Why study Computer Vision?

- Images and movies are everywhere
- Fast-growing collection of useful applications
  - building representations of the 3D world from pictures
  - automated surveillance (who's doing what)
  - movie post-processing
  - face finding
- Various deep and attractive scientific mysteries
  - how does object recognition work?
- Greater understanding of human vision

Properties of Vision

- One can "see the future"
  - Cricketers avoid being hit in the head
    - There's a reflex --- when the right eye sees something going left, and the left eye sees something going right, move your head fast.
  - Gannets pull their wings back at the last moment
    - Gannets are diving birds; they must steer with their wings, but wings break unless pulled back at the moment of contact.
    - Area of target over rate of change of area gives time to contact.

Properties of Vision

- 3D representations are easily constructed
  - There are many different cues.
  - Useful
    - to humans (avoid bumping into things; planning a grasp; etc.)
    - in computer vision (build models for movies).
  - Cues include
    - multiple views (motion, stereopsis)
    - texture
    - shading

Properties of Vision

- People draw distinctions between what is seen
  - “Object recognition”
  - This could mean “is this a fish or a bicycle?”
  - It could mean “is this George Washington?”
  - It could mean “is this poisonous or not?”
  - It could mean “is this slippery or not?”
  - It could mean “will this support my weight?”
  - Great mystery
    - How to build programs that can draw useful distinctions based on image properties.

Part I: The Physics of Imaging

- How images are formed
  - Cameras
    - What a camera does
    - How to tell where the camera was
  - Light
    - How to measure light
    - What light does at surfaces
    - How the brightness values we see in cameras are determined
  - Color
    - The underlying mechanisms of color
    - How to describe it and measure it

Part II: Early Vision in One Image

- Representing small patches of image
  - For three reasons
    - We wish to establish correspondence between (say) points in different images, so we need to describe the neighborhood of the points
    - Sharp changes are important in practice --- known as "edges"
    - Representing texture by giving some statistics of the different kinds of small patch present in the texture.
      - Tigers have lots of bars, few spots
      - Leopards are the other way

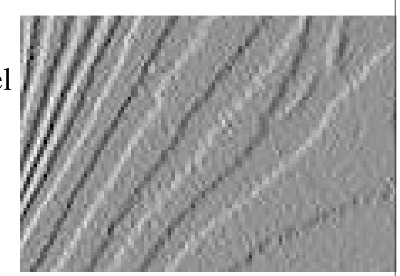
Representing an image patch

- Filter outputs
  - essentially form a dot-product between a pattern and an image, while shifting the pattern across the image
  - strong response -> image locally looks like the pattern
  - e.g. derivatives measured by filtering with a kernel that looks like a big derivative (bright bar next to dark bar)

Convolve this image

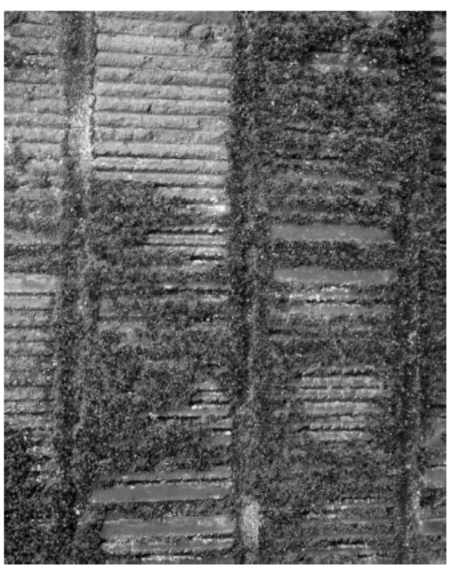
With this kernel

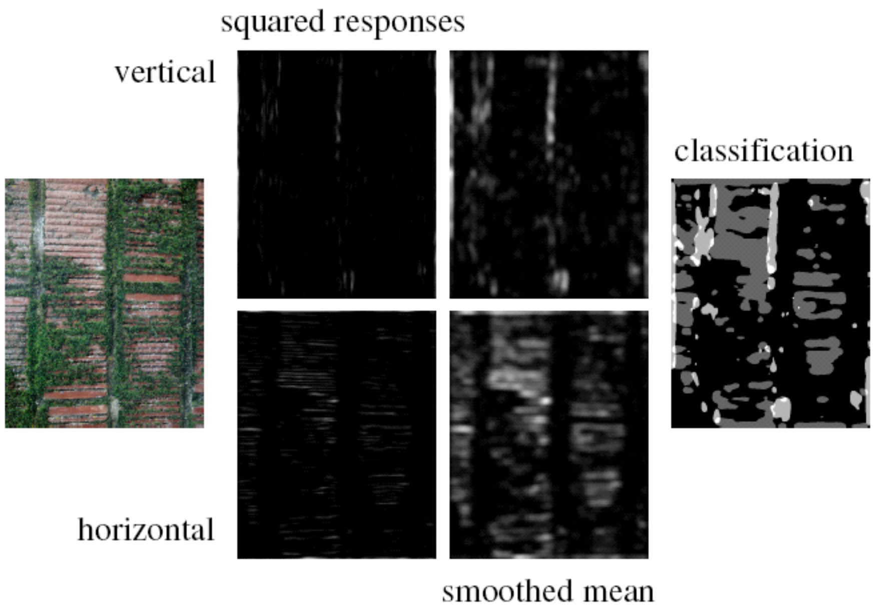
To get this



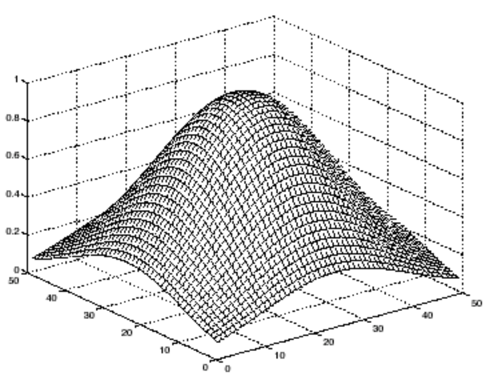
Texture

- Many objects are distinguished by their texture
  - Tigers, cheetahs, grass, trees
- We represent texture with statistics of filter outputs
  - For tigers, bar filters at a coarse scale respond strongly
  - For cheetahs, spots at the same scale
  - For grass, long narrow bars
  - For the leaves of trees, extended spots
- Objects with different textures can be segmented
- The variation in textures is a cue to shape



Shape from texture



Part III: Early Vision in Multiple Images

- The geometry of multiple views
  - Where could it appear in camera 2 (3, etc.) given it was here in 1 (1 and 2, etc.)?
- Stereopsis
  - What we know about the world from having 2 eyes
- Structure from motion
  - What we know about the world from having many eyes
    - or, more commonly, our eyes moving.

Part IV: Mid-Level Vision

- Finding coherent structure so as to break the image or movie into big units
  - Segmentation:
    - Breaking images and videos into useful pieces
    - E.g. finding video sequences that correspond to one shot
    - E.g. finding image components that are coherent in internal appearance
  - Tracking:
    - Keeping track of a moving object through a long sequence of views

Part V: High Level Vision (Geometry)

- The relations between object geometry and image geometry
  - Model based vision
    - find the position and orientation of known objects
  - Smooth surfaces and outlines
    - how the outline of a curved object is formed, and what it looks like
  - Aspect graphs
    - how the outline of a curved object moves around as you view it from different directions
  - Range data

Part VI: High Level Vision (Probabilistic)

- Using classifiers and probability to recognize objects
  - Templates and classifiers
    - how to find objects that look the same from view to view with a classifier
  - Relations
    - break up objects into big, simple parts, find the parts with a classifier, and then reason about the relationships between the parts to find the object.
  - Geometric templates from spatial relations
    - extend this trick so that templates are formed from relations between much smaller parts

3D Reconstruction from multiple views

- Multiple views arise from
  - stereo
  - motion
- Strategy
  - "triangulate" from distinct measurements of the same thing
- Issues
  - Correspondence: which points in the images are projections of the same 3D point?
  - The representation: what do we report?
  - Noise: how do we get stable, accurate reports

Part VII: Some Applications in Detail

- Finding images in large collections
  - searching for pictures
  - browsing collections of pictures
- Image based rendering
  - often very difficult to produce models that look like real objects
    - surface weathering, etc., create details that are hard to model
    - Solution: make new pictures from old

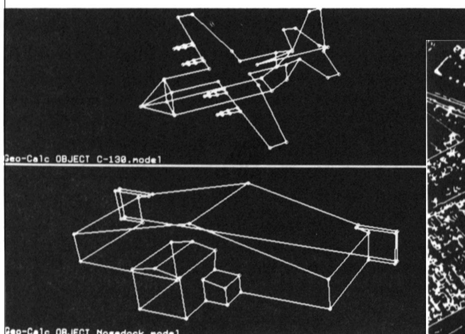
Some applications of recognition

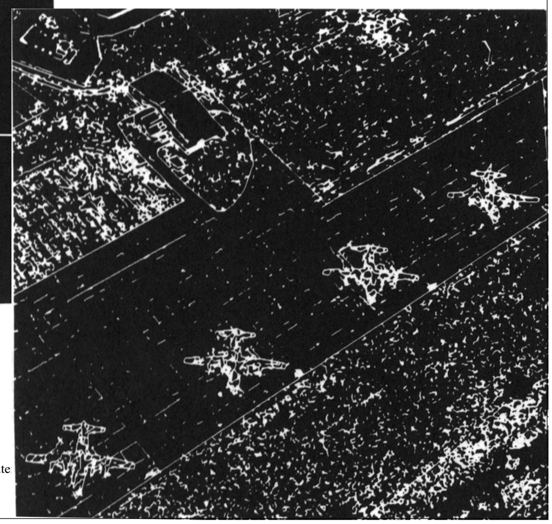
- Digital libraries
  - Find me the pic of JFK and Marilyn Monroe embracing
  - NCMEC
- Surveillance
  - Warn me if there is a mugging in the grove
- HCI
  - Do what I show you
- Military
  - Shoot this, not that

What are the problems in recognition?

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

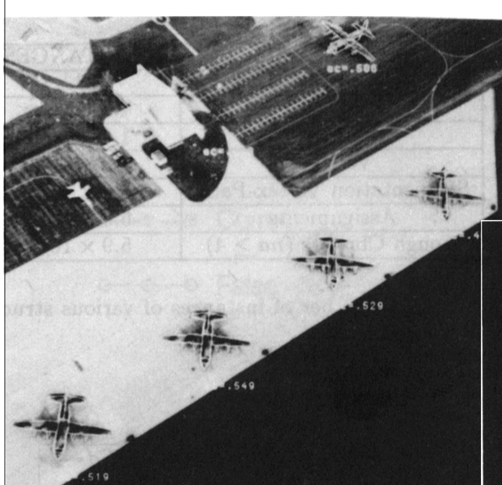
History

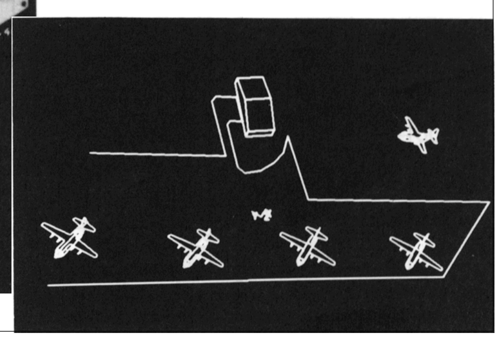




Compute

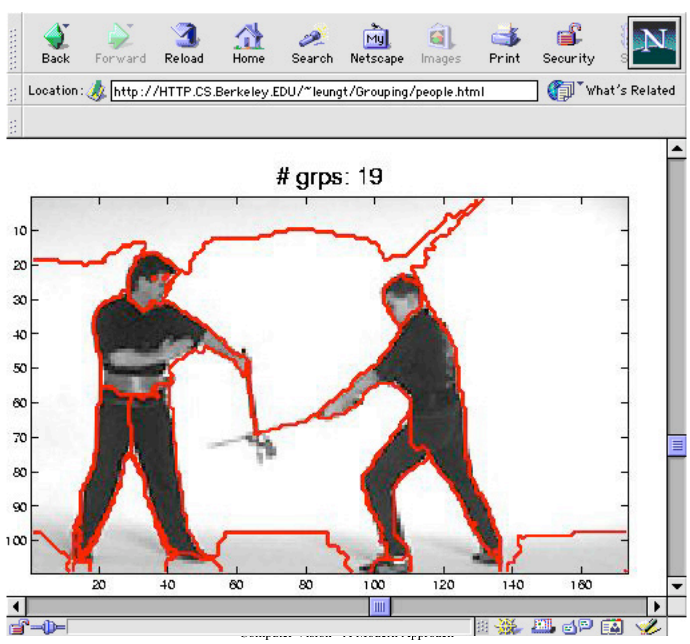
History-II

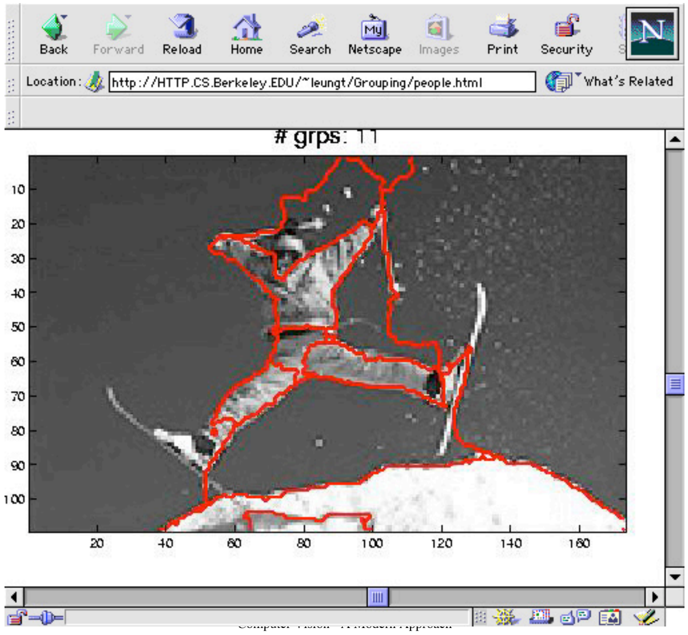


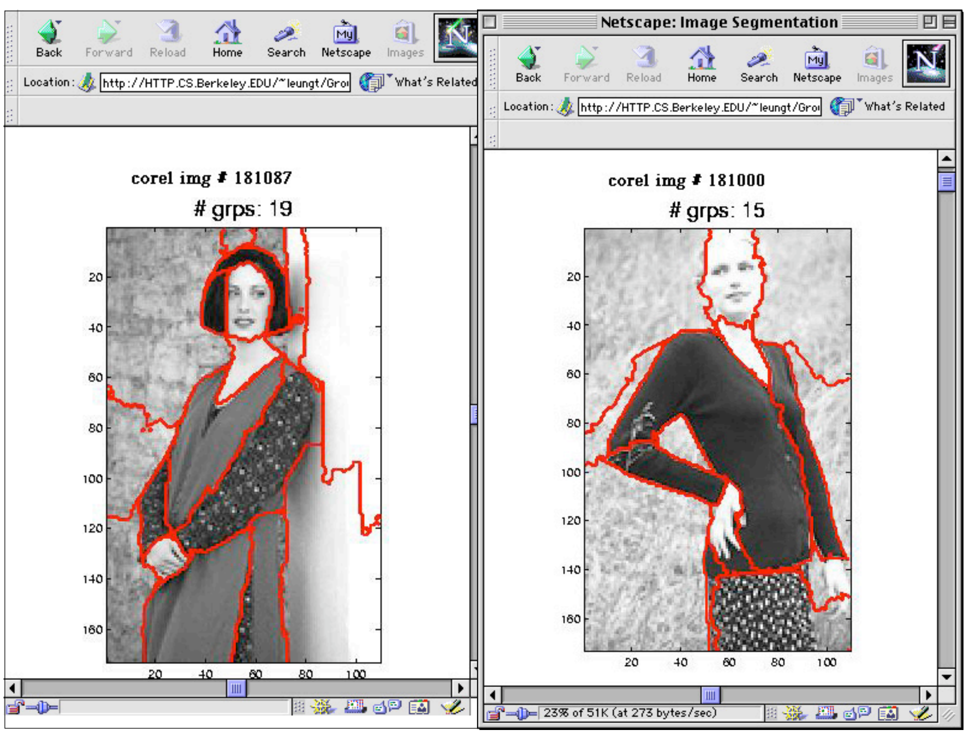


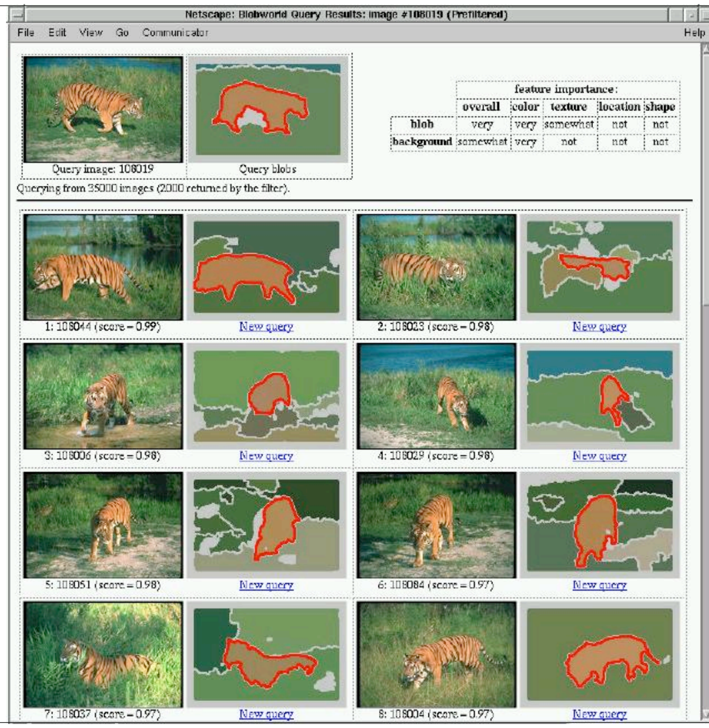
Segmentation

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



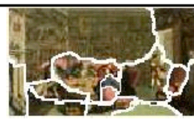
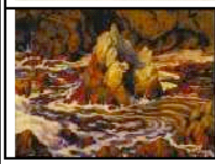

Back
Forward
Reload
Home
Search
Netscape
Images
Print
Security
Location:
http://HTTP.CS.Berkeley.EDU/~leungt/Grouping/people.html
What's Related
grps: 11

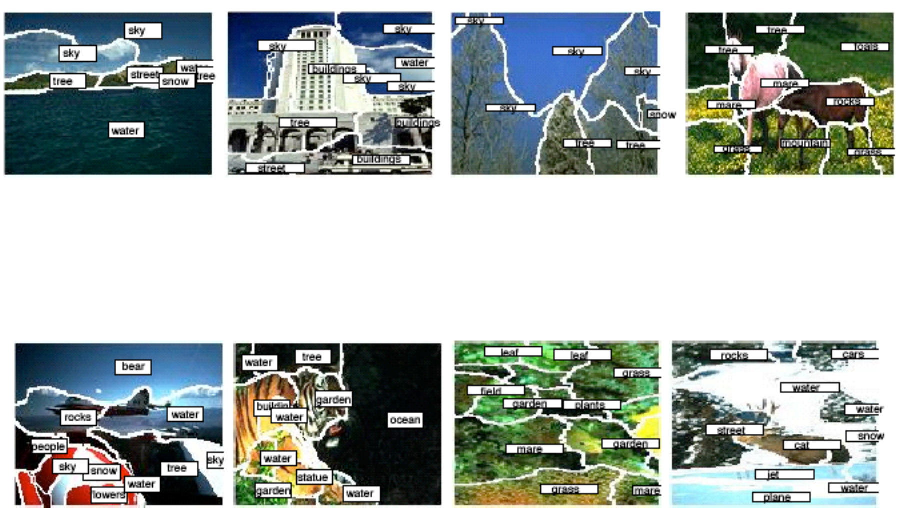

Netscape: Image Segmentation
corel img # 181087
grps: 19
corel img # 181000
grps: 15
23% of 51K (at 273 bytes/sec)



large importance attached fact old dutch century more command whale ship was per son was divided officer word means fat cutter time made days was general vessel whale hunting concern british title old dutch official present rank such more good american officer boat night watch ground command ship deck grand political sea men mast way professional superior

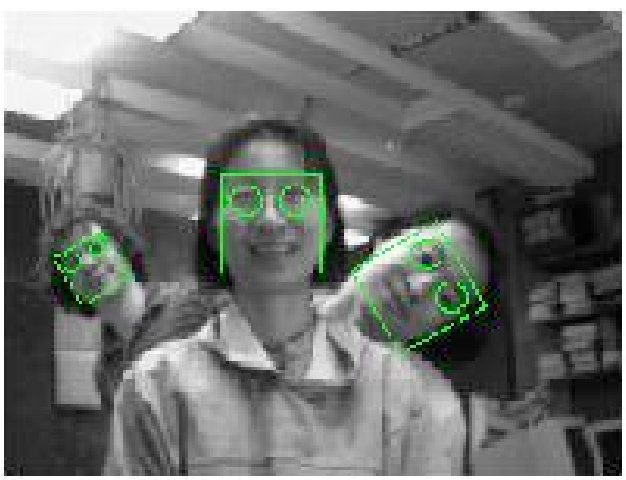
"The large importance attached to the harpooneer's vocation is evinced by the fact, that originally in the old Dutch Fishery, two centuries and more ago, the command of a whale-ship was not wholly lodged in the person now called the captain, but was divided between him and an officer called the Specksynder. Literally this word means Fat-Cutter; usage, however, in time made it equivalent to Chief Harpooneer. In those days, the captain's authority was restricted to the navigation and general management of the vessel; while over the whale-hunting department and all its concerns, the Specksynder or Chief Harpooneer reigned supreme. In the British Greenland Fishery, under the corrupted title of Specksioneer, this old Dutch official is still retained, but his former dignity is sadly abridged. At present he ranks simply as senior Harpooneer; and as such, is but one of the captain's more inferior subalterns. Nevertheless, as upon the good conduct of the harpooneers the success of a whaling voyage largely depends, and since …"

		Associated Words KUSATSU SERIES STATION TOKAIDO TOKAIDO GOJUSANTSUGI PRINT HIROSHIGE Predicted Words (rank order) tokaido print hiroshige object artifact series ordering gojusantsugi station facility arrangement minakuchi sakanoshita maisaka a
		Associated Words SYNTAX LORD PRINT ROWLANDSON Predicted Words (rank order) rowlandson print drawing life_form person object artifact expert art creation animal graphic_art painting structure view
		Associated Words DRAWING ROCKY SEA SHORE Predicted Words (rank order) print hokusai kunisada object artifact huge process natural_process district administrative_district state_capital rises



Matching templates

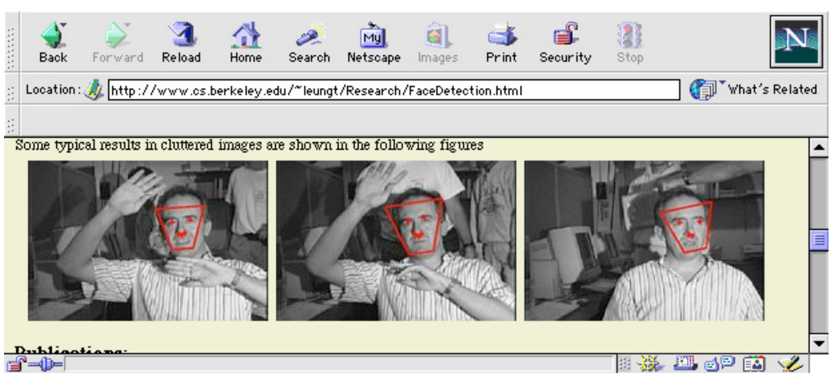
- Some objects are 2D patterns
  - e.g. faces
- Build an explicit pattern matcher
  - discount changes in illumination by using a parametric model
  - changes in background are hard
  - changes in pose are hard



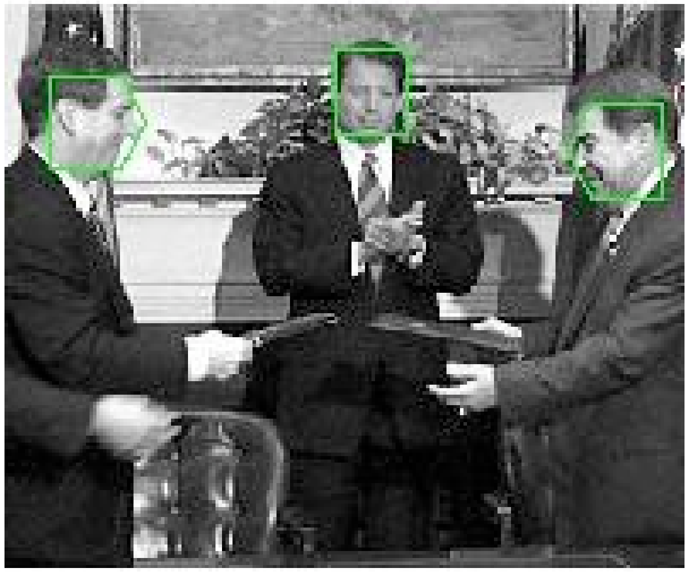
http://www.ri.cmu.edu/projects/project_271.html

Relations between templates

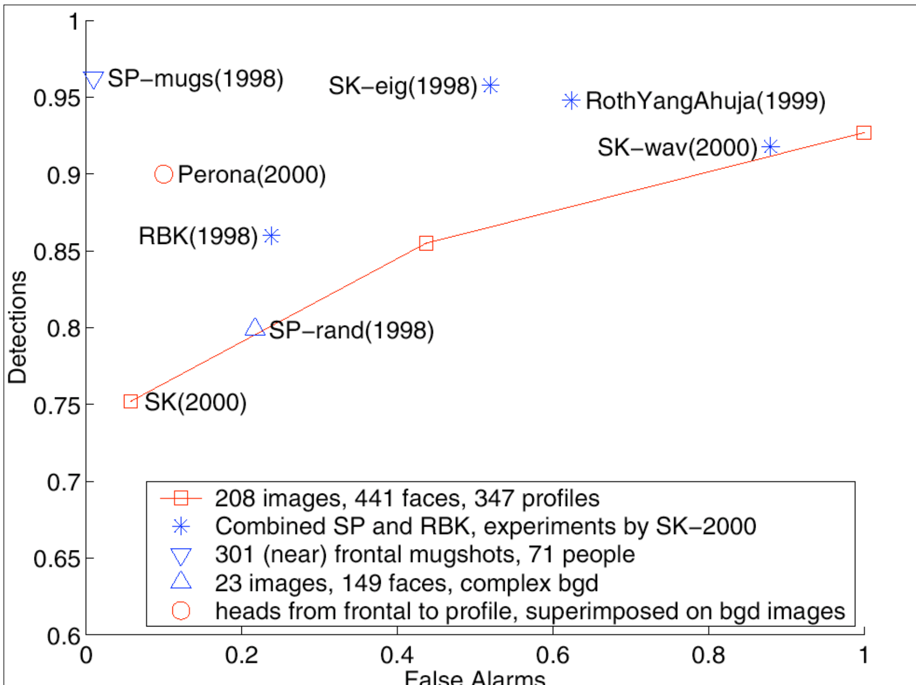
- e.g. find faces by
  - finding eyes, nose, mouth
  - finding assembly of the three that has the "right" relations



Computer Vision - A Modern Approach



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



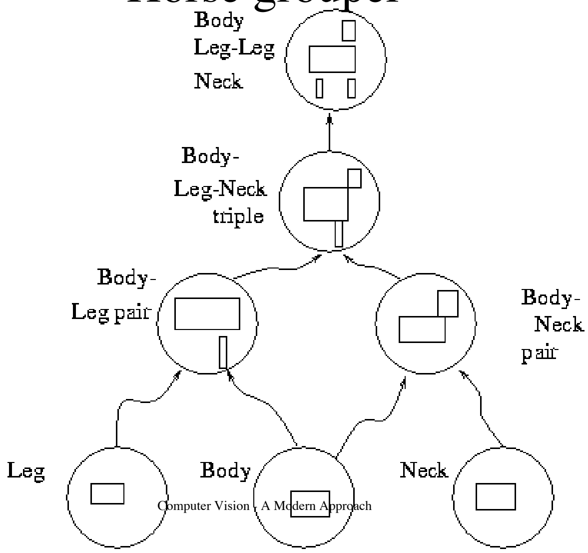
Representing the 3D world

- Assemblies of primitives
  - fit parametric forms
  - Issues
    - what primitives?
    - uniqueness of representation
    - few objects are actual primitives
- Indexed collection of images
  - use interpolation to predict appearance between images
  - Issues
    - occlusion is a mild nuisance
    - structuring the collection can be tricky

People

- Skin is characteristic; clothing hard to segment
  - hence, people wearing little clothing
- Finding body segments:
  - finding skin-like (color, texture) regions that have nearly straight, nearly parallel boundaries
- Grouping process constructed by hand, tuned by hand using small dataset.
- When a sufficiently large group is found, assert a person is present

Horse grouper



Computer Vision - A Modern Approach

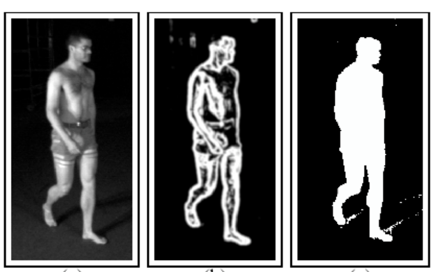
Returned data set

Computer Vision - A Modern Approach

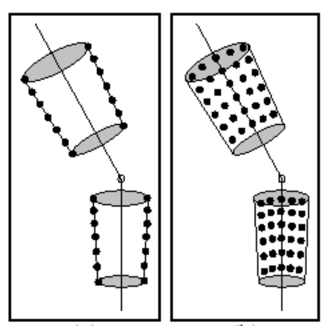
Tracking

- 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

The nasty likelihood



(a) (b) (c)



(a) (b)

About this class

- Image formation, relation between images and the world
- Features: color, edges, corners, lines, motion flow, histograms
- Segmentation and grouping
- Tracking, motion estimation
- Recognition and learning