

CS543 - Computer Vision

D.A. Forsyth

Administrative matters

- Instructor: D.A. Forsyth, 3310 Siebel Hall, daf@uiuc.edu
- Office hours: 14h30-15h30 Tue, Thur, or walk in
- TA: Ali Farhadi, afarhad2@uiuc.edu
- Evaluation: Homeworks, projects

Computer Vision

A MODERN APPROACH

FORSYTH ■ PONCE



Topics

- Overview, Cameras, Shading
- Color, Texture
- Edges, Corners, Segmentation
- Structure and Motion
- Features, Matching and Recognition
- People (finding, tracking, understanding)

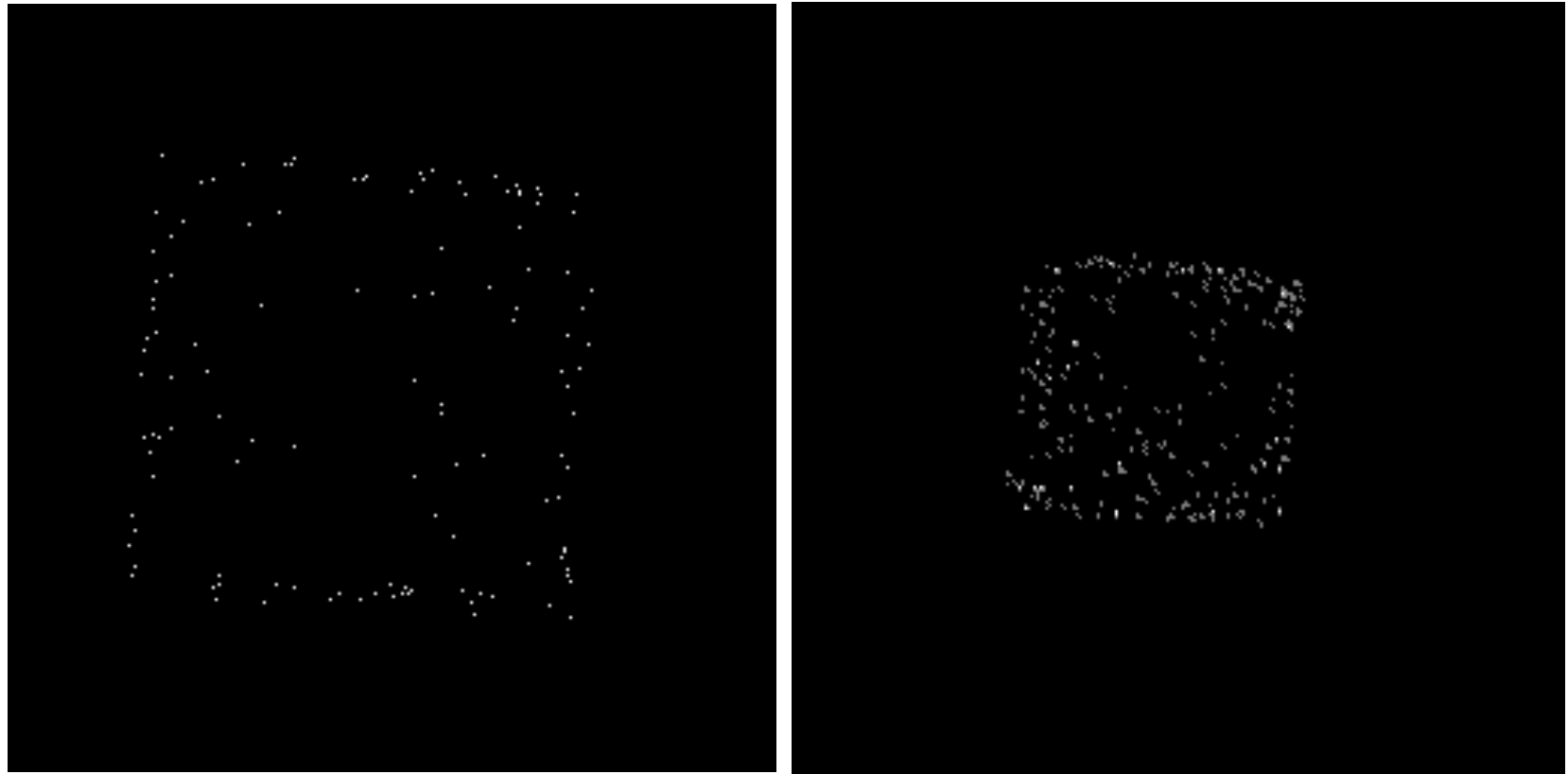
Example problems

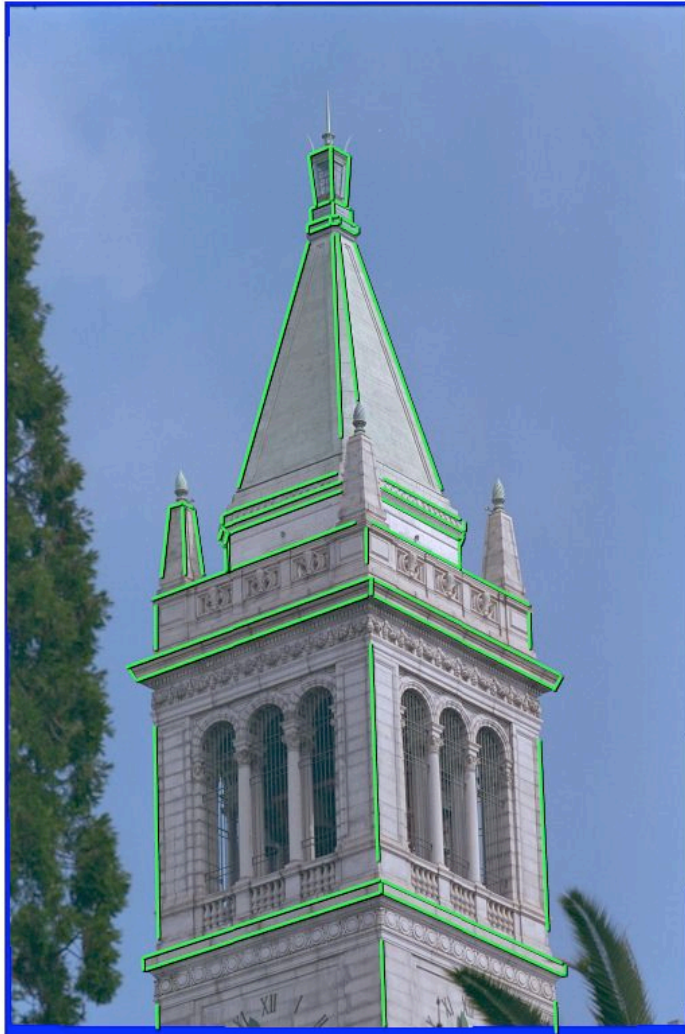
- **Obstacle avoidance**
 - A cricketer avoids being hit in the head (->) (<-)
 - the gannet pulls its wings in in time, by measuring time to contact
- **Reconstructing representations of the 3D world**
 - from multiple views
 - from shading
 - from structural models, etc
- **Recognition**
 - draw distinctions between what is seen
 - is it soggy?
 - will it eat me?
 - can I eat it?
 - is it a cat?
 - is it my cat?

3D Reconstruction

- Cues:
 - variation in appearance in multiple views
 - stereo
 - motion
 - (possibly) shading
 - (possibly) contour
 - texture
 - rich body of geometric theory
- Issues:
 - correspondence
 - noise
 - the nature of the representation

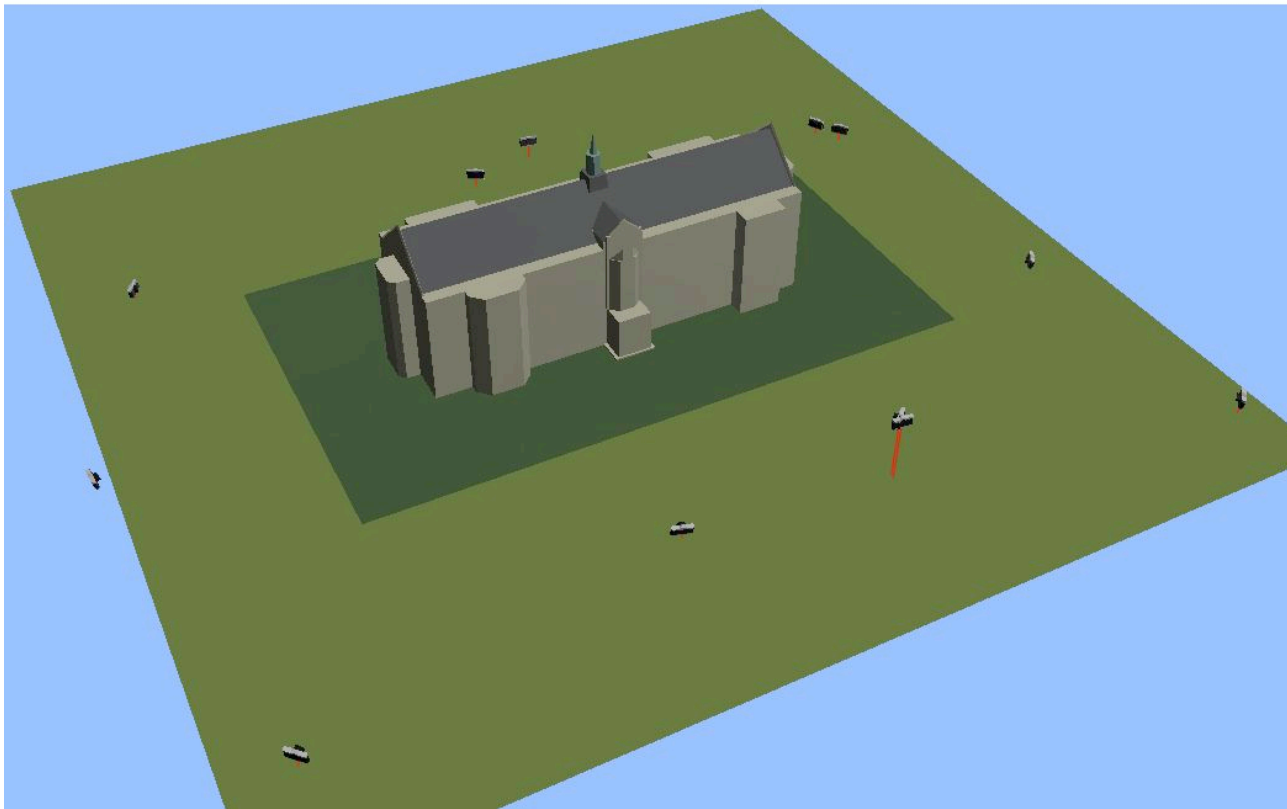
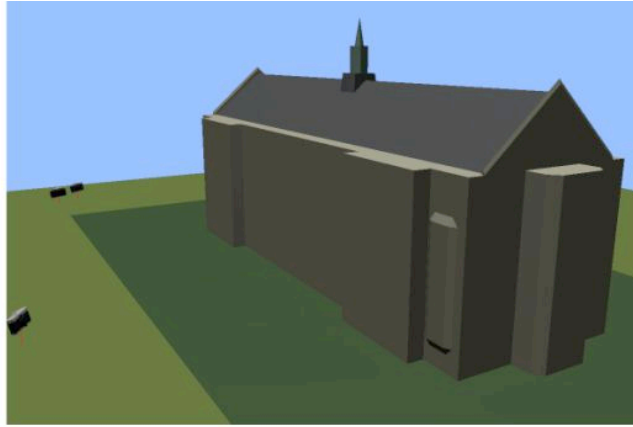
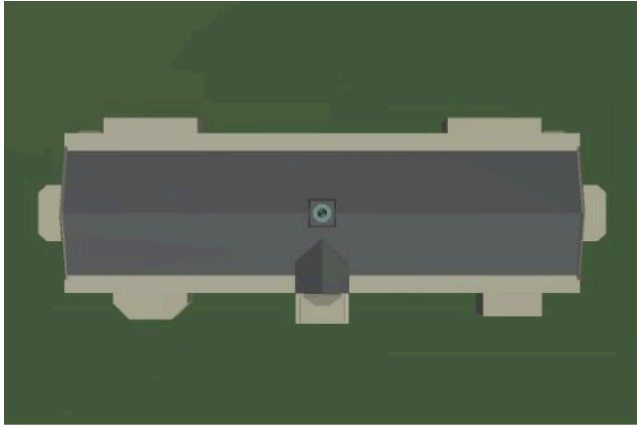
Structure from motion



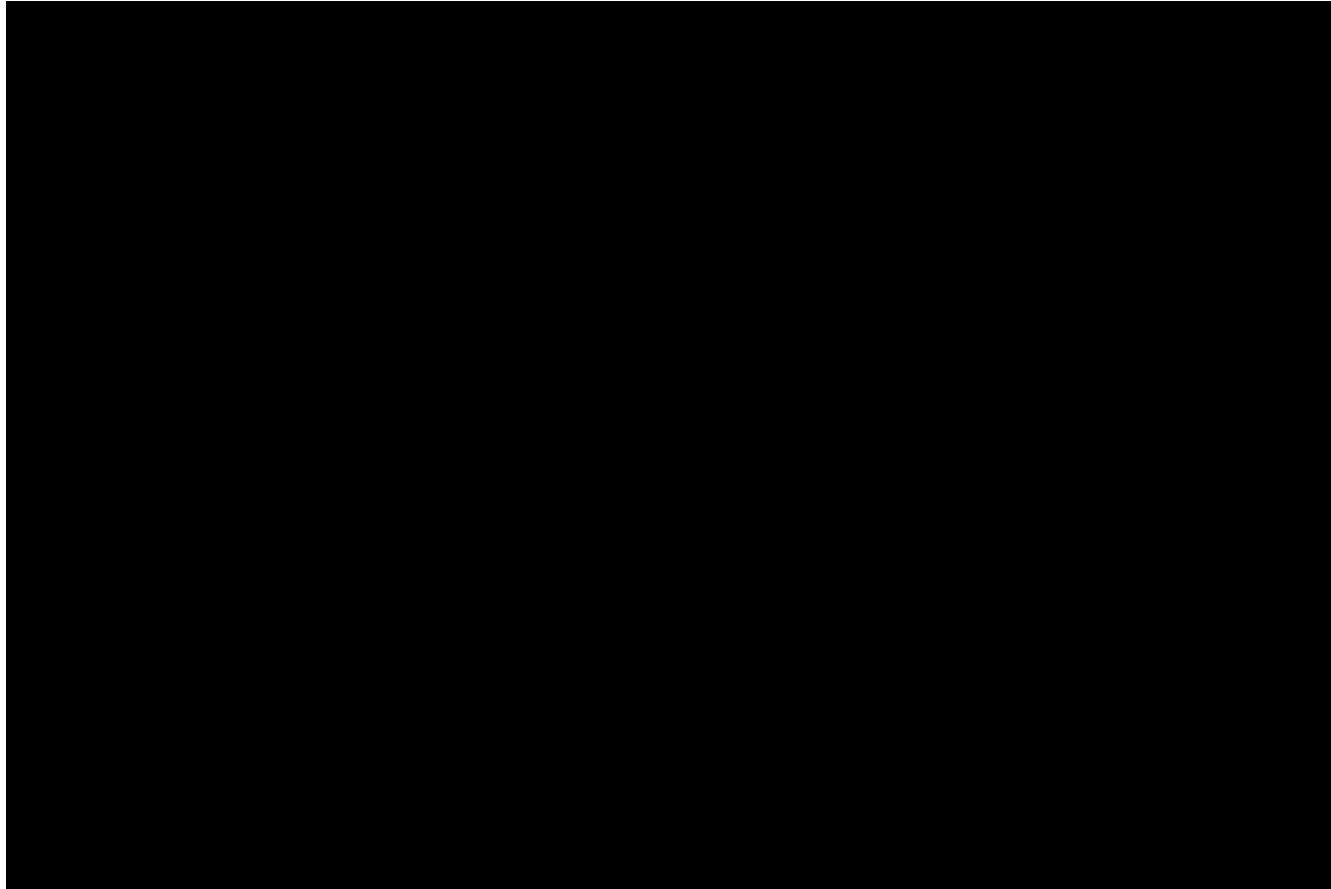


Work by Paul Debevec and Jitendra Malik









White, Crane, Forsyth 2007

Mosaics



Figure from “Quicktime VR - An Image-Based Approach to Virtual Environment Navigation”
S.E. Chen, SIGGRAPH 95

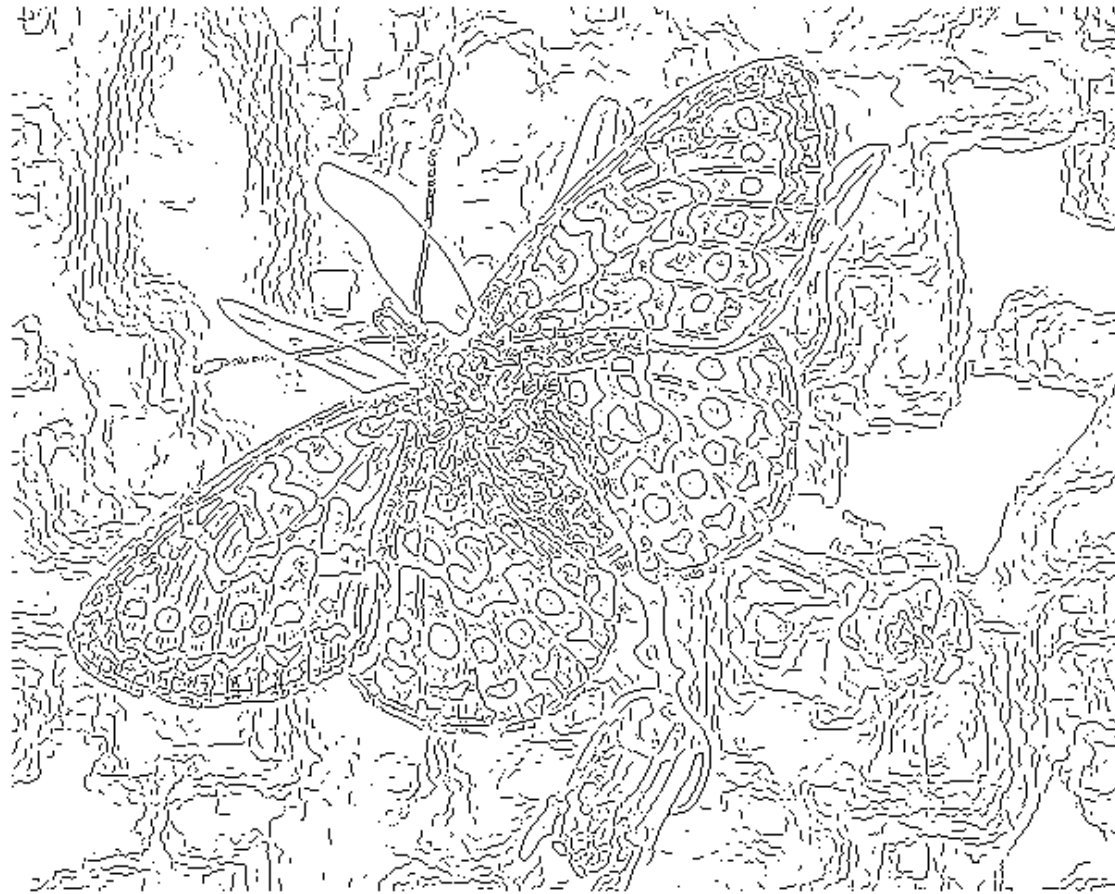
Correspondence

- Local representation of image properties make things easier
 - identify points which are easily localised
 - corners
 - which lie on edges
 - compare with points in next image
 - points which “look similar” may well match
 - search radius is constrained by geometry
 - in ways we will not discuss

Edges

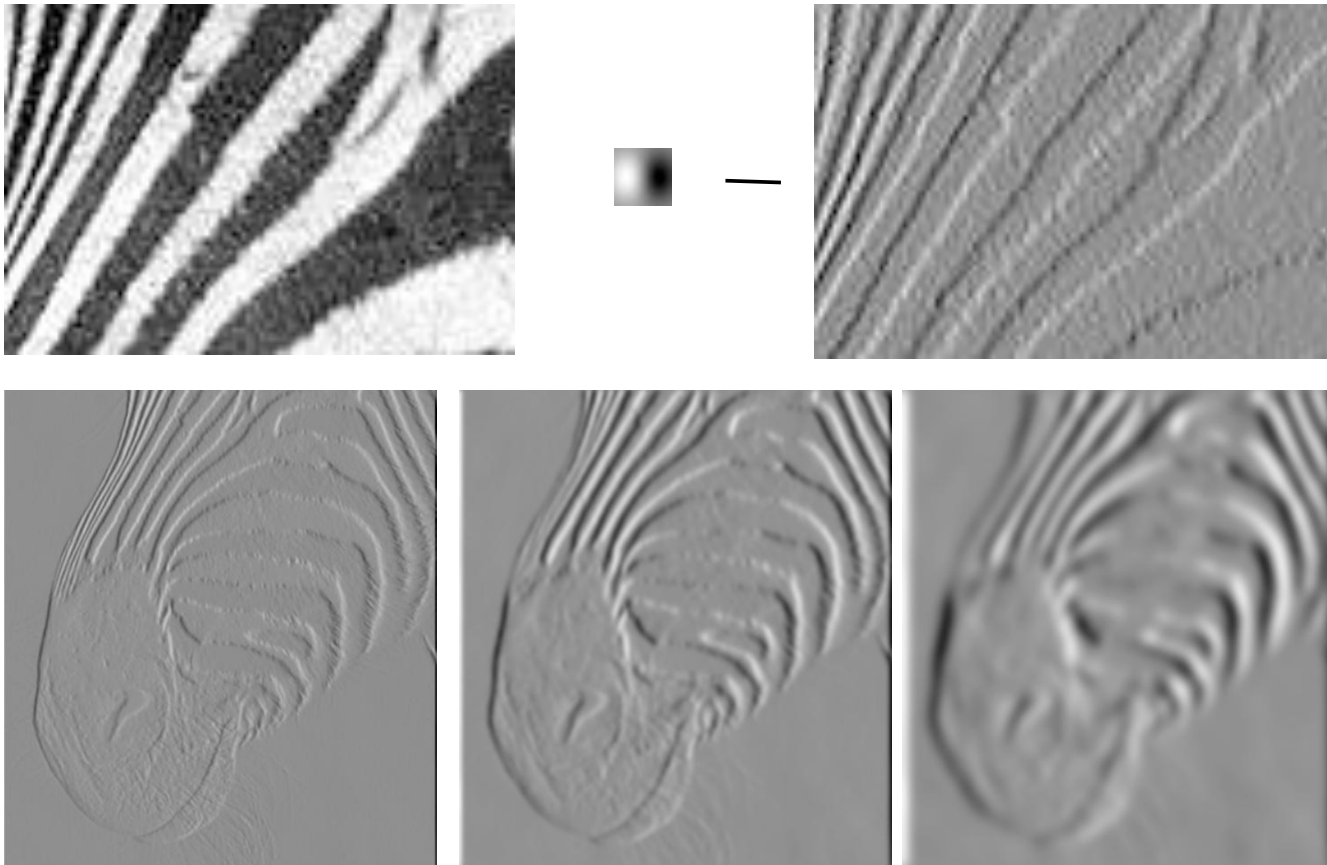


Edges



fine scale; high contrast threshold

Scale and Filtering



Scale gets coarser ->

Edges



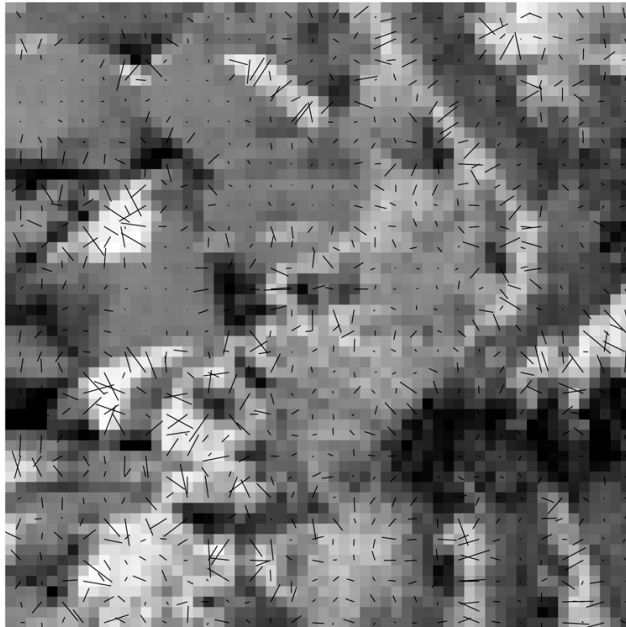
coarse scale; high contrast threshold

Edges

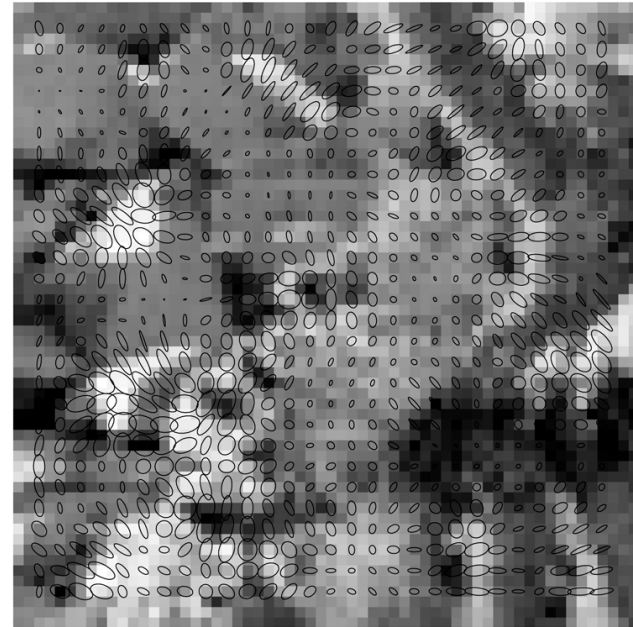


coarse scale; low contrast threshold

Finding corners



Edge orientations at points in an image



Covariance of gradient vectors over larger windows
Strong gradient and strong covariance suggests a corner

Local Representations

- What do edge responses look like nearby?
 - SIFT features
- What is the “general pattern” of grey levels?
 - statistics of filters

Lowe's SIFT features

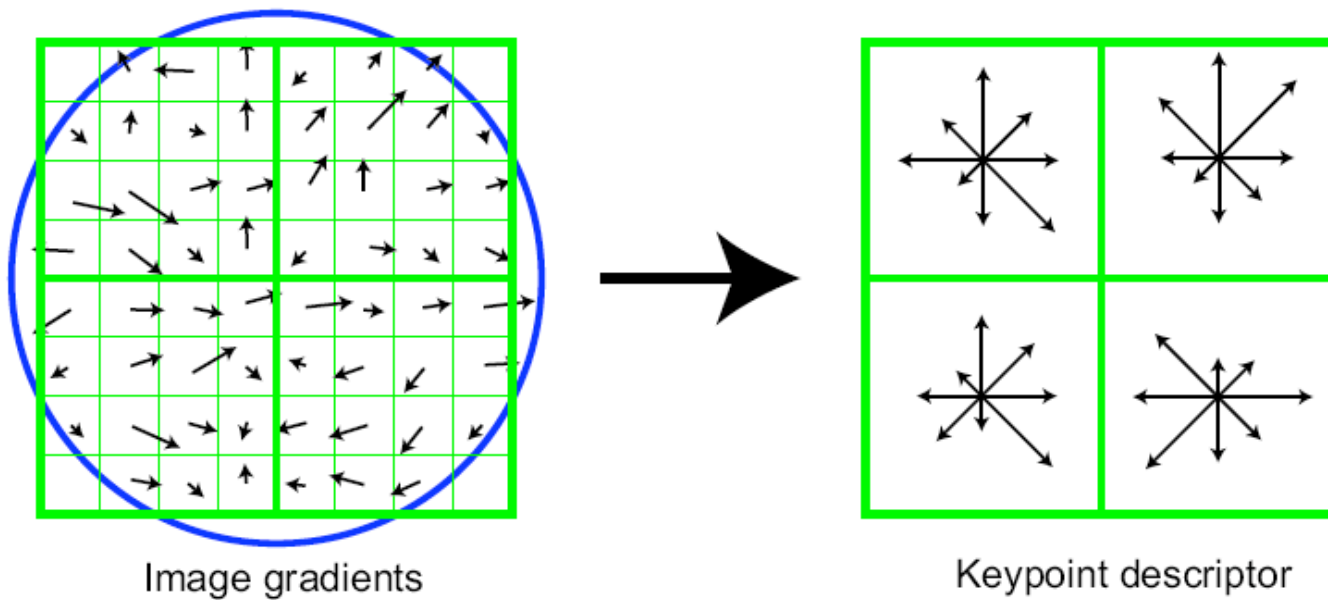


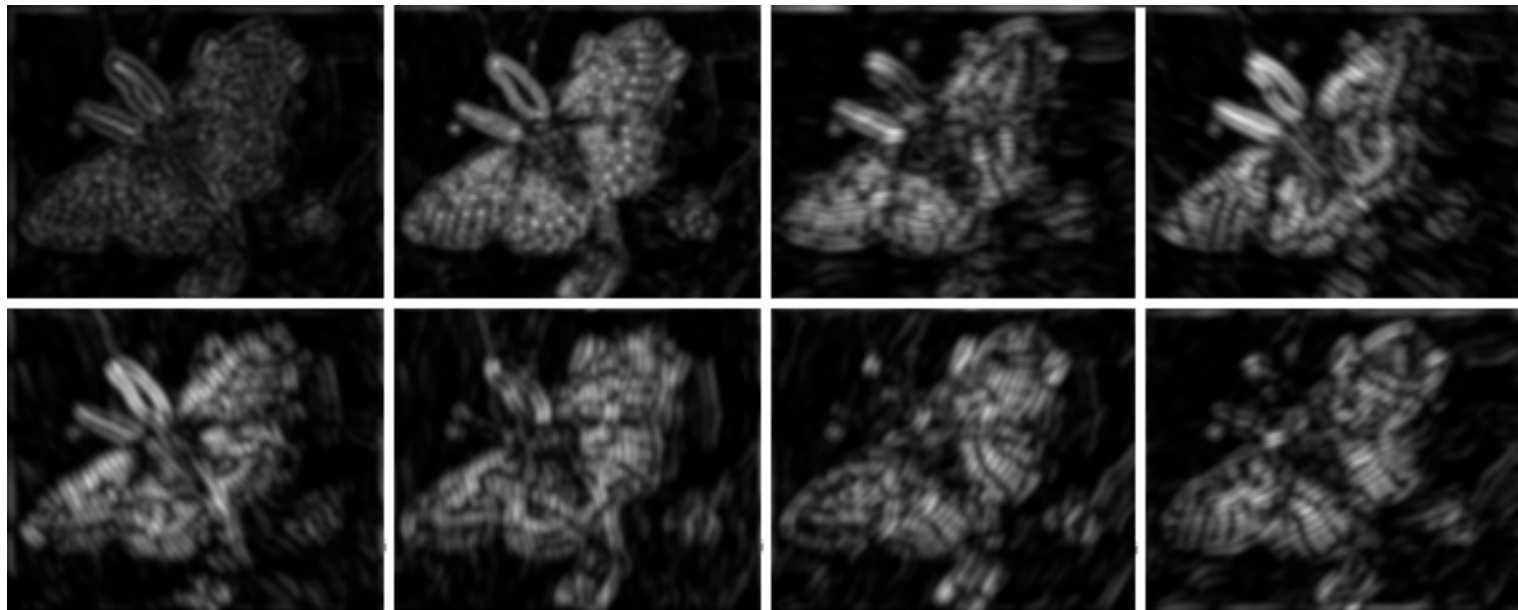
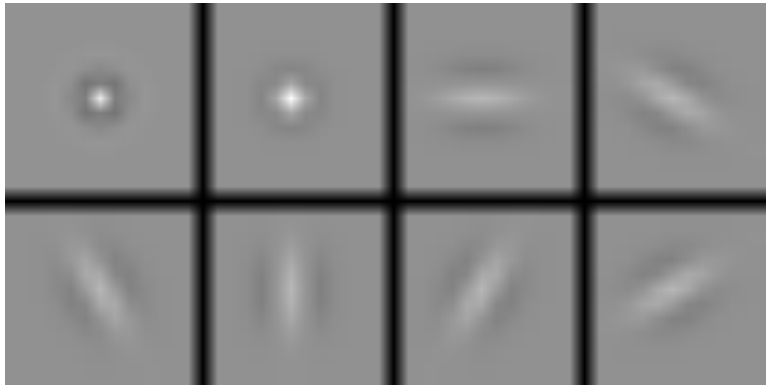
Fig 7 from:

Distinctive image features from scale-invariant keypoints

David G. Lowe, *International Journal of Computer Vision*, 60, 2 (2004), pp. 91-110.

Filter representations

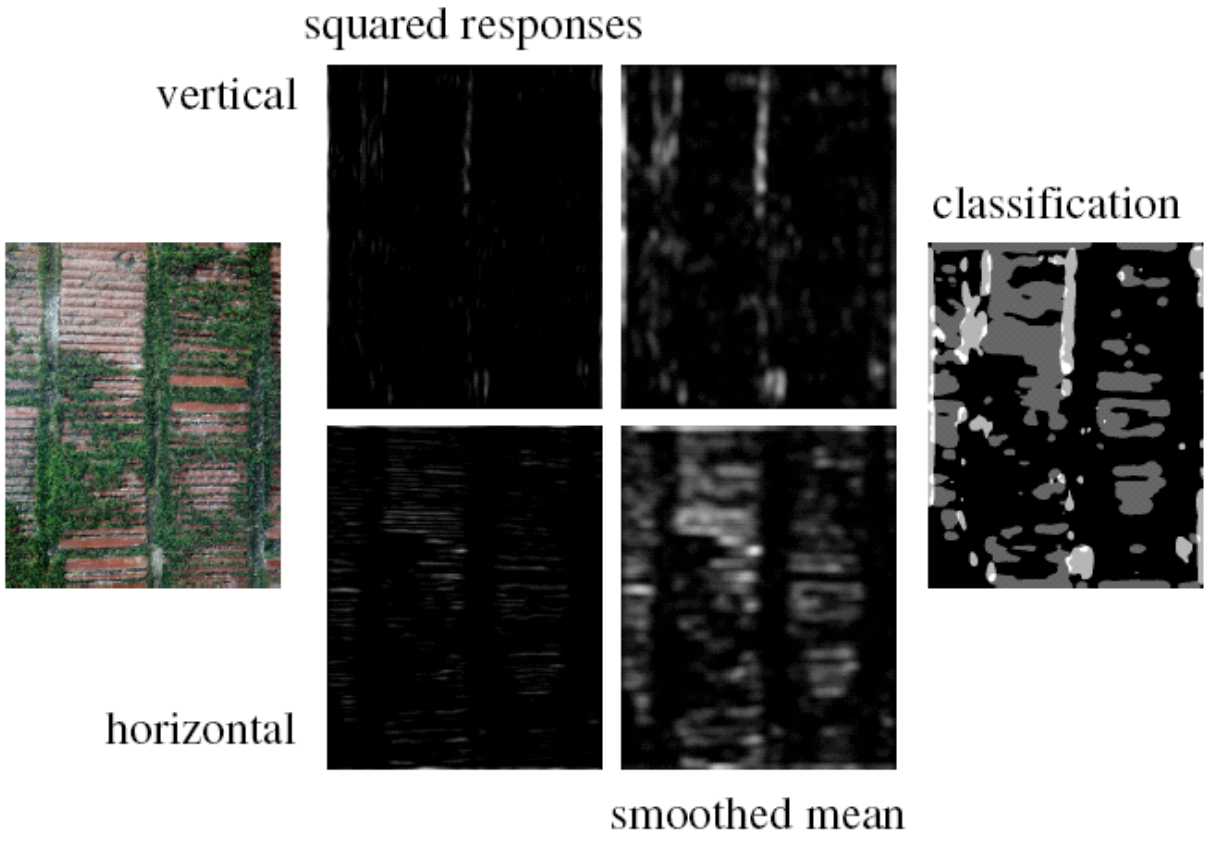
- Filters
 - A dot-product between a pattern and an image,
 - shifting the pattern across the image
 - strong response -> image locally looks like the pattern
 - e.g. gradients
 - e.g. spots and bars



Texture

- Describe by statistics of filter outputs
- Issues
 - describing textures
 - for region segmentation
 - inferring shape from texture
 - synthesizing textures





Shape from texture

- The foreshortening of textures gives us a cue to shape
 - But how?
 - Details remain rather murky



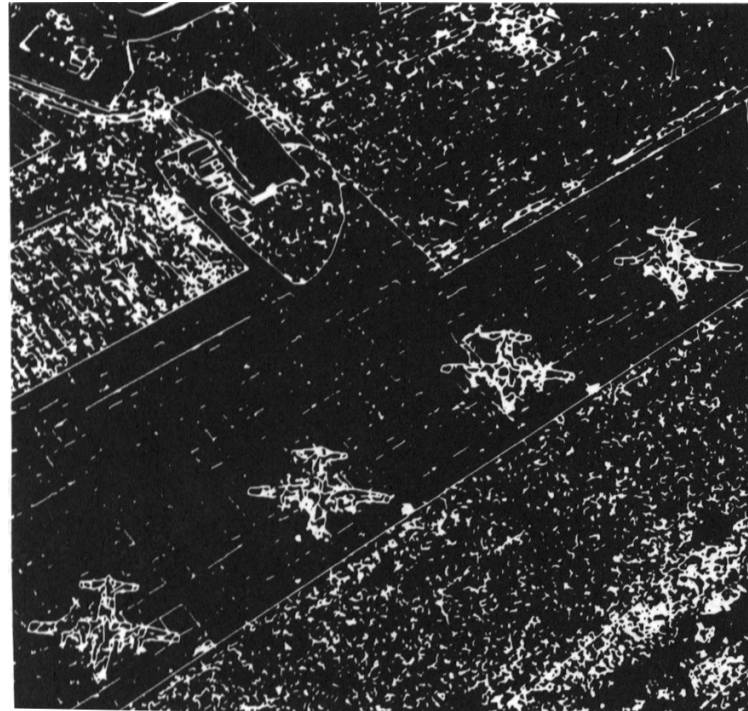
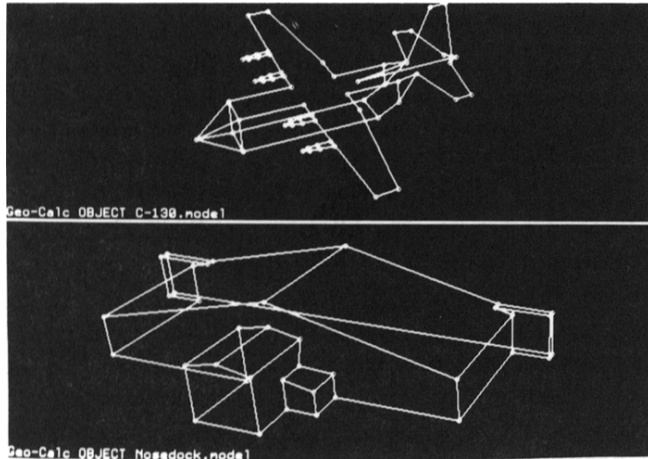
Some applications of recognition

- Digital libraries
 - Find me the pic of JFK and Marilyn Monroe embracing
 - NCMEC
- Surveillance
 - Warn me if there is a mugging in the grove
- HCI
 - Do what I show you
- Military
 - Shoot this, not that

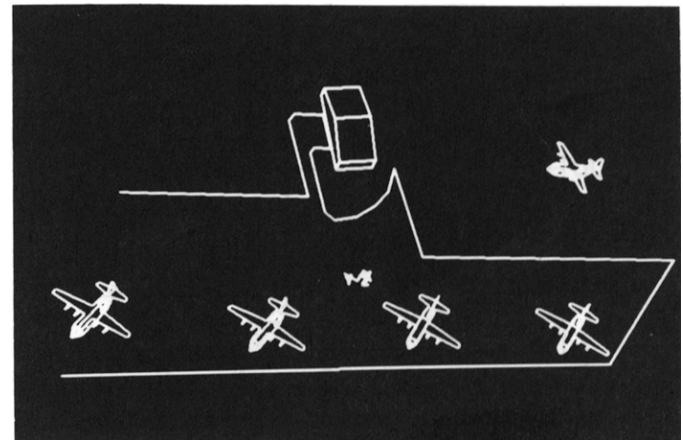
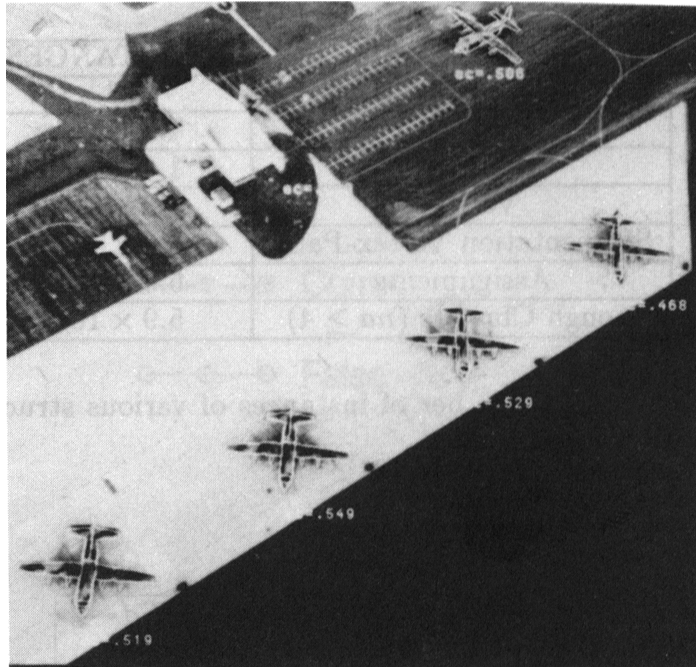
What are the problems in recognition?

- Which bits of image should be recognised together?
 - Segmentation.
- How can objects be recognised without focusing on detail?
 - Abstraction.
- How can objects with many free parameters be recognised?
 - No popular name, but it's a crucial problem anyhow.
- How do we structure very large modelbases?
 - again, no popular name; abstraction and learning come into this

History



History-II




Segmentation


- Which image components “belong together”?
- Belong together=lie on the same object
- Cues
 - similar colour
 - similar texture
 - not separated by contour
 - form a suggestive shape when assembled

Netscape: Blobworld Query Results: image #108019 (Preferred)

File Edit View Go Communicator Help



Query image: 108019




Query blobs

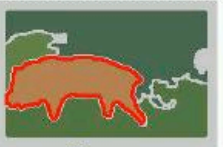
feature importance:

	overall	color	texture	location	shape
blob	very	very	somewhat	not	not
background	somewhat	very	not	not	not


Querying from 35000 images (2000 returned by the filter).




1: 108044 (score = 0.99)




[New query](#)




2: 108023 (score = 0.98)




[New query](#)




3: 108006 (score = 0.98)




[New query](#)




4: 108029 (score = 0.98)




[New query](#)




5: 108051 (score = 0.98)




[New query](#)




6: 108084 (score = 0.97)




[New query](#)



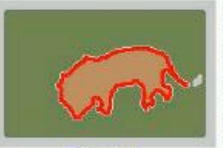
7: 108037 (score = 0.97)



[New query](#)



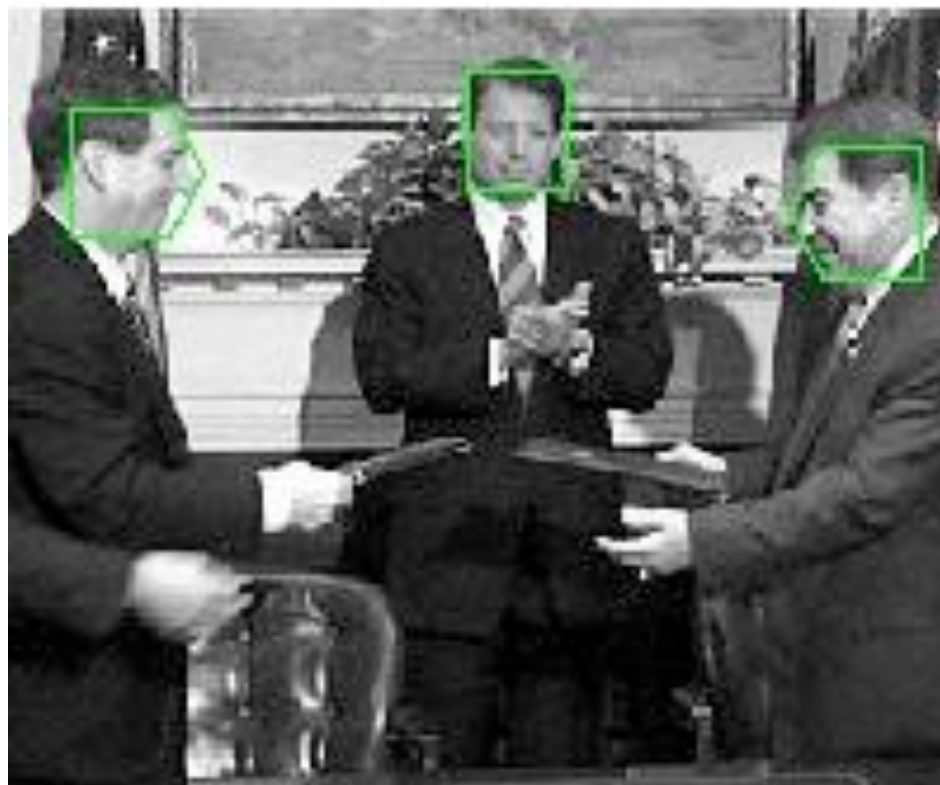
8: 108004 (score = 0.97)



[New query](#)

Matching templates

- Some objects are 2D patterns
 - e.g. faces
- Build an explicit pattern matcher
 - discount changes in illumination by using a parametric model
 - changes in background are hard
 - changes in pose are hard



http://www.ri.cmu.edu/projects/project_320.html

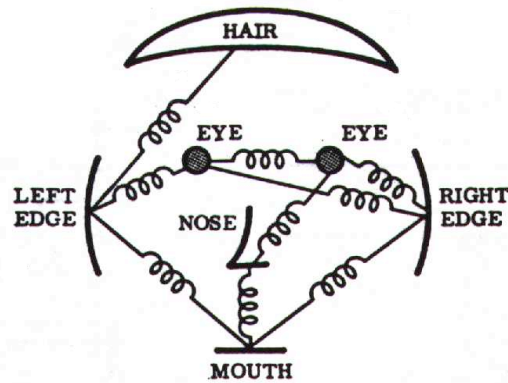


http://www.ri.cmu.edu/projects/project_271.html

Relations between templates

- e.g. find faces by
 - finding eyes, nose, mouth
 - finding assembly of the three that has the “right” relations

Constellations of parts



Fischler & Elschlager 1973

Yuille '91

Brunelli & Poggio '93

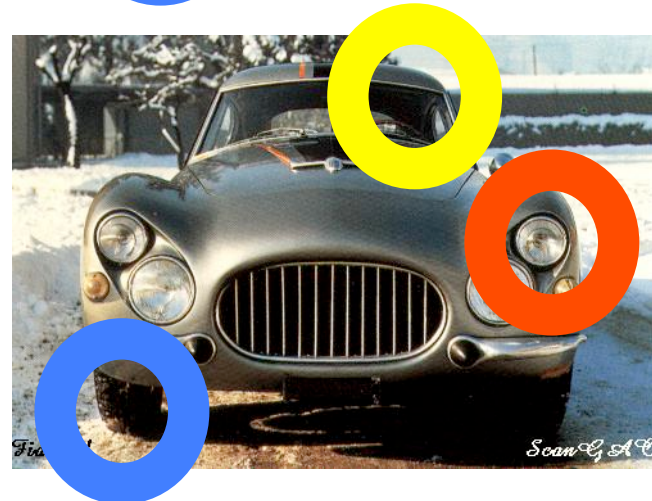
Lades, v.d. Malsburg et al. '93

Cootes, Lanitis, Taylor et al. '95

Amit & Geman '95, '99

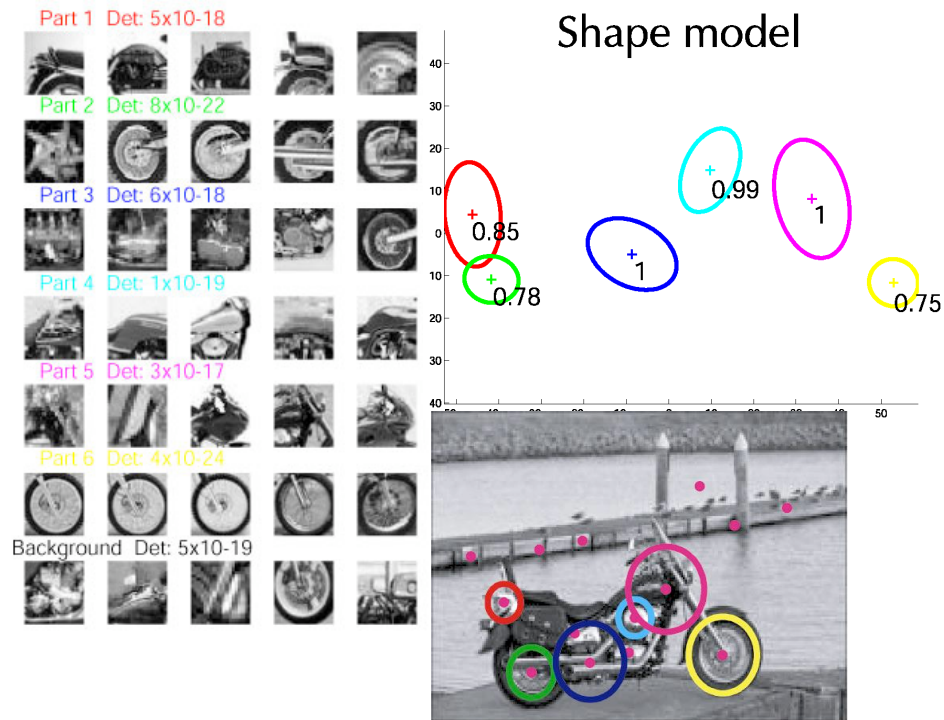
Perona et al. '95, '96, '98, '00

Agarwal & Roth '02



Typical models

Motorbikes



Spotted cats

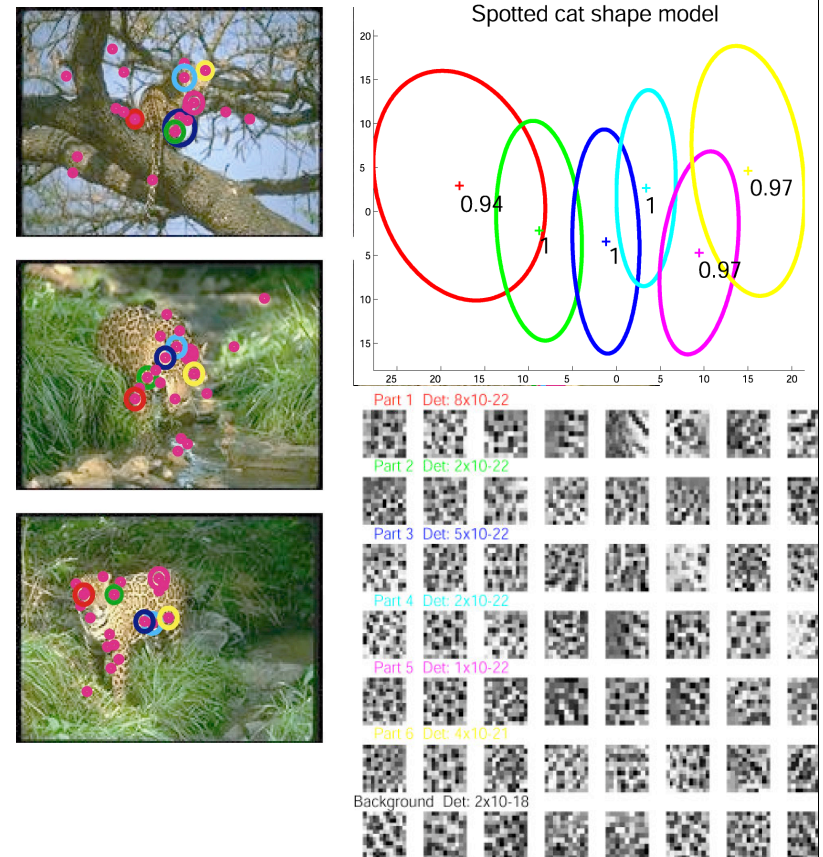


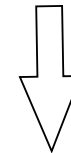
Figure after Fergus et al, 03; see also Fergus et al, 04

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



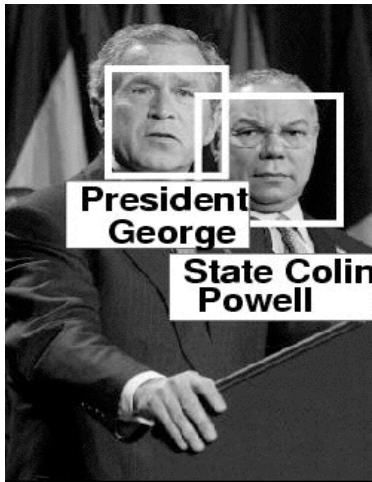
Data examples



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)



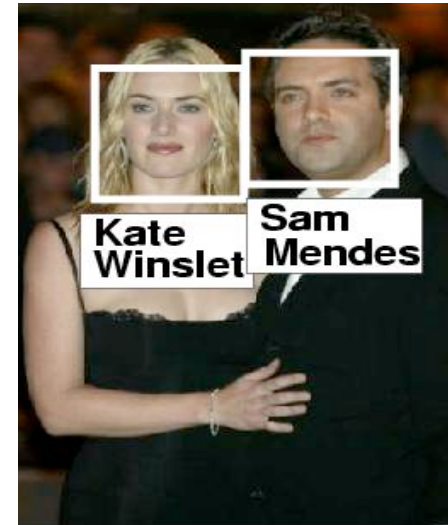
President George W. Bush waves as he leaves the White House for a day trip to North Carolina, July 25, 2002. A White House spokesman said that Bush would be compelled to veto Senate legislation creating a new department of homeland security unless changes are made. (Kevin Lamarque/Reuters)



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

Tracking

- Use a model to predict next position and refine using next image
- Model:
 - simple dynamic models (second order dynamics)
 - kinematic models
 - etc.
- Face tracking and eye tracking now work rather well
- People tracking is hard, but important and do-able

