

Introduction

Selim Aksoy

Department of Computer Engineering

Bilkent University

saksoy@cs.bilkent.edu.tr

What is computer vision?

“What does it mean, to see? The plain man's answer (and Aristotle's, too) would be, to know what is where by looking.”

-- David Marr, Vision (1982)

- Automatic understanding of images and video
 - Computing properties of the 3D world from visual data (measurement).
 - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (perception and interpretation).

Adapted from Trevor Darrell, UC Berkeley,
Alyosha Efros, Carnegie Mellon

Why study computer vision?

- As image sources multiply, so do applications
 - Relieve humans of boring, easy tasks
 - Enhance human abilities: human-computer interaction, visualization
 - Perception for robotics / autonomous agents
 - Organize and give access to visual content
- Goals of vision research:
 - Give machines the ability to understand scenes.
 - Aid understanding and modeling of human vision.
 - Automate visual operations.

Adapted from Trevor Darrell, UC Berkeley

Why study computer vision?



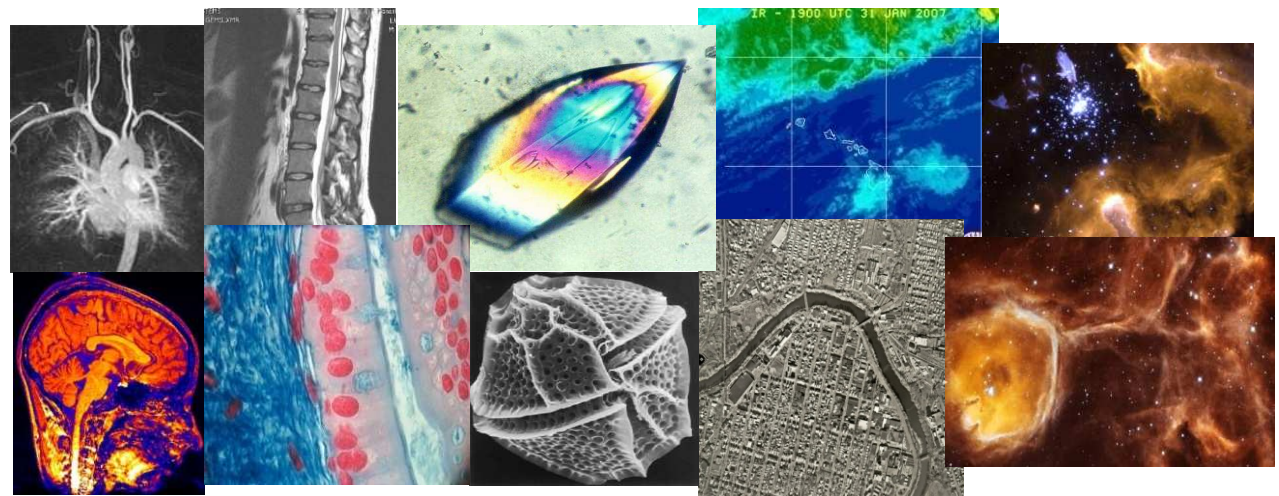
Personal photo albums



Movies, news, sports

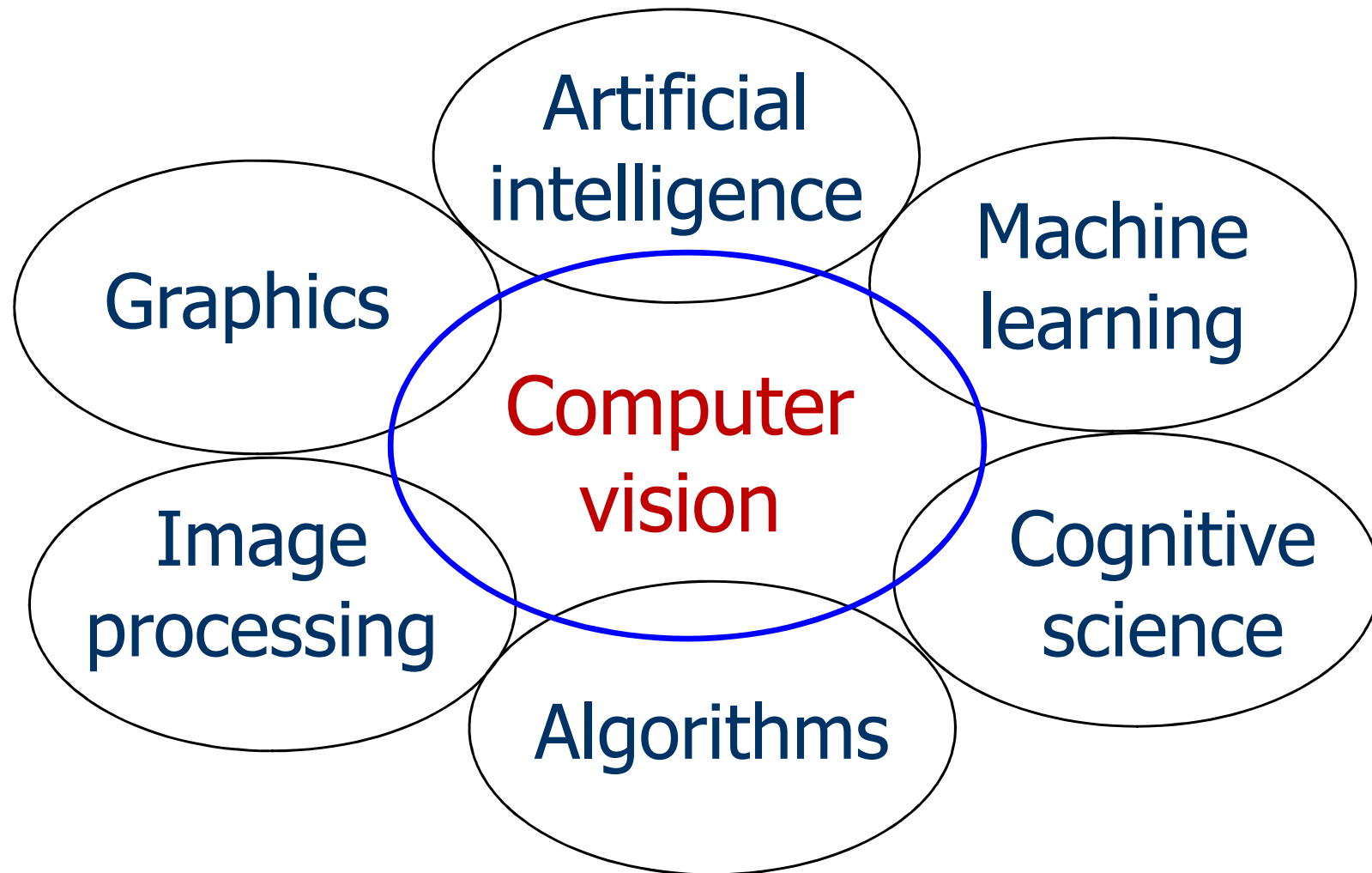


Surveillance and security



Medical and scientific images

Related disciplines



Applications

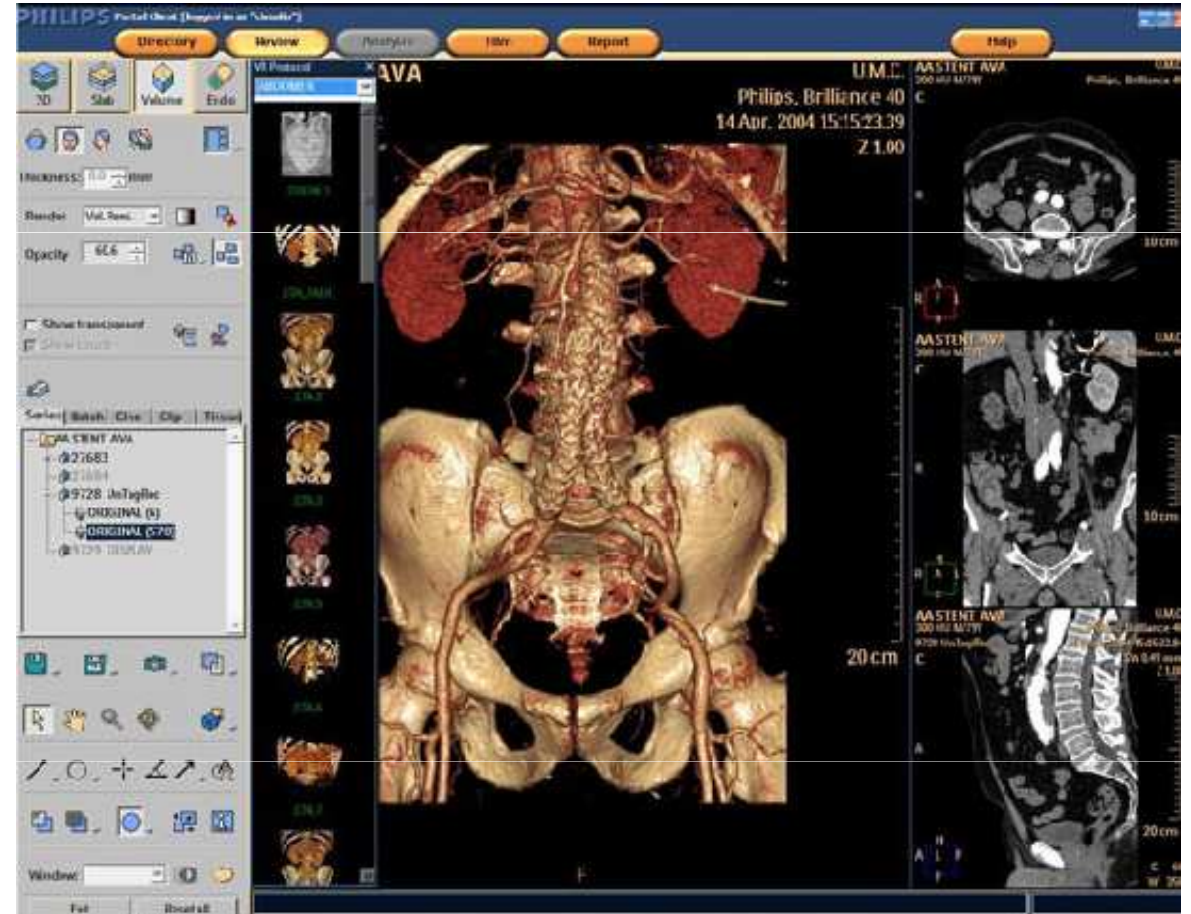
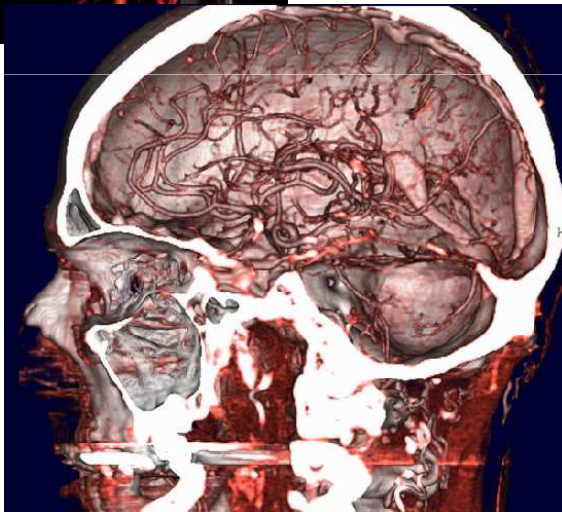
- Medical image analysis
- Security
 - Biometrics
 - Surveillance
 - Tracking
 - Target recognition
- Remote sensing
- Robotics
- Industrial inspection, quality control
- Document analysis
- Multimedia
- Assisted living
- Human-computer interfaces
- ...

Medical image analysis



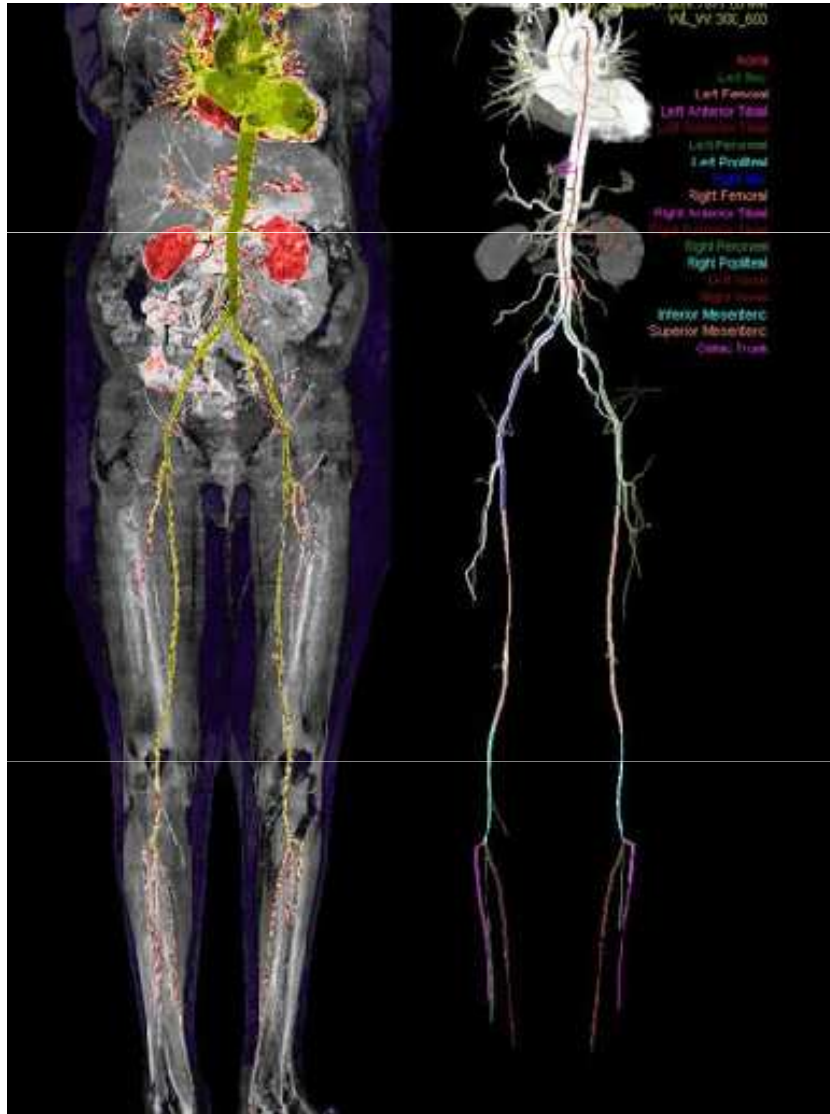
<http://www.clarontech.com>

Medical image analysis

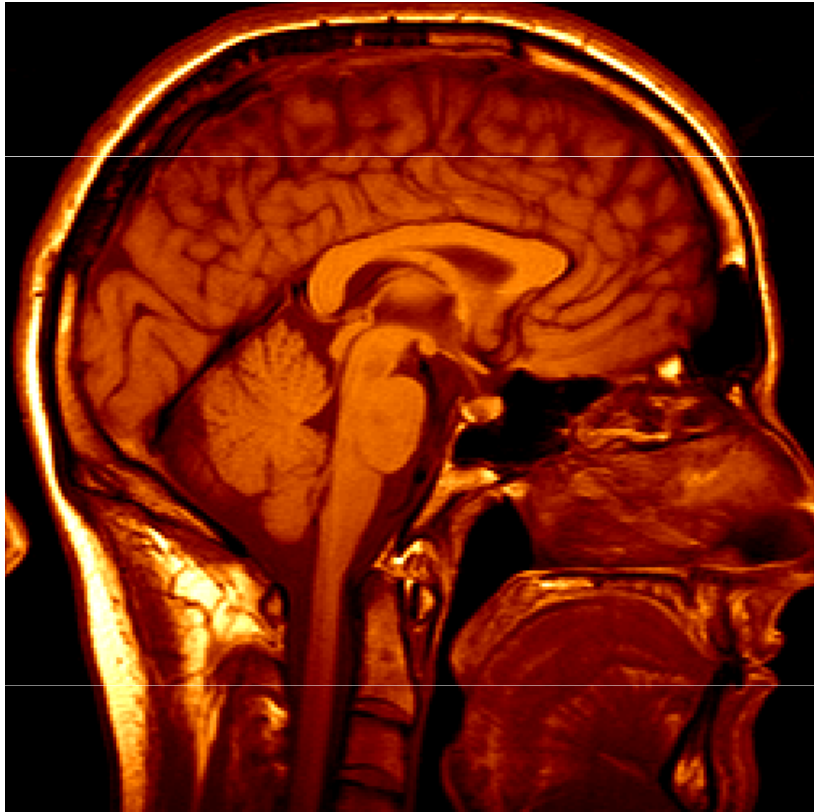


<http://www.clarontech.com>

Medical image analysis



Medical image analysis

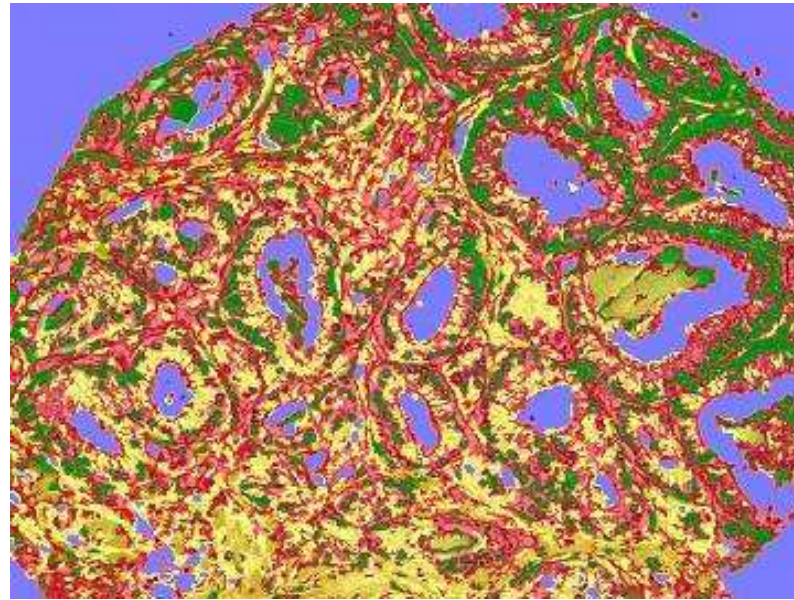
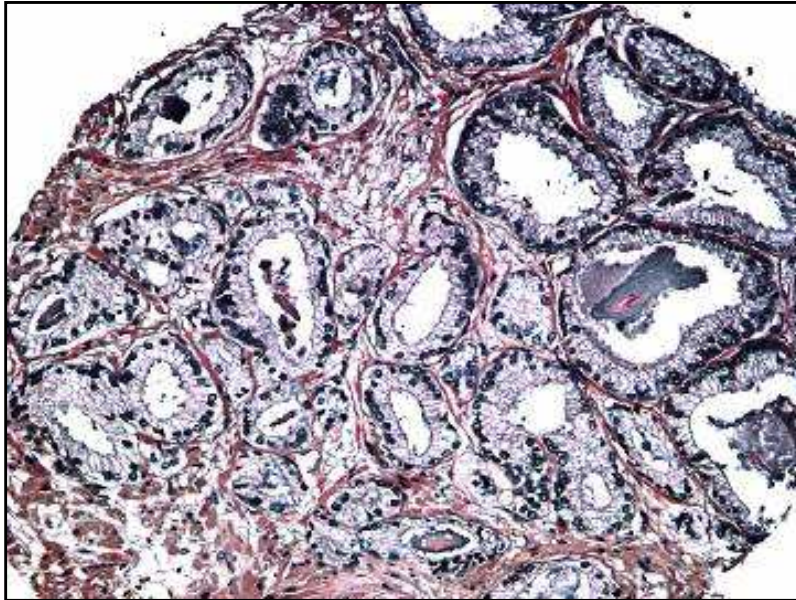


3D imaging: MRI, CT

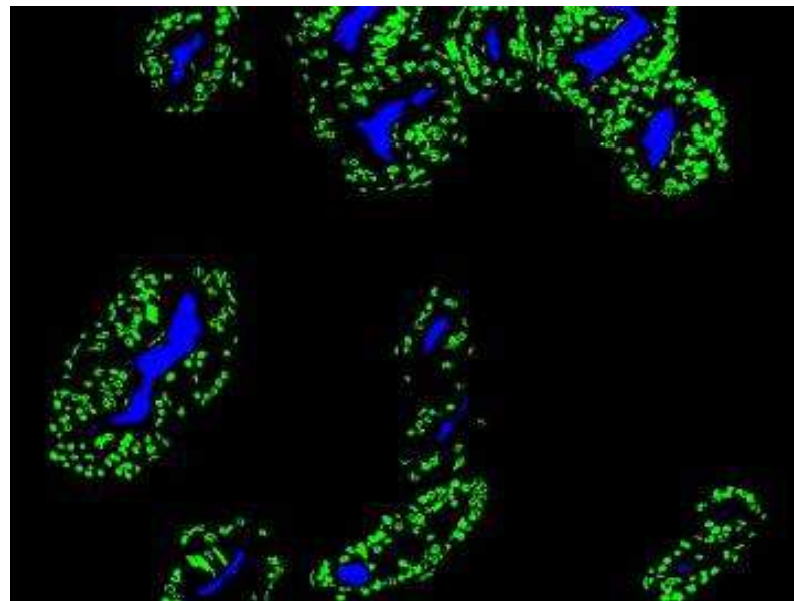
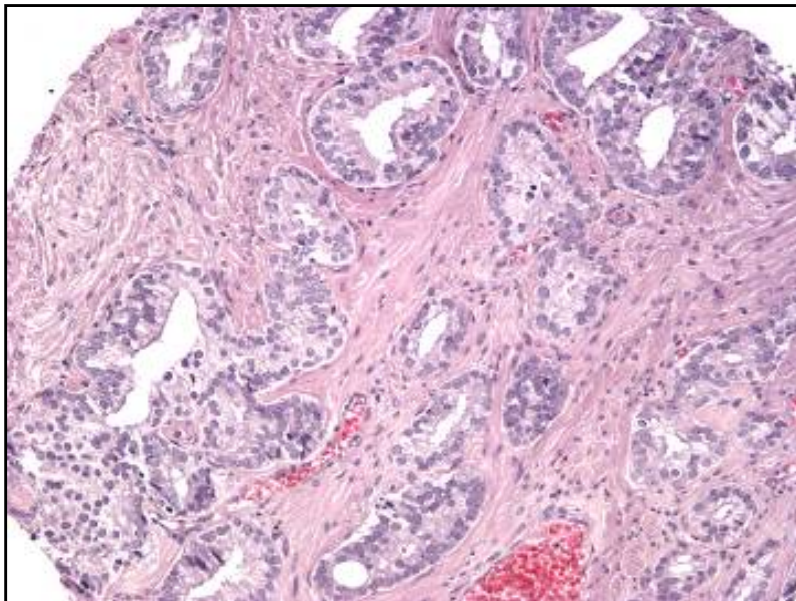


Image guided surgery
Grimson et al., MIT

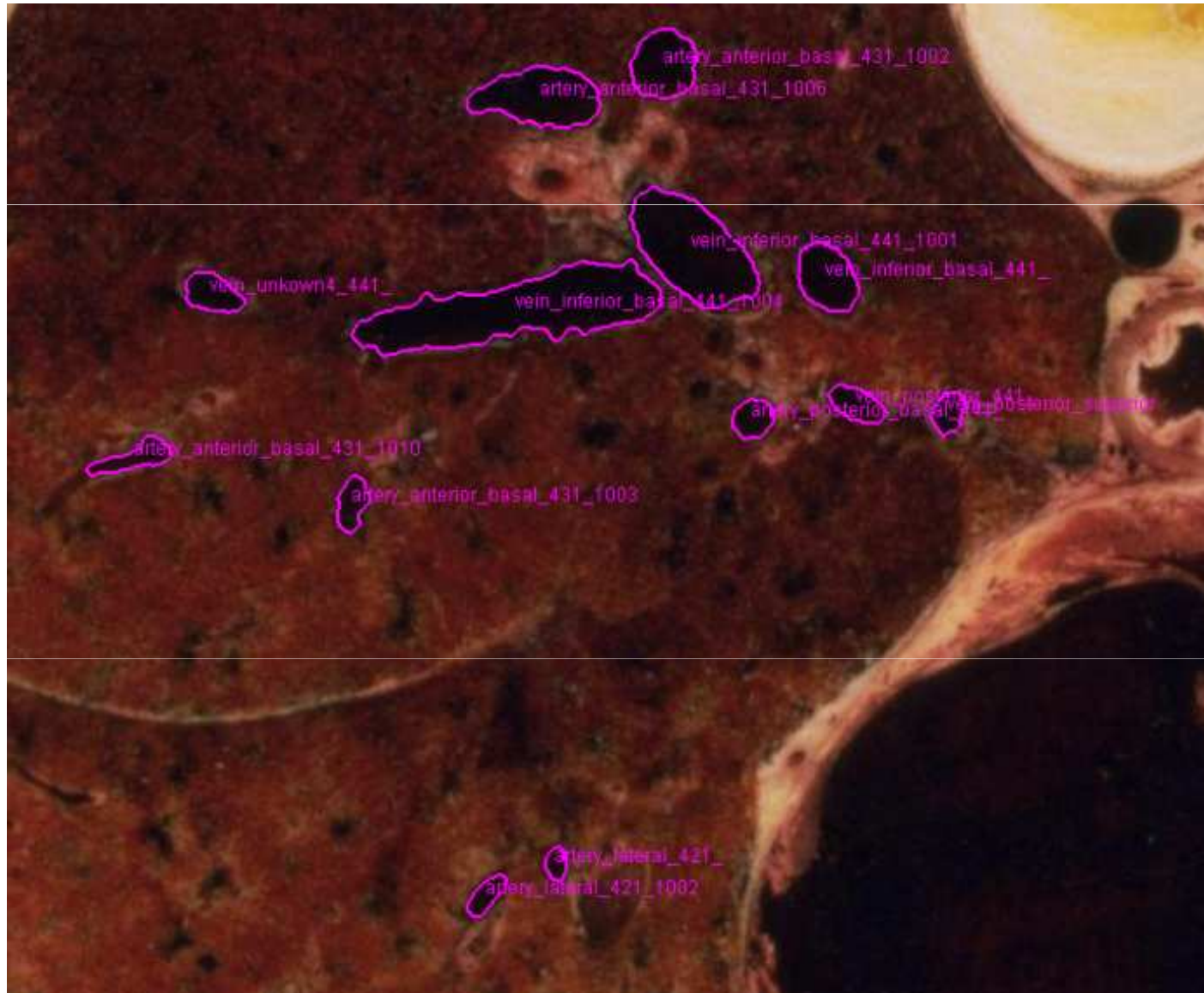
Medical image analysis



Cancer
detection
and
grading



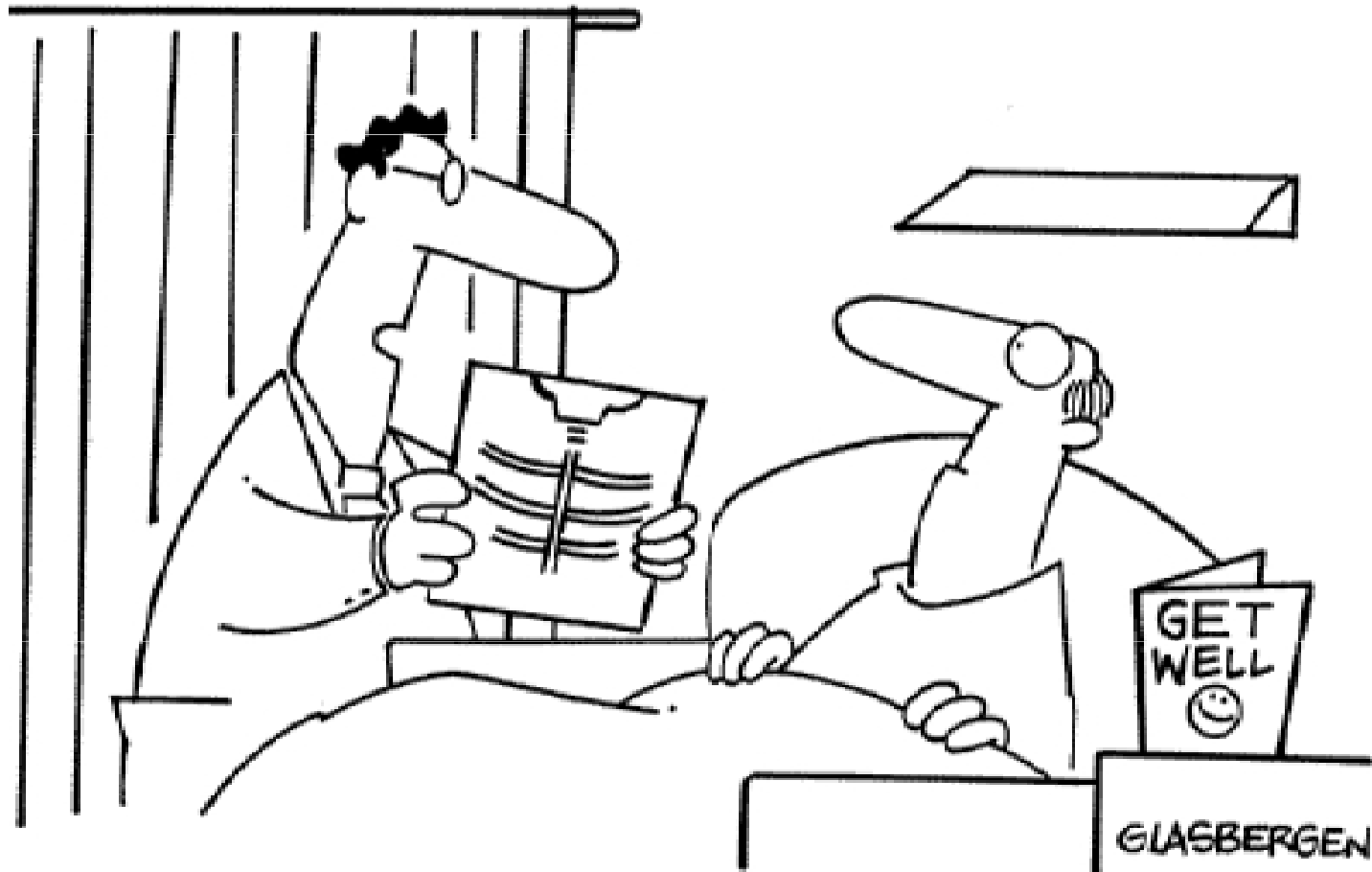
Medical image analysis



Slice of
lung

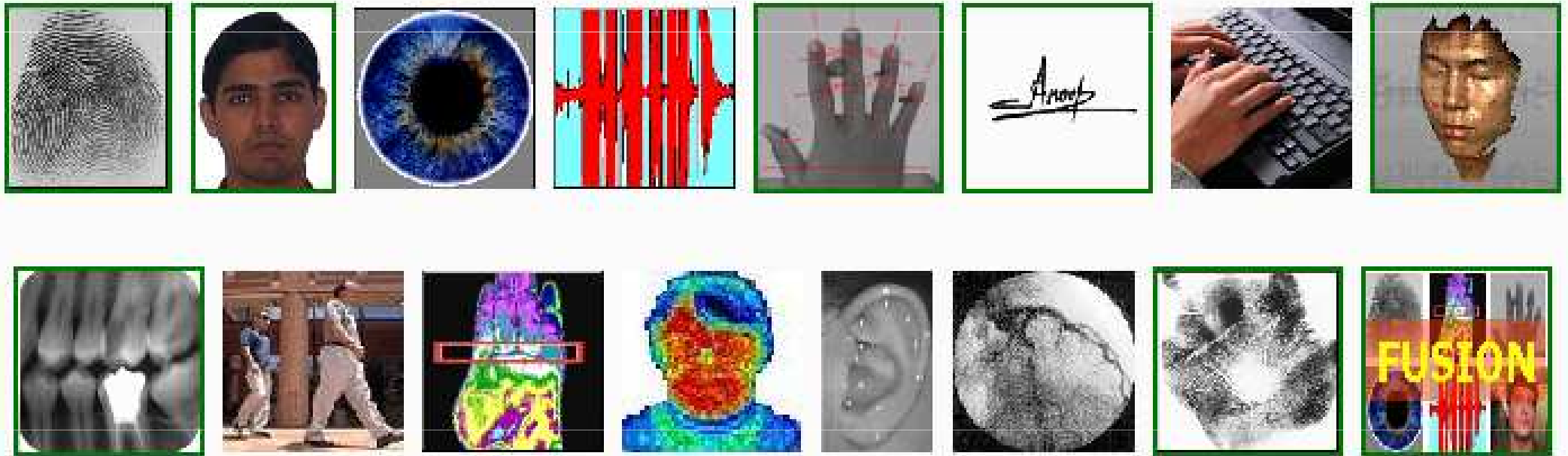
Medical image analysis

© 2000 Randy Glasbergen. www.glasbergen.com



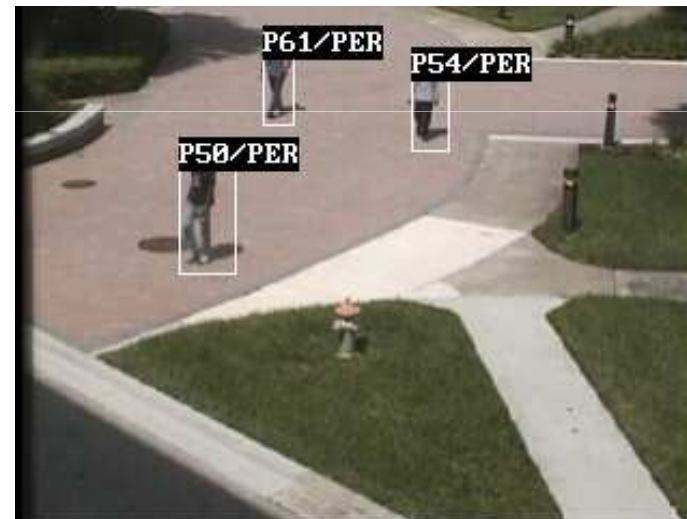
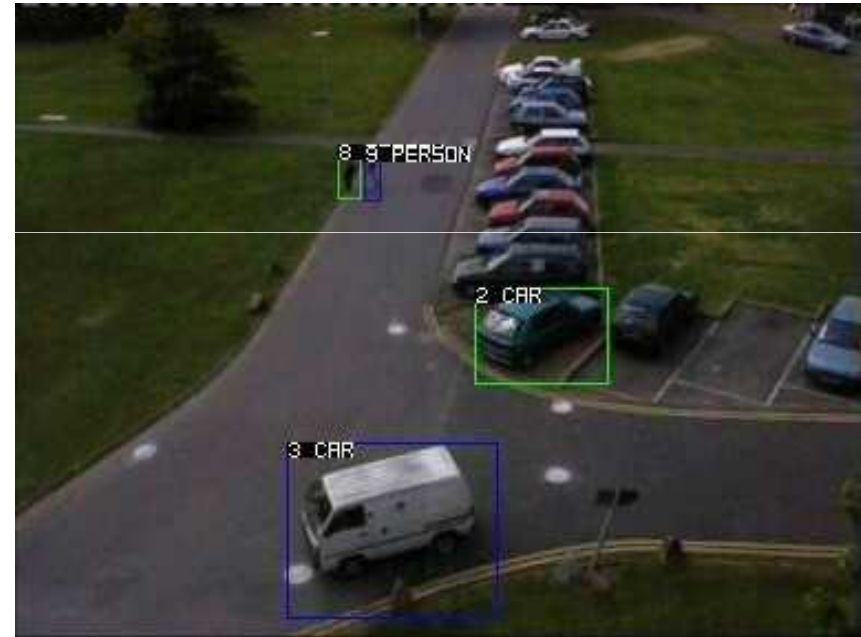
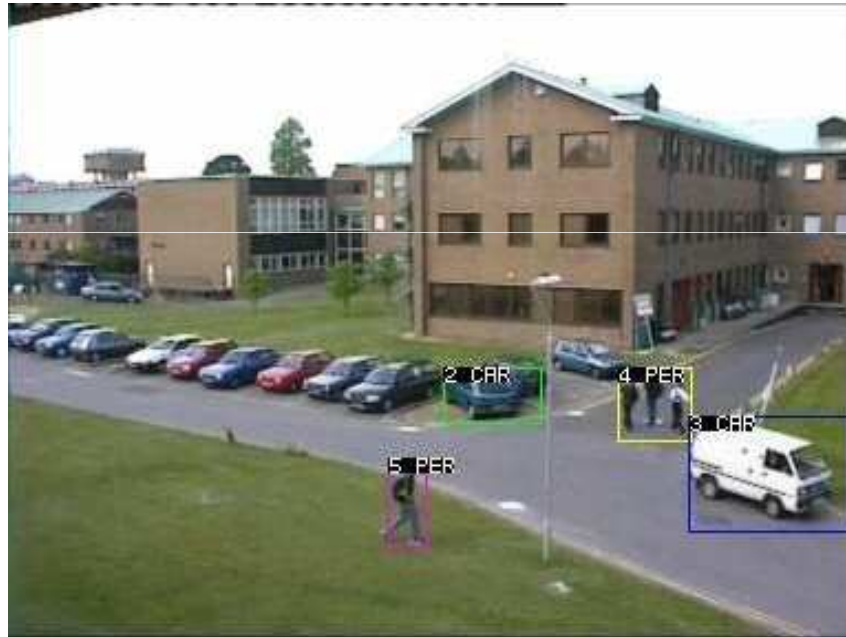
**“Your x-ray showed a broken rib,
but we fixed it with Photoshop.”**

Biometrics

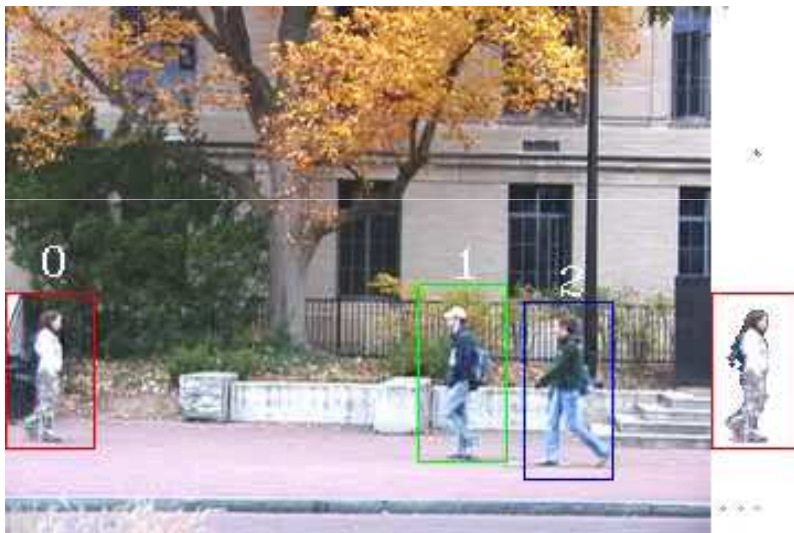


Adapted from Anil Jain, Michigan State

Surveillance and tracking



Surveillance and tracking



Adapted from Octavia Camps, Penn State

Surveillance and tracking



Adapted from Martial Hebert, CMU

Surveillance and tracking



Generating traffic patterns

Surveillance and tracking



Tracking in UAV videos

Adapted from Martial Hebert, CMU, and
Masaharu Kobashi, U of Washington

Smart cars

[▶▶ manufacturer products](#)[consumer products ◀◀](#)

Our Vision. Your Safety.



rear looking camera

forward looking camera

side looking camera

EyeQ Vision on a Chip

[> read more](#)

Vision Applications



Road, Vehicle, Pedestrian Protection and more

[> read more](#)

AWS Advance Warning System

[> read more](#)

News

- ▶ [Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System](#)
- ▶ [Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end](#)

[> all news](#)

Events

- ▶ [Mobileye at Equip Auto, Paris, France](#)
- ▶ [Mobileye at SEMA, Las Vegas, NV](#)

[> read more](#)

Adapted from CSE 455, U of Washington

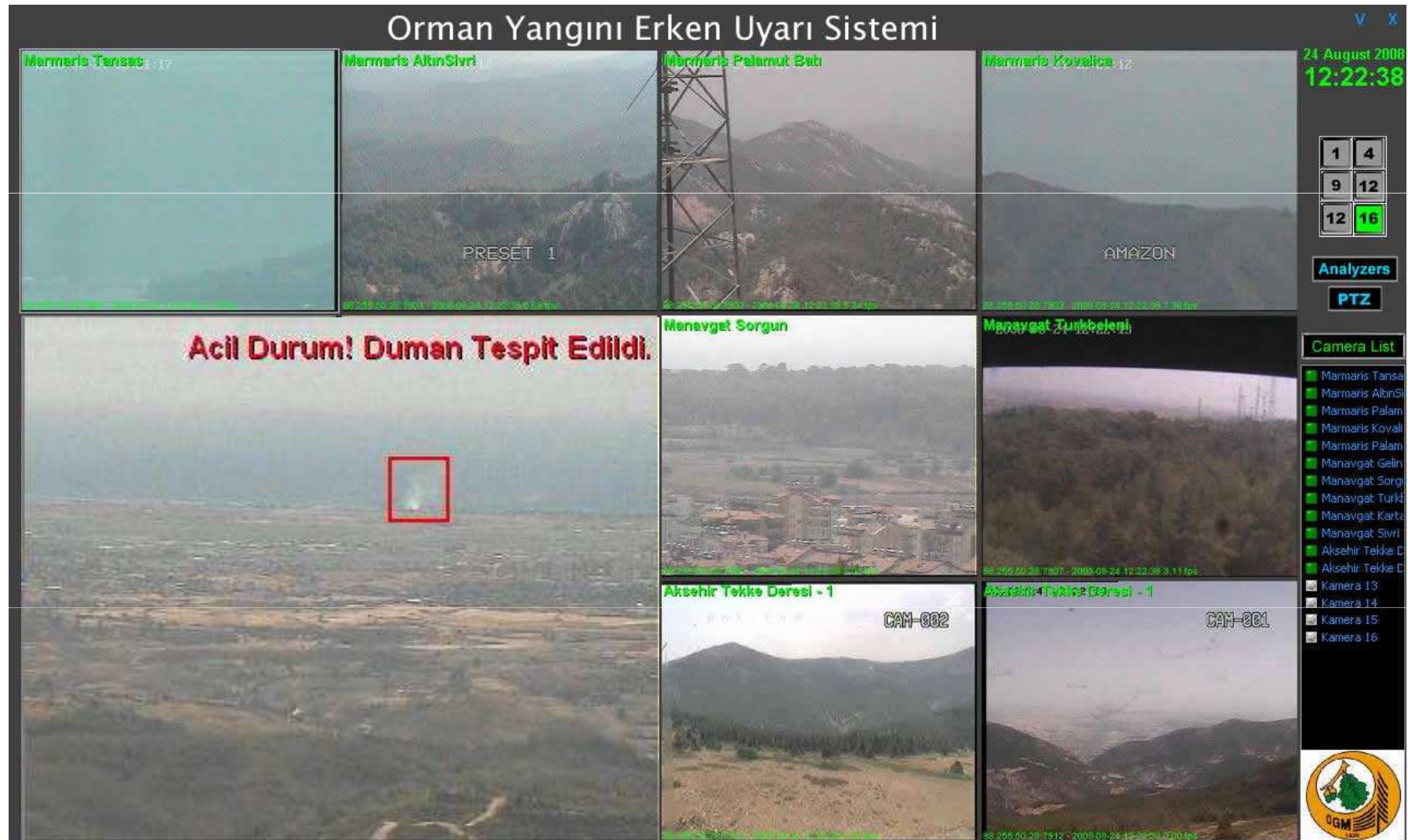
Vehicle and pedestrian protection



Lane departure warning, collision warning, traffic sign recognition, pedestrian recognition, blind spot warning

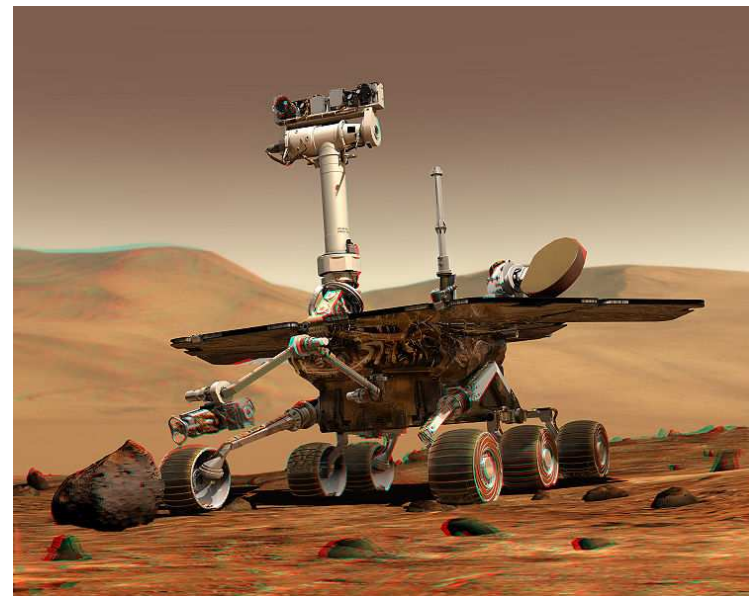
<http://www.mobileye-vision.com>

Forest fire monitoring system



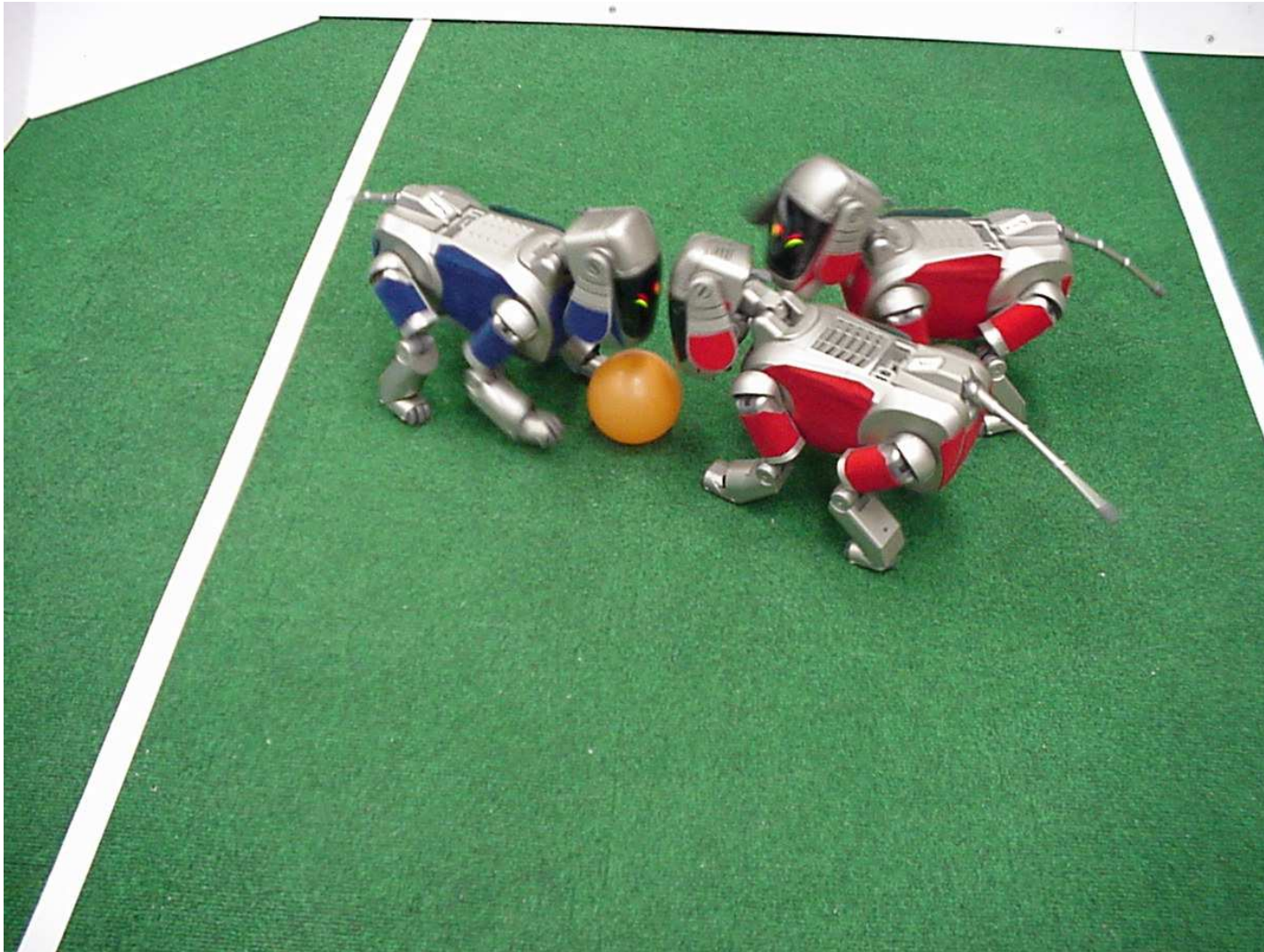
Early warning of forest fires

Robotics



Adapted from CSE 455, U of Washington

Robotics



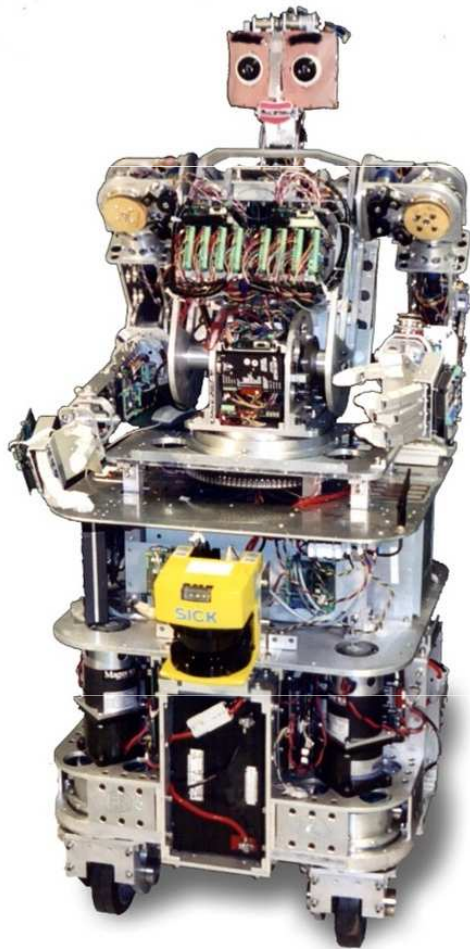
Adapted from Steven Seitz, U of Washington

Autonomous navigation



<http://www.darpa.mil/grandchallenge/index.asp>
http://en.wikipedia.org/wiki/DARPA_Grand_Challenge

Autonomous navigation



Michigan State University



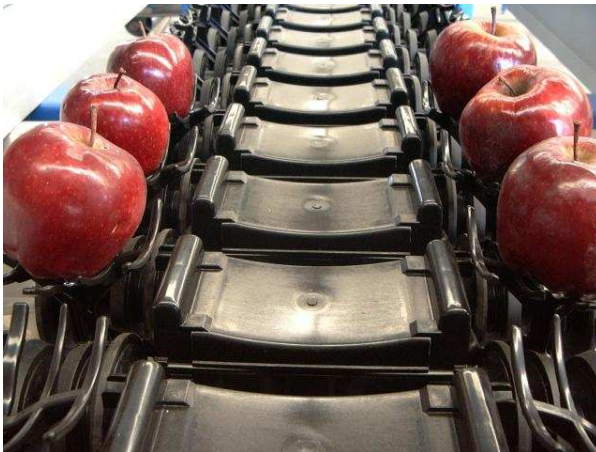
General Dynamics Robotics Systems
<http://www.gdrs.com>

Face detection and recognition



Adapted from CSE 455, U of Washington
27

Industrial automation



Automatic fruit sorting

Industrial automation



Industrial robotics;
bin picking

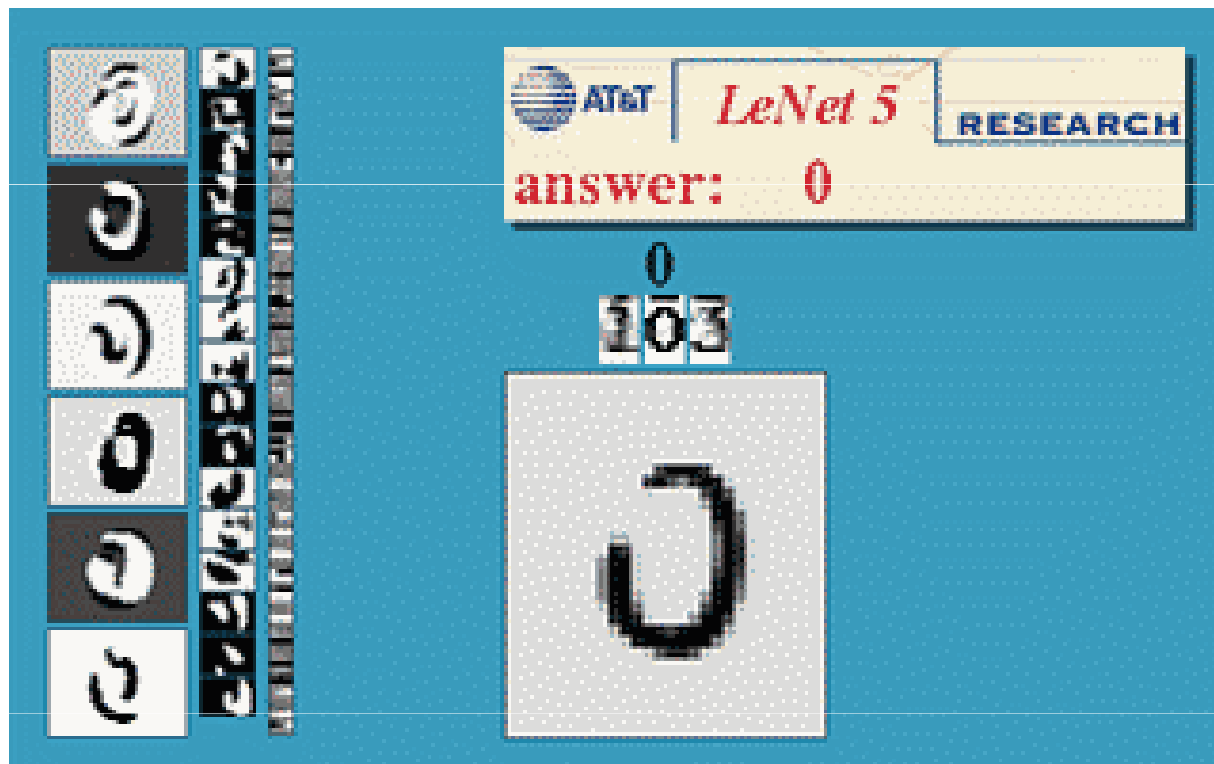
<http://www.braintech.com>

Postal service automation



General Dynamics Robotics Systems
<http://www.gdrs.com>

Optical character recognition



Digit recognition, AT&T labs
<http://www.research.att.com/~yann>



License plate recognition

Adapted from Steven Seitz, U of Washington

Document analysis

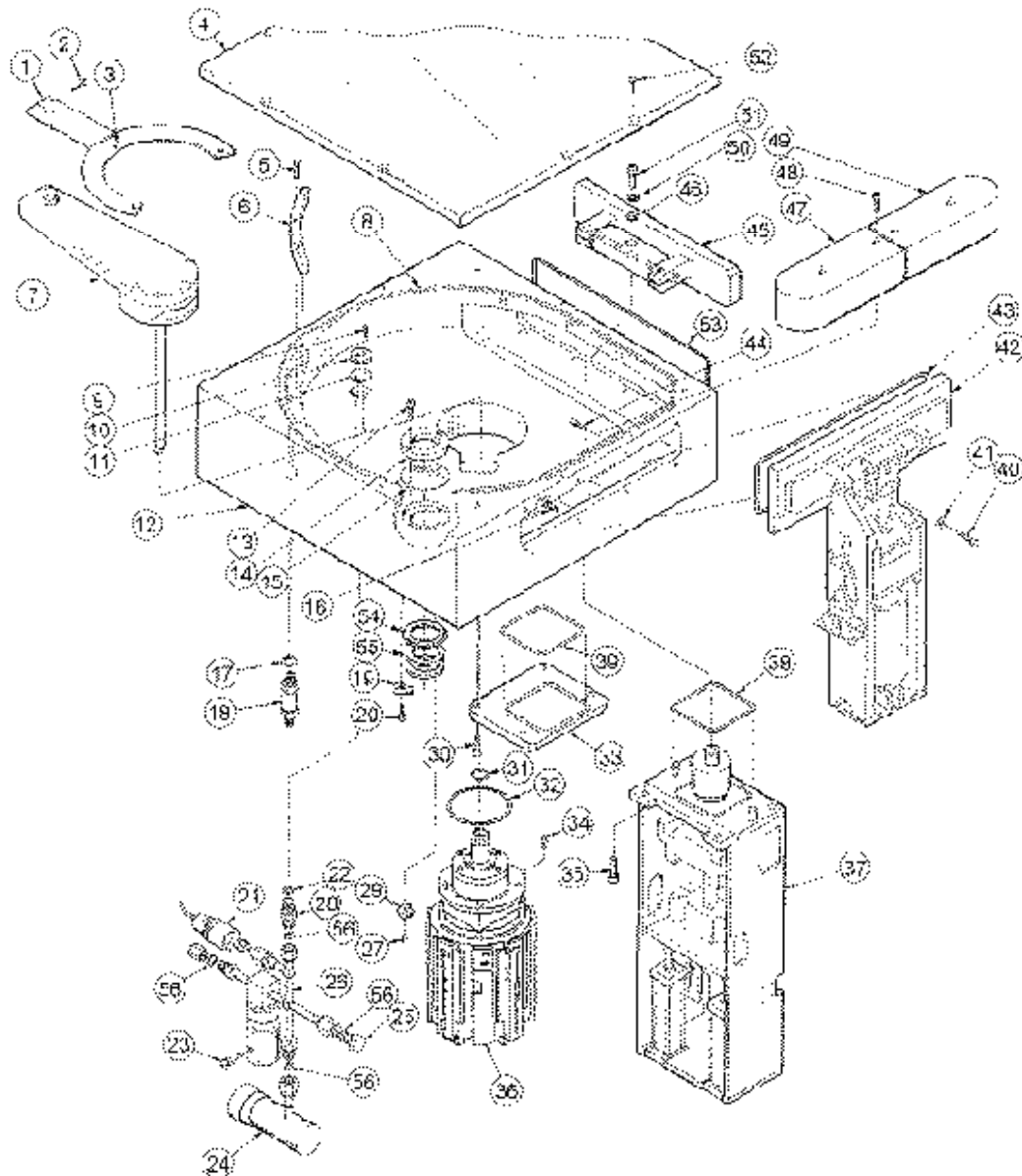
儘眼望遠極，
佰程無窮哩。
壹物明域現，
以迺吾後脊！

I looked as hard as I could see,
beyond 100 plus infinity
an object of bright intensity
– it was the back of me!

Figure 1.5: (Left) Chinese characters and (right) English equivalent. Is it possible that a machine could automatically translate one into the other? Chinese characters and poem courtesy of John Weng.

Adapted from Shapiro and Stockman

Document analysis



**CMS/Fisher
HealthCare**

Blood Bank / Dry/Imitates



Model 145 Isotemp® Dry Bath Incubator

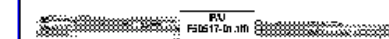
**Holds 1 in 4 heating blocks with choice of 11 well sizes.
Maintains every sample in within $\pm 0.1^\circ\text{C}$ at temperature.**

Single sample wells are shaped so that a 10 mm block delivers same amount of heat to all parts of the sample tube. No temperature gradient exists in the bath. The heater is a ceramic, flat array, in a double cast in epoxy with drilled cooling wells. Sample tubes rest on a plate to plus, to prevent localized heating. A low cost, stainless steel, 10 mm plate on a thick, uniform heat distribution plate in the back of the bath. Plate is built to last, 5.5 mm, dry bath maintains cleanest probes because tubes are easily.

Ambient to 125°C (25°C to 125°C) Control that temperature controlled from 25°C to 125°C for extreme reactions. Inactivated of core. Bathing, close, cross-matching and the other, decomposition. Dimensions: 8.1 x 15.9 x 4" H, 10.5 x 28 x 11 mm. With 11 mm and plug. Heating blocks sold separately (see lower right).

Electrical Requirements	Cat. No.	Each
100V, 50/60 Hz, 100W	11-215-100	223.58
200V, 50/60 Hz, 100W	11-215-100	223.58

Maximum temperature: 125°C (with 11 mm plug)
Minimum temperature: 25°C



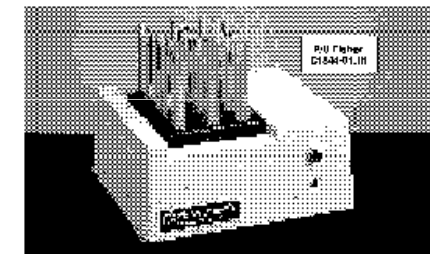
Incu-Block® Partial Immersion Thermometers

For all standard bath, in blocks and water baths. Critical temperatures 25°C , 30°C , 37°C , 58°C are marked with arrows. Available with shallow, normal, contamination and Teller® marking. Total length: 178 mm. In the case of 35 mm.

Range, $^\circ\text{C}$	Div., $^\circ\text{C}$	Yellow Case	Cat. No.	Each
35-57	0.5	N	14-992	45.22
35-57	0.07	Var	14-999	46.17

More Thermometers

For more thermometers, including digital types,
see page 952



Model 147 Isotemp® Dry Bath

Holds single heating block with choice of 11 well sizes.

Single sample wells are shaped so that a 10 mm block delivers same amount of heat to all parts of the sample tube. No temperature gradient exists in the bath. The heater is a ceramic, flat array, in a double cast in epoxy with drilled cooling wells. Sample tubes rest on a plate to plus, to prevent localized heating. A low cost, stainless steel, 10 mm plate on a thick, uniform heat distribution plate in the back of the bath. Plate is built to last, 5.5 mm, dry bath maintains cleanest probes because tubes are easily.

Ambient to 125°C (25°C to 125°C) Control that temperature controlled from 25°C to 125°C for extreme reactions. Inactivated of core. Bathing, close, cross-matching and the other, decomposition. Dimensions: 8.1 x 15.9 x 4" H, 10.5 x 28 x 11 mm. With 11 mm and plug. Heating blocks sold separately (see below).

Electrical Requirements	Cat. No.	Each
100V, 50/60 Hz, 100W	11-215-102	223.58

Interchangeable Heating Blocks for Isotemp® Dry Baths

For Models 145 and 147 Dry Baths. Composed of brass and steel alloy, alloy (chrome) stainless steel. Dimensions: 2.1 x 3.0 x 1.0 (H x W x D) (mm).

For 11-215-100 and 147 Dry Baths. Composed of brass and steel alloy, alloy (chrome) stainless steel. Dimensions: 2.1 x 3.0 x 1.0 (H x W x D) (mm).

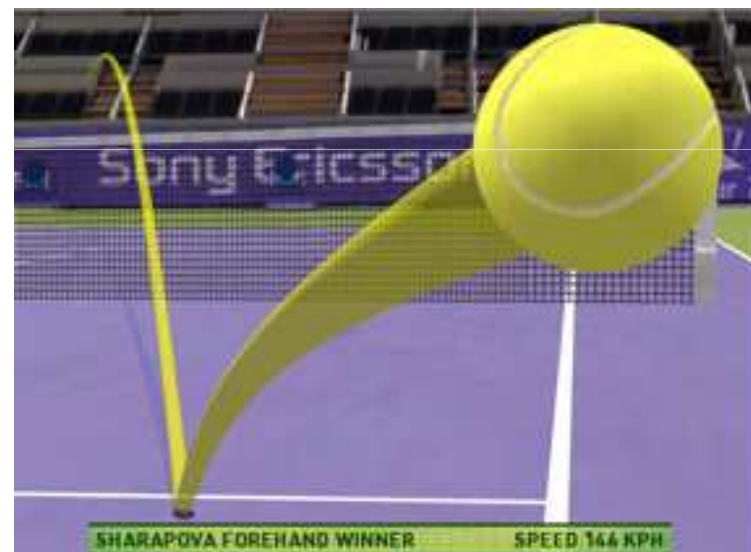
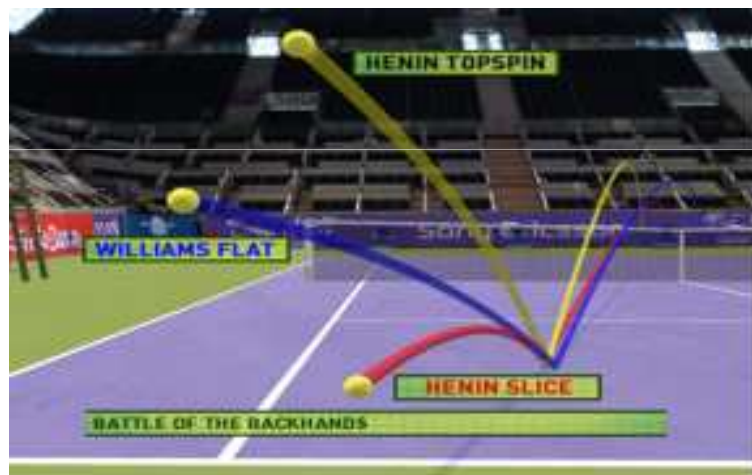
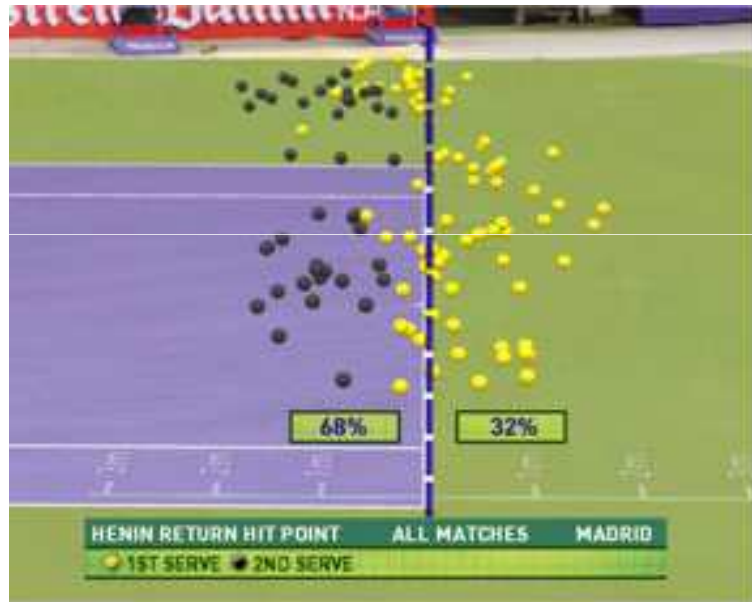
The 11-215-100 block is specifically designed to hold twenty 9.5 mm. Bathing, close, cross-matching and the other, decomposition. Dimensions: 8.1 x 15.9 x 4" H, 10.5 x 28 x 11 mm. With 11 mm and plug. Heating blocks sold separately (see below).

Tube, mm	Wells/Block	Cat. No.	Each
9.5	35	11-215-105	21.18
10	31	11-215-107	21.18
12	20 (10 mm)	11-215-109	21.18
12	12	11-215-100	21.18
12	12	11-215-101	21.18
12	12	11-215-111	21.18
12	12	11-215-113	21.18
12	12	11-215-128	21.18
12	12	11-215-115	21.18
12	12	11-215-117	21.18
12	12	11-215-119	21.18

For more thermometers, including digital types,
see page 952

To Order Call: 1-800-645-1560 Fax: 1-800-645-1560

Sports video analysis



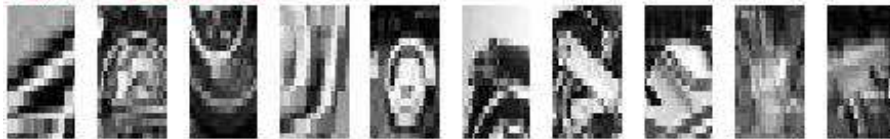
Tennis review system

Scene classification

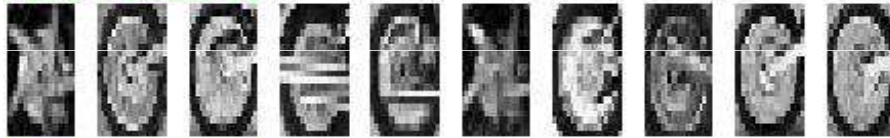


Object recognition

Part 1 – Det:5e-18



Part 2 – Det:8e-22



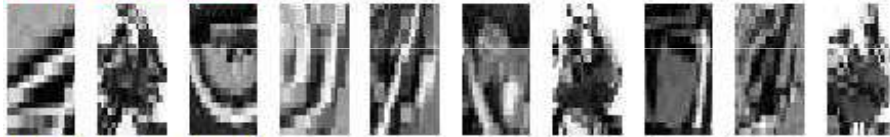
Part 3 – Det:6e-18



Part 4 – Det:1e-19



Part 5 – Det:3e-17



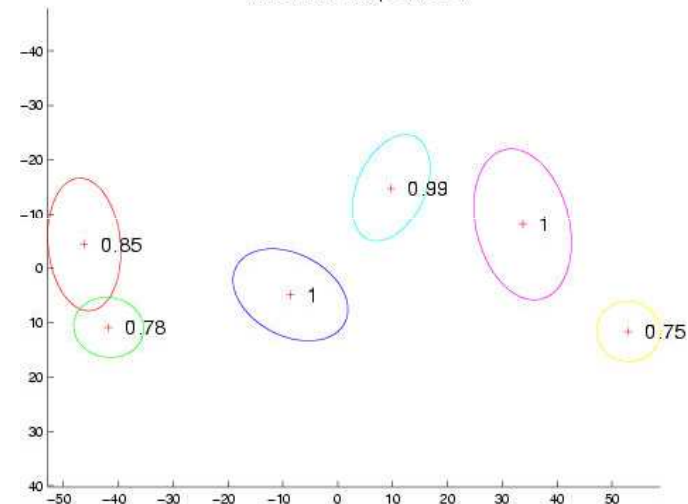
Part 6 – Det:4e-24



Background – Det:5e-19



Motorbike shape model



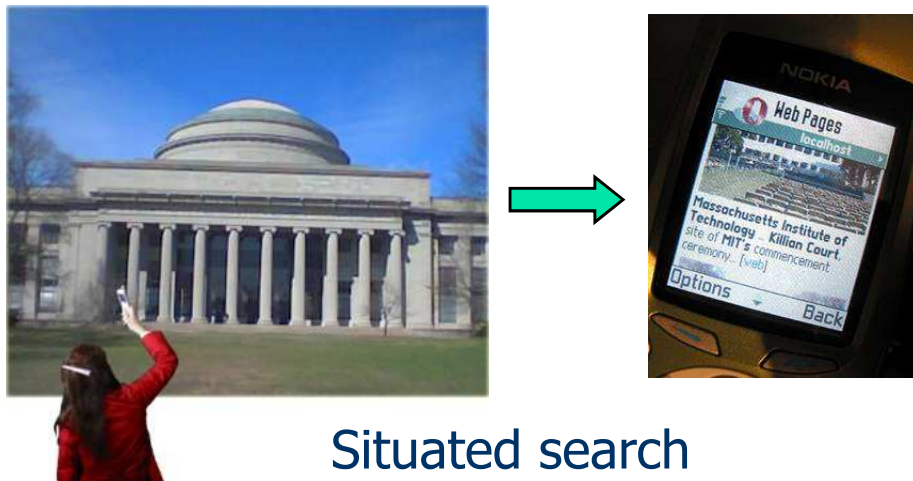
Object recognition



Lincoln, Microsoft Research



kooaba



Situated search
Yeh et al., MIT



Google Goggles
Bing Vision

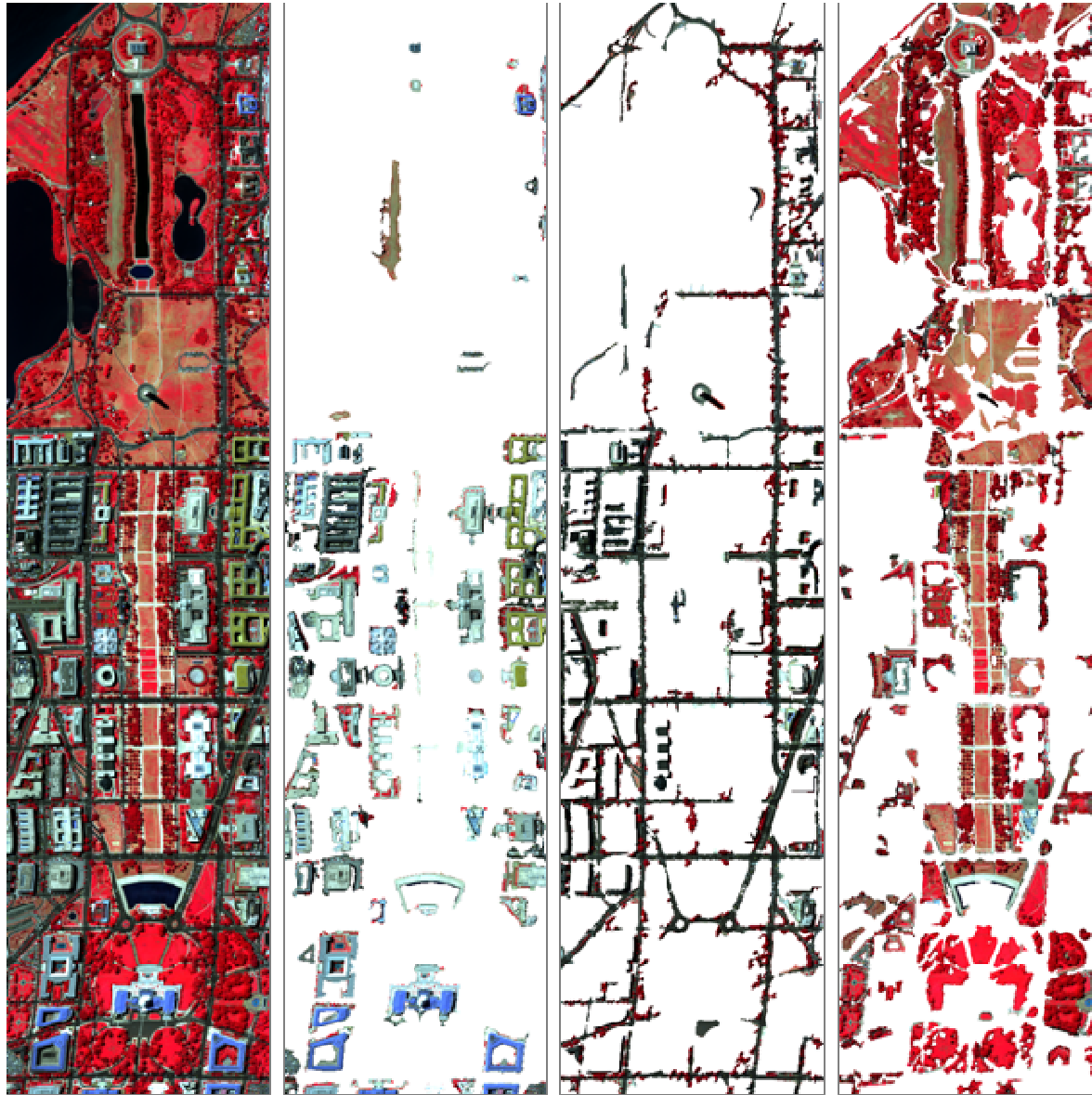
Land cover classification



Land cover classification



Object recognition

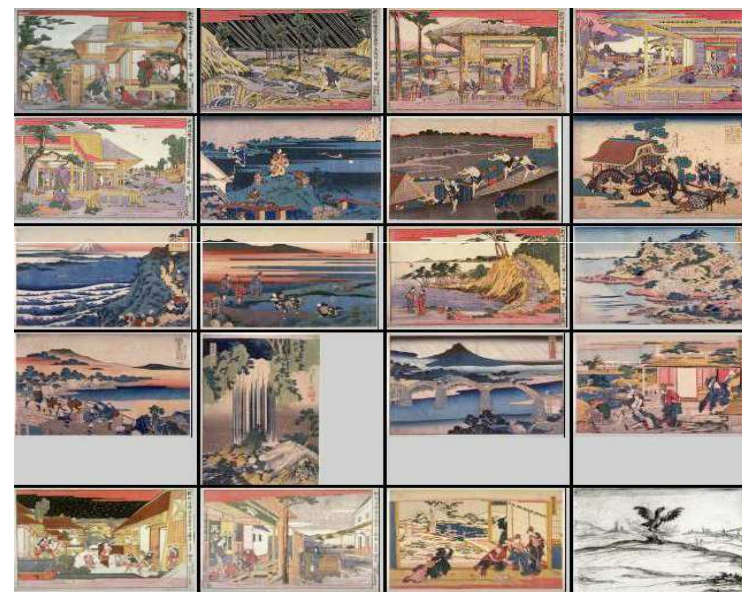
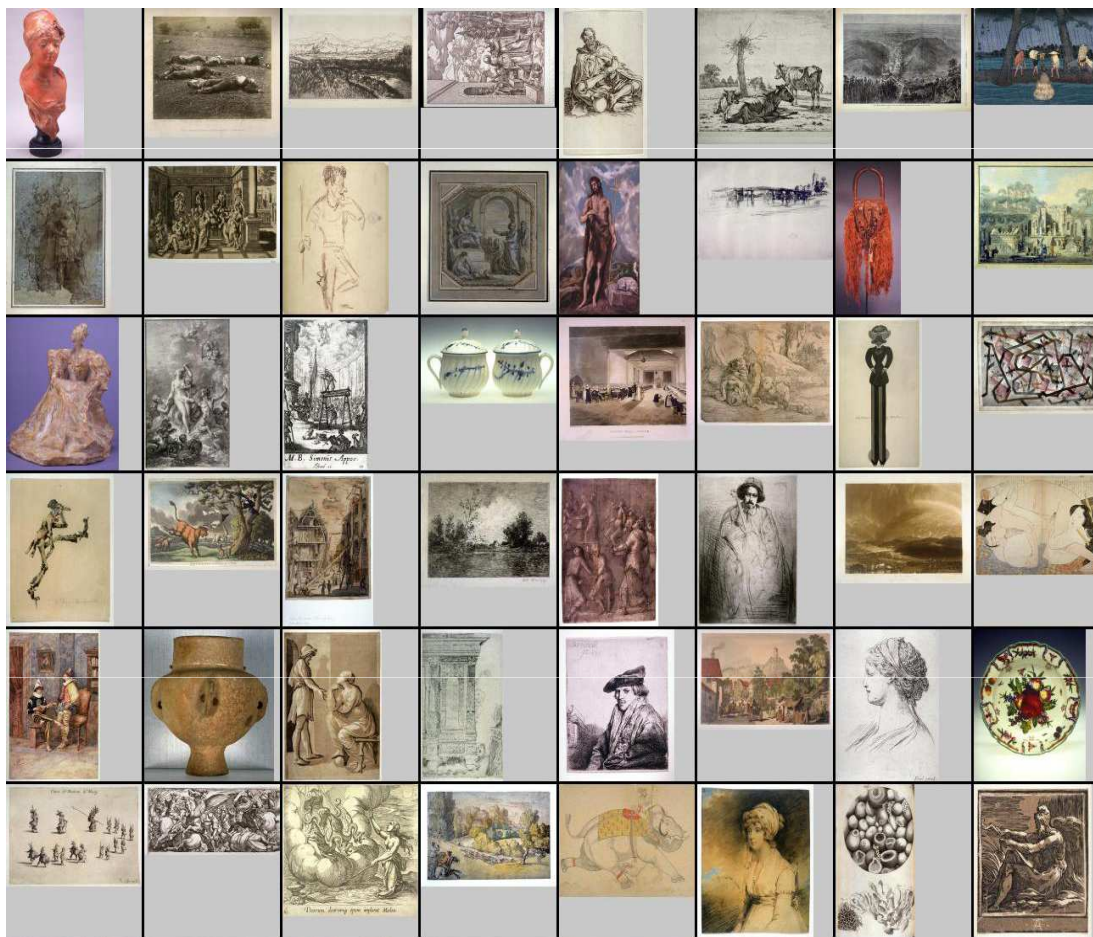


Object recognition



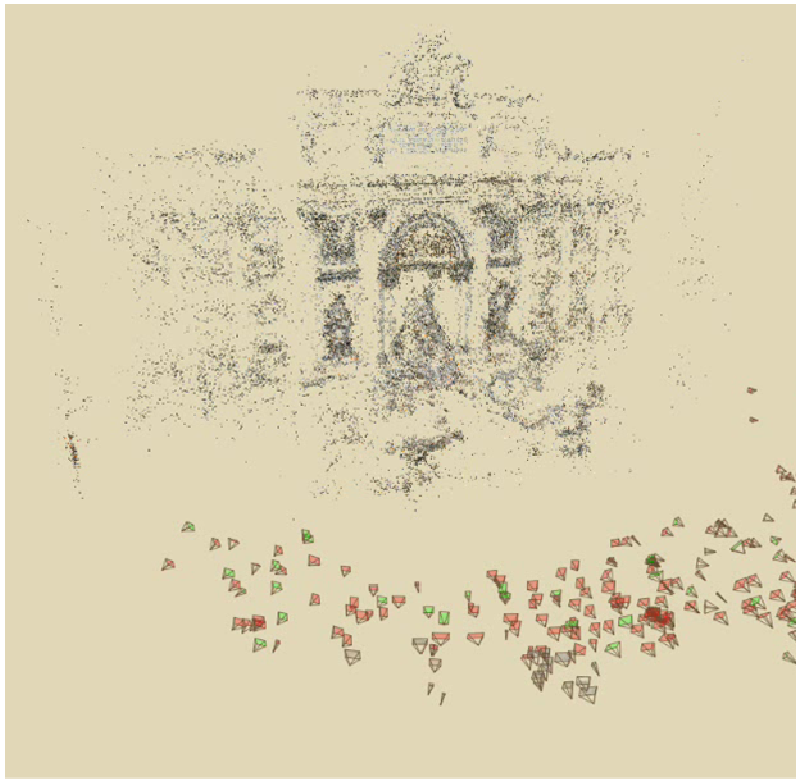
Recognition of buildings and building groups

Organizing image archives



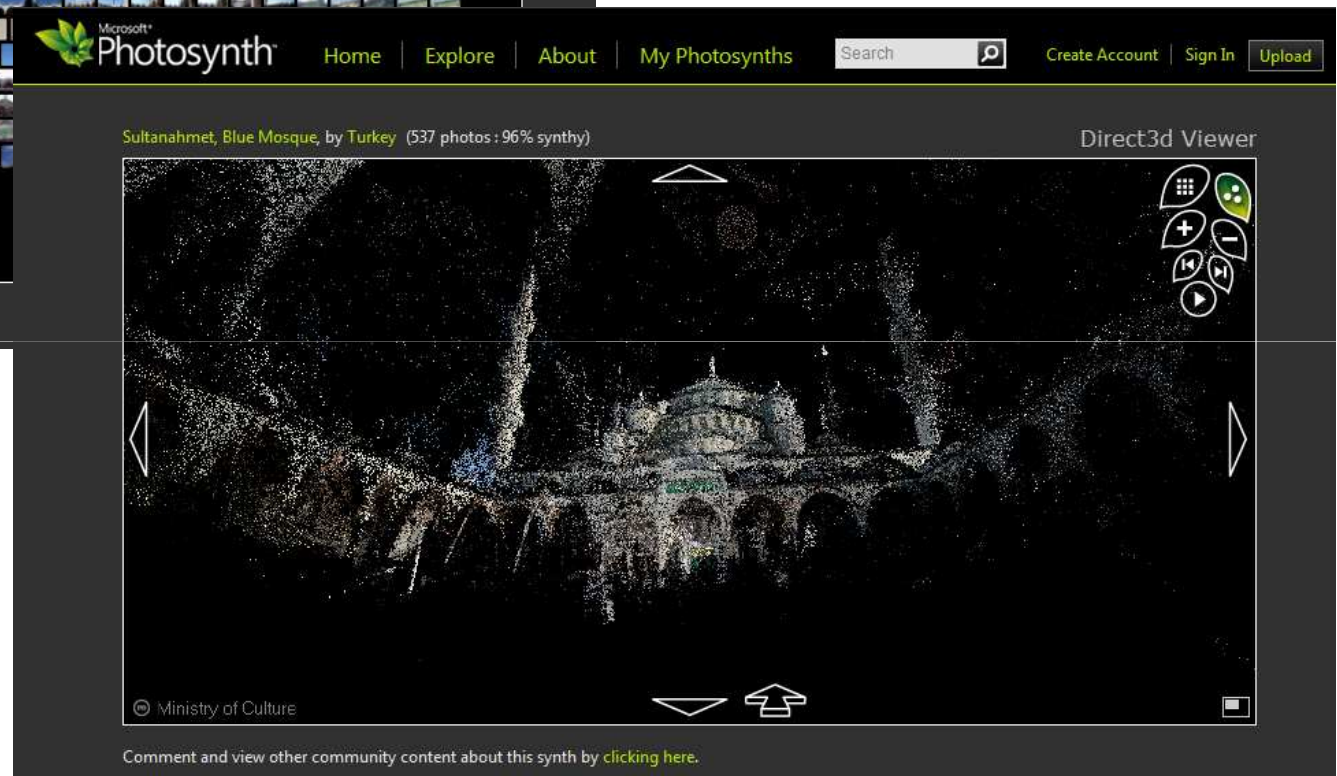
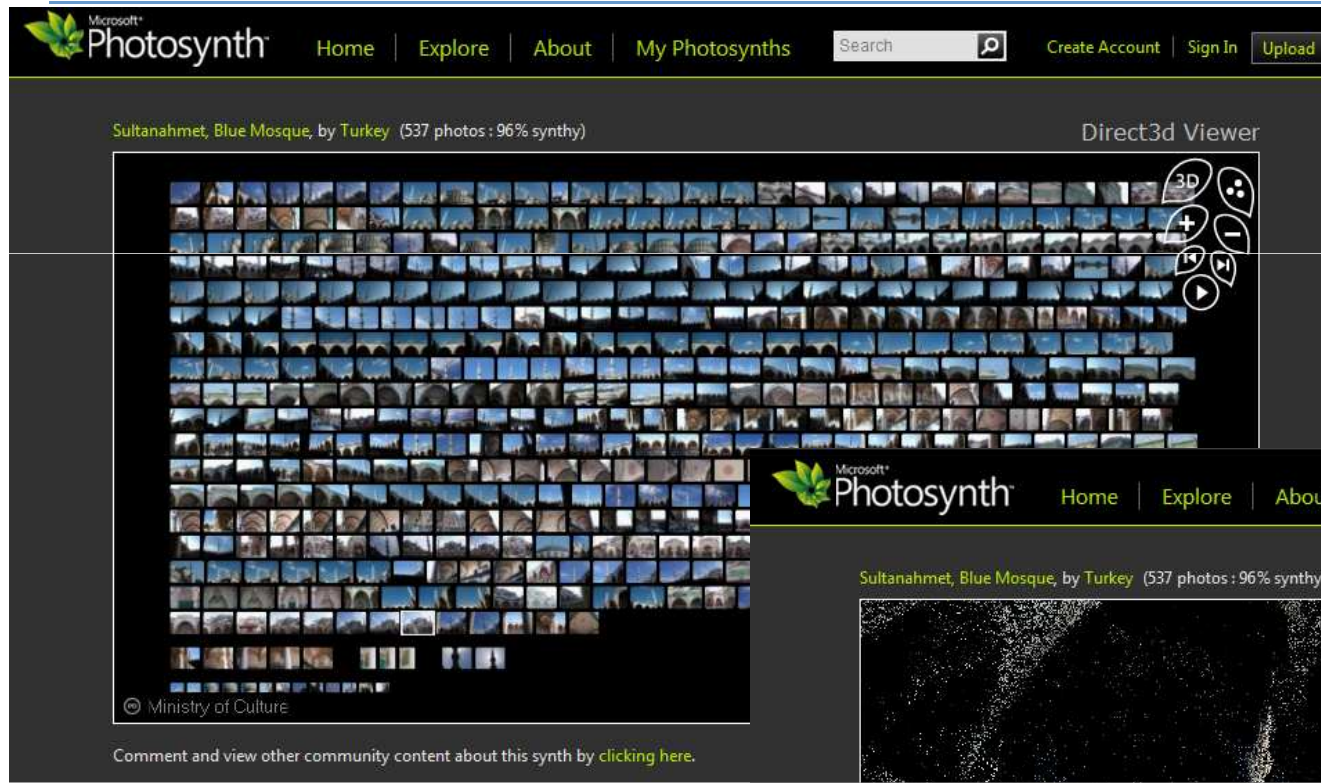
Adapted from Pinar Duygulu, Bilkent University

Photo tourism: exploring photo collections

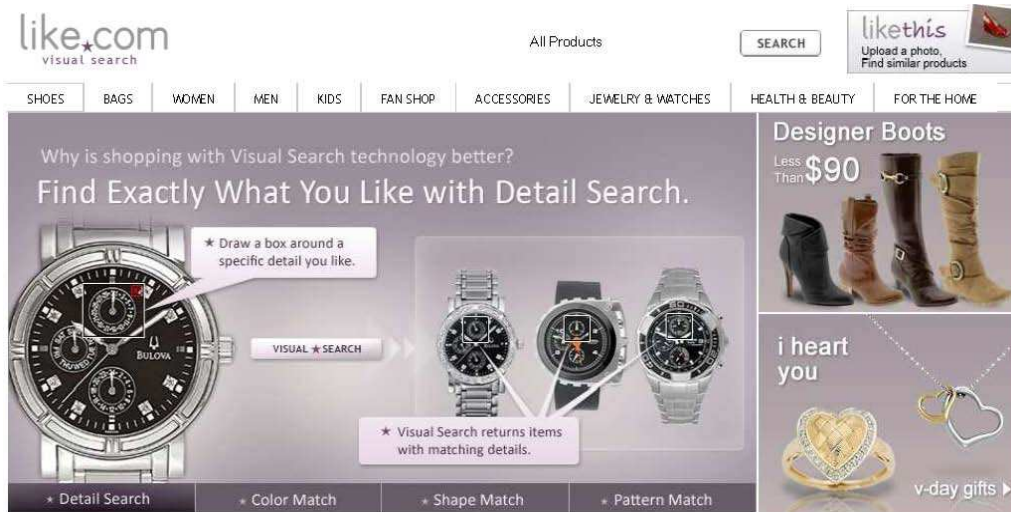


Building 3D scene models from individual photos

Photosynth



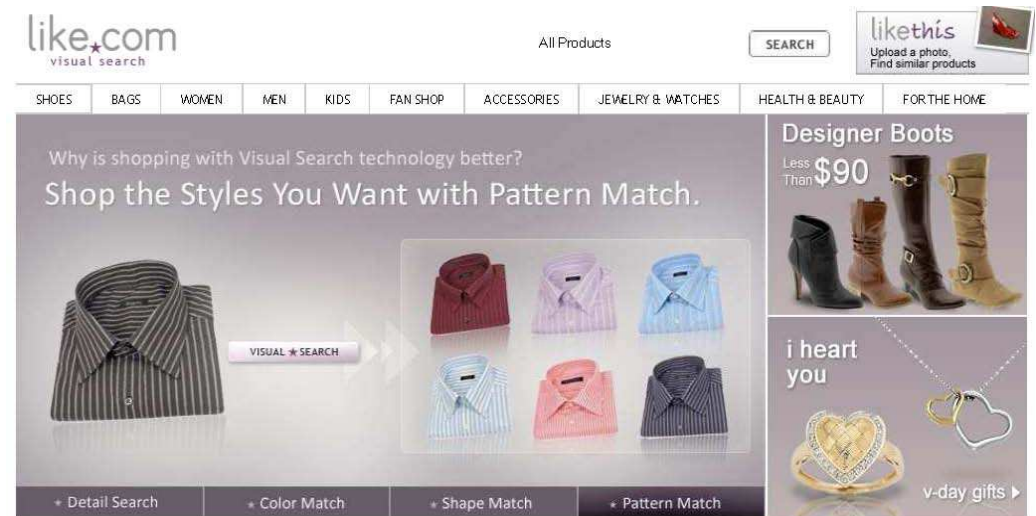
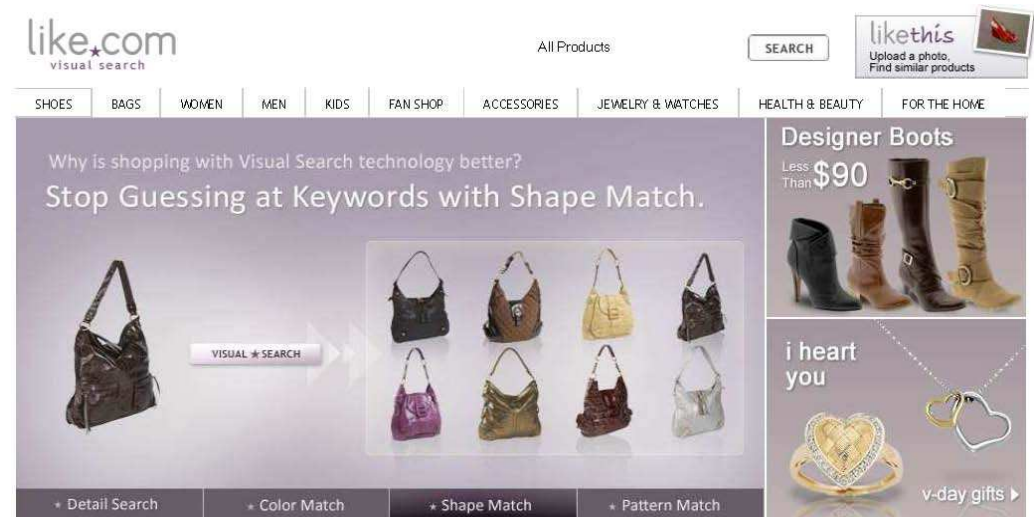
Content-based retrieval



Welcome! Click [VISUAL * SEARCH](#) to start visual shopping.



Online shopping catalog search



<http://www.like.com>

3D scanning and reconstruction



Adapted from Linda Shapiro, U of Washington

Earth viewers (3D modeling)



Motion capture



Visual effects



Adapted from CSE 455, U of Washington

Motion capture



Microsoft's XBox Kinect



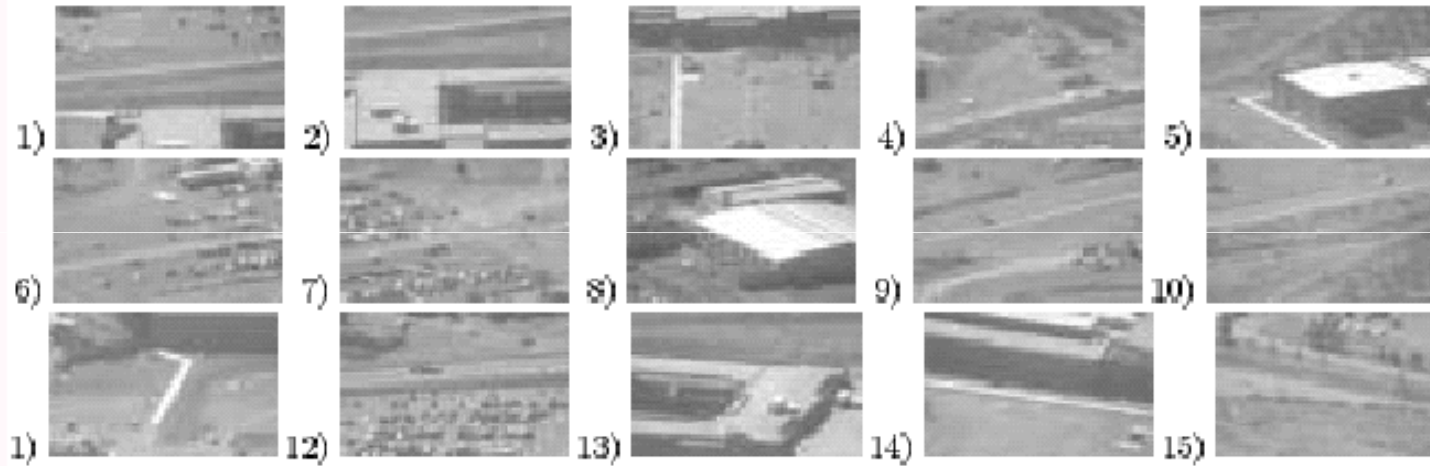
Adapted from CSE 455, U of Washington

Mozaic



Adapted from David Forsyth, UC Berkeley

Mozaic



Adapted from David Forsyth, UC Berkeley

Critical issues

- What information should be extracted?
- How can it be extracted?
- How should it be represented?
- How can it be used to aid analysis and understanding?

Perception and grouping

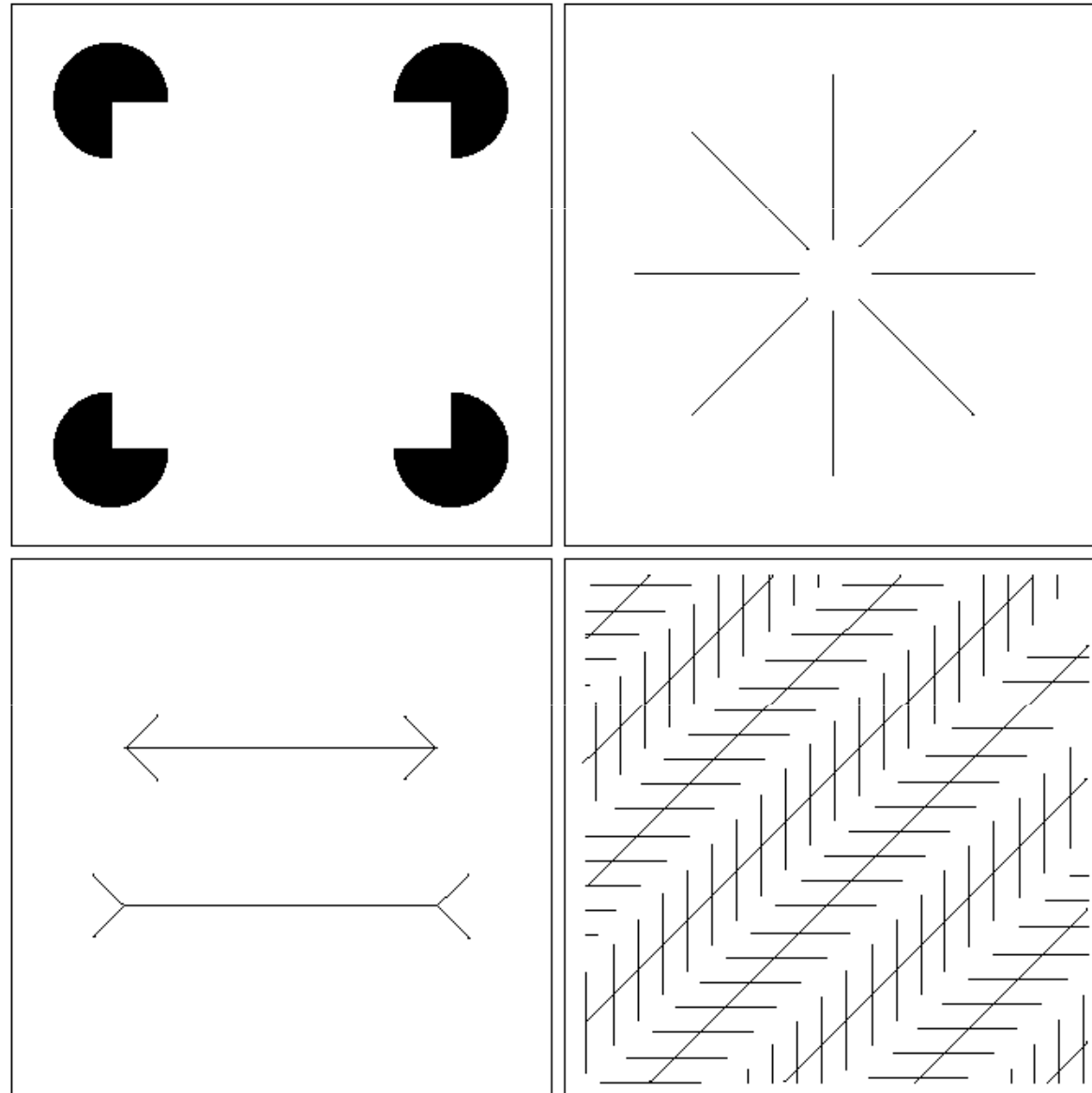


Subjective
contours

Perception and grouping

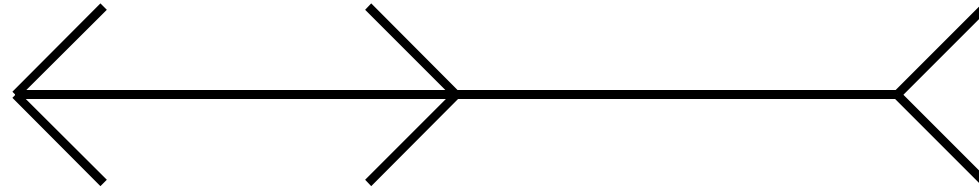
a b
c d

FIGURE 2.9 Some well-known optical illusions.

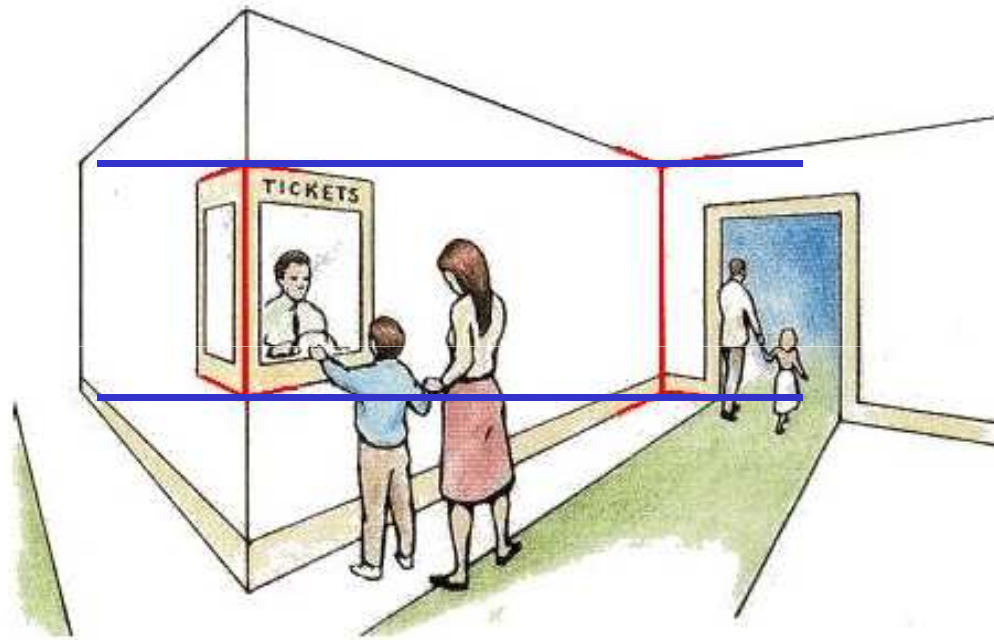


Adapted from Gonzales and Woods

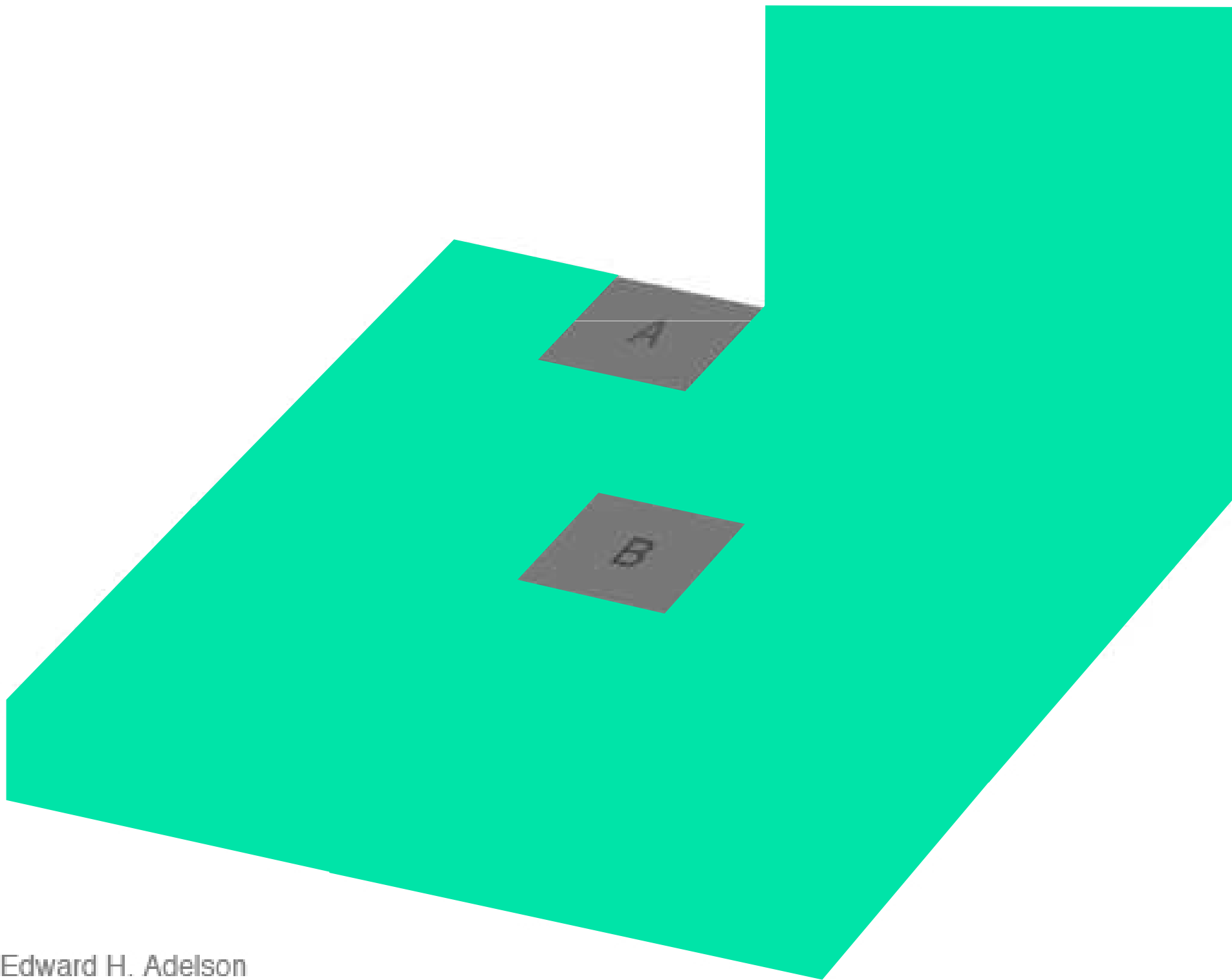
Perception and grouping

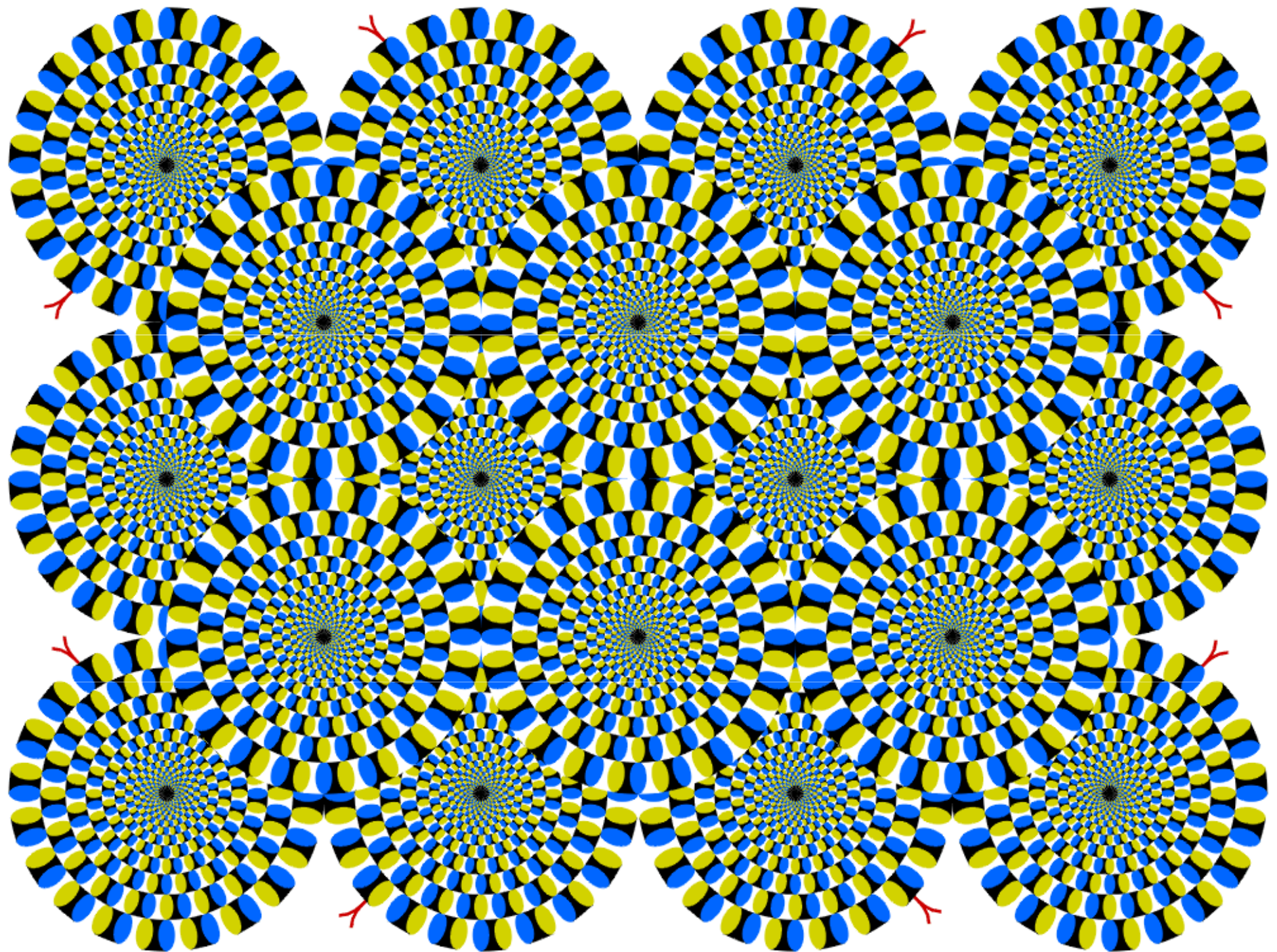


Müller-Lyer Illusion



Adapted from Alyosha Efros, Carnegie Mellon





Perception and grouping

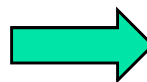


Occlusion

* types of
“junctions” give
cues about
surfaces,
occlusion, and
light.

I. Rock, The Logic of Perception, 1983.

Adapted from Michael Black, Brown University

[illegible]

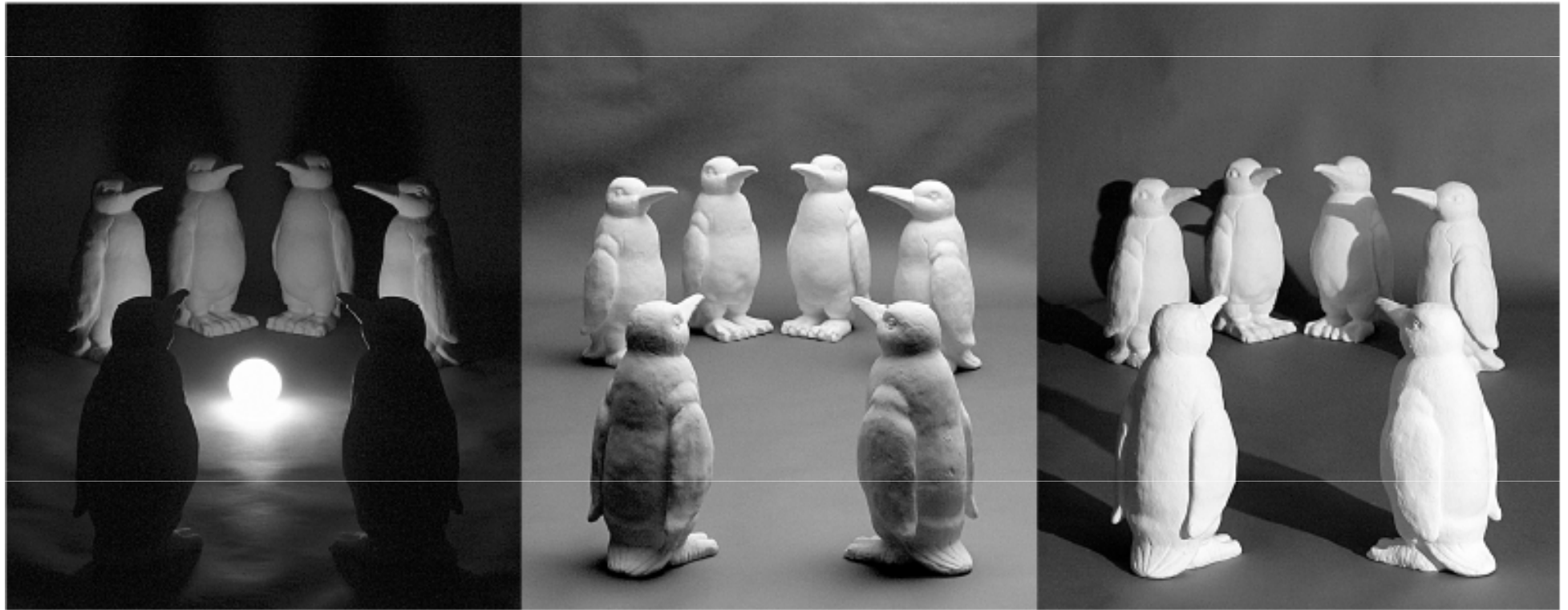
Challenges 1: view point variation



Michelangelo 1475-1564

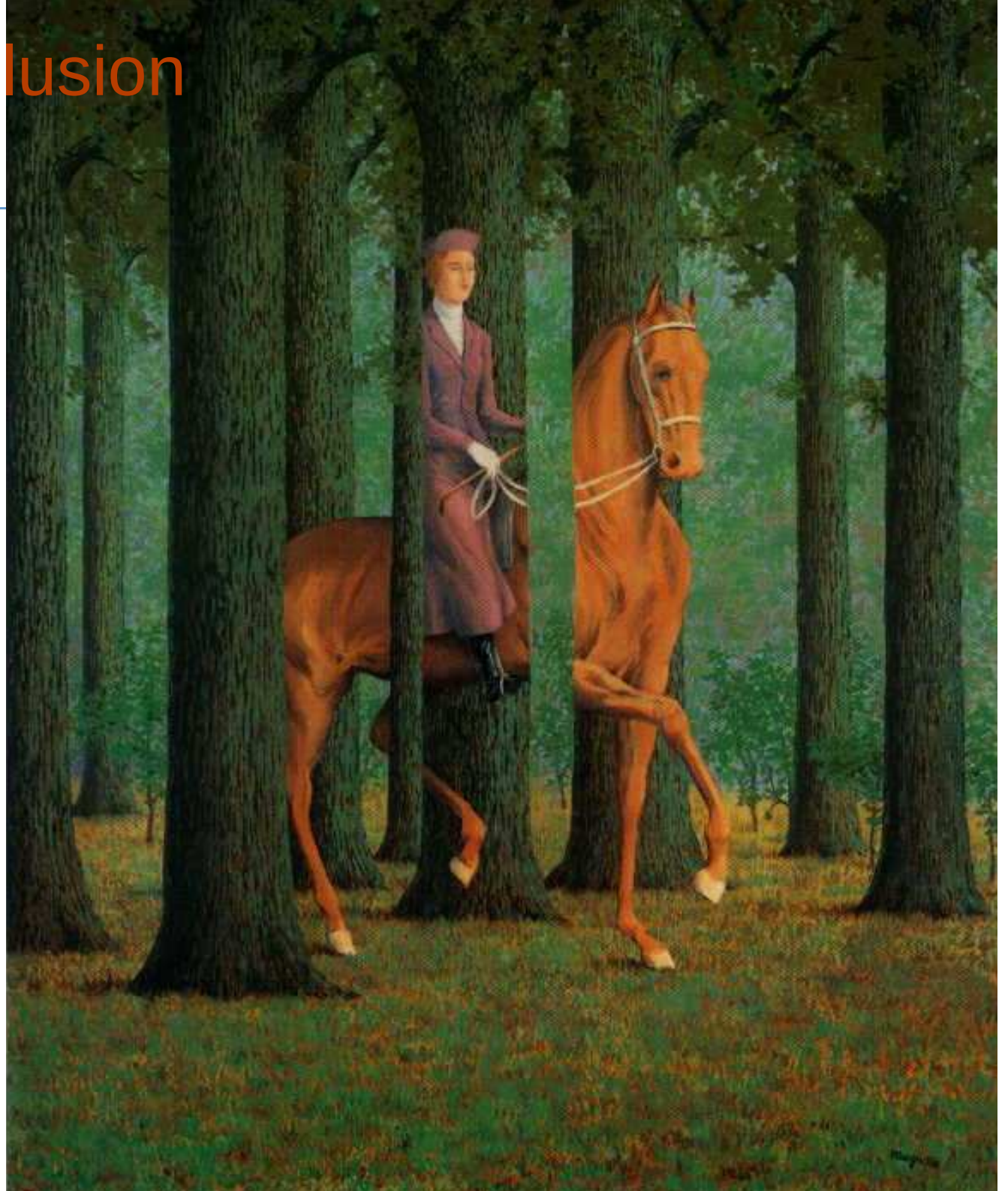


Challenges 2: illumination



Challenges 3: occlusion

Magritte, 1957



Adapted from L. Fei-Fei,
R. Fergus, A. Torralba

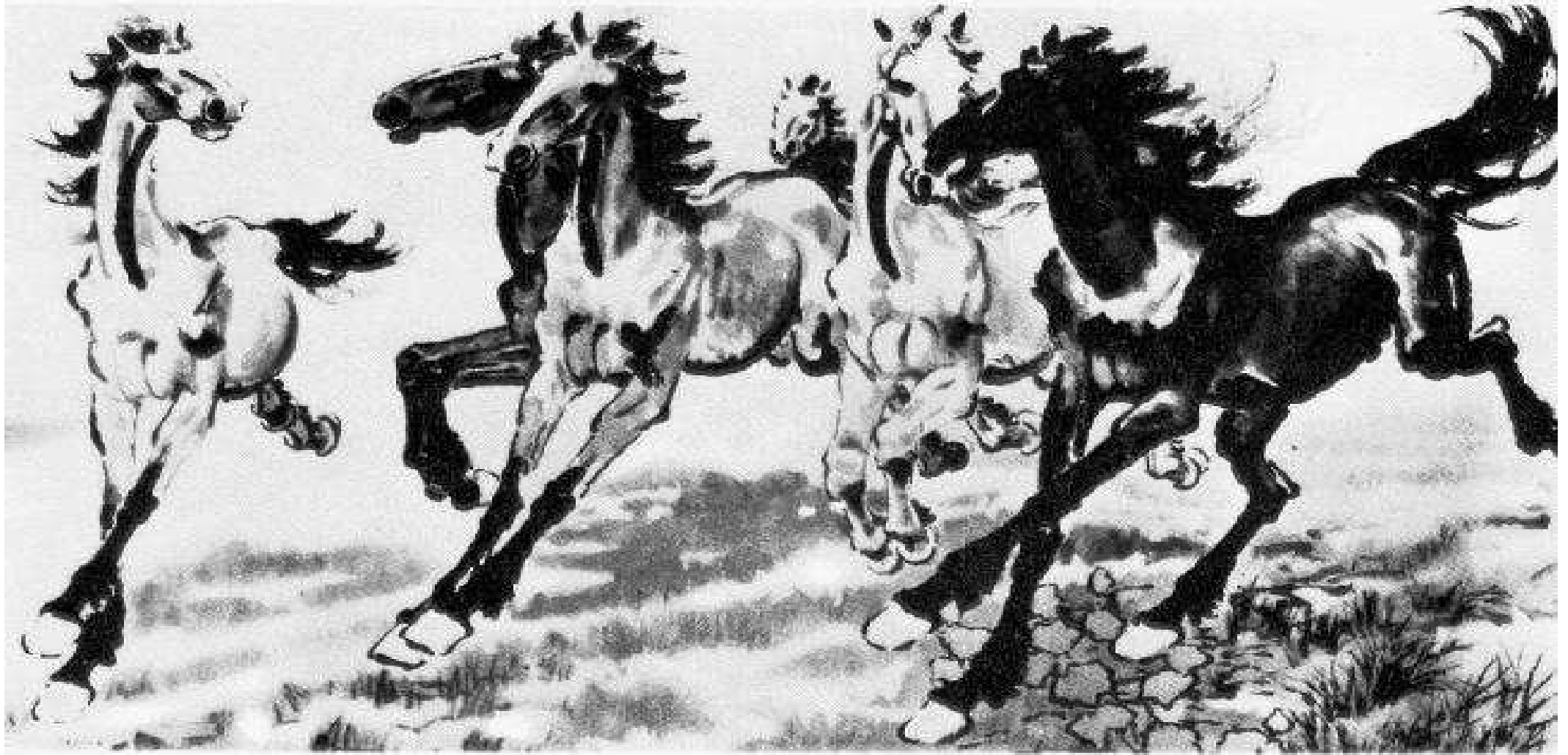
CS 484, Spring 2012

Challenges 4: scale



Adapted from L. Fei-Fei,
R. Fergus, A. Torralba

Challenges 5: deformation



Xu, Beihong 1943

Challenges 6: background clutter



Adapted from Fei-Fei Li

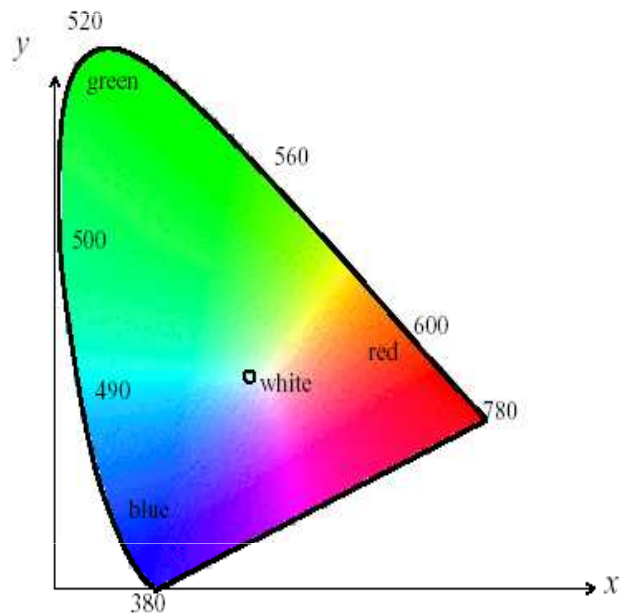
Challenges 7: intra-class variation



Recognition

- How can different cues such as color, texture, shape, motion, etc., can be used for recognition?
 - Which parts of image should be recognized together?
 - How can objects be recognized without focusing on detail?
 - How can objects with many free parameters be recognized?
 - How do we structure very large model bases?

Color



Adapted from Martial Hebert, CMU

Texture



Color, texture, and proximity



Adapted from Fei-Fei Li

Segmentation

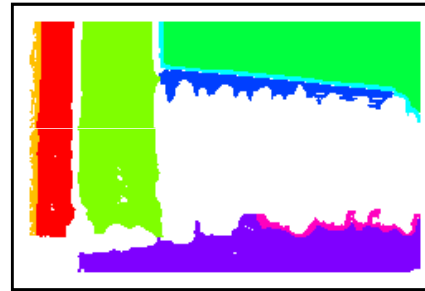
Original Images



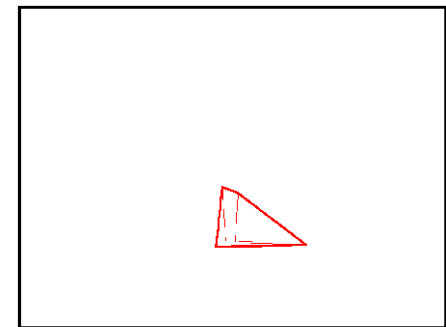
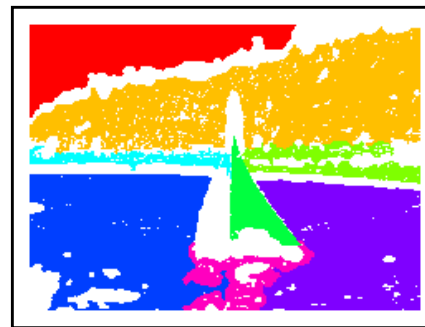
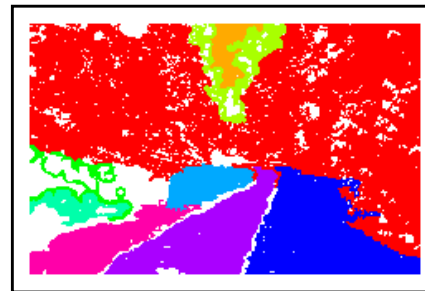
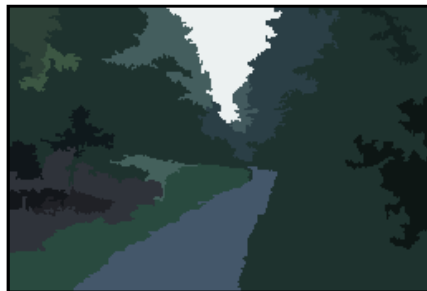
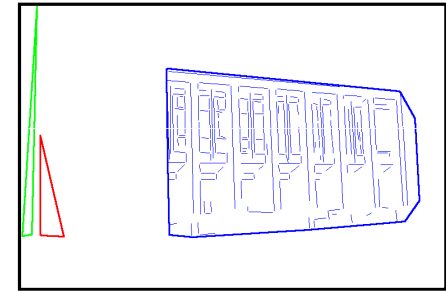
Color Regions



Texture Regions

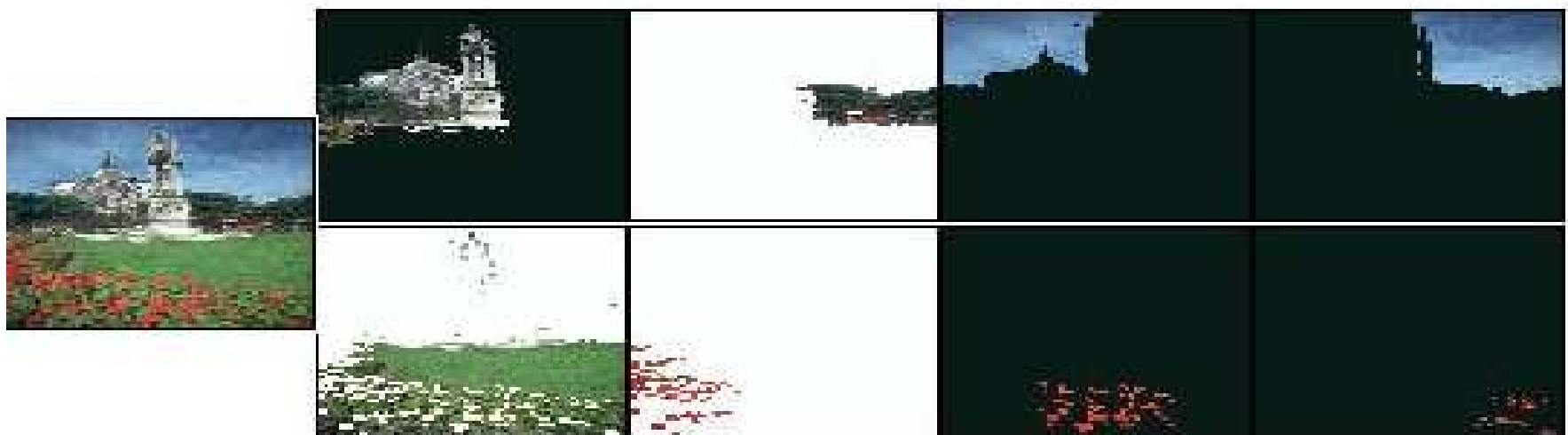
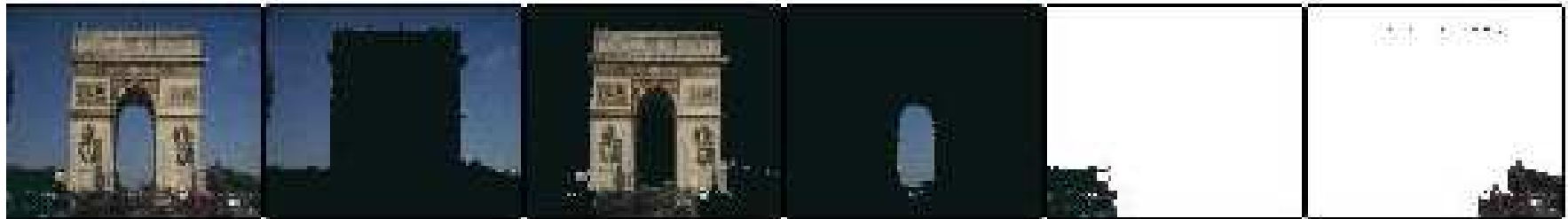


Line Clusters

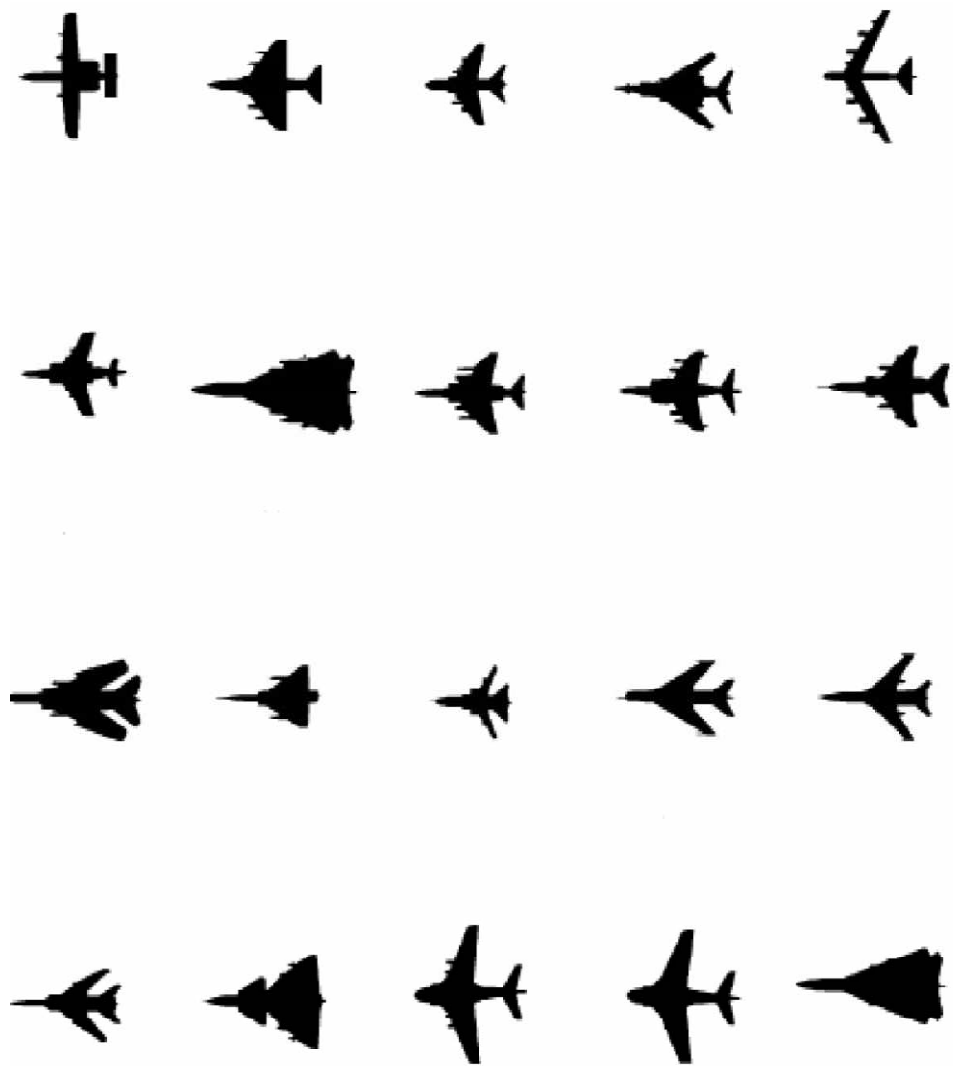


Adapted from Linda Shapiro, U of Washington

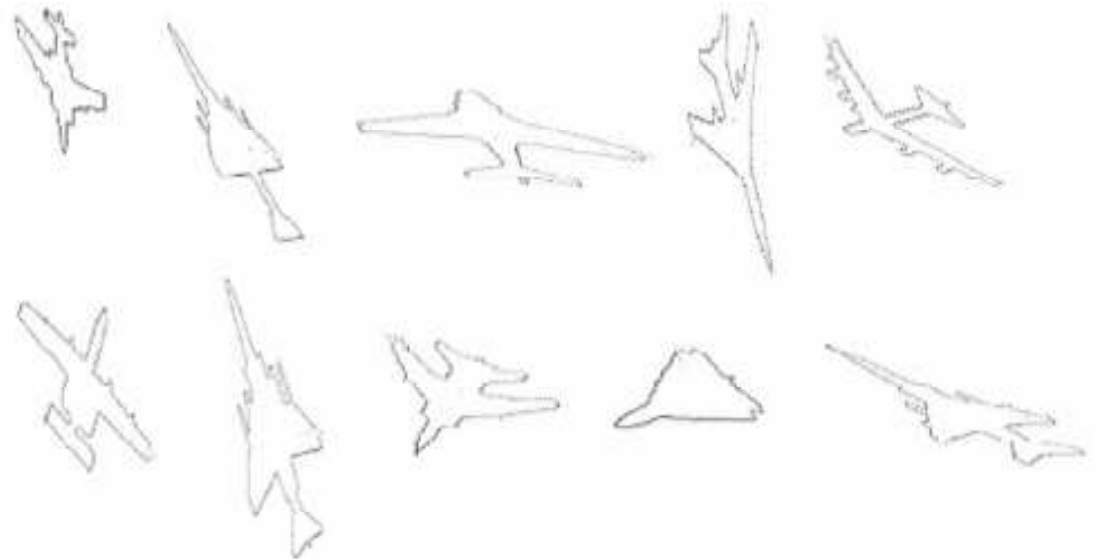
Segmentation



Shape



Model database



Recognized objects

Adapted from Enis Cetin, Bilkent University

Motion



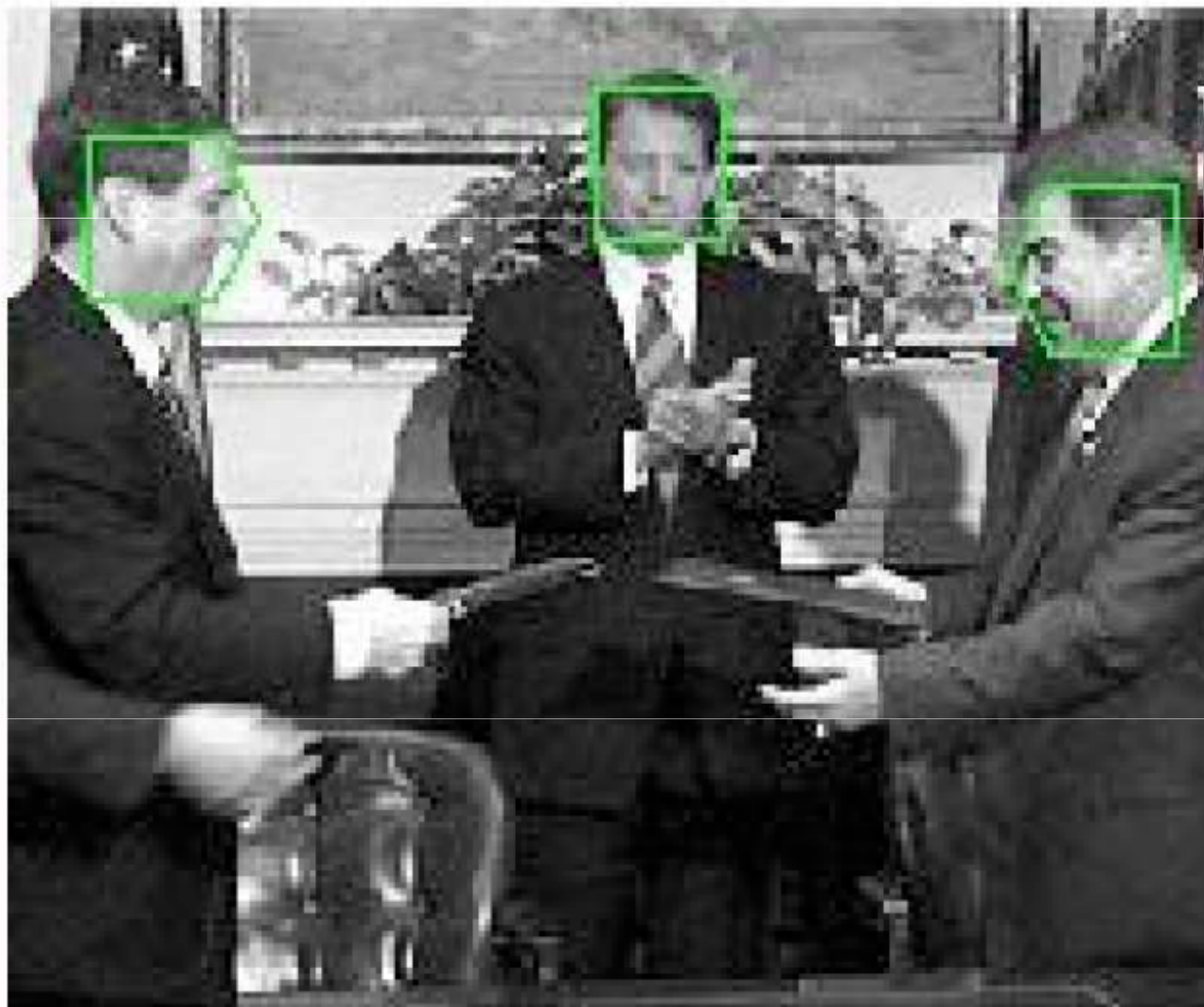
Adapted from Michael Black, Brown University

Detection



Adapted from David Forsyth, UC Berkeley

Detection

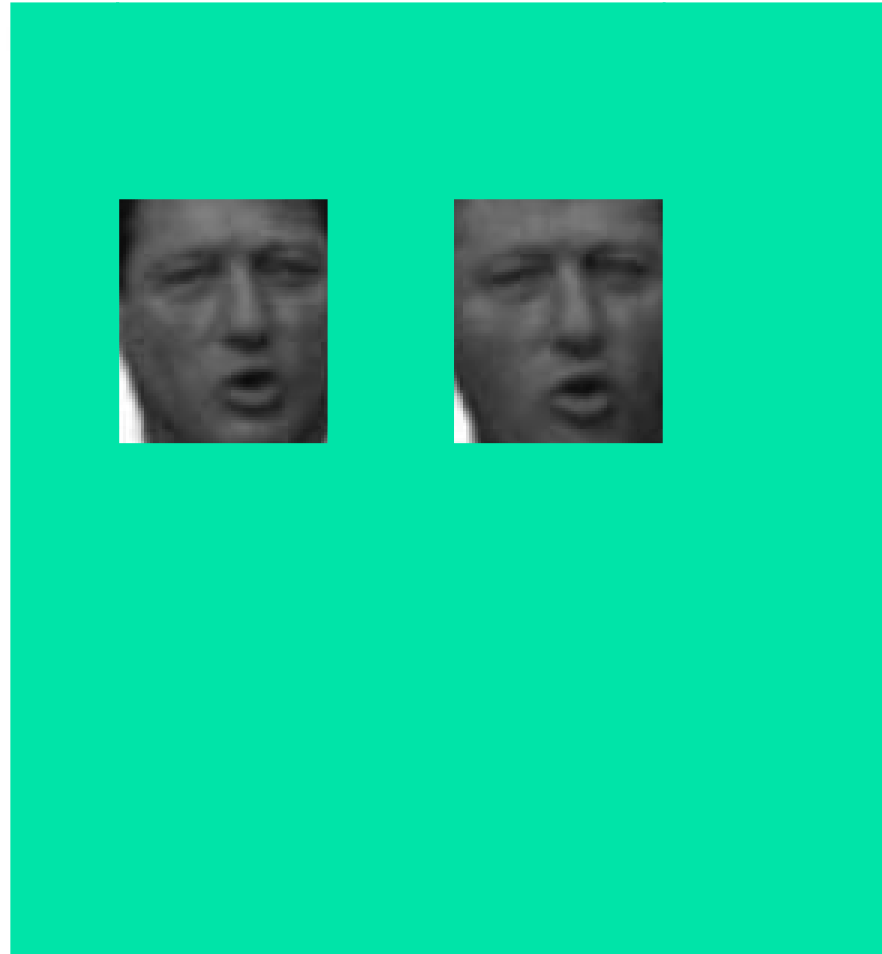


Adapted from David Forsyth, UC Berkeley

Detection

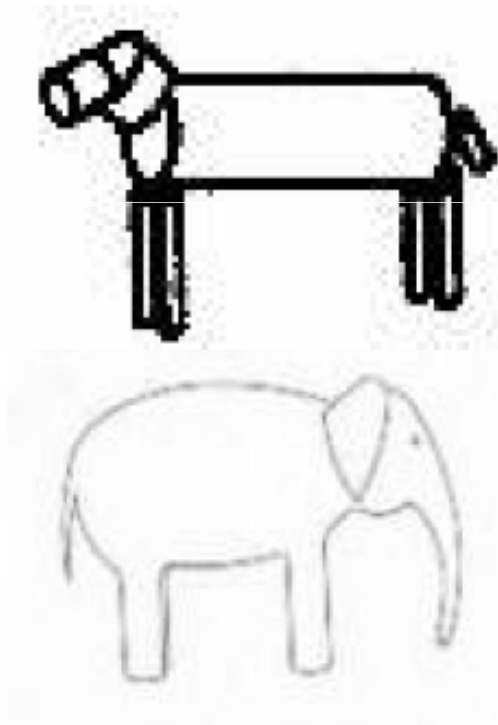
What are our
“models”?

How good
are they?

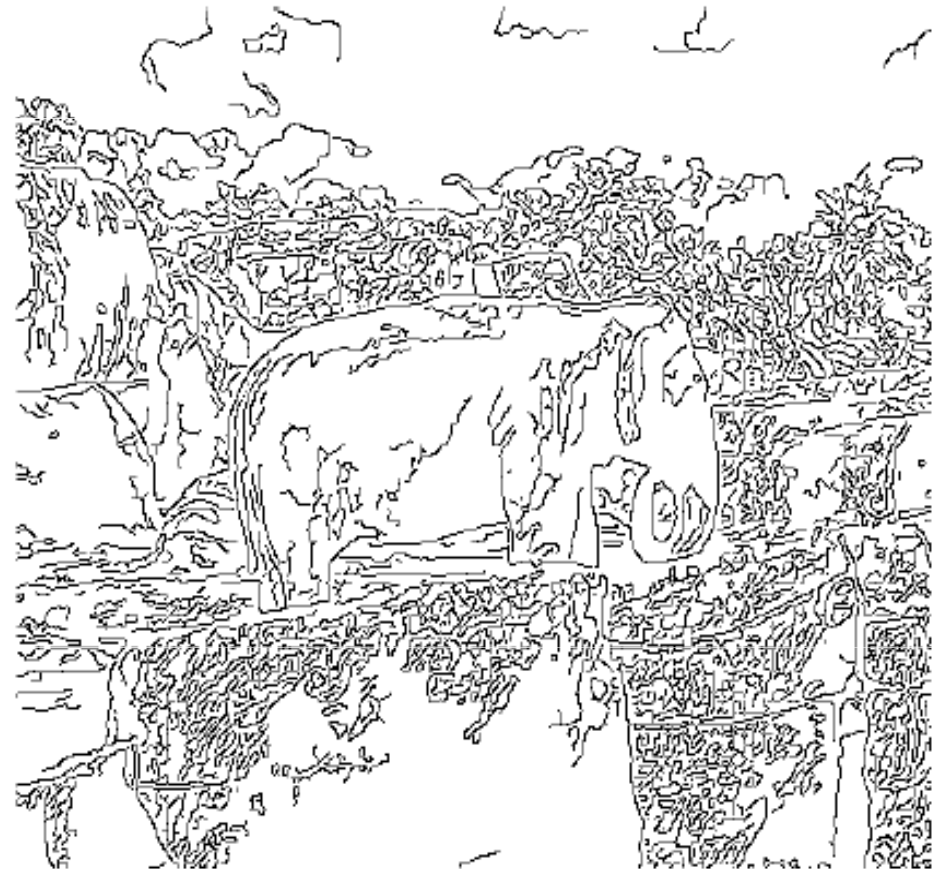


Adapted from Michael Black, Brown University

Recognition

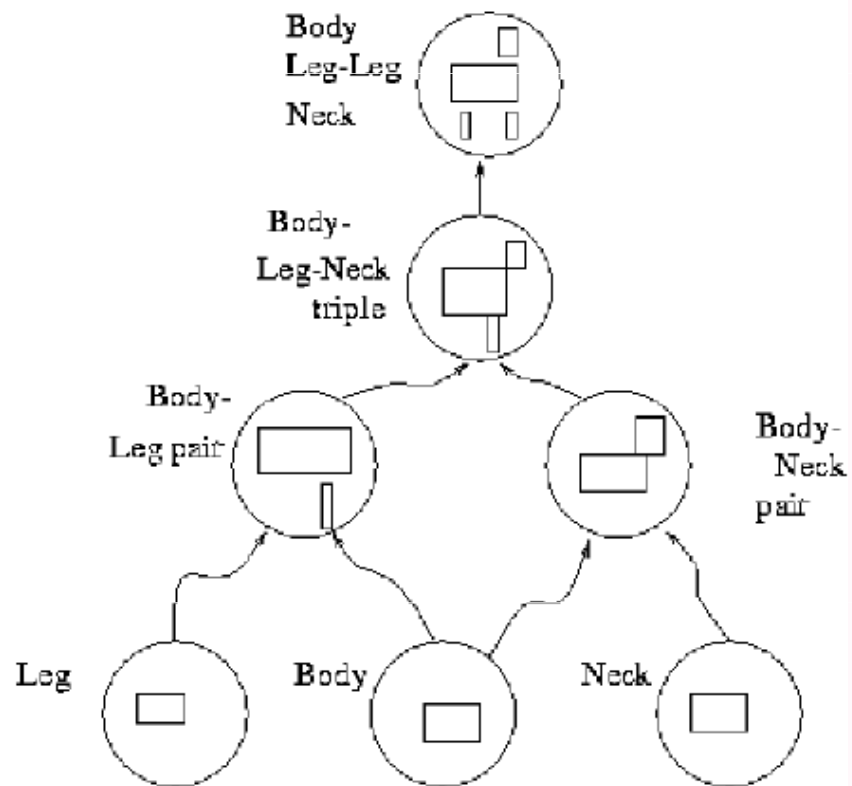


Match “model” to
measurements?



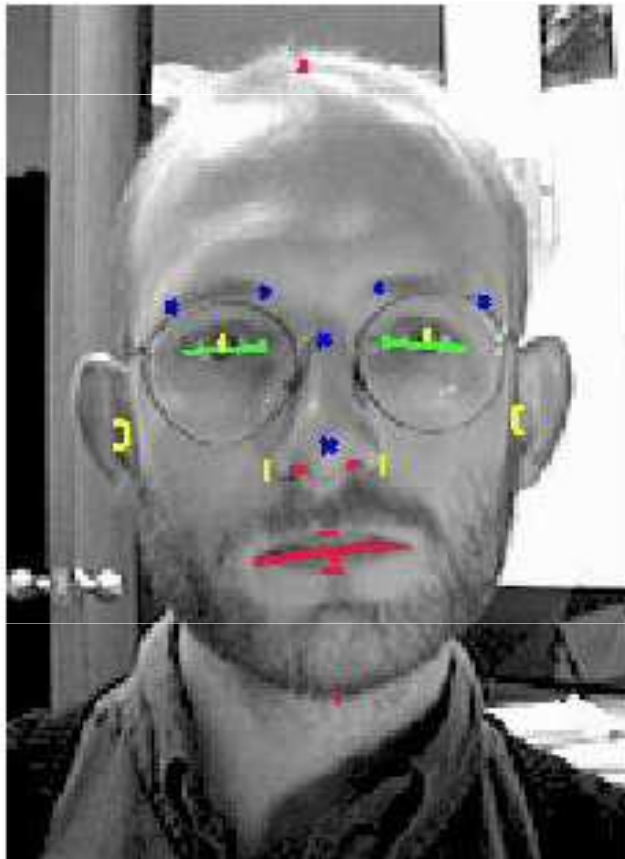
Adapted from Michael Black, Brown University

Recognition

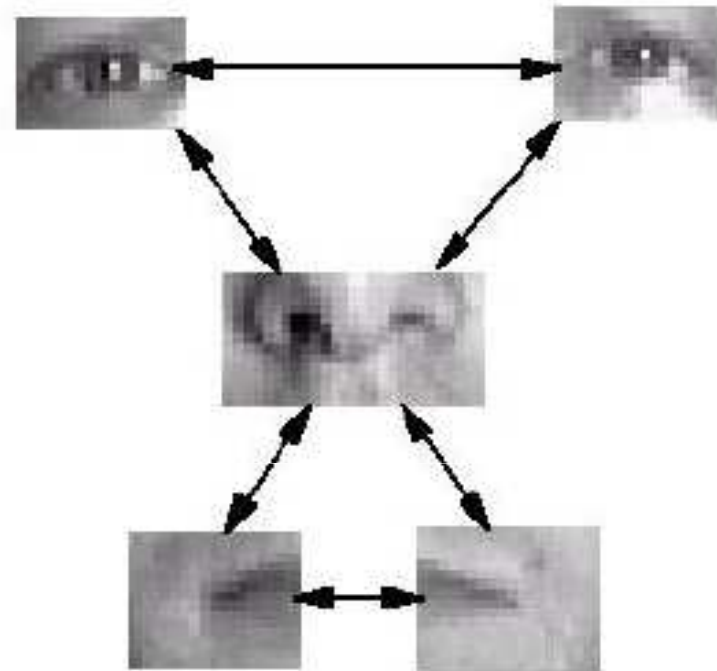


Adapted from David Forsyth, UC Berkeley

Parts and relations



Patch Model



<http://www.research.ibm.com/ecvg/biom/facereco.html>

Adapted from Michael Black, Brown University

Parts and relations



How flexible are the spatial relations of the parts?

Adapted from Michael Black, Brown University

Context



Adapted from Antonio Torralba, MIT

Context



Adapted from Antonio Torralba, MIT

Context



Adapted from Derek Hoiem, CMU

Context



Adapted from
Derek Hoiem, CMU

Stages of computer vision

- Low-level

image → image

- Mid-level

↑ image → features / attributes
↓ Image analysis / image understanding

- High-level

features → “making sense”, recognition

Low-level

sharpening
→



←
blurring

Adapted from Linda Shapiro, U of Washington

Low-level



original image

Canny



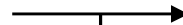
edge image

Mid-level



edge image

ORT



data
structure



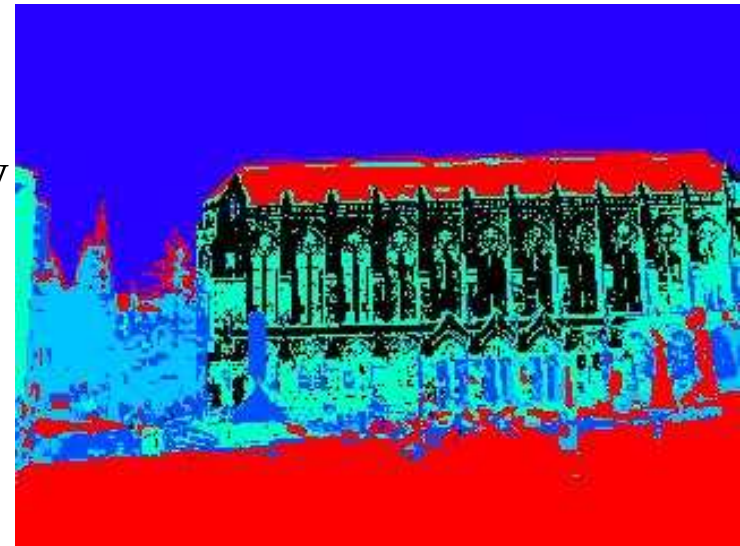
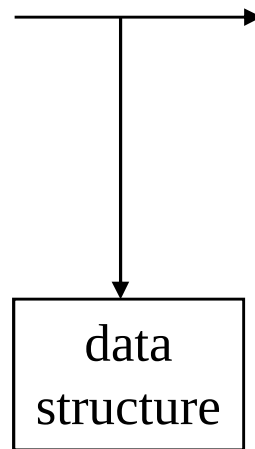
circular arcs and line segments

Mid-level



original color image

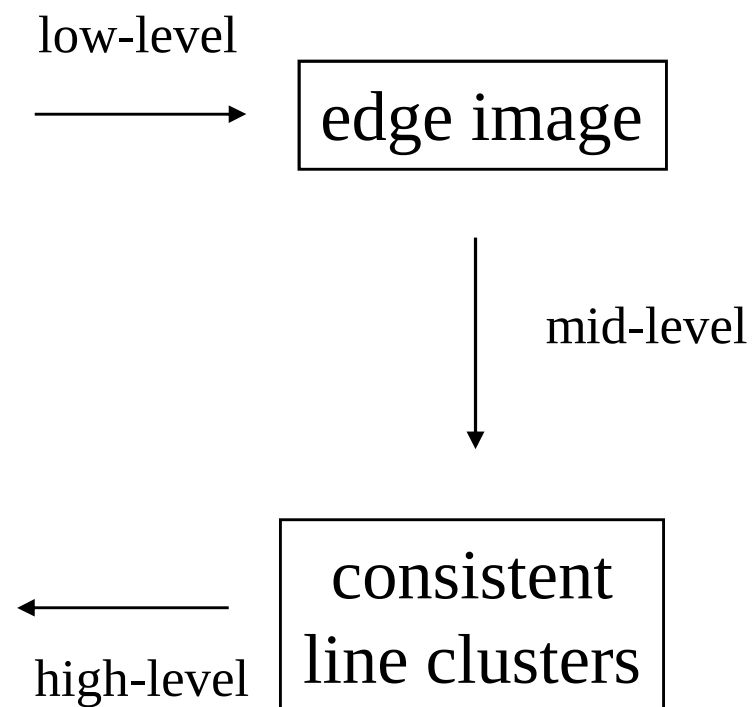
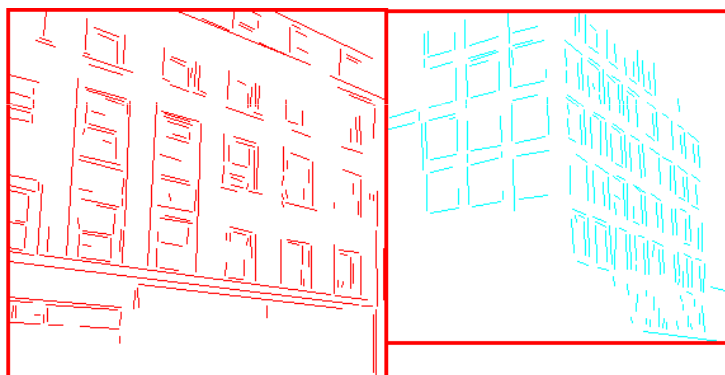
K-means
clustering
(followed by
connected
component
analysis)



regions of homogeneous color

Adapted from Linda Shapiro, U of Washington

Low-level to high-level



Adapted from Linda Shapiro, U of Washington