# Introduction to Computer Vision

Michael J. Black

## Object Recognition

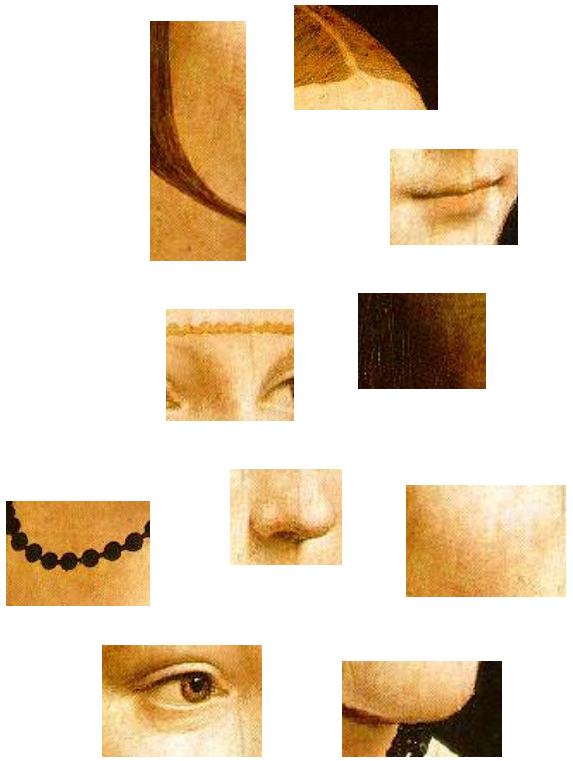
# News

- All assignments graded.
- Project due date is Dec 16 and it is a hard deadline!
- I have to submit grades by Dec 18.

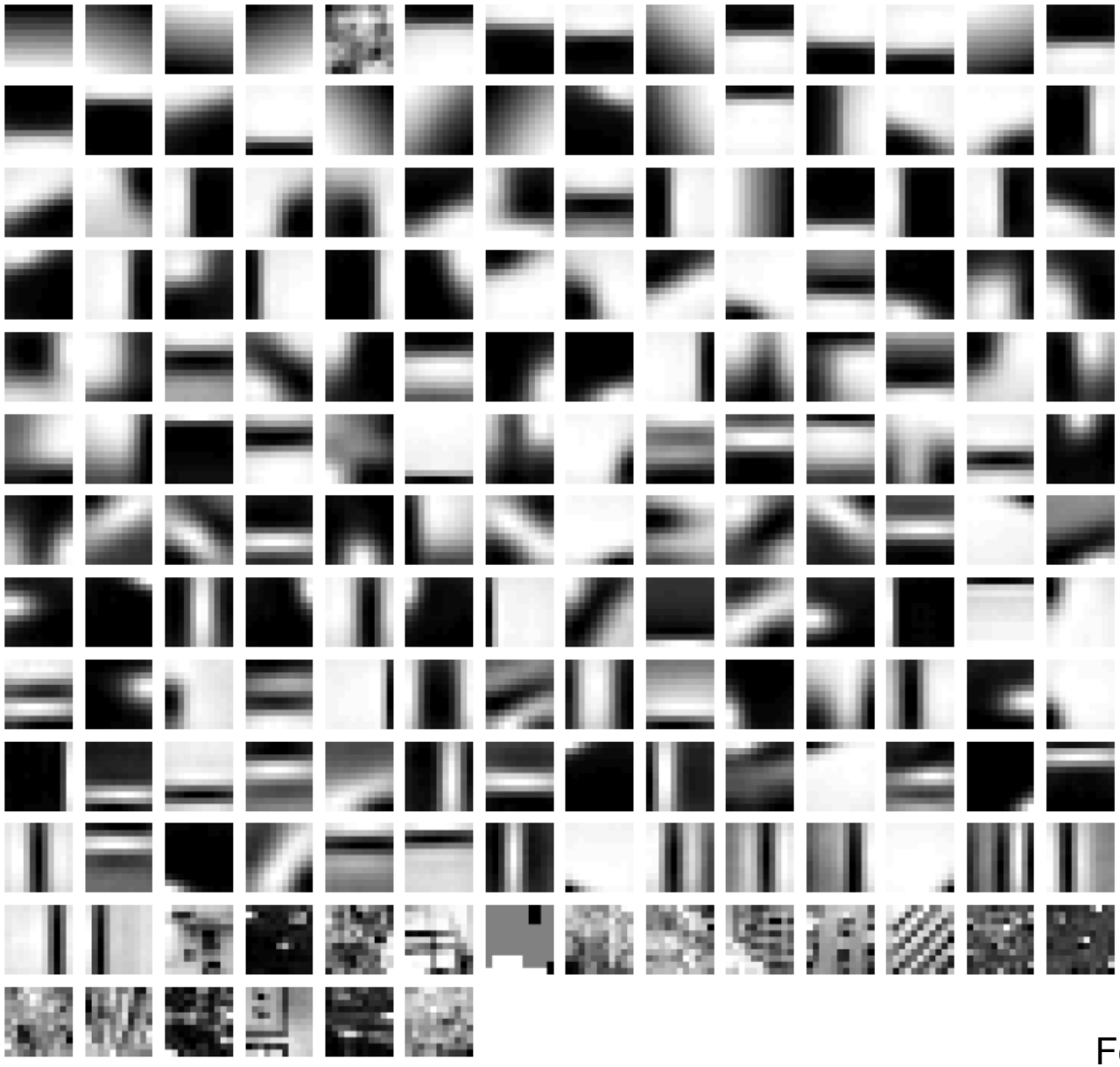
# Three main issues

- Representation
  - How to represent an object category
- Learning
  - How to form the classifier, given training data
- Recognition
  - How the classifier is to be used on novel data

# 1.Feature detection and representation

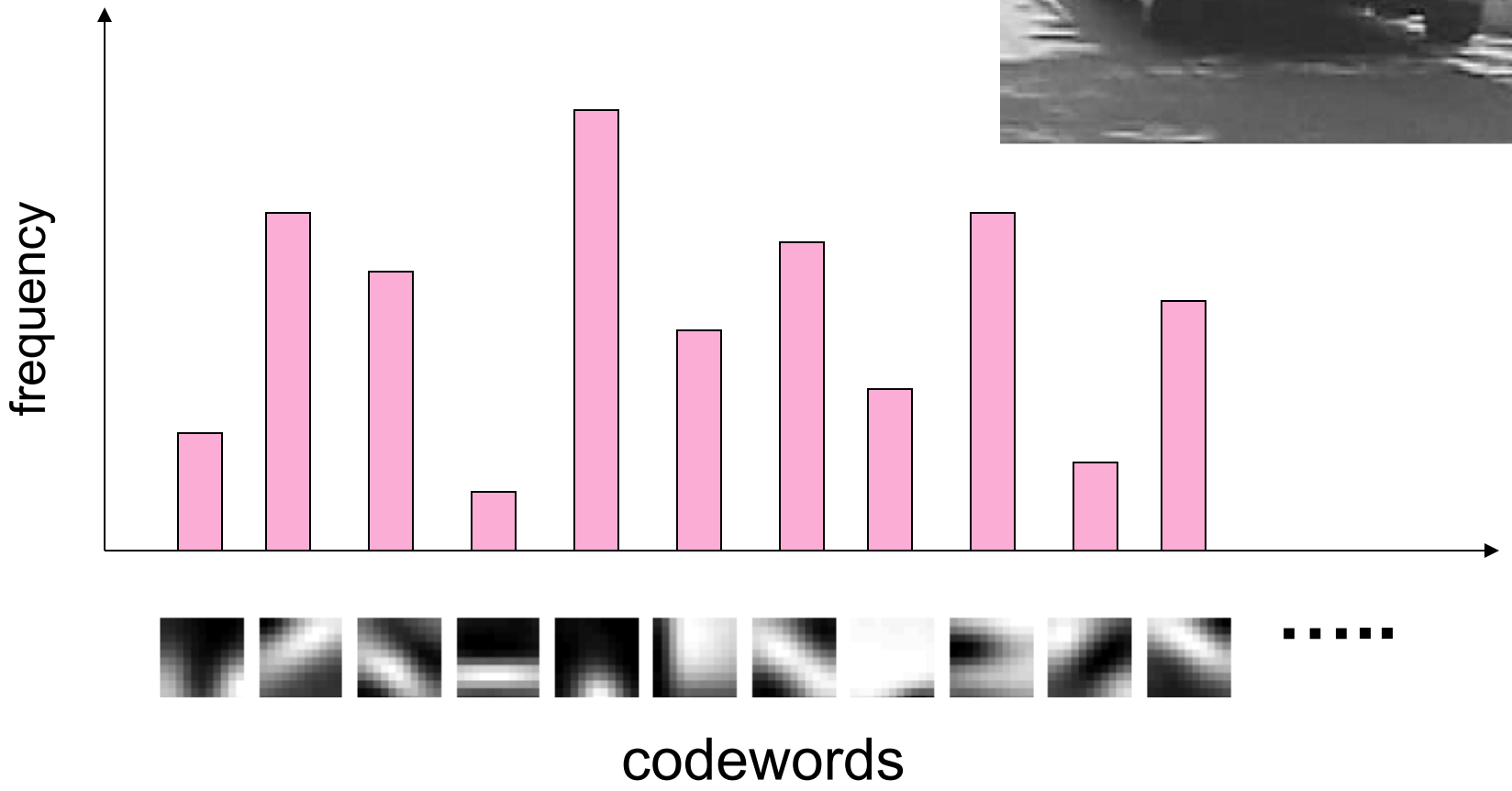


# 2. Codewords dictionary formation



Fei-Fei et al. 2005

# 3. Image representation

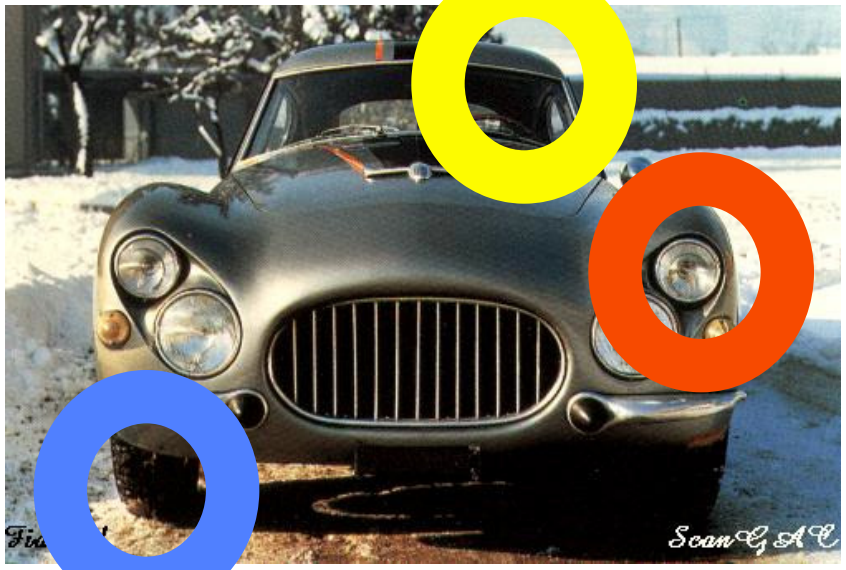


# Problem with bag-of-words

- All have equal probability for bag-of-words methods
- Location information is important

# Model: Parts and Structure



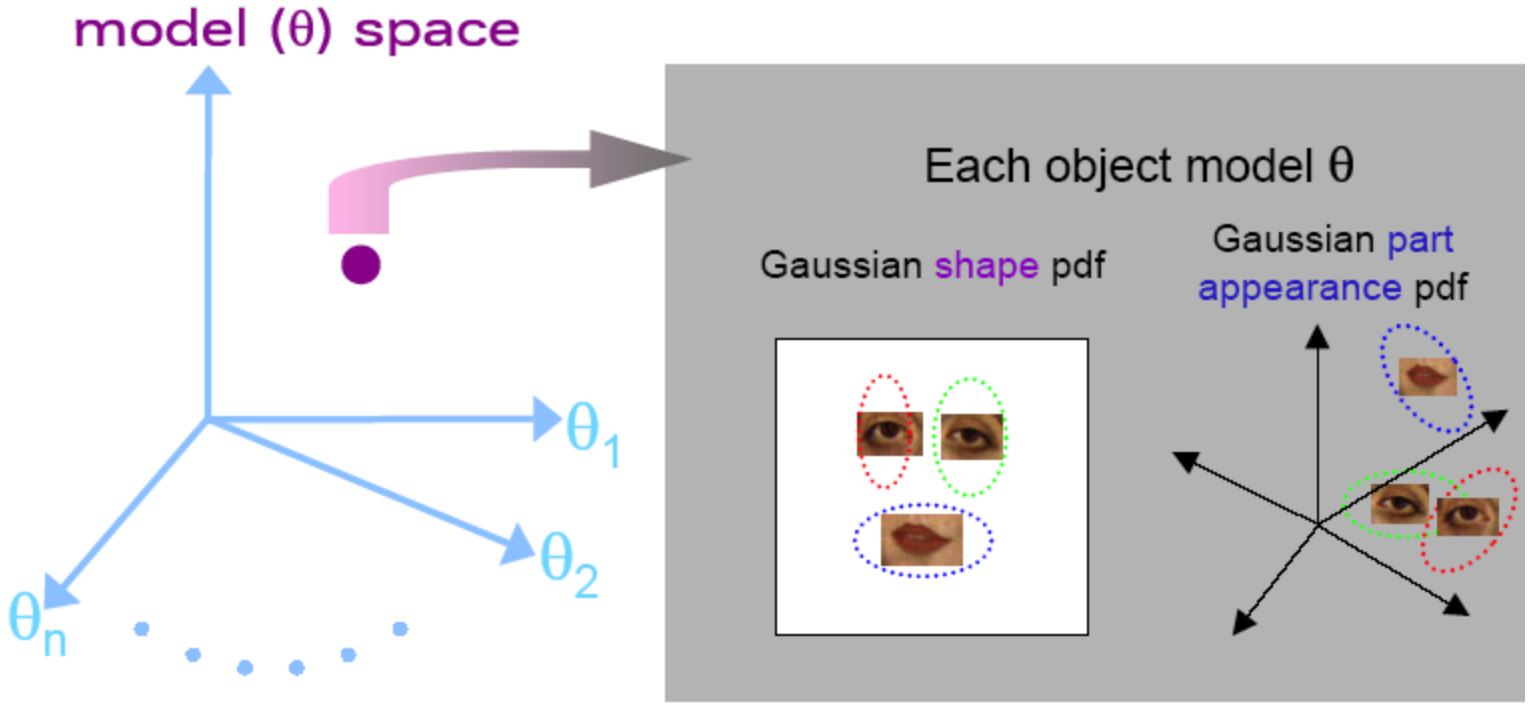


# Representation of appearance

- Needs to handle intra-class variation
  - Task is no longer matching of descriptors
  - Implicit variation (VQ to get discrete appearance)
  - Explicit model of appearance (e.g. Gaussians in SIFT space)


- Dependency structure
  - Often assume each part's appearance is independent
  - Common to assume independence with location

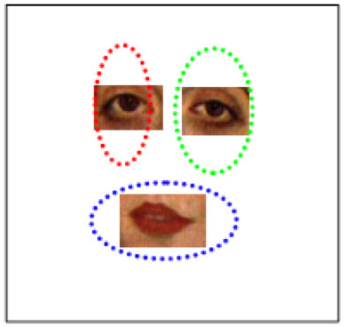
# Representing Objects



## Foreground model

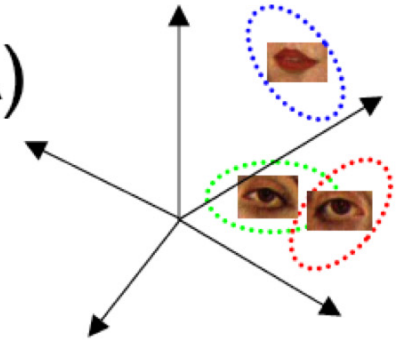
(X)

Gaussian shape pdf

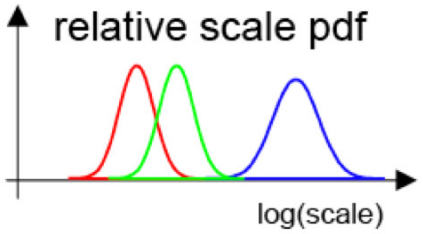


(A)

Gaussian part appearance pdf

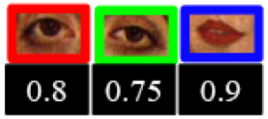


Gaussian relative scale pdf




(S)

Prob. of detection

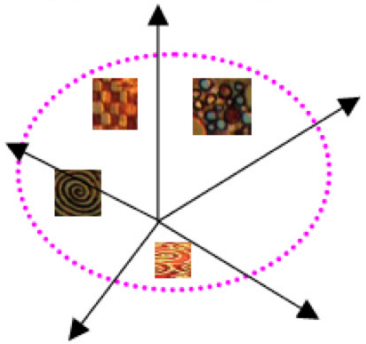



(D)

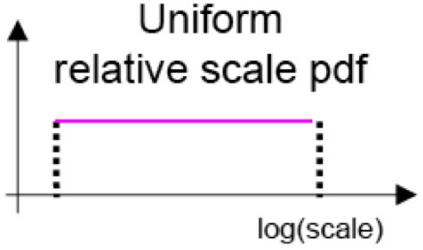
## Clutter model

Uniform shape pdf

Gaussian background appearance pdf



Uniform relative scale pdf




Poission pdf on # detections

# Demo (2)

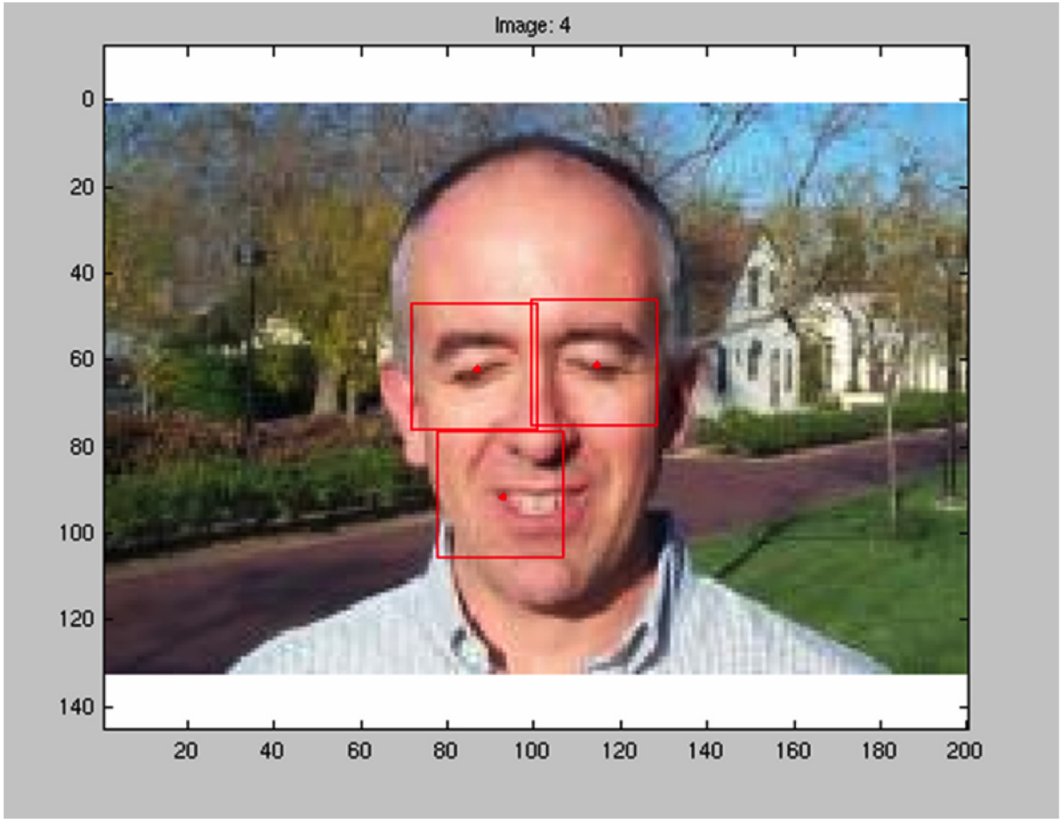
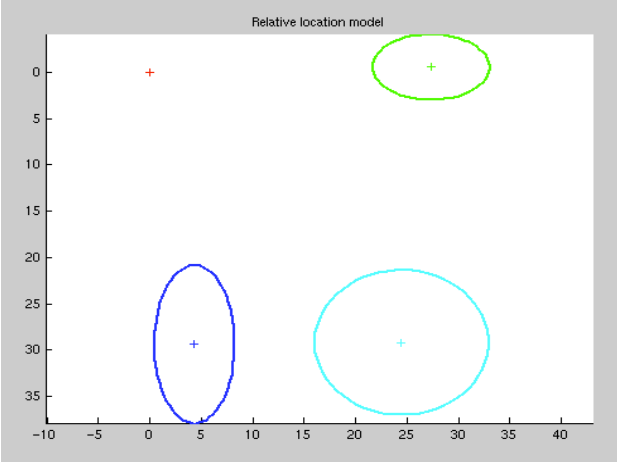
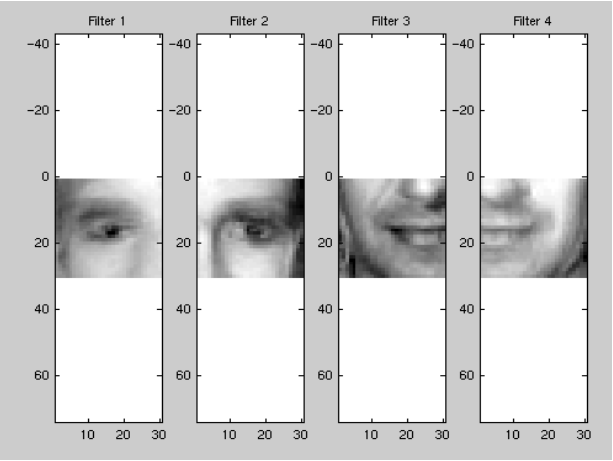
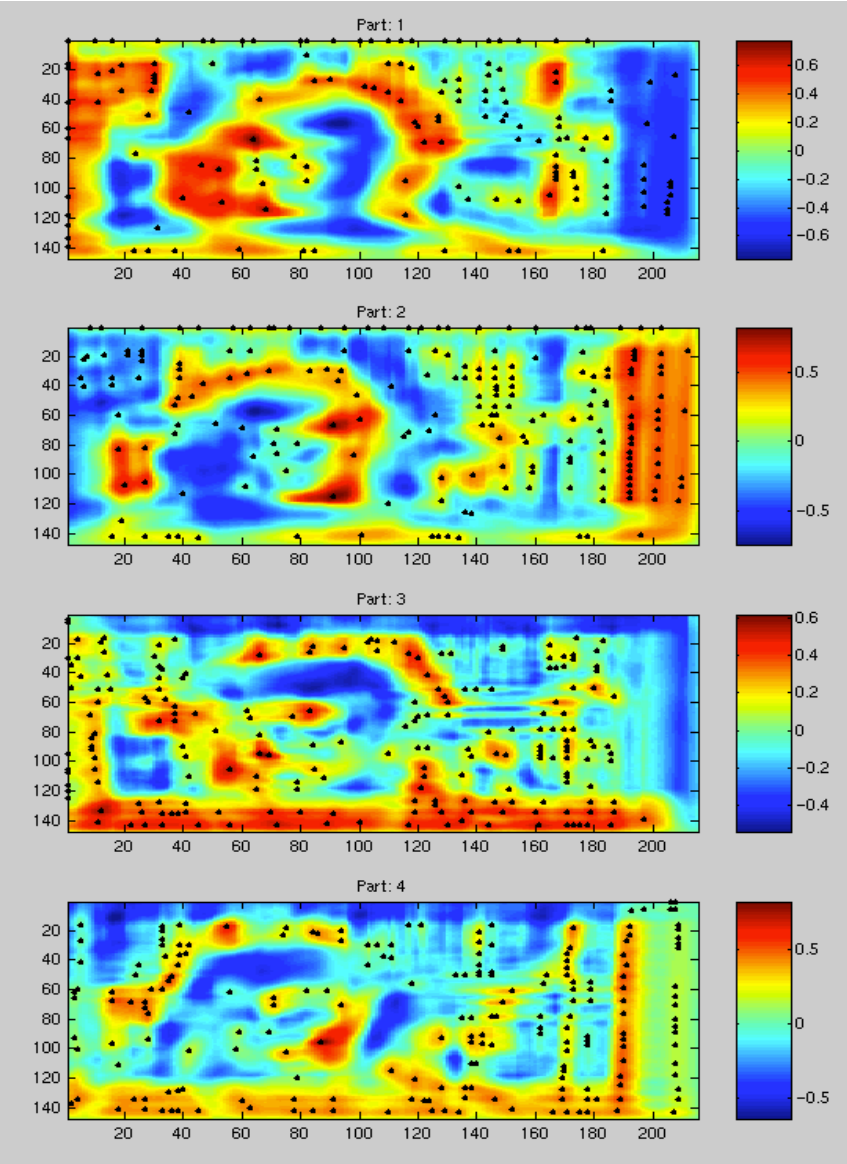


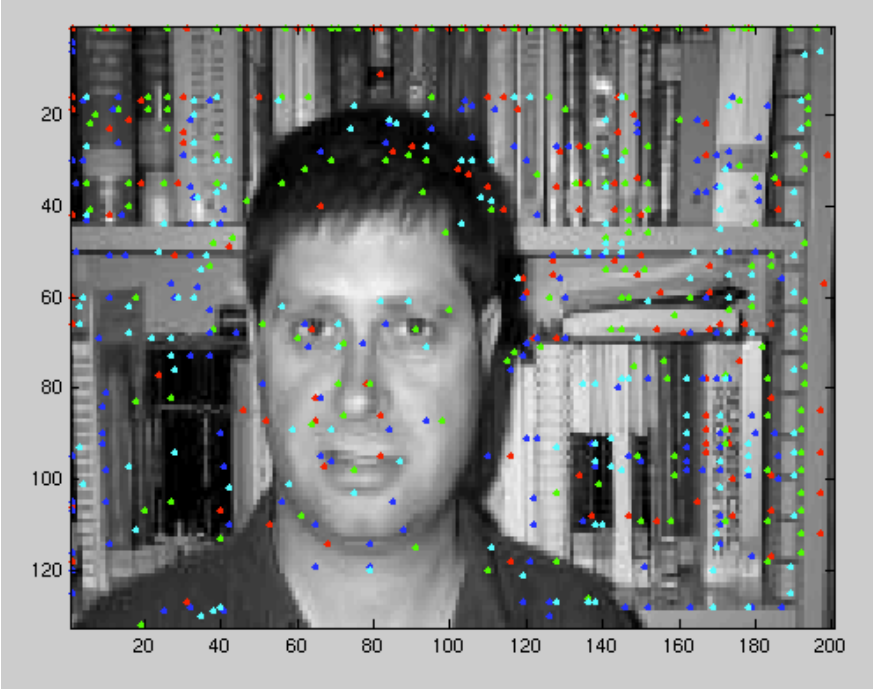
Image: 4
0
20
40
60
80
100
120
140
20
40
60
80
100
120
140
160
180
200


Relative location model
0
5
10
15
20
25
30
35
−10
−5
0
5
10
15
20
25
30
35
40

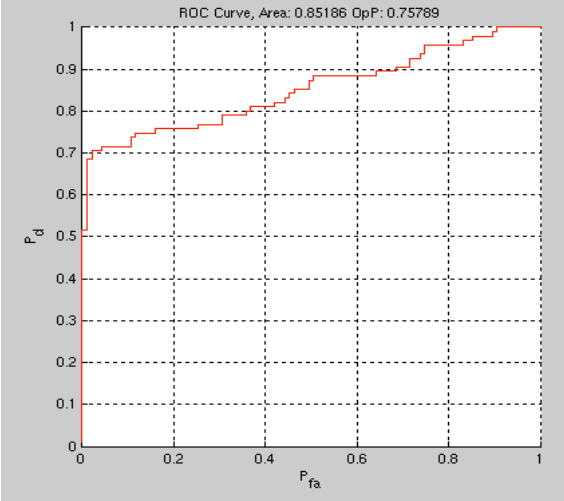

Filter 1
Filter 2
Filter 3
Filter 4
−40
−20
0
20
40
60
10
20
30

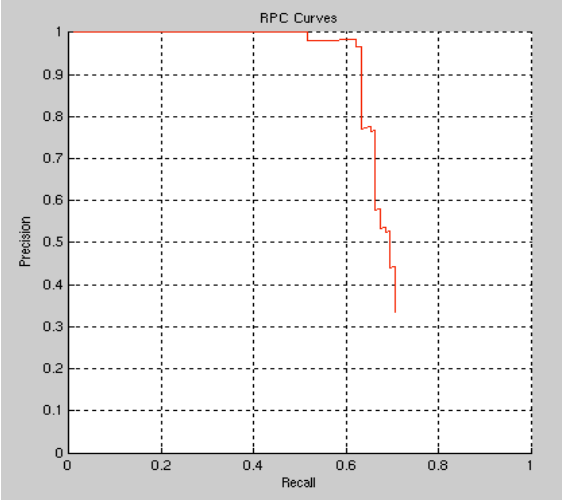
# Demo (3)



# Demo (4)


ROC Curve, Area: 0.85186 OpP: 0.75789
$P_d$
$P_{fa}$


RPC Curves
Precision
Recall

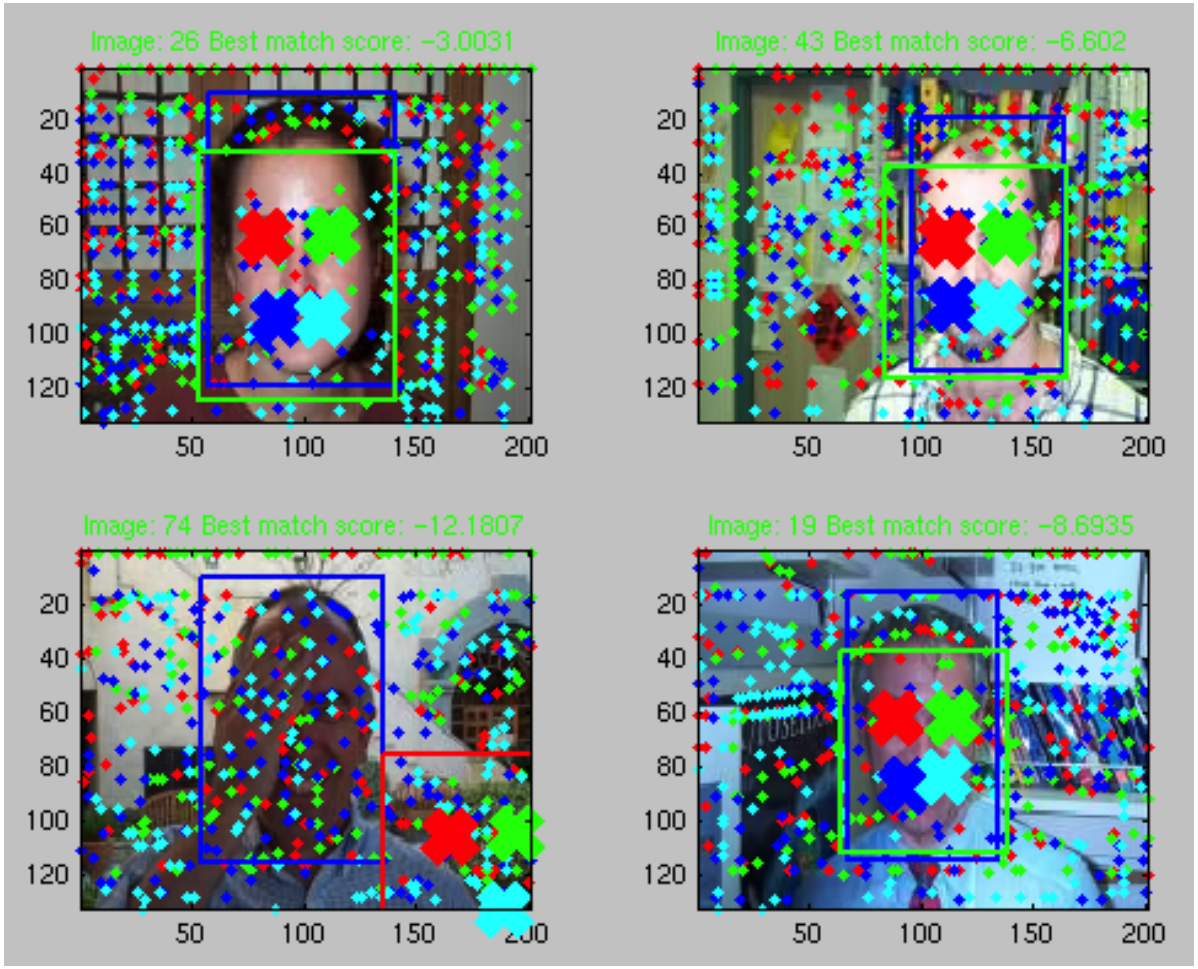
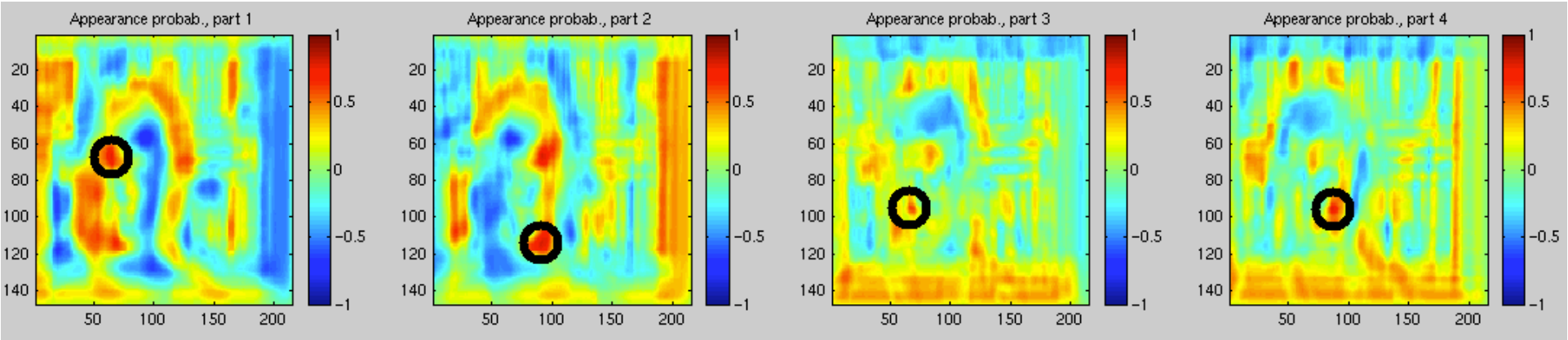
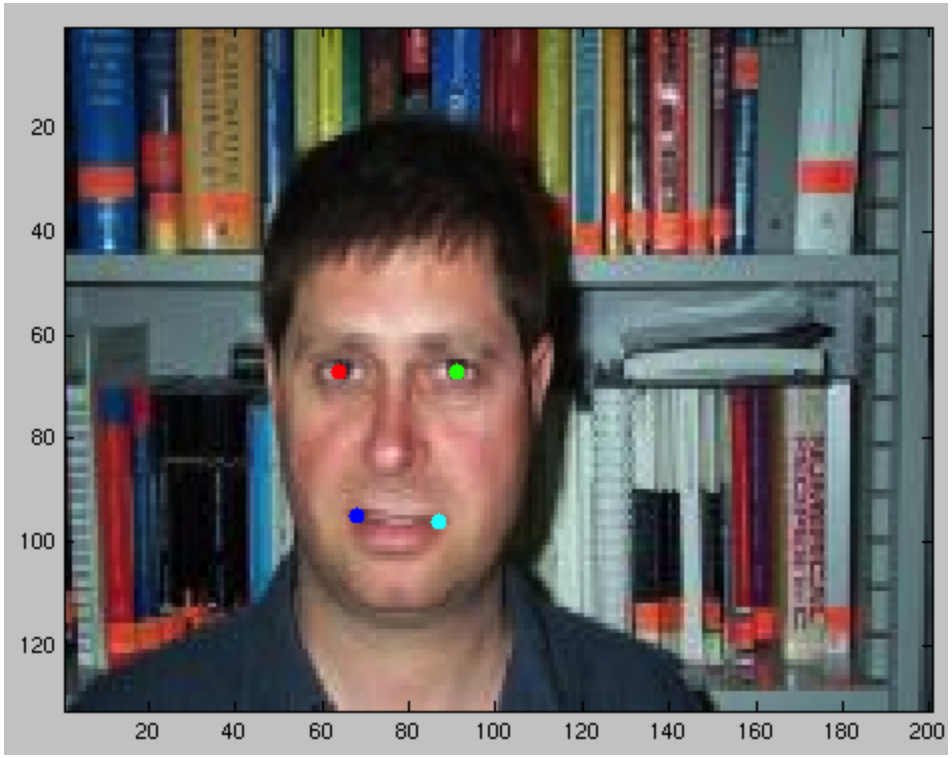
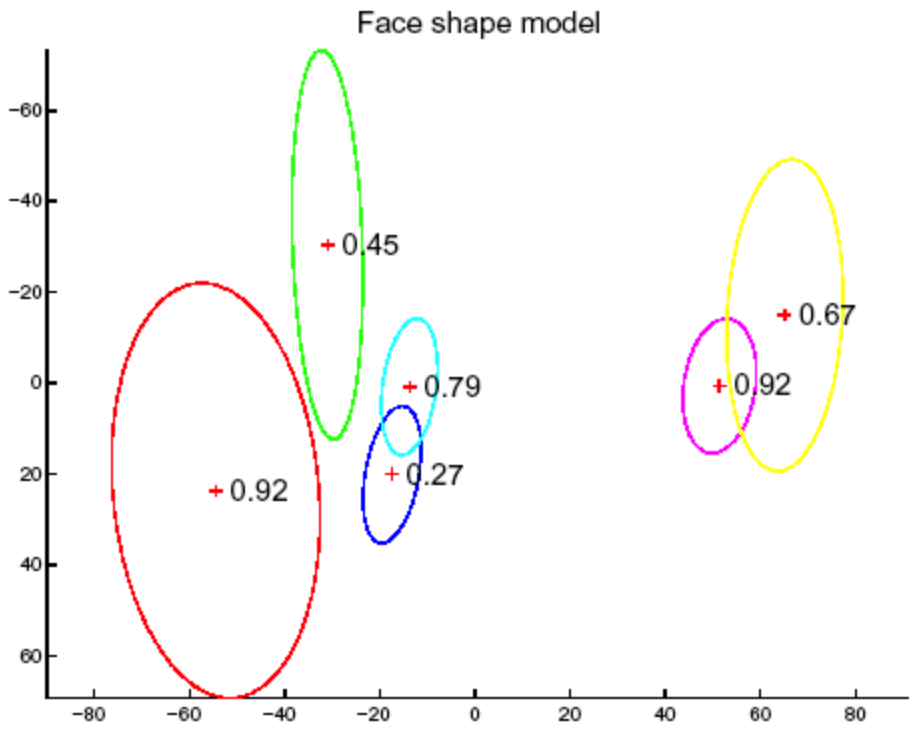


Image: 26 Best match score: −3.0031
Image: 43 Best match score: −6.602
Image: 74 Best match score: −12.1807
Image: 19 Best match score: −8.6935

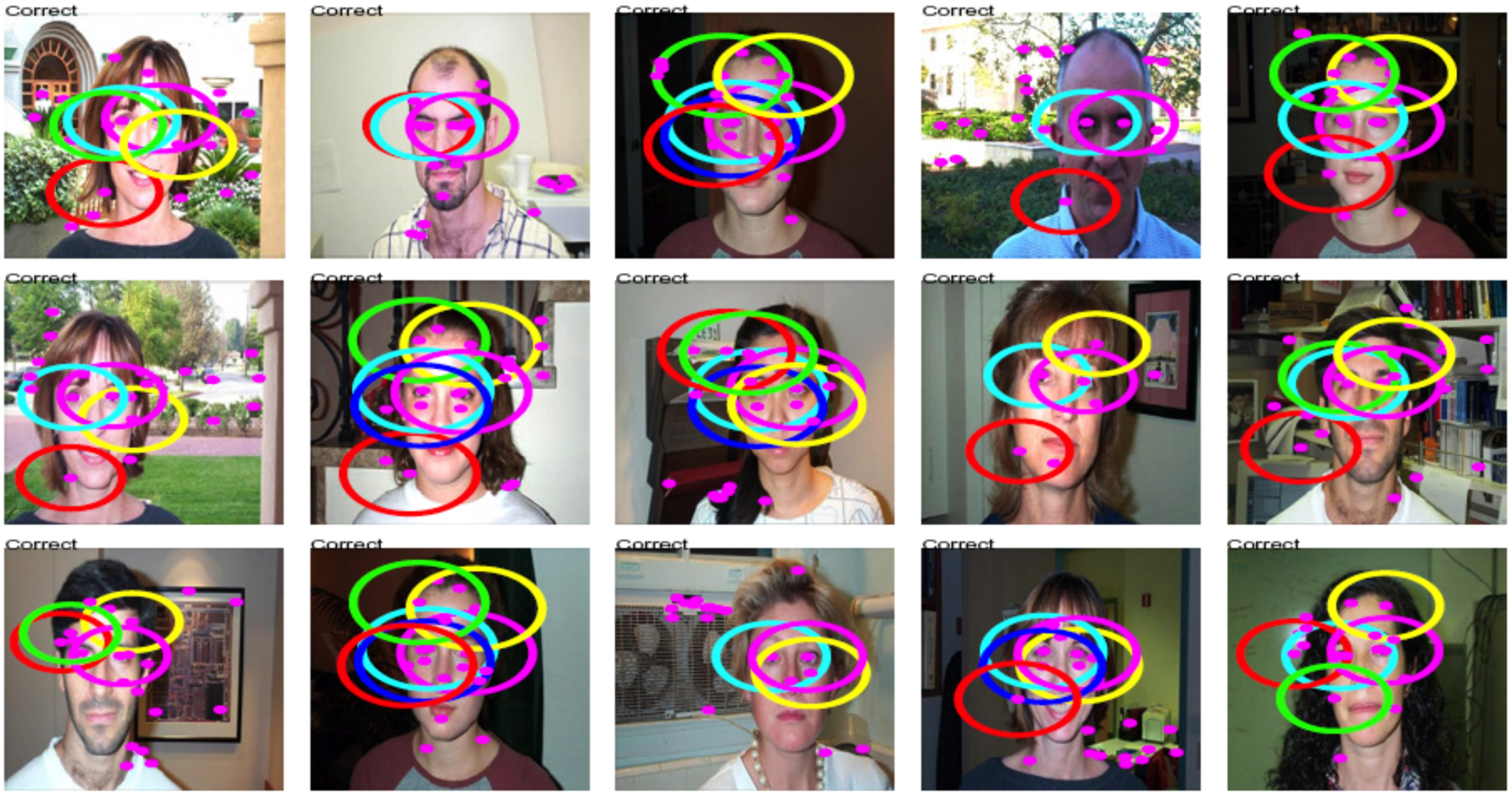
# Demo: efficient methods

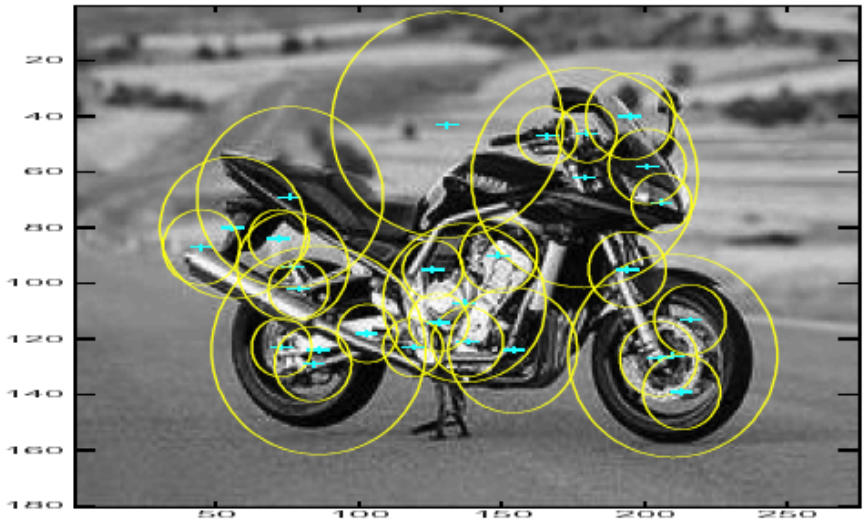


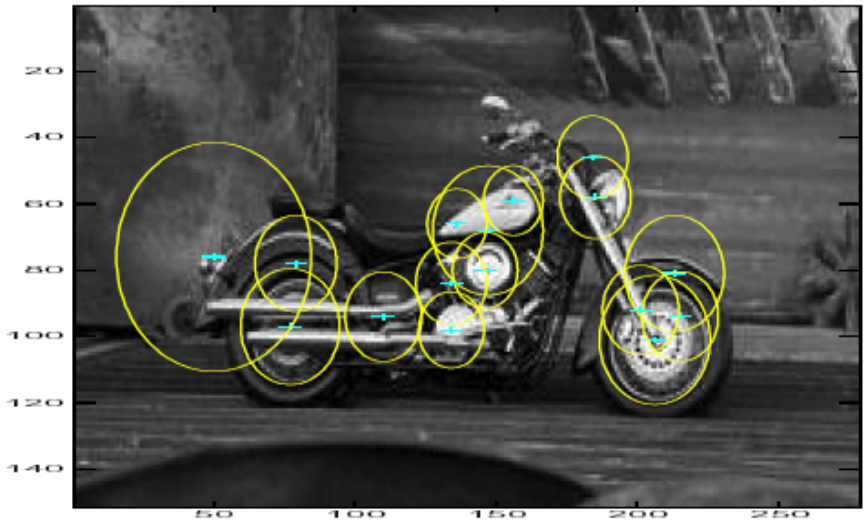



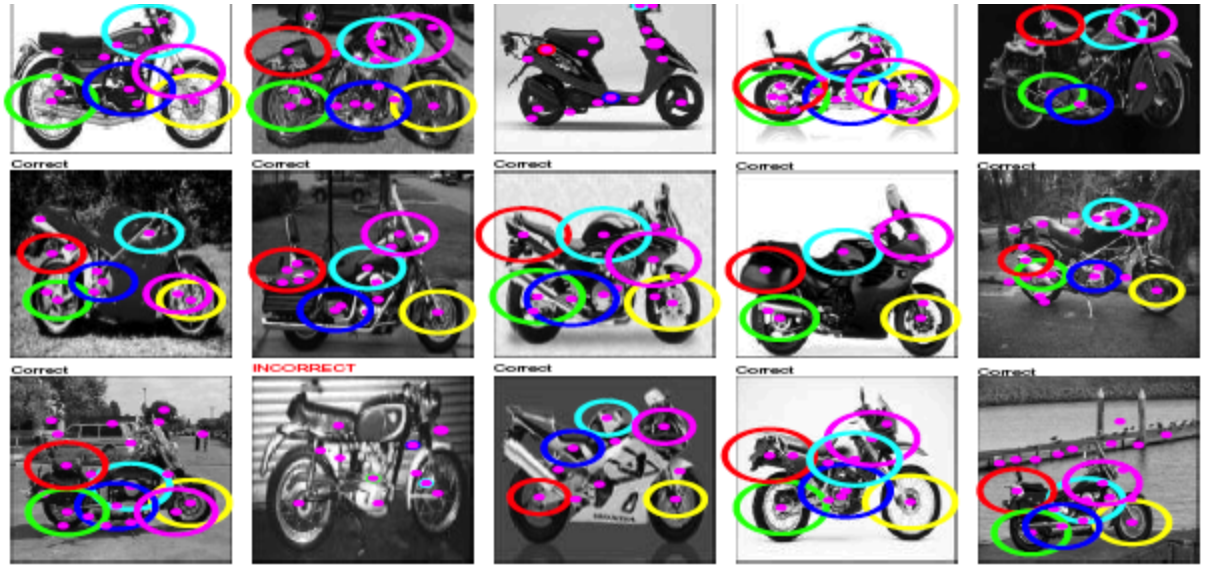

Face shape model
0.45
0.79
0.27
0.92
0.67
0.92
−60
−40
−20
0
20
40
60
−80
−60
−40
−20
0
20
40
60
80


Part 1 – Det:5e−21
Part 2 – Det:2e−28
Part 3 – Det:1e−36
Part 4 – Det:3e−26
Part 5 – Det:9e−25
Part 6 – Det:2e−27
Background – Det:2e−19


Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct
Correct


20
40
60
80
100
120
140
160
180
50
100
150
200
250


20
40
60
80
100
120
140
50
100
150
200
250


Correct
Correct
Correct
Correct
Correct
Correct
INCORRECT
Correct
Correct
Correct

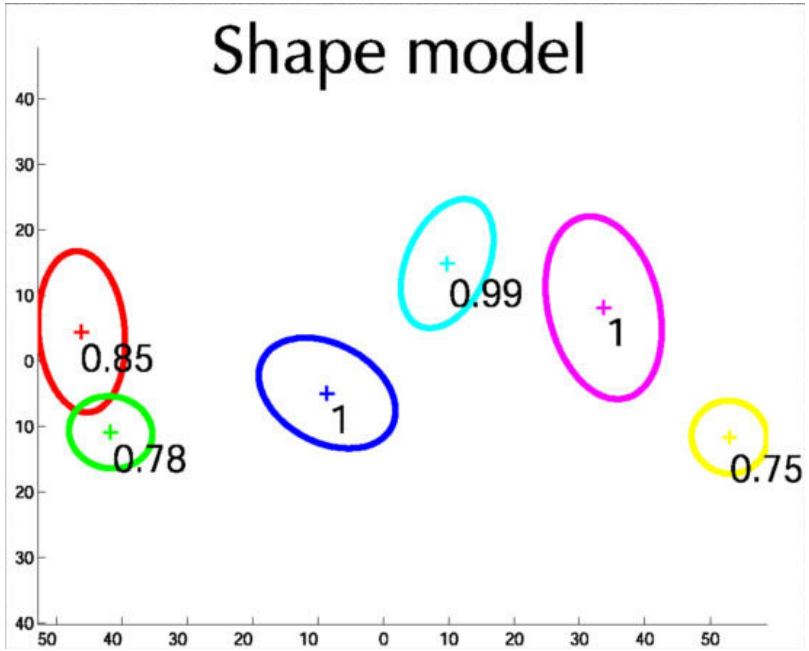
Learn parts from examples.

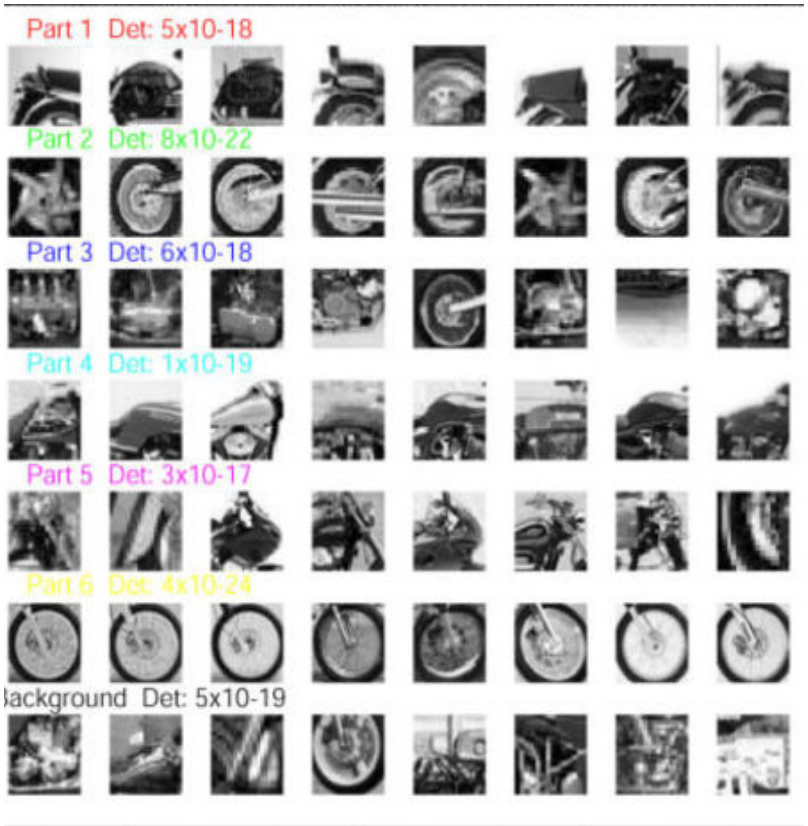
Find interesting points (structure tensor), find similar ones, use PCA to model them.


From: Rob Fergus http://www.robots.ox.ac.uk/%7Efergus/

# Shape

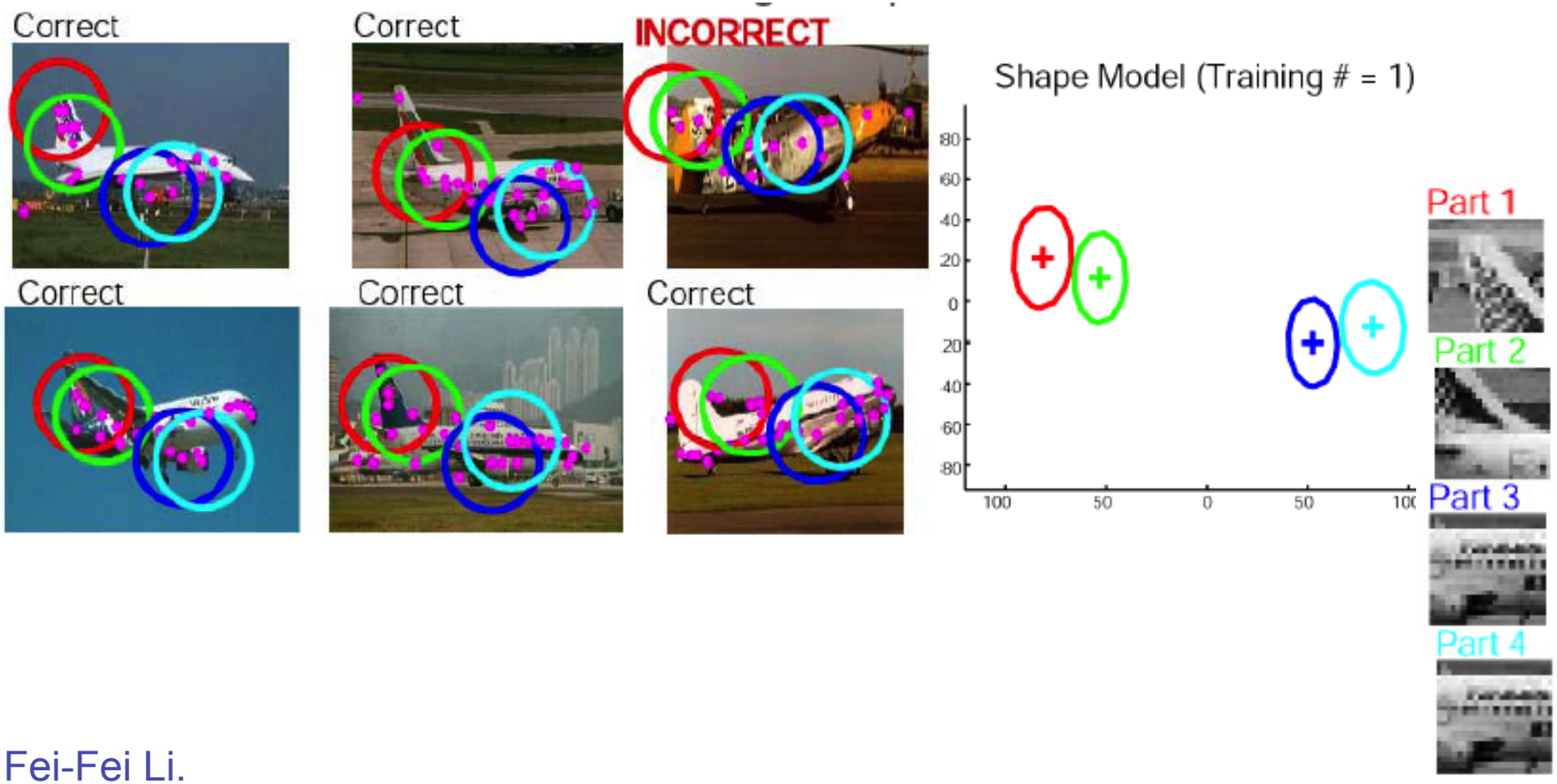
Given a “vocabulary” of parts, learn a model of their spatial relationships




From: Rob Fergus http://www.robots.ox.ac.uk/%7Efergus/

# Recognizing Objects



# ~100 Things We've Learned

Pinhole camera
Perspective projection
Orthographic projection
Weak perspective
PCA
Eigenvalues
Eigenvectors
Inpainting
Markov random field
Particle filter
Image statistics
Continuation method
Graduated non-convexity
MAP estimate

Gaussian pyramid
Laplacian pyramid
Matlab
Linear filtering
Convolution
Gaussian
Gradient
Dimensionality reduction
Monte Carlo sampling
Convolution
Correlation

Projection
Finite differences
Steerable filter
Gradient magnitude
DoG
Template matching
Normalized correlation
SSD
Subspaces
Basis image
SVD
Eigenfaces
Histogram

# ~100 Things We've Learned

Random variable
Marginalize
Expectation
Statistical independence
Conditional independence
Joint probability
Conditional probability
Bayes Theorem
Likelihood
Prior
Classifier
Tracking
Regularization
Stereo

Posterior
Covariance
Structure tensor
Mahalanobis distance
Whitening
Denoising
Motion field
Optical flow
Taylor series
Brightness constancy
OFCE
Aperture problem
Outliers
Rectification
Epipole

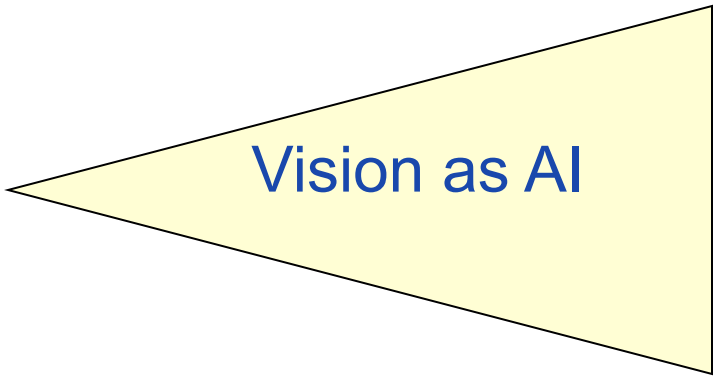
Affine
Least squares
Generative model
Warping
Interpolation
Super resolution
Occlusion
Robust statistics
Influence function
Breakdown point
Gradient descent
Annealing
Discontinuities
Binocular disparity

# ~100 Things We've Learned

So are we done?

# Timeline

1975-1985

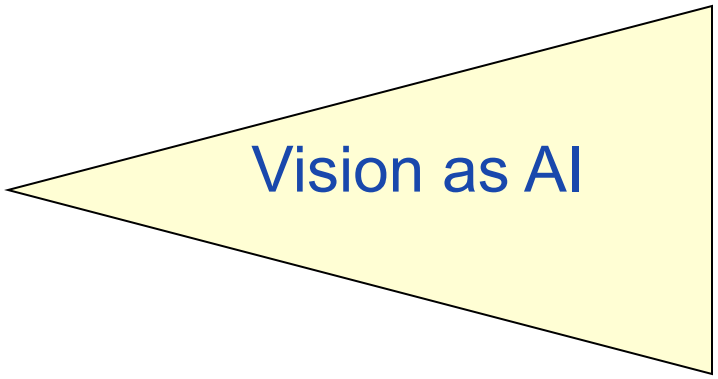



Early view (50's-60's): Minsky thought the vision sub-problem of AI could be solved by a single PhD student in a single summer. Done. Move on.

Lofty goals and early excitement.

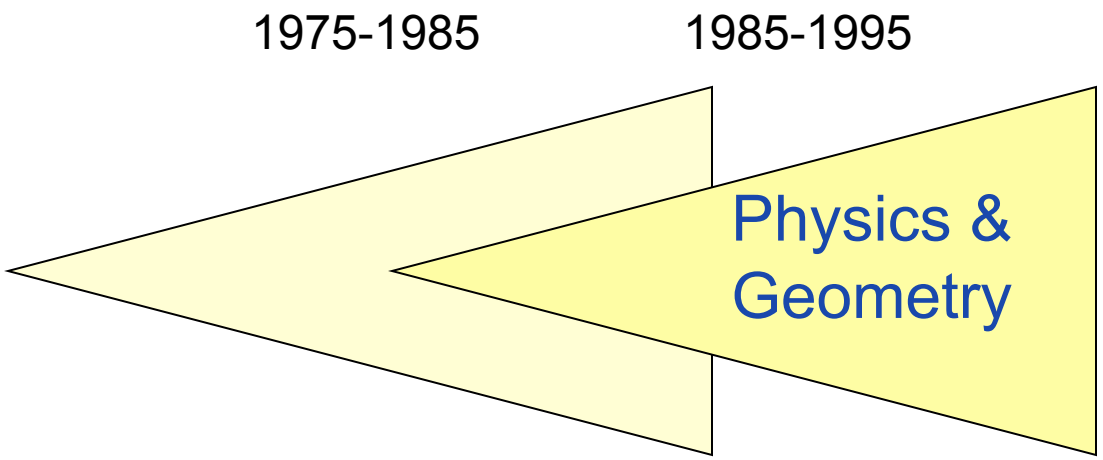
# Timeline

1975-1985




Shattered dreams and early disappointment.

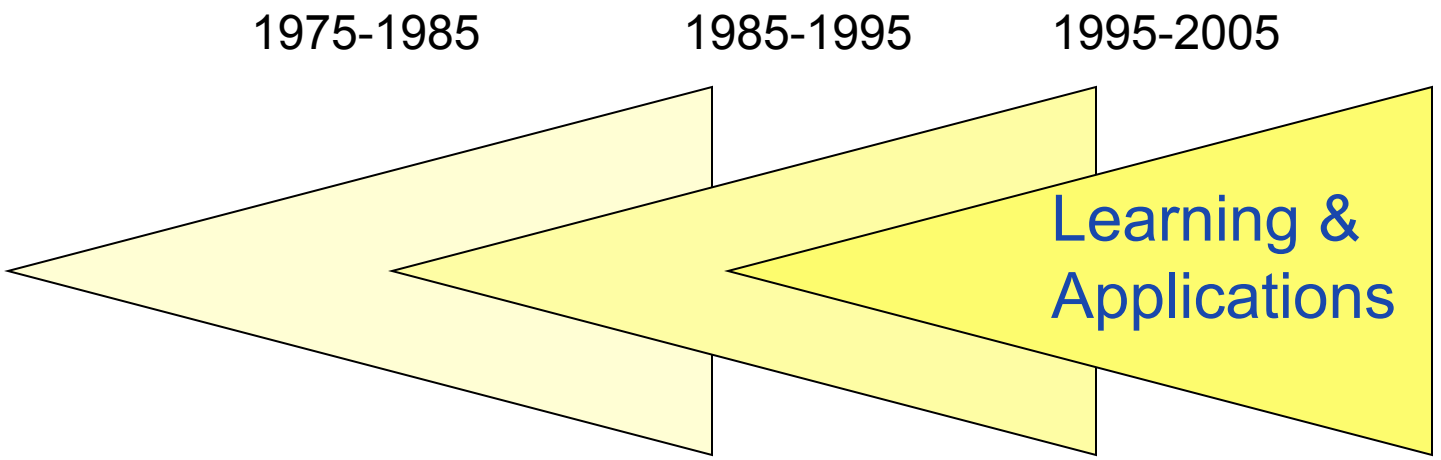
# Timeline




Regroup, focus on the basics

* metric reconstruction, quantitative evaluation.
* optimization methods.

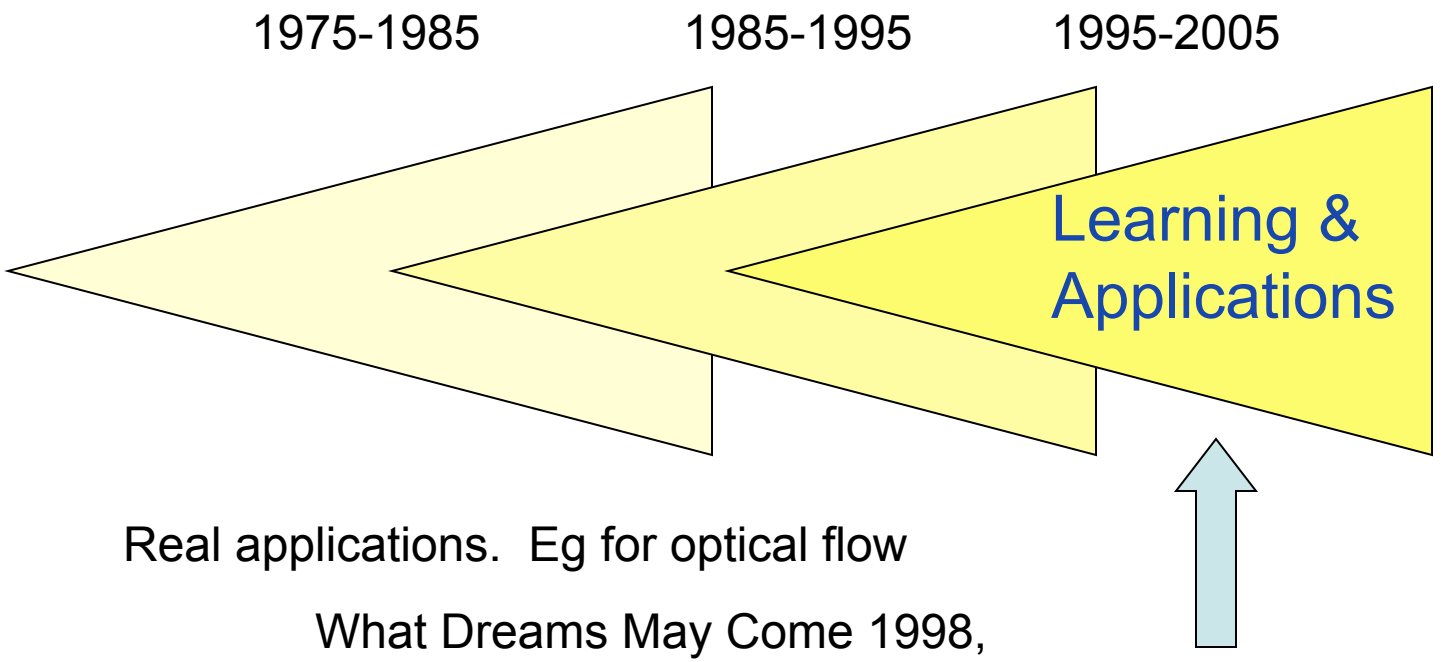
# Timeline




Trends: big disks, digital cameras, Firewire, fast processors, desktop video.

Machine learning provides a new grounding.

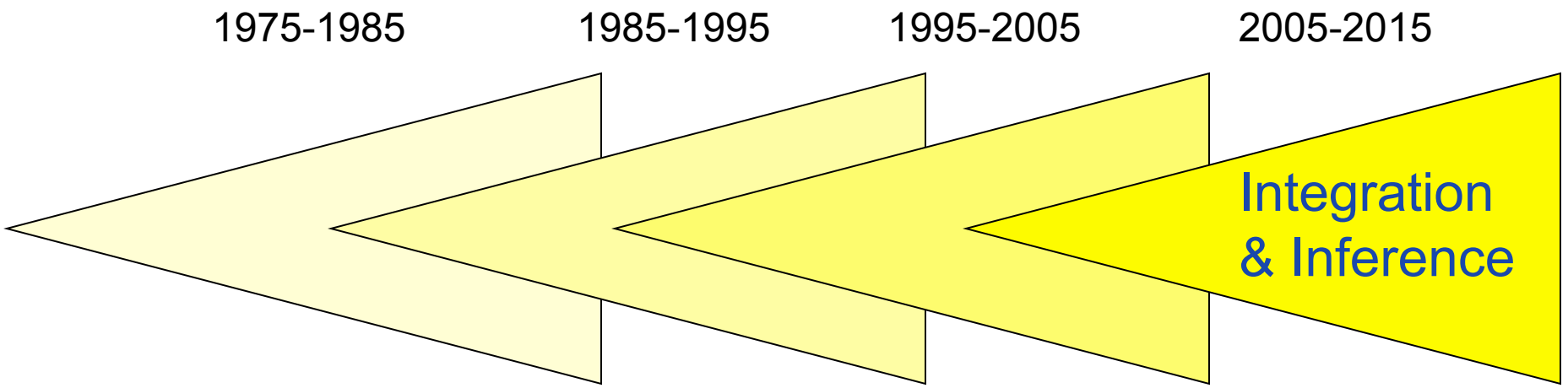
# Timeline




Real applications. Eg for optical flow

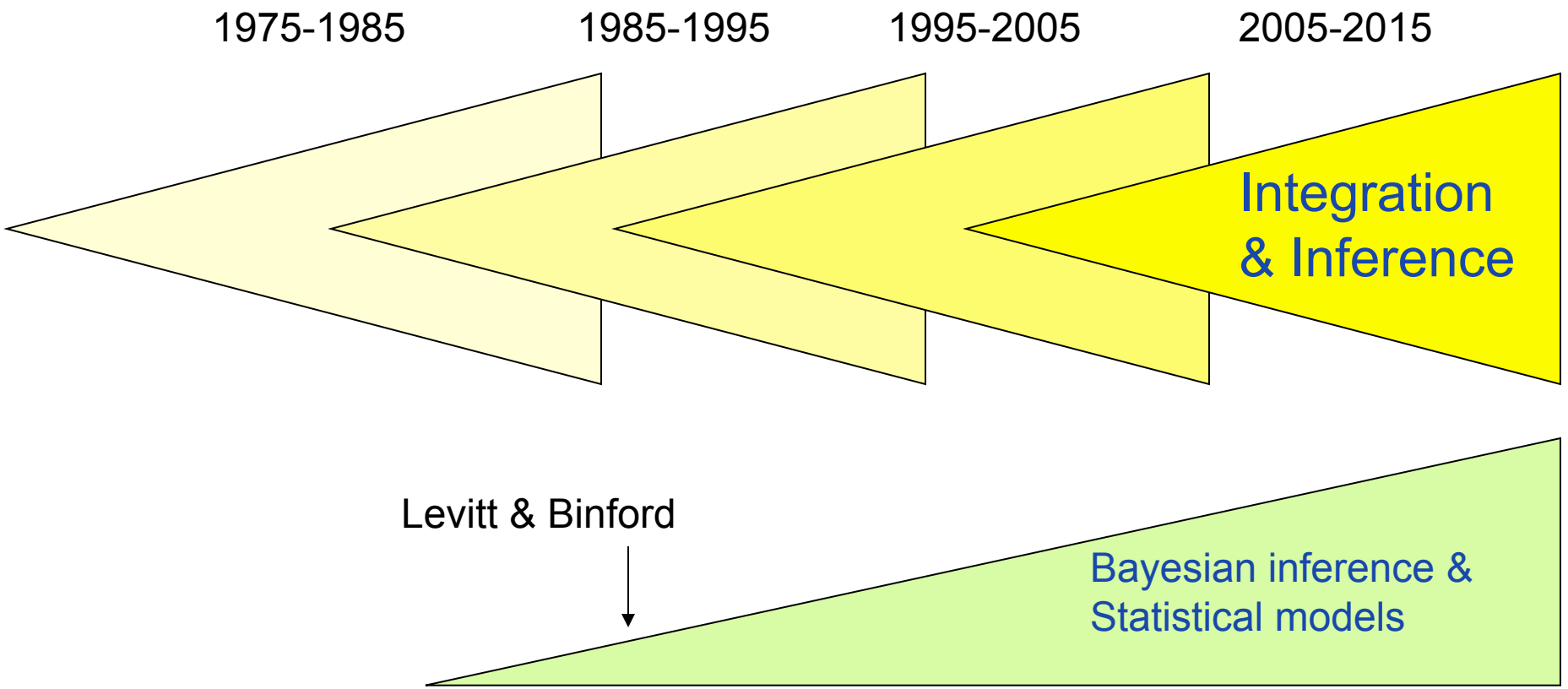
What Dreams May Come 1998,
The Matrix 1999

# Timeline



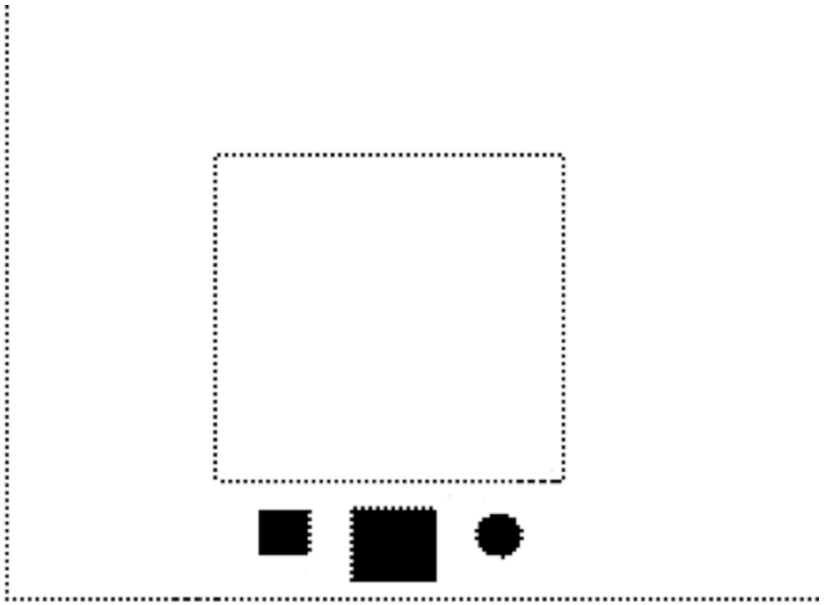

Return to some of the early goals with new tools.

# Timeline


1975-1985
1985-1995
1995-2005
2005-2015
Integration
& Inference
Levitt & Binford
Bayesian inference &
Statistical models

# What is still far off?

Motion interpretation.



<Play>

Heider&Simmel, 1944

* Here “vision” problem is trivial but explanation is hard.