portion sizes and providing more nutrition information, as it and other companies face growing concern and even lawsuits due to rising obesity rates. In May of this year, San Francisco attorney **OUT Stephen Joseph OUT**, shown above, sought to ban Oreo cookies in California – a suit that was withdrawn less than two weeks later. Photo by Tim Wimborne/Reuters REUTERS/Tim Wimborne

Classifier	labels correct	IN correct	OUT correct
Baseline	67%	100%	0%
EM Labeling with Language Model	76%	95%	56%
MM Labeling with Language Model	84%	87%	76%

successful of all the invertebrate groups. They have exploited an incredible array of habitats and because of their small size (some are truly microscopic) most go totally unnoticed. Many live freely in the soil, but there is also a vast array of species that live as parasites on plants or animals. This SEM image of a Peacock mite shows an animal less than 0.5 mm long.

Photo credit: Betsvill ARC (http://emu.arsusda.gov/default.html) USDA