Computer Vision

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

- 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 algorithms that can draw useful conclusions based on image properties.

Main topics

- Shape (and motion) recovery

  "What is the 3D shape of what I see?"
- Segmentation

  "What belongs together?"
- Tracking

  "Where does something go?"
- Recognition

  "What is it that I see?"

Main topics

- Camera & Light
  - Geometry, Radiometry, Color
- Digital images
  - Filters, edges, texture, optical flow
- Shape (and motion) recovery
  - Multi-view geometry
  - Stereo, motion, photometric stereo, …
- Segmentation
  - Clustering, model fitting, probalistic
- Tracking
  - Linear dynamics, non-linear dynamics
- Recognition
  - templates, relations between templates

Camera and lights

- 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

Digital images

- 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

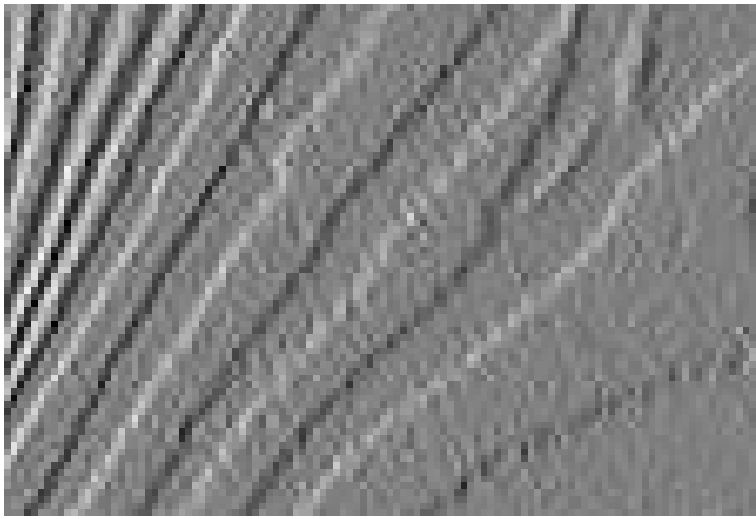
Representing an image patch

- 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

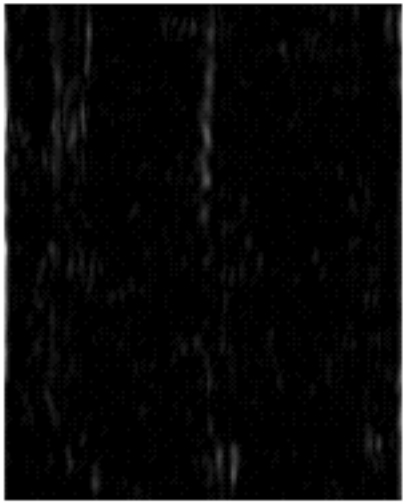


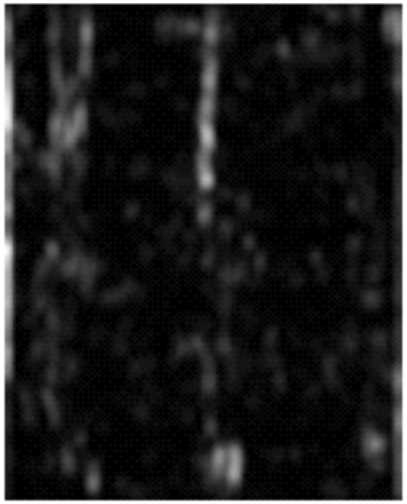
Texture

- 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

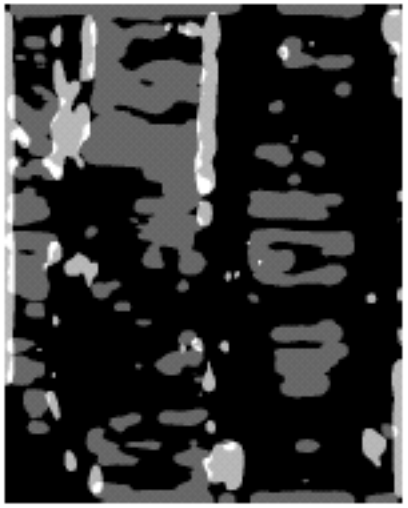
squared responses

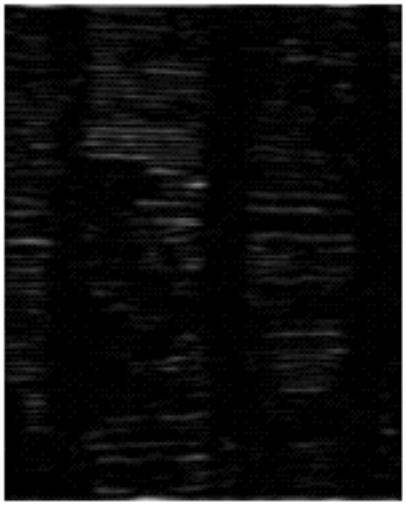
vertical

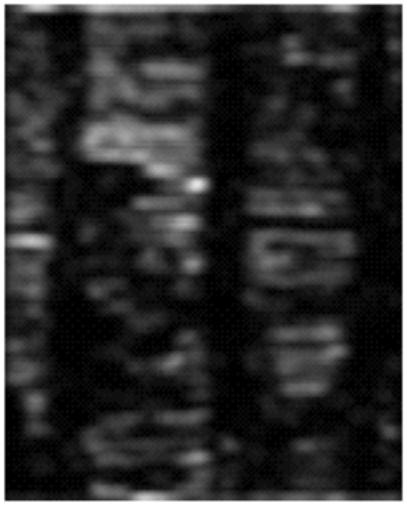




classification





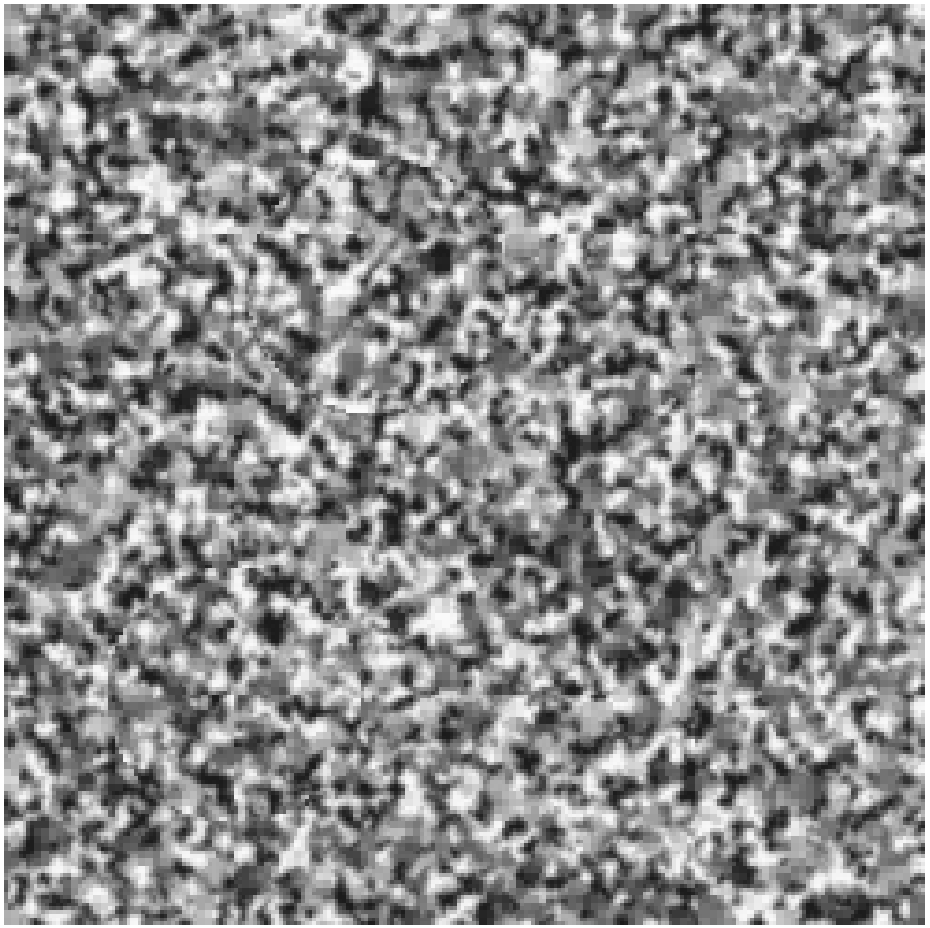


horizontal

smoothed mean

Optical flow

Where do pixels move?

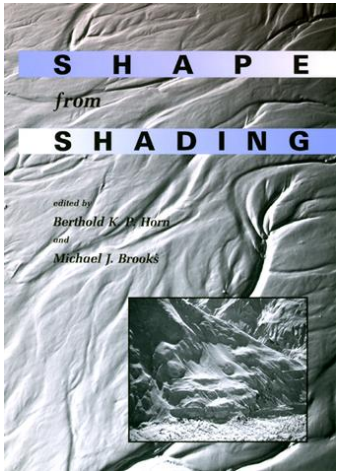


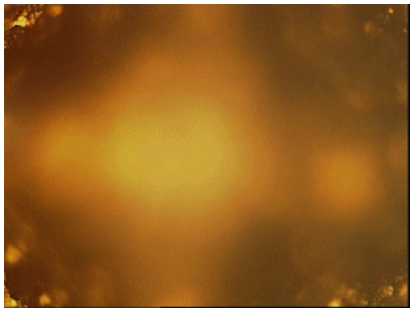
Movie special effects

Compute camera motion from point motion

Shape from …

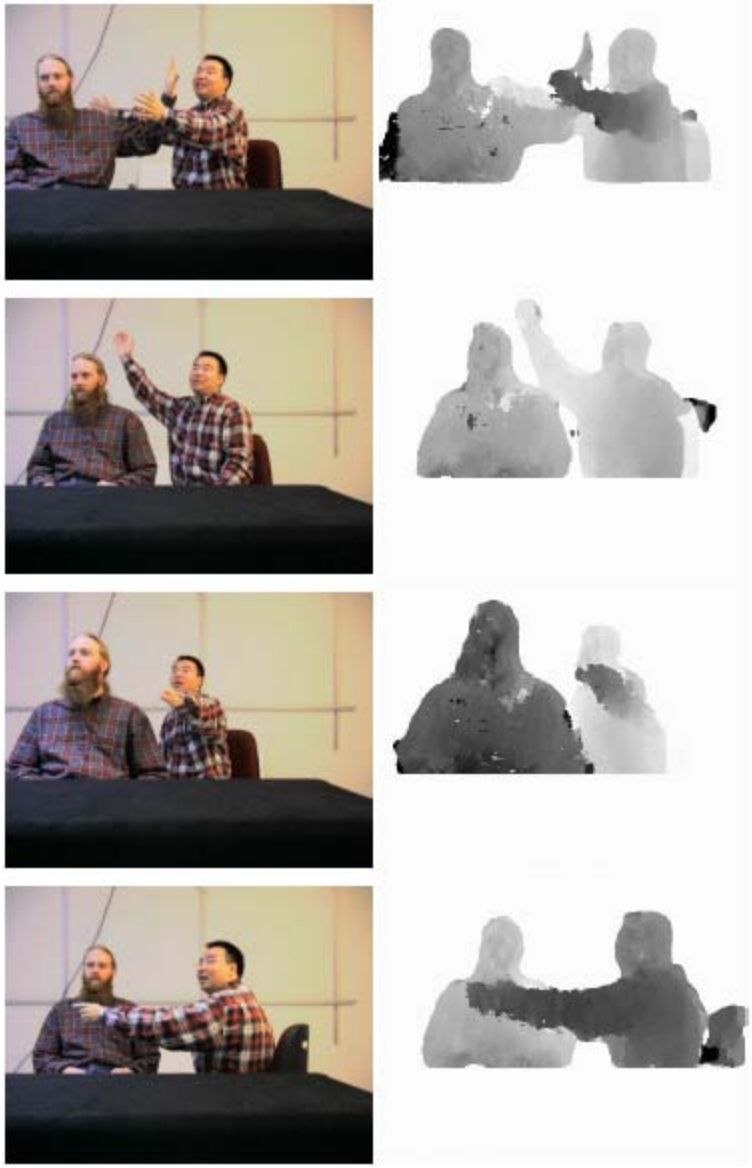
many different approaches/cues





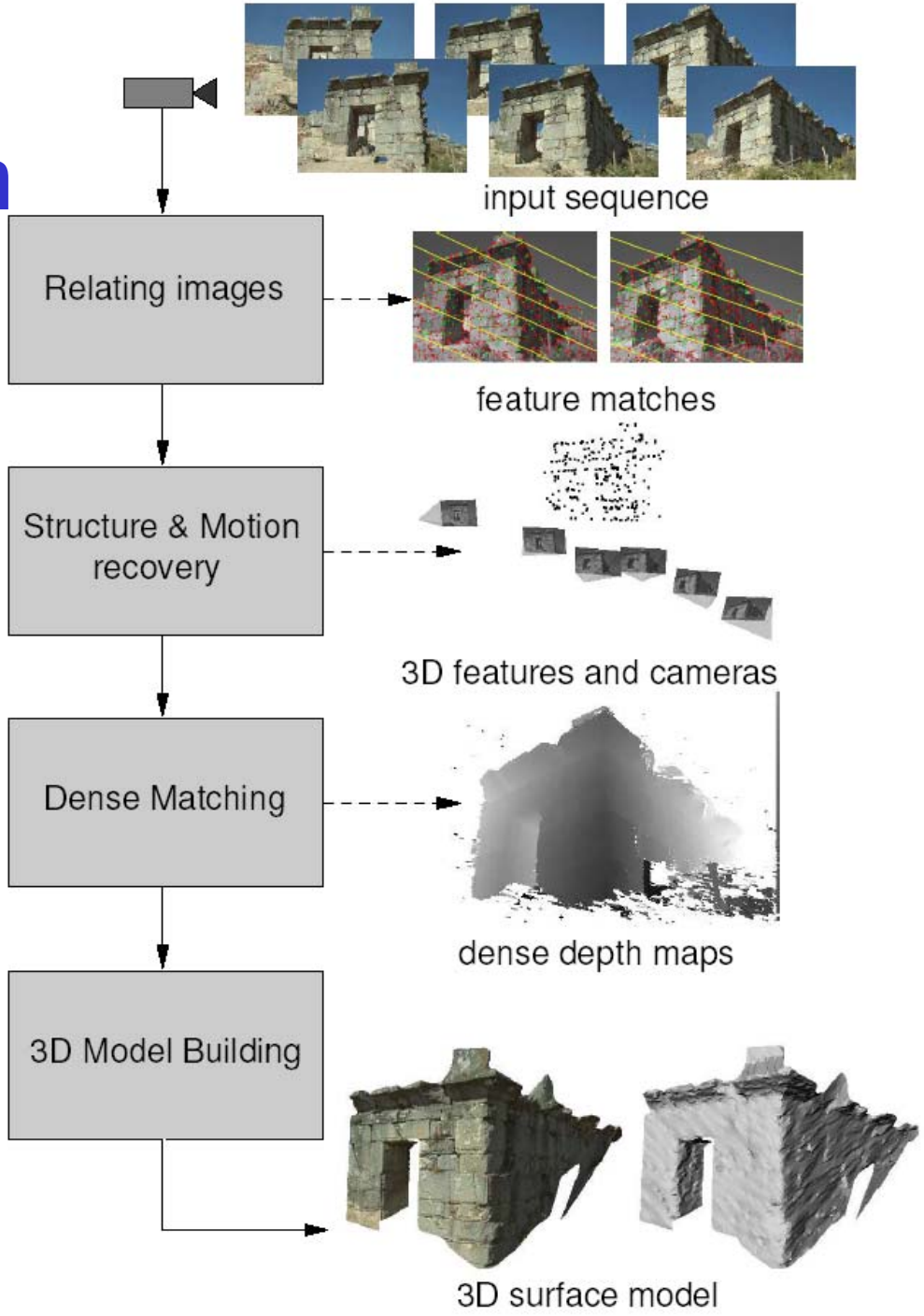
Motion, texture, shading, …

Real-time stereo *on GPU*



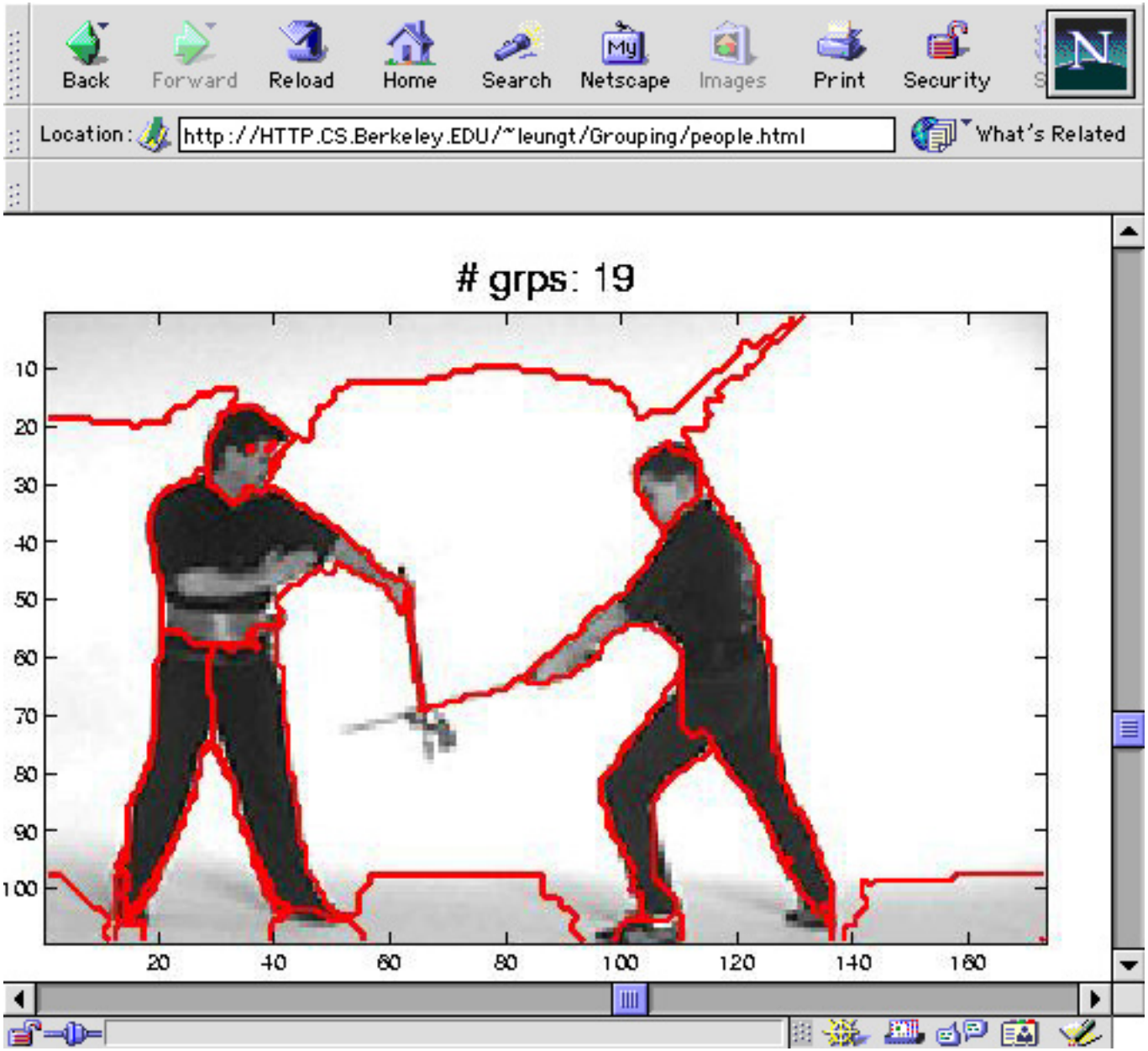
- Background differencing
- Stereo matching
- Depth reconstruction

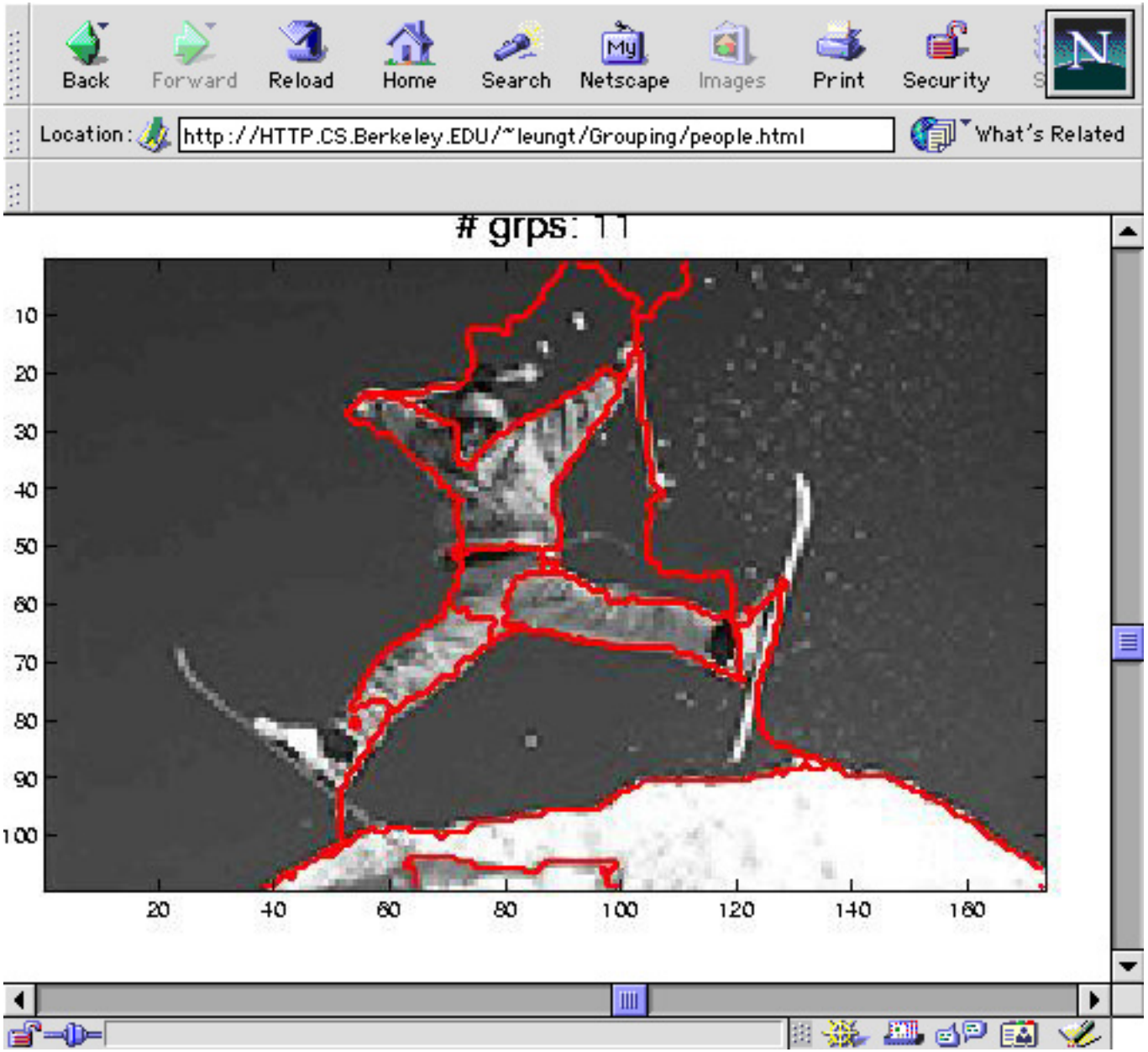
Structure from Motion



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


Back
Forward
Reload
Home
Search
Netscape
Images
Print
Security
Location :
http://HTTP.CS.Berkeley.EDU/~leungt/Grouping/people.html
What's Related
grps: 19
10
20
30
40
50
60
70
80
90
100
20
40
60
80
100
120
140
160

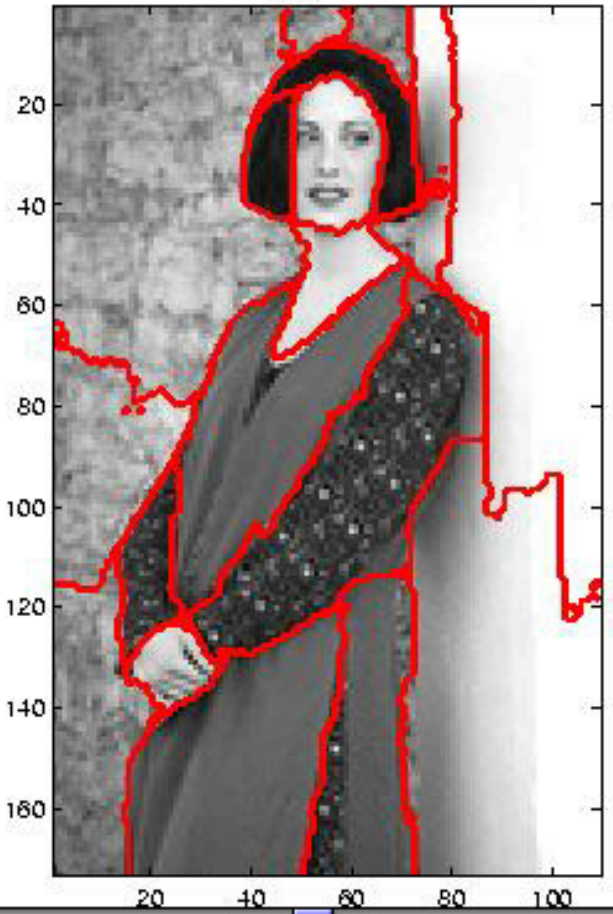

Back
Forward
Reload
Home
Search
Netscape
Images
Print
Security
Location :
http://HTTP.CS.Berkeley.EDU/~leungt/Grouping/people.html
What's Related
grps: 11
10
20
30
40
50
60
70
80
90
100
20
40
60
80
100
120
140
160

Back Forward Reload Home Search Netscape Images

Location : http://HTTP.CS.Berkeley.EDU/~leungt/Gro What's Related

corel img # 181087

grps: 19



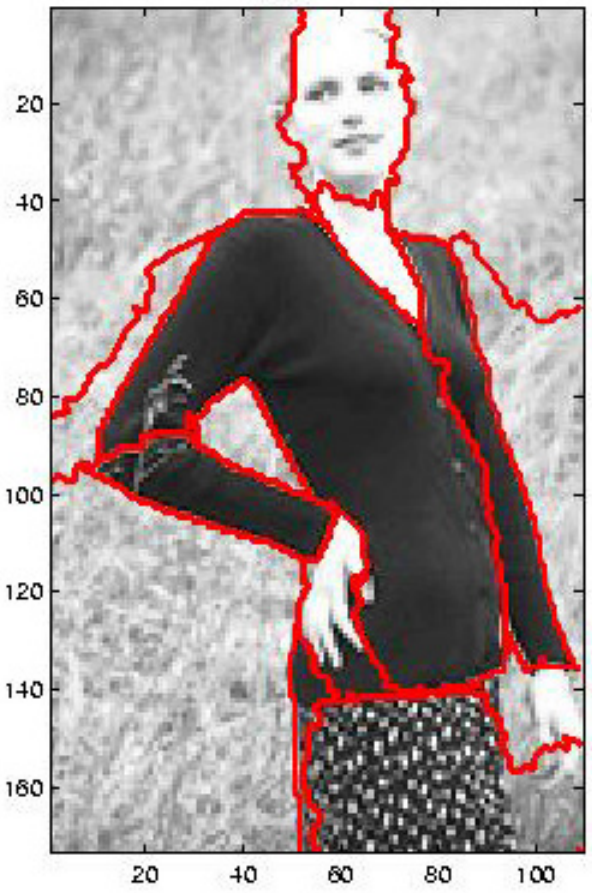
Netscape: Image Segmentation

Back Forward Reload Home Search Netscape Images

Location : http://HTTP.CS.Berkeley.EDU/~leungt/Gro What's Related

corel img # 181000

grps: 15



23% of 51K (at 273 bytes/sec)

CBIR

Content Based Image Retrieval

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

File Edit View Go Communicator Help

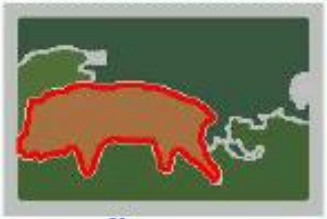
Query image: 108019

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

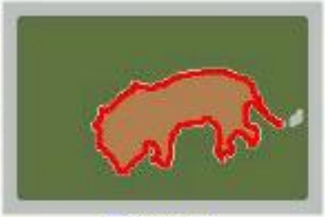
6: 108084 (score = 0.97)

New query

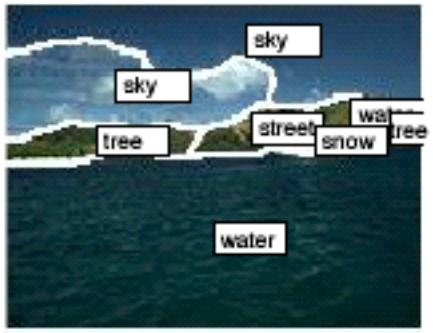
7: 108037 (score = 0.97)

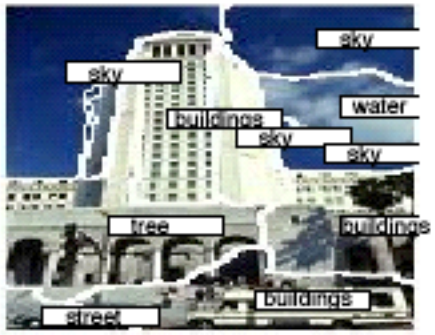
New query

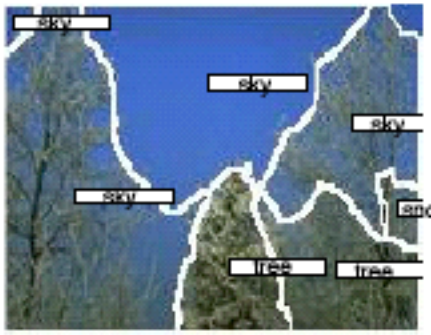
8: 108004 (score = 0.97)



New query

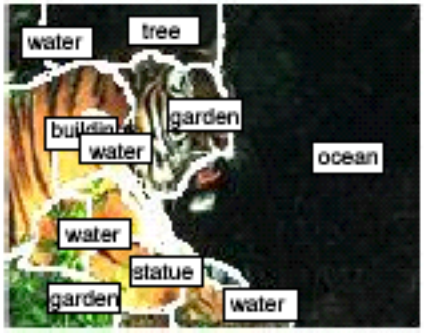

sky
sky
tree
street
snow
water

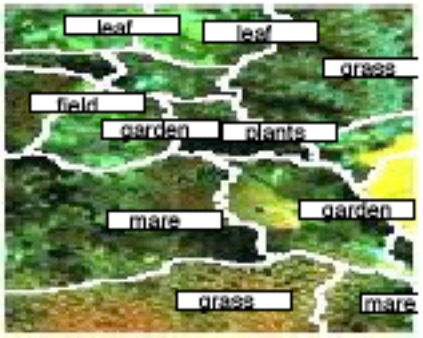

sky
sky
buildings
water
sky
sky
tree
buildings
buildings
street


sky
sky
sky
sky
snow
tree
tree

tree
tree
mare
mare
rocks
mountain

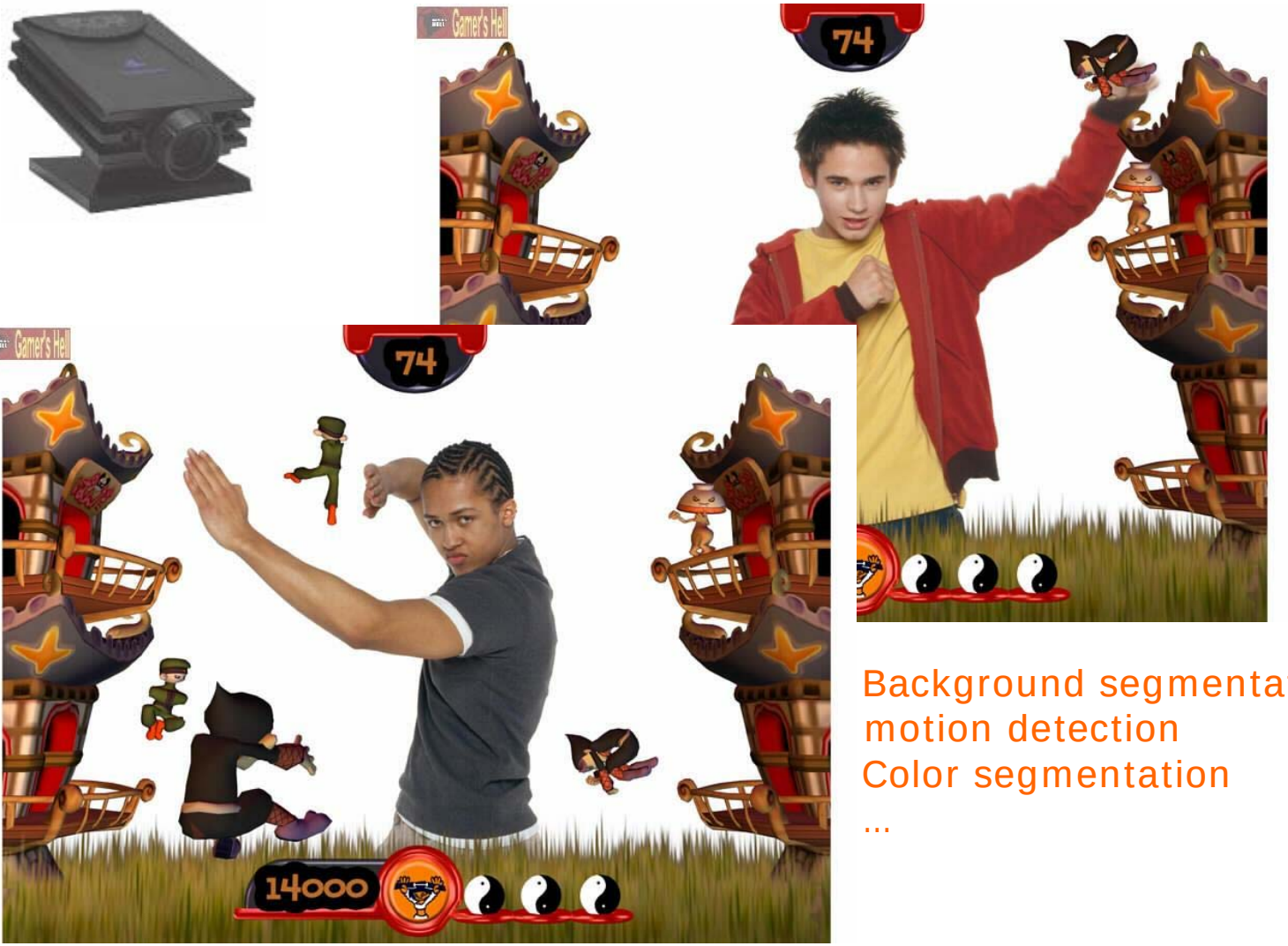
bear
rocks
water
people
sky
snow
tree
water
lowers


water
tree
garden
water
ocean
water
statue
garden
water


leaf
leaf
grass
garden
plants
mare
garden
grass
mare

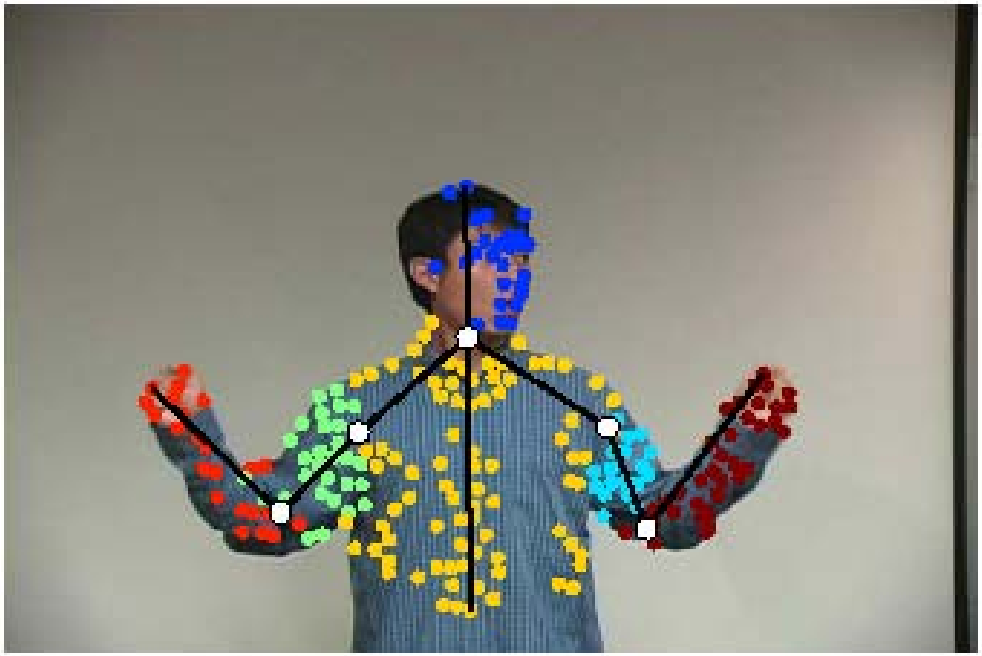
rocks
water
water
street
snow
cat
jet
water
plane

Sony's Eye Toy: Computer Vision for the masses

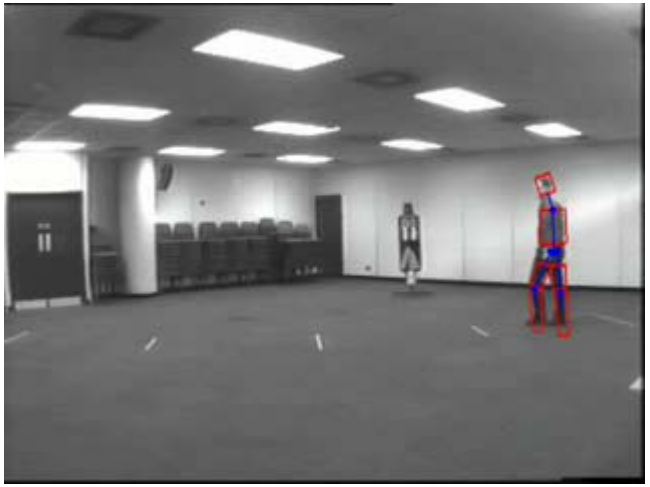


Background segmentation/ motion detection
Color segmentation
…

Also motion segmentation, etc.



Tracking examples



Object recognition

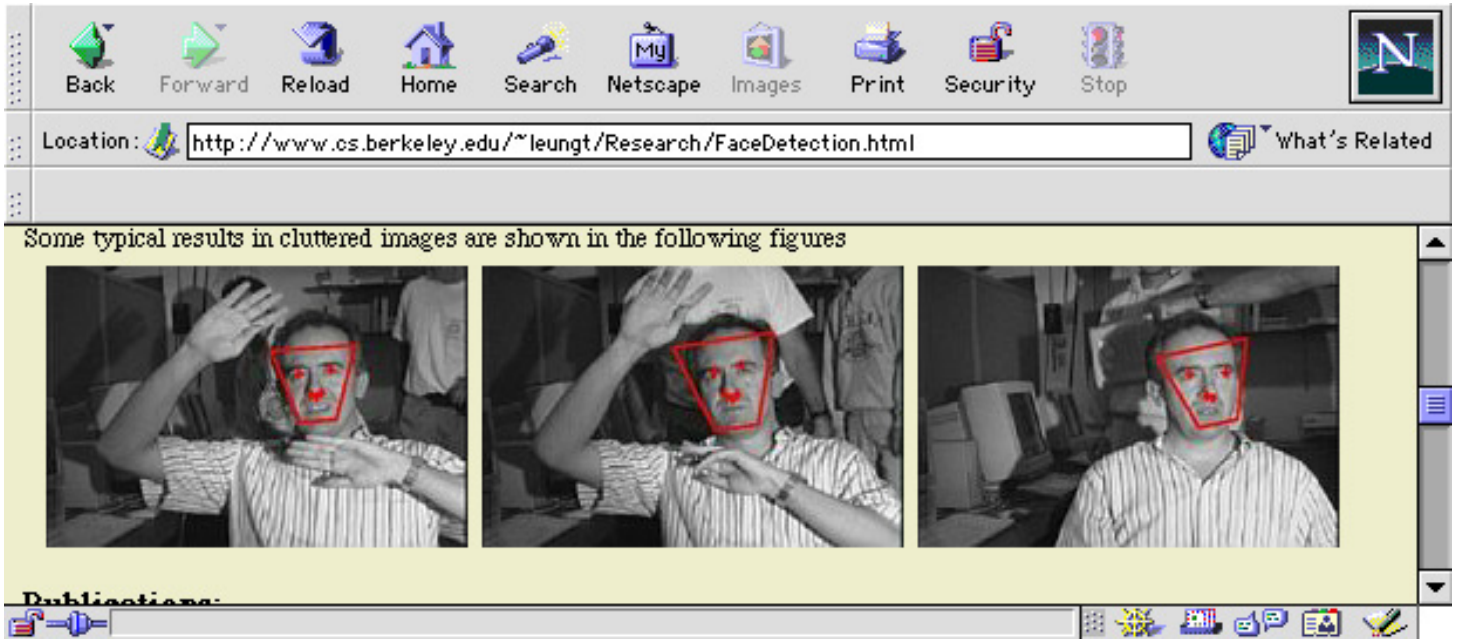
Image-based recognition

(Nayar et al. ‘96)

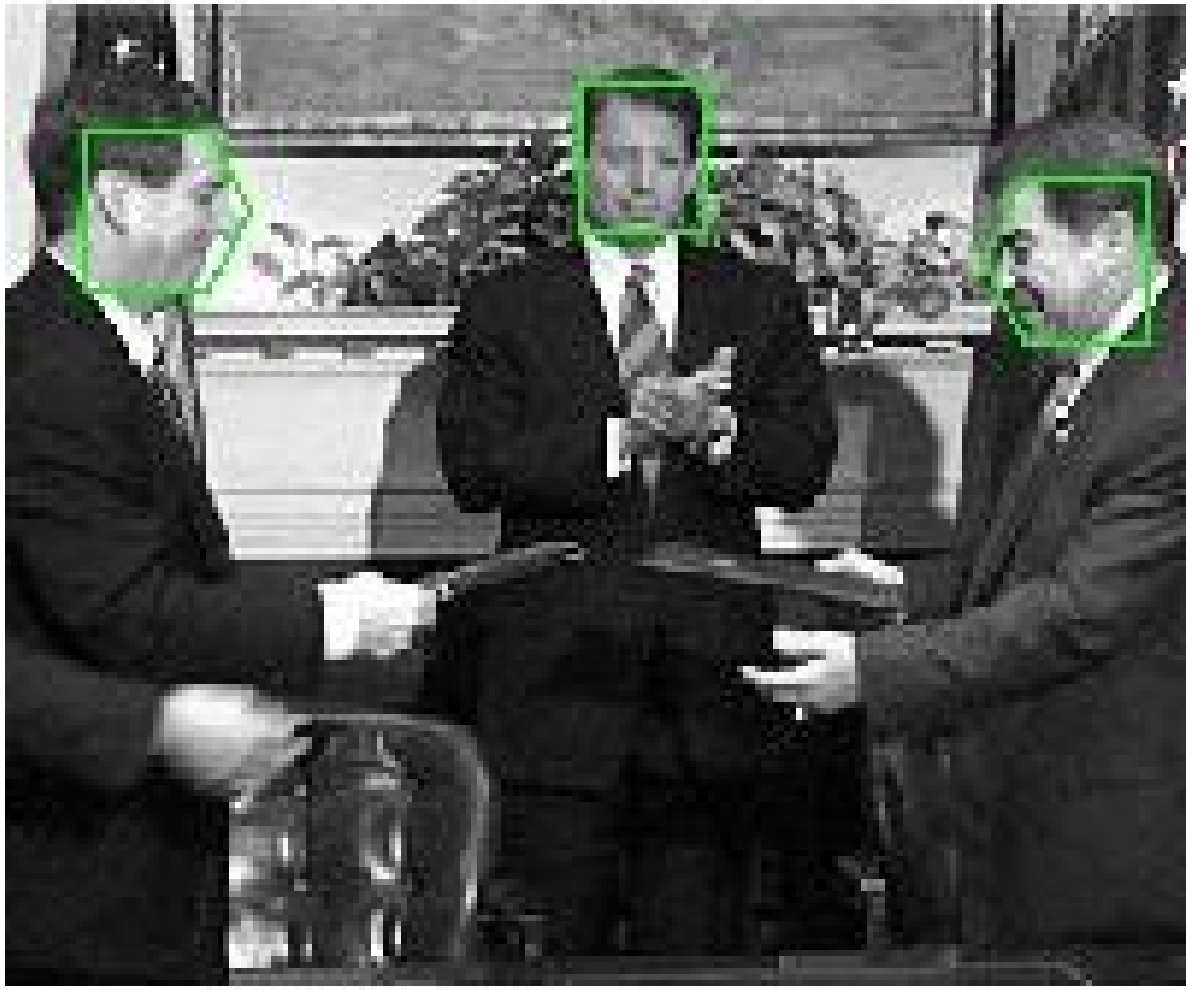
Object recognition using templates and relations

- Find bits and pieces, see if it fits together in a meaningful way

e.g. nose, eyes, …



Face detection



http://vasc.ri.cmu.edu/cgi-bin/demos/findface.cgi