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

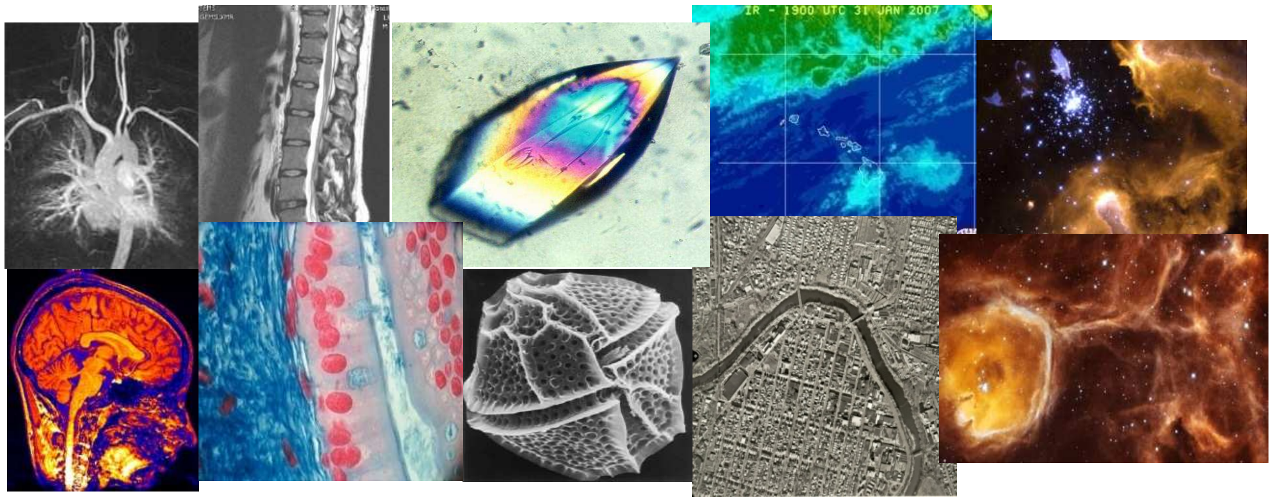
Why study computer vision?

Personal photo albums

Movies, news, sports

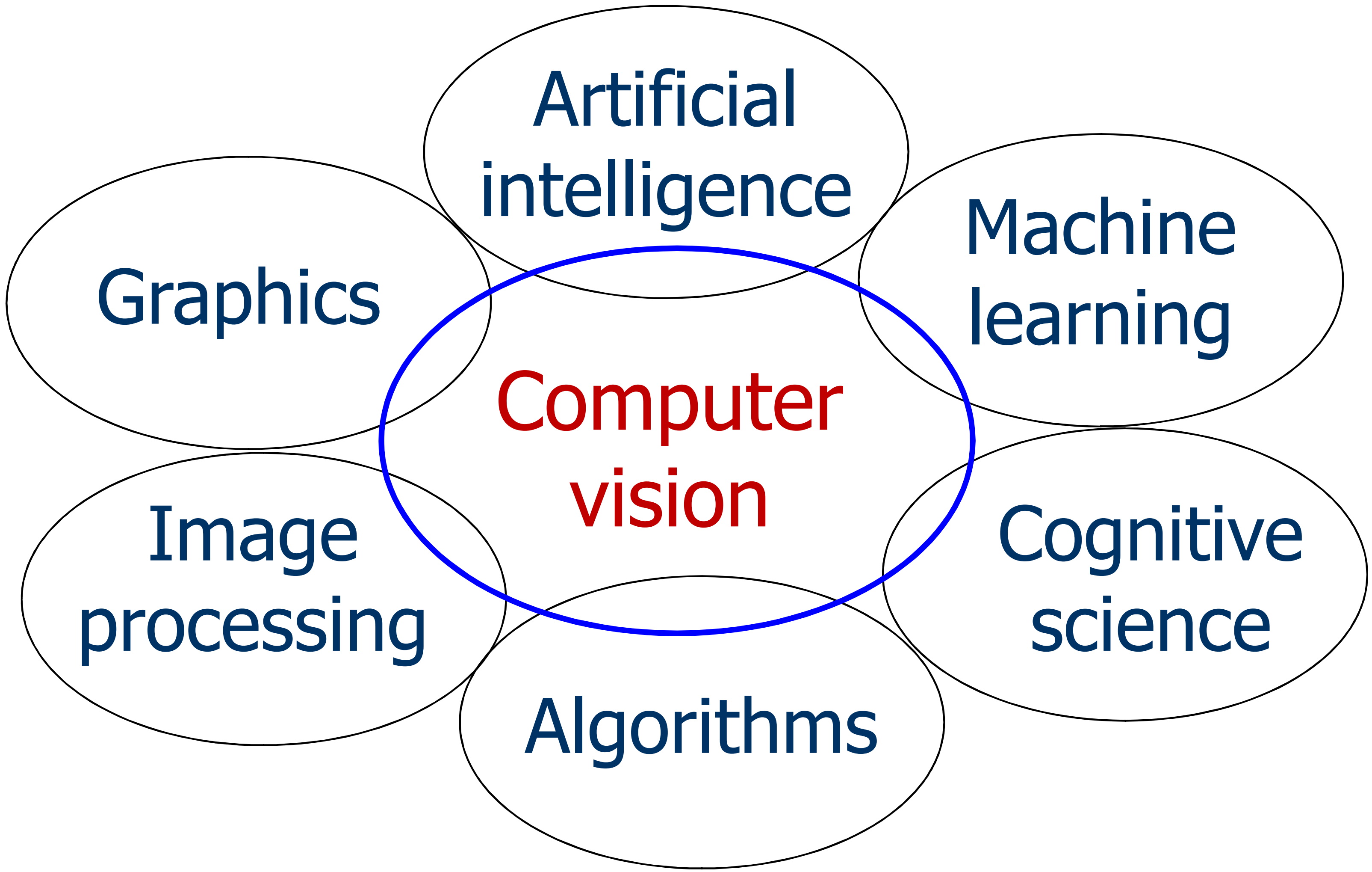
Google Image Search Picasa flickr webshots picsearch YouTube Broadcast Yourself

Surveillance and security



Medical and scientific images

Related disciplines



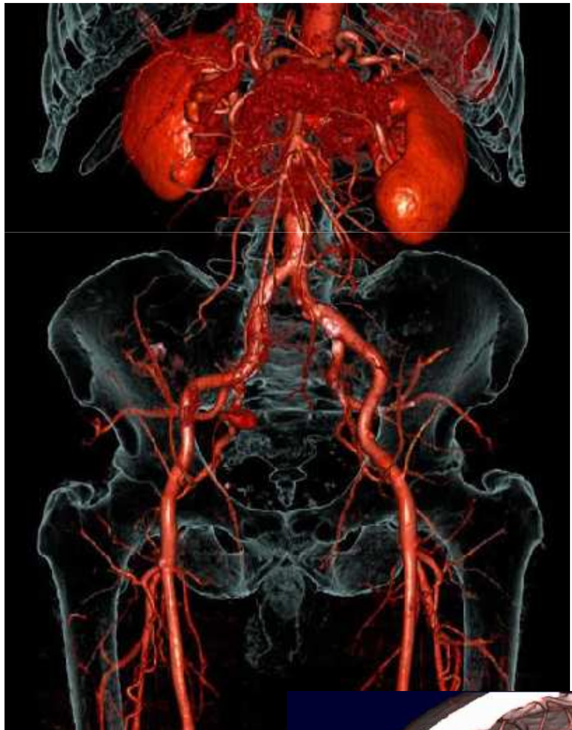
Applications

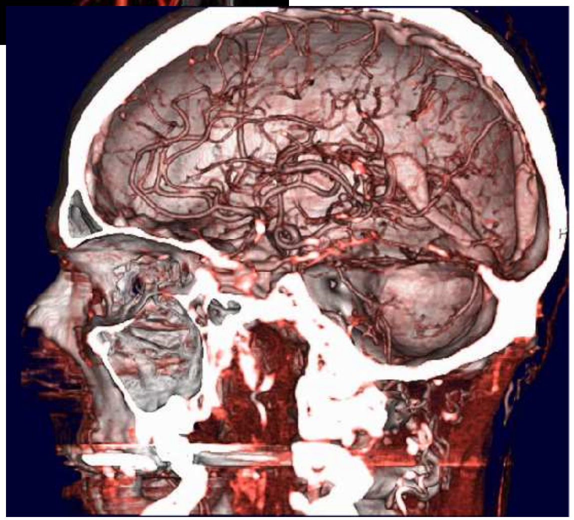
- Medical image analysis
- Security
  - Biometrics
  - Surveillance
  - Tracking
  - Target recognition
- Remote sensing
- Robotics
- Industrial inspection, quality control
- Document analysis
- Multimedia
- Assisted living
- Human-computer interfaces
- …

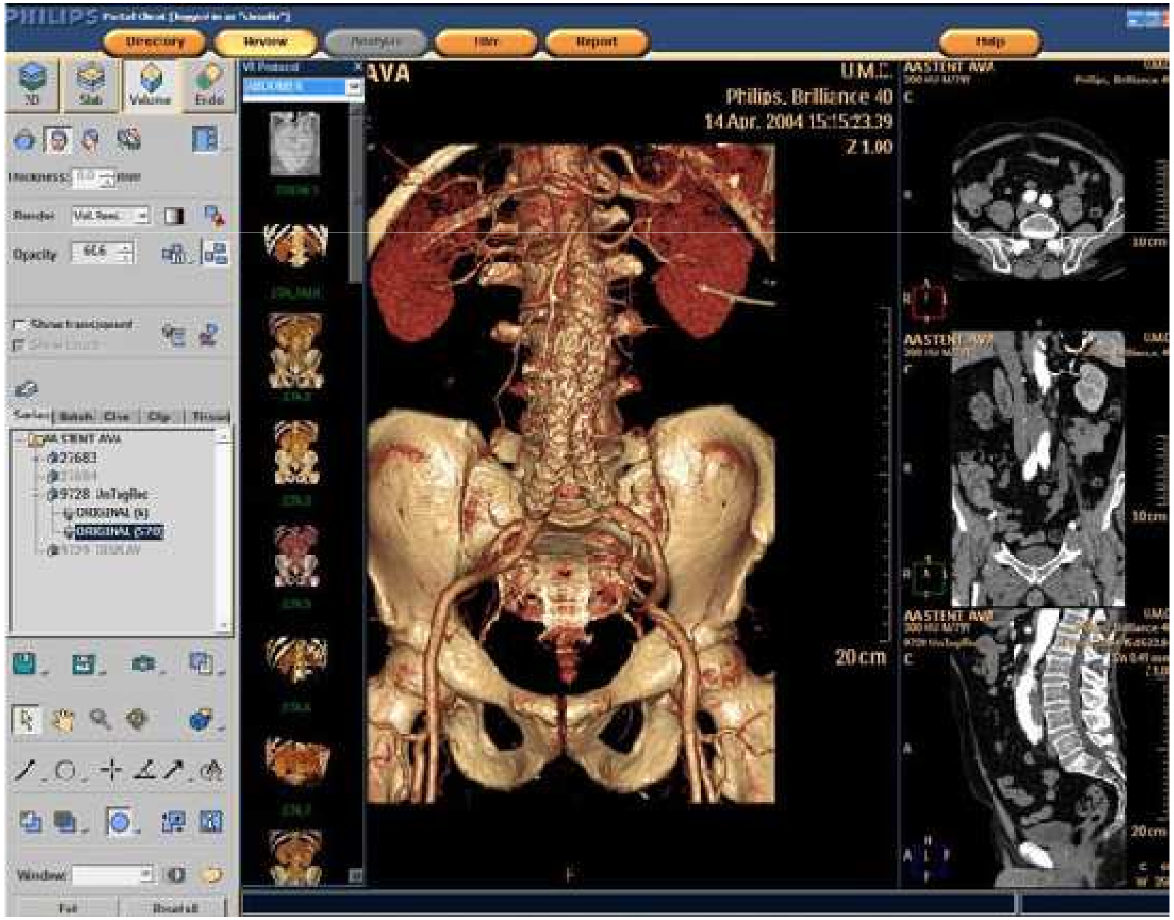
Medical image analysis

http://www.clarontech.com

Medical image analysis

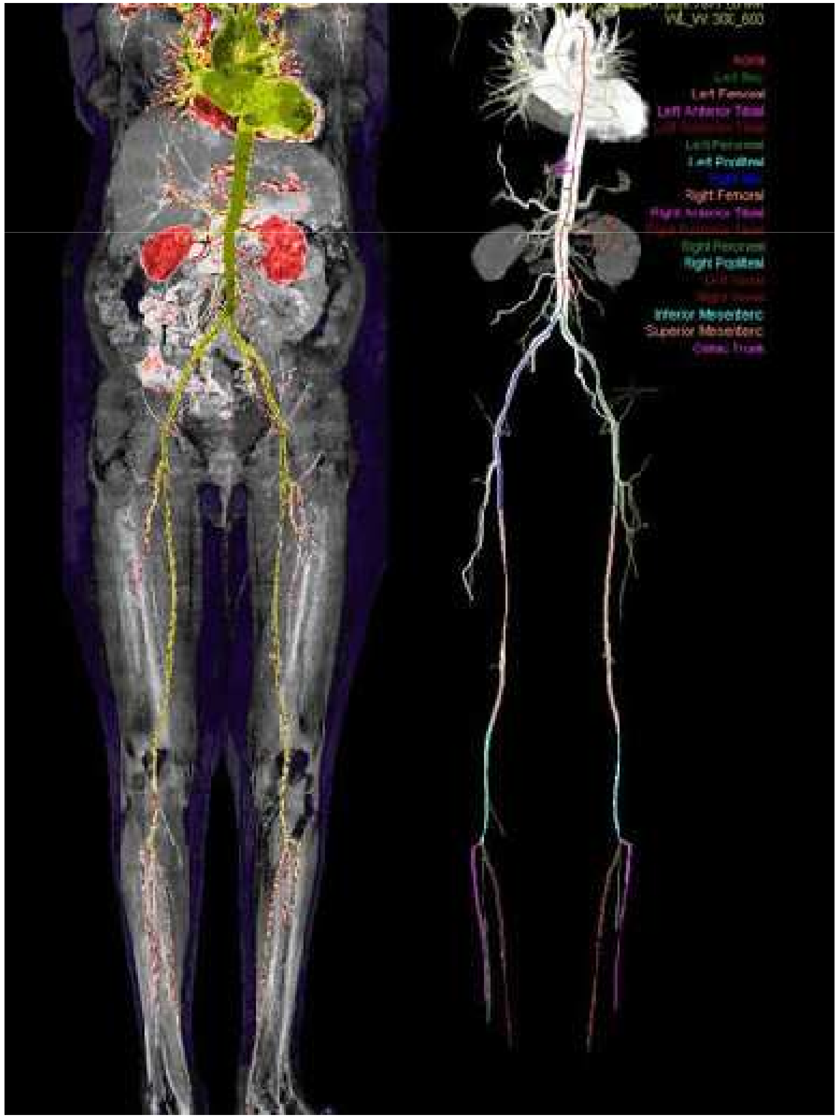


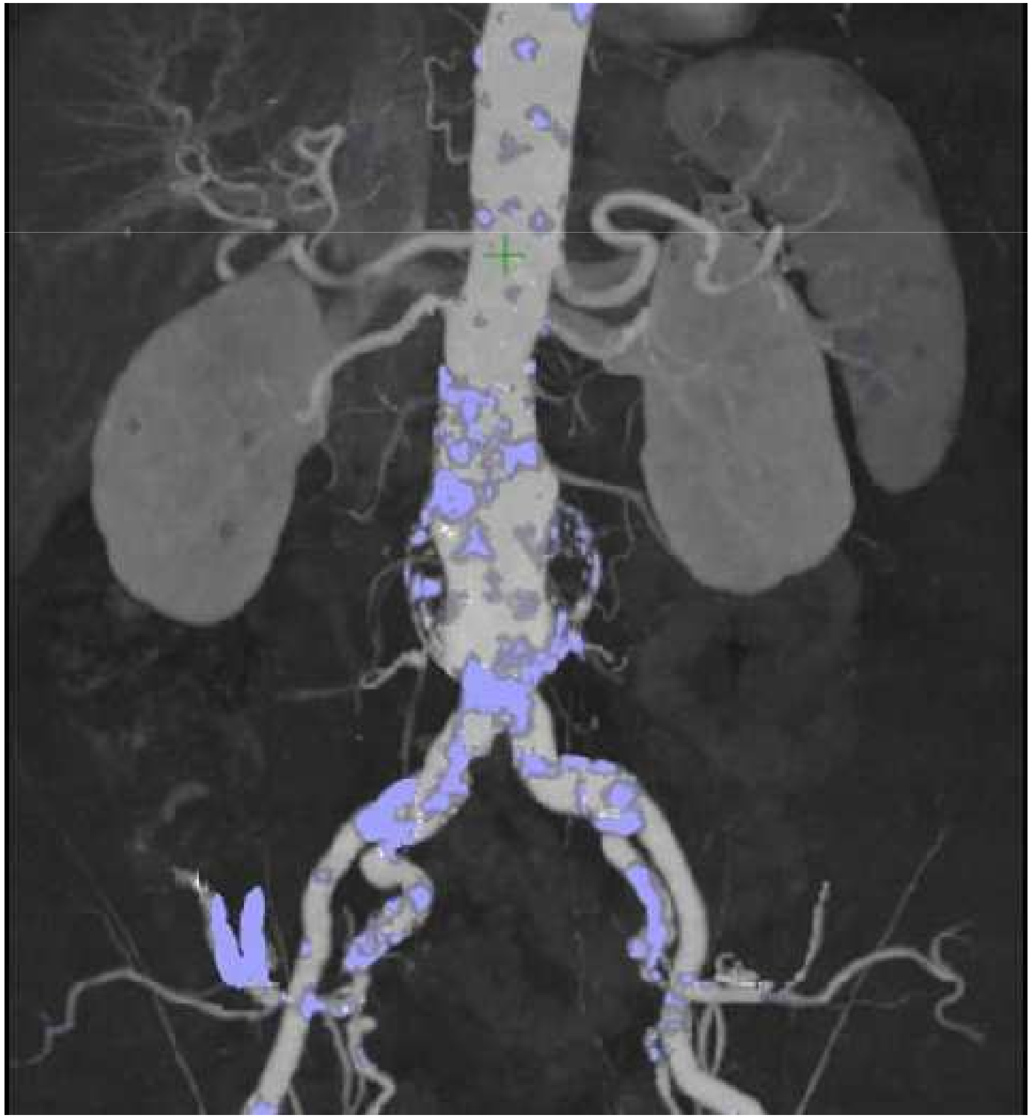




http://www.clarontech.com

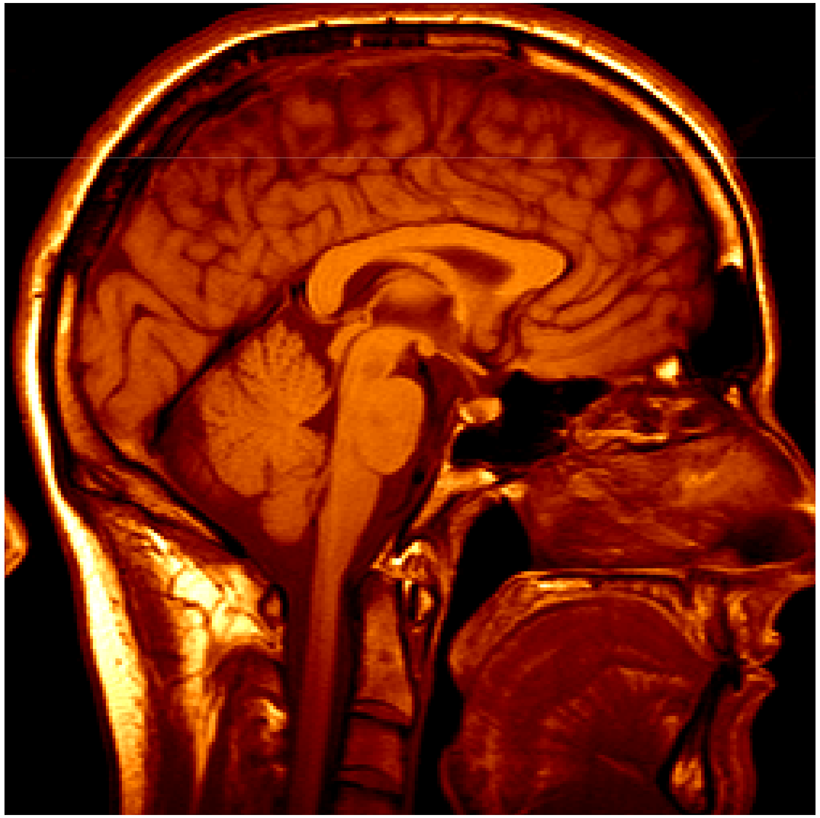
Medical image analysis





http://www.clarontech.com

Medical image analysis

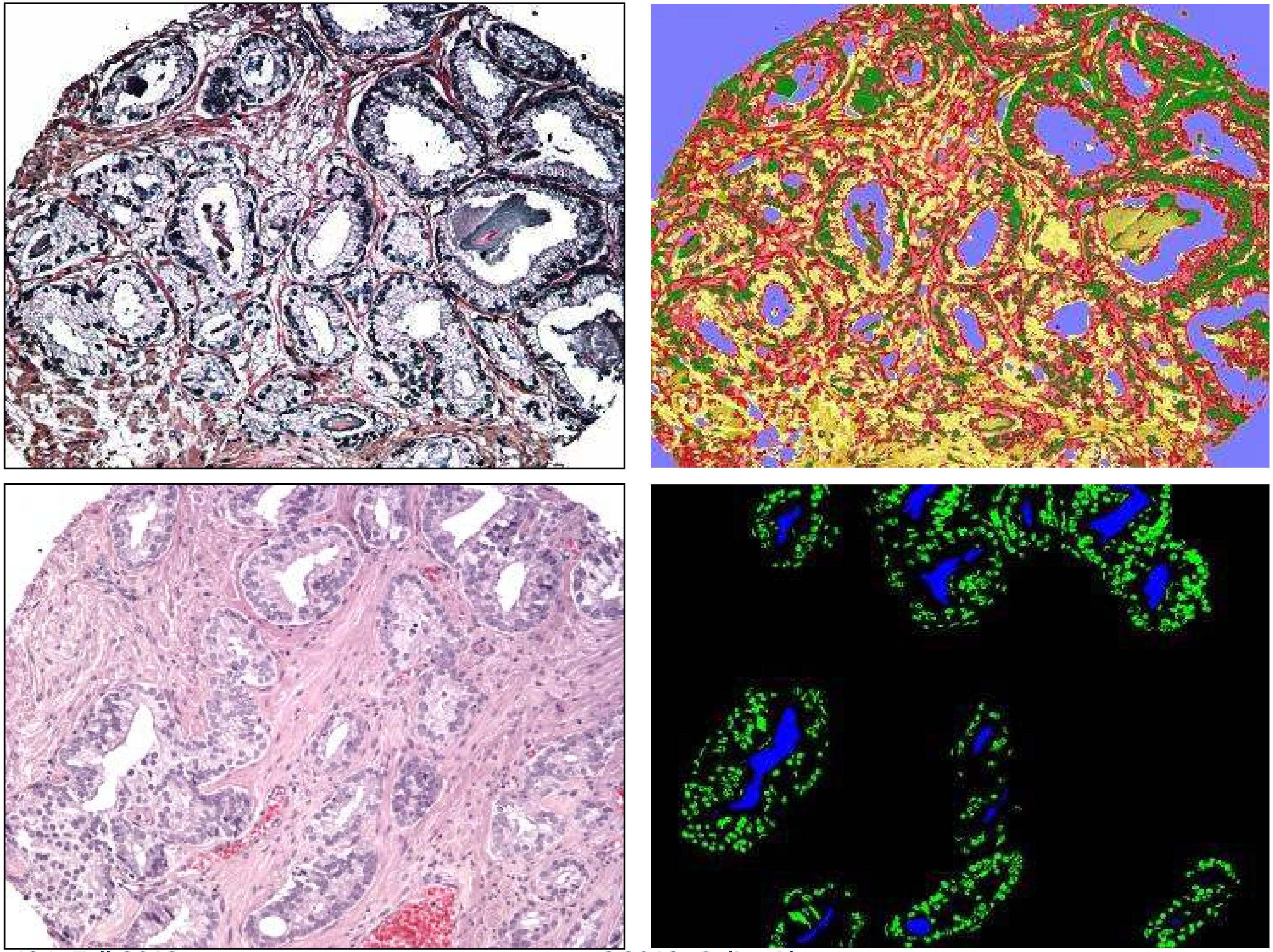


3D imaging: MRI, CT

Image guided surgery
Grimson et al., MIT

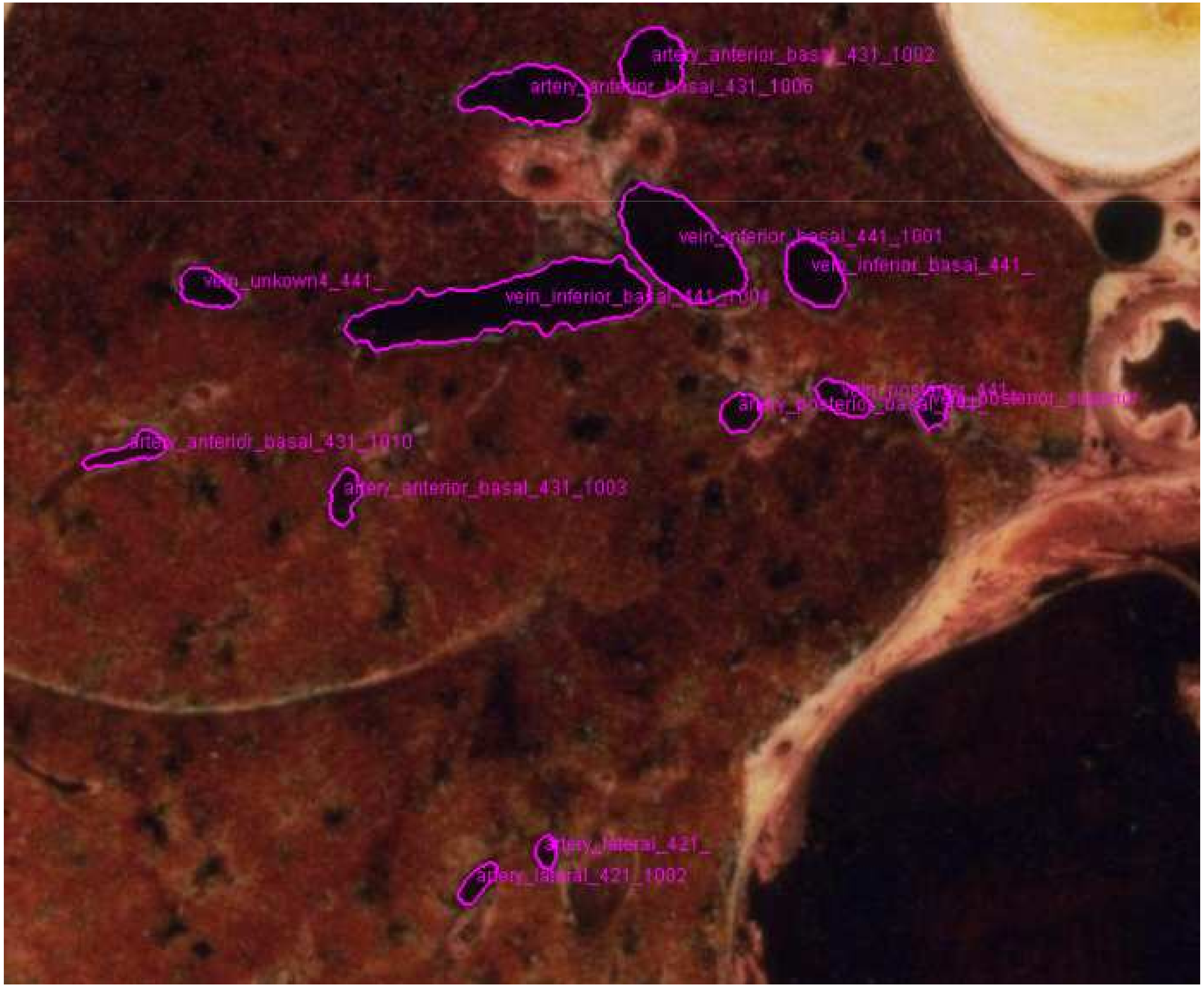
Adapted from CSE 455, U of Washington

Medical image analysis



Cancer detection and grading

Medical image analysis



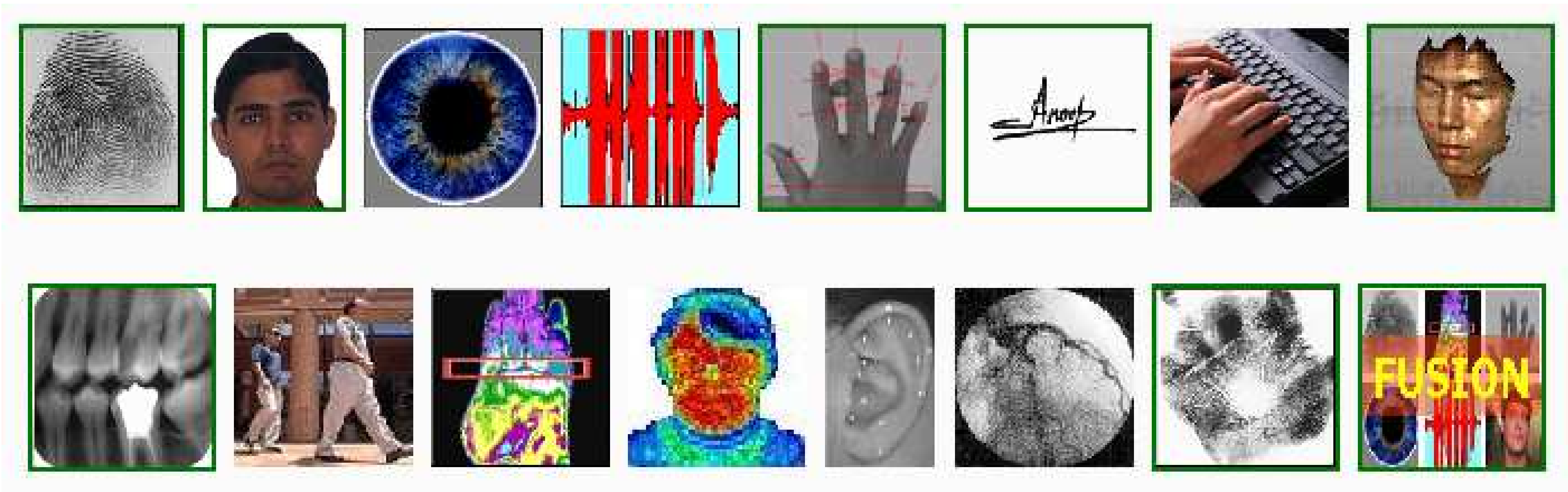
Slice of lung

Adapted from Linda Shapiro, U of Washington

Medical image analysis

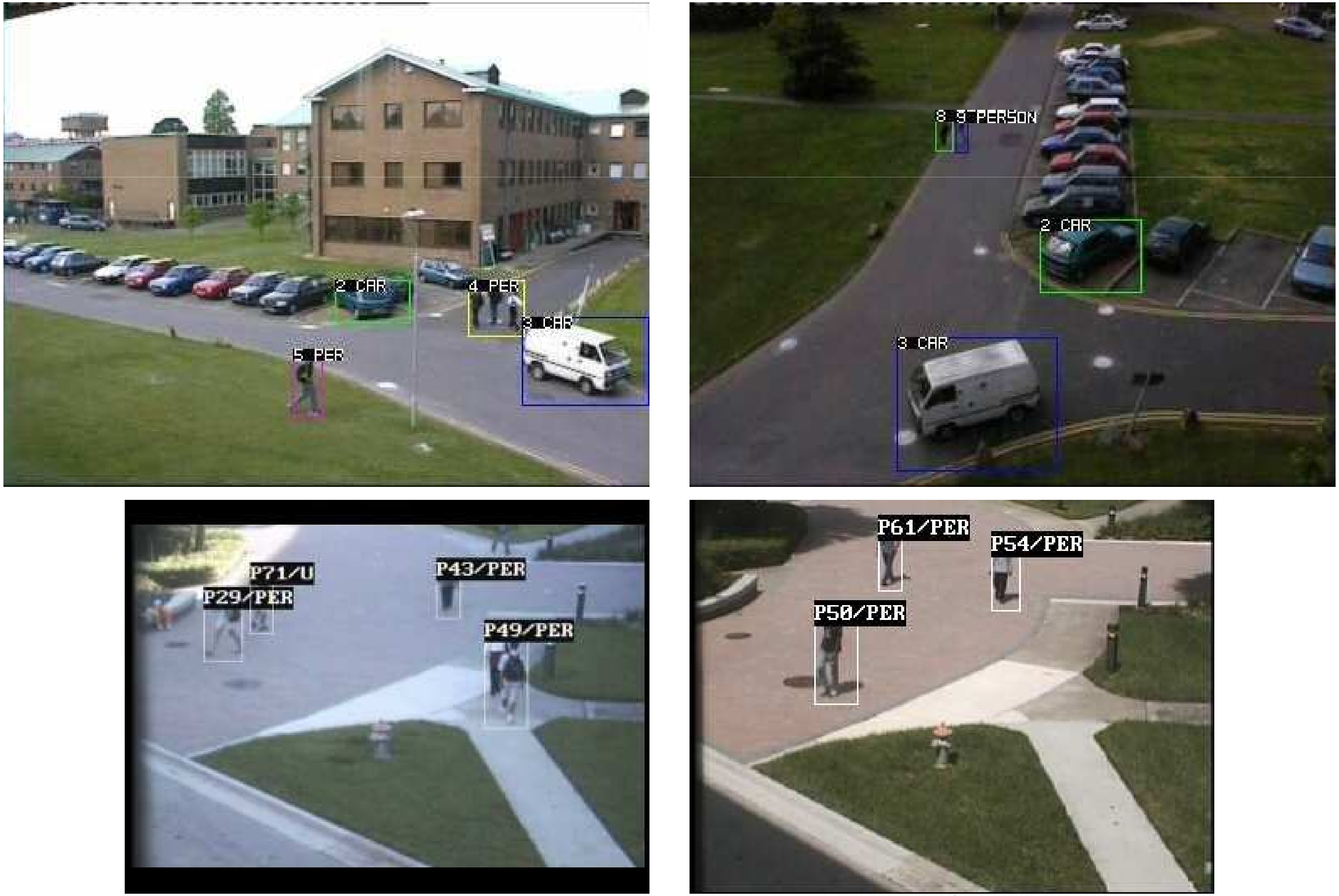
"Your x-ray showed a broken rib, but we fixed it with Photoshop."

Biometrics



Adapted from Anil Jain, Michigan State

Surveillance and tracking



University of Central Florida, Computer Vision Lab

Surveillance and tracking

Adapted from Octavia Camps, Penn State

Surveillance and tracking

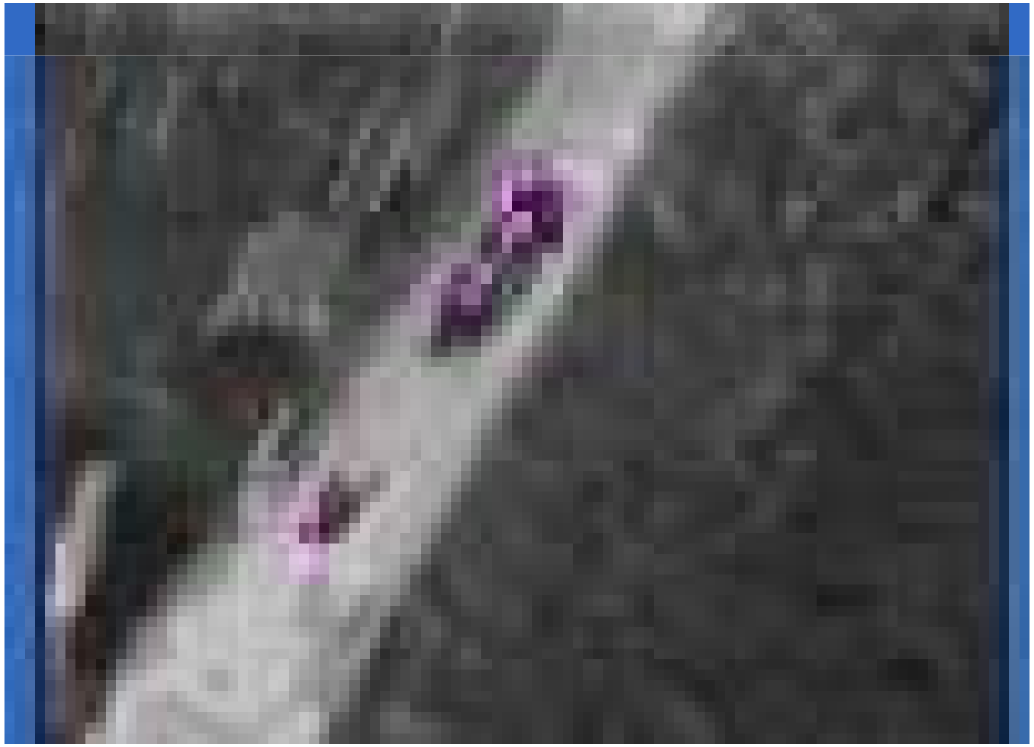
Adapted from Martial Hebert, CMU

Surveillance and tracking

Generating traffic patterns

University of Central Florida, Computer Vision Lab

Surveillance and tracking



Tracking in UAV videos

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

Smart cars

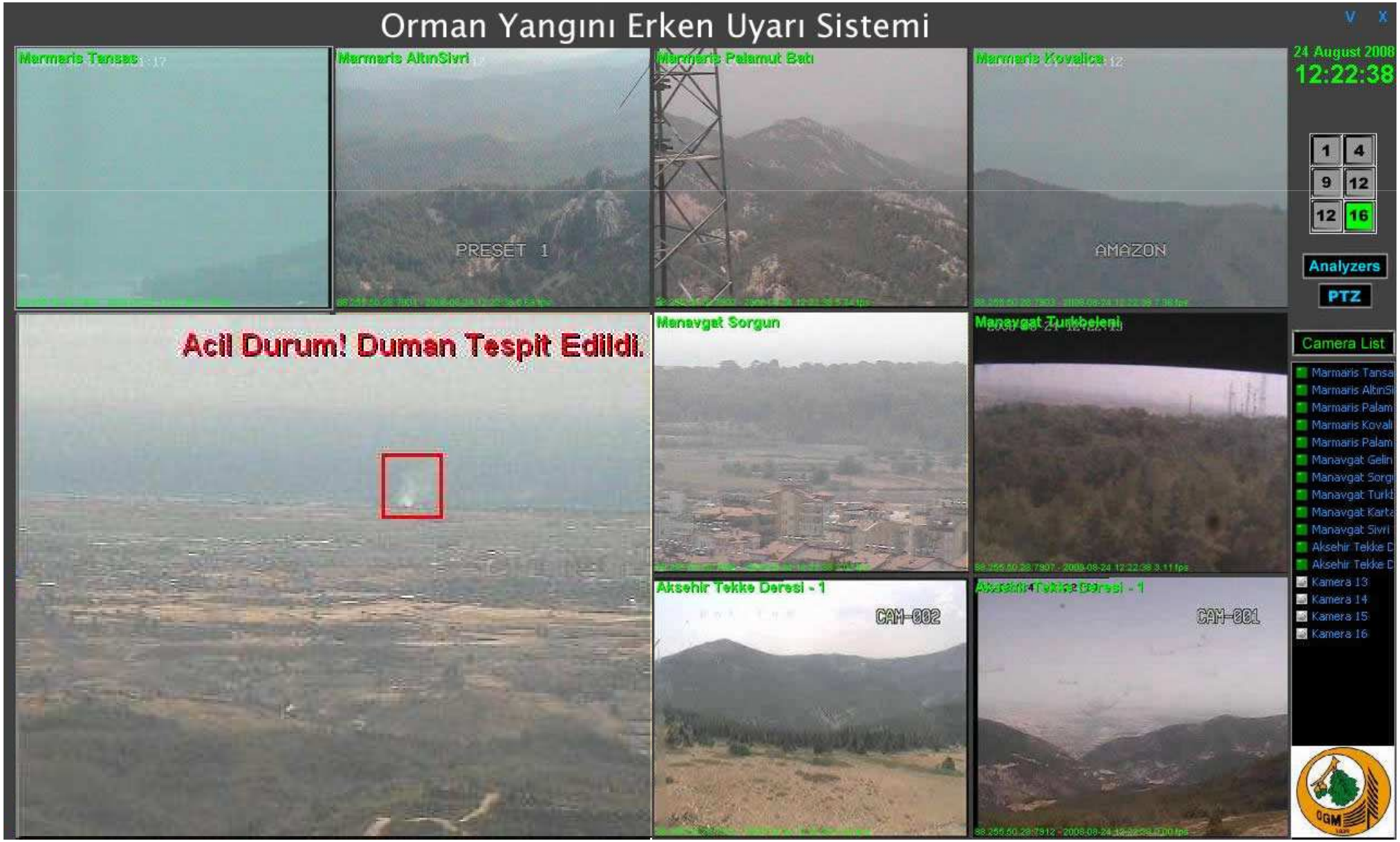
Adapted from CSE 455, U of Washington

Vehicle and pedestrian protection


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

http://www.mobileye-vision.com

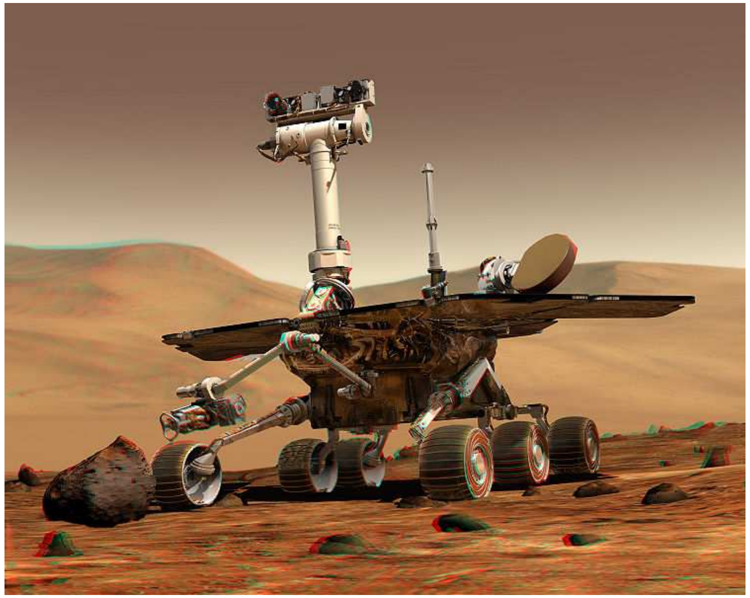
Forest fire monitoring system



Early warning of forest fires

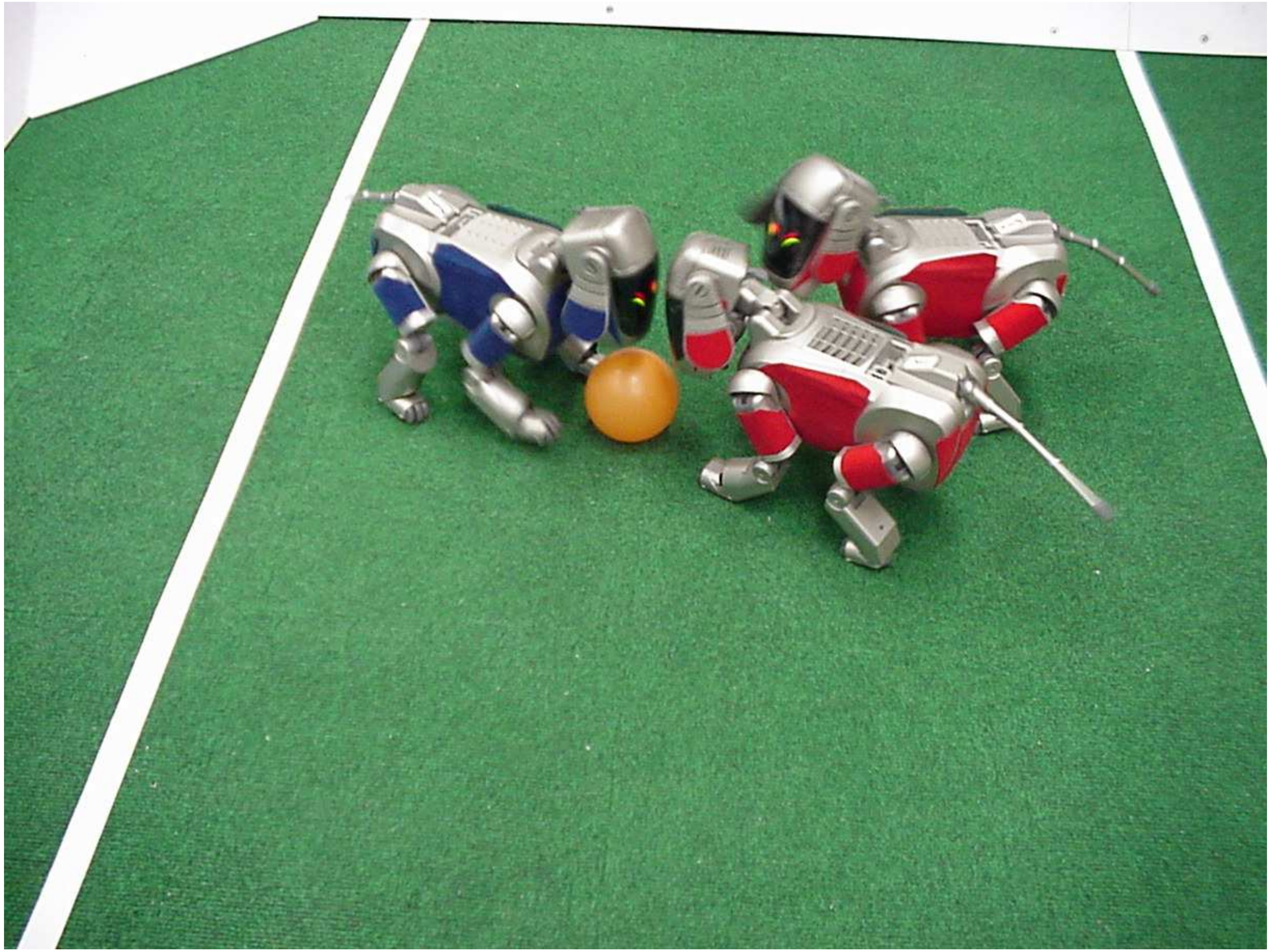
Adapted from Enis Cetin, Bilkent University

Robotics



Adapted from CSE 455, U of Washington

Robotics



Adapted from Steven Seitz, U of Washington

Autonomous navigation

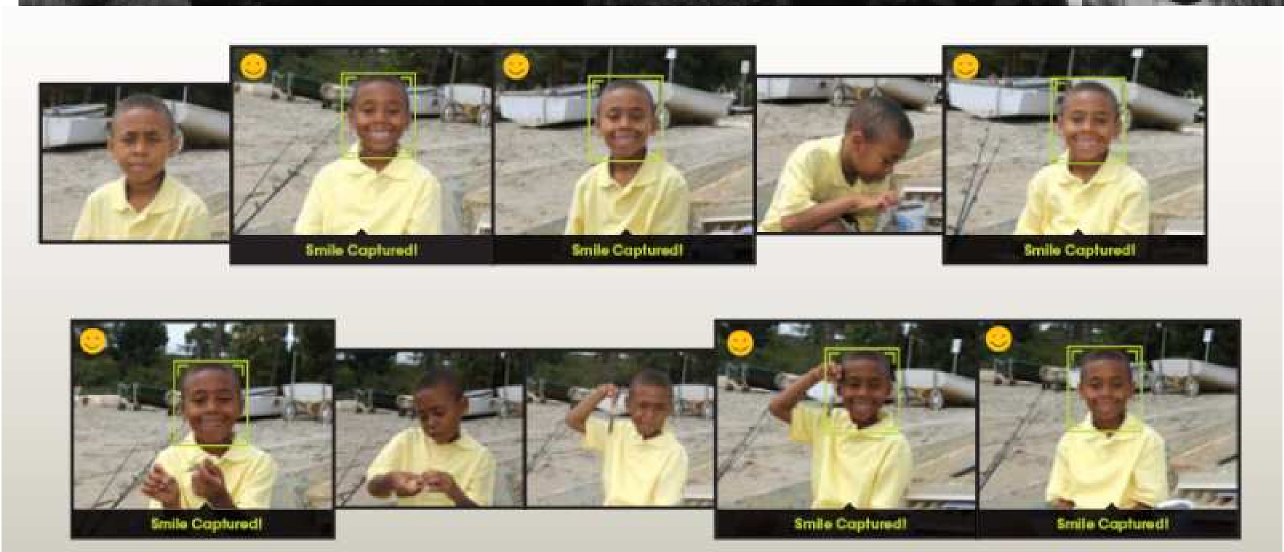
http://www.darpa.mil/grandchallenge/index.asp
http://en.wikipedia.org/wiki/DARPA_Grand_Challenge

Autonomous navigation

Michigan State University

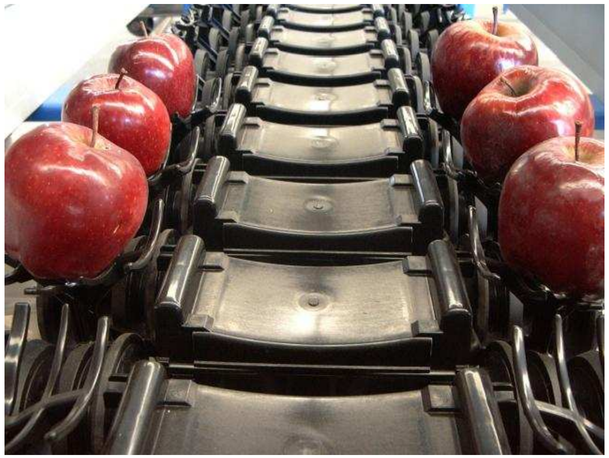
General Dynamics Robotics Systems
http://www.gdrs.com

Face detection and recognition



Adapted from CSE 455, U of Washington

Industrial automation



Automatic fruit sorting

Color Vision Systems
http://www.cvs.com.au

Industrial automation

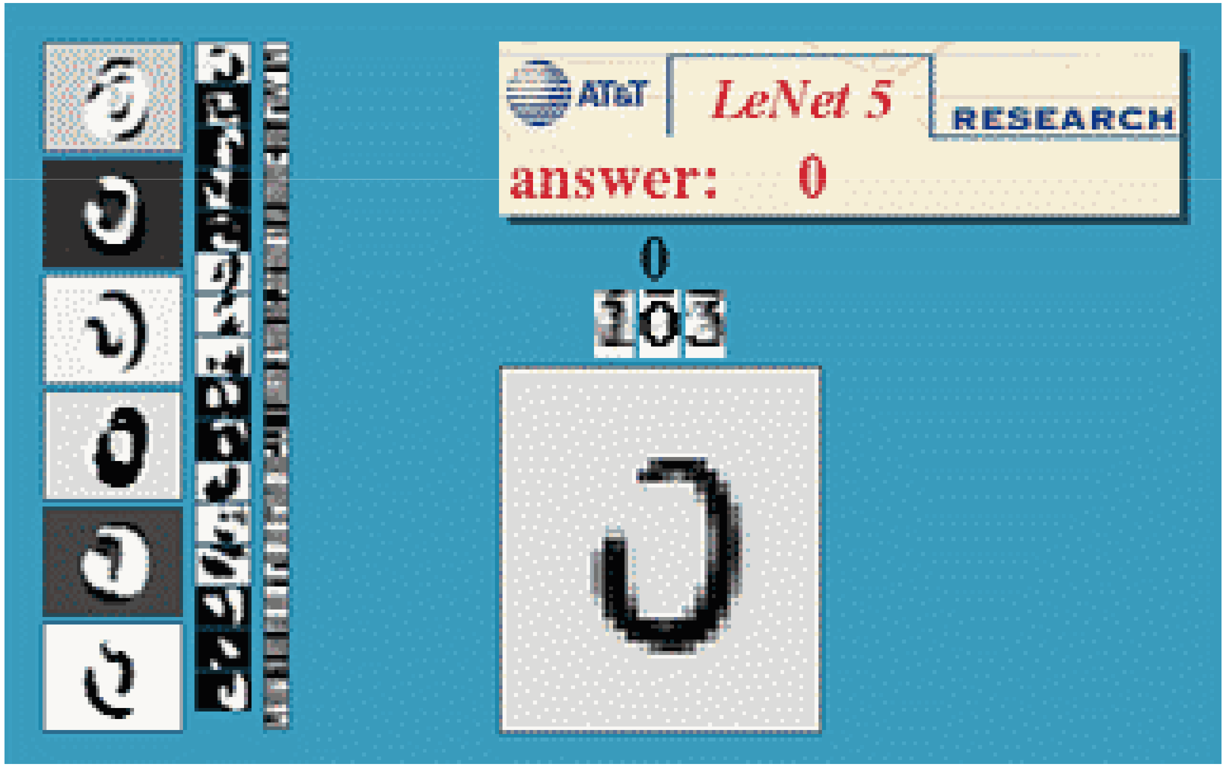
Industrial robotics;
bin picking

http://www.braintech.com

Postal service automation

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

Optical character recognition

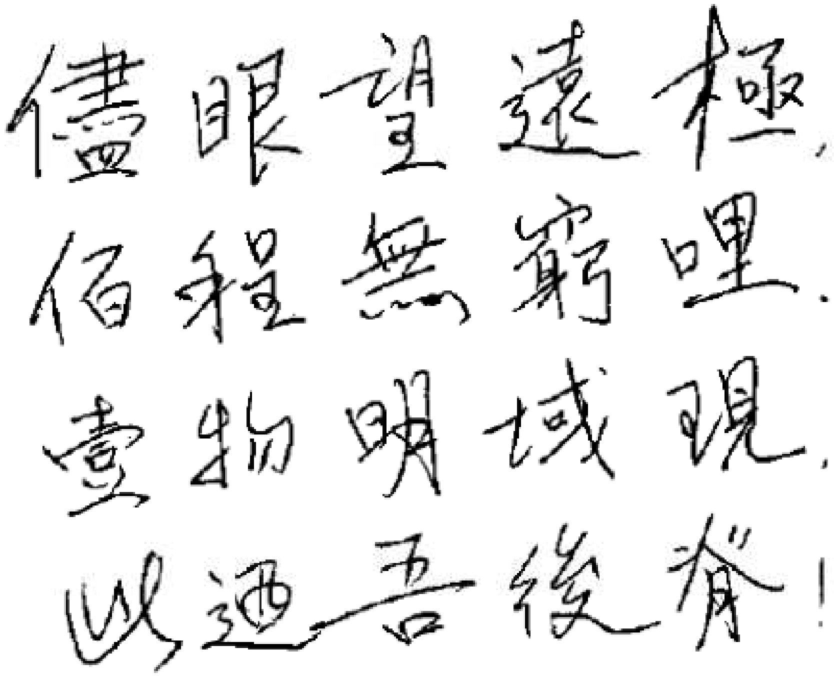


Digit recognition, AT&T labs
http://www.research.att.com/~yann

License place recognition

Adapted from Steven Seitz, U of Washington

Document analysis

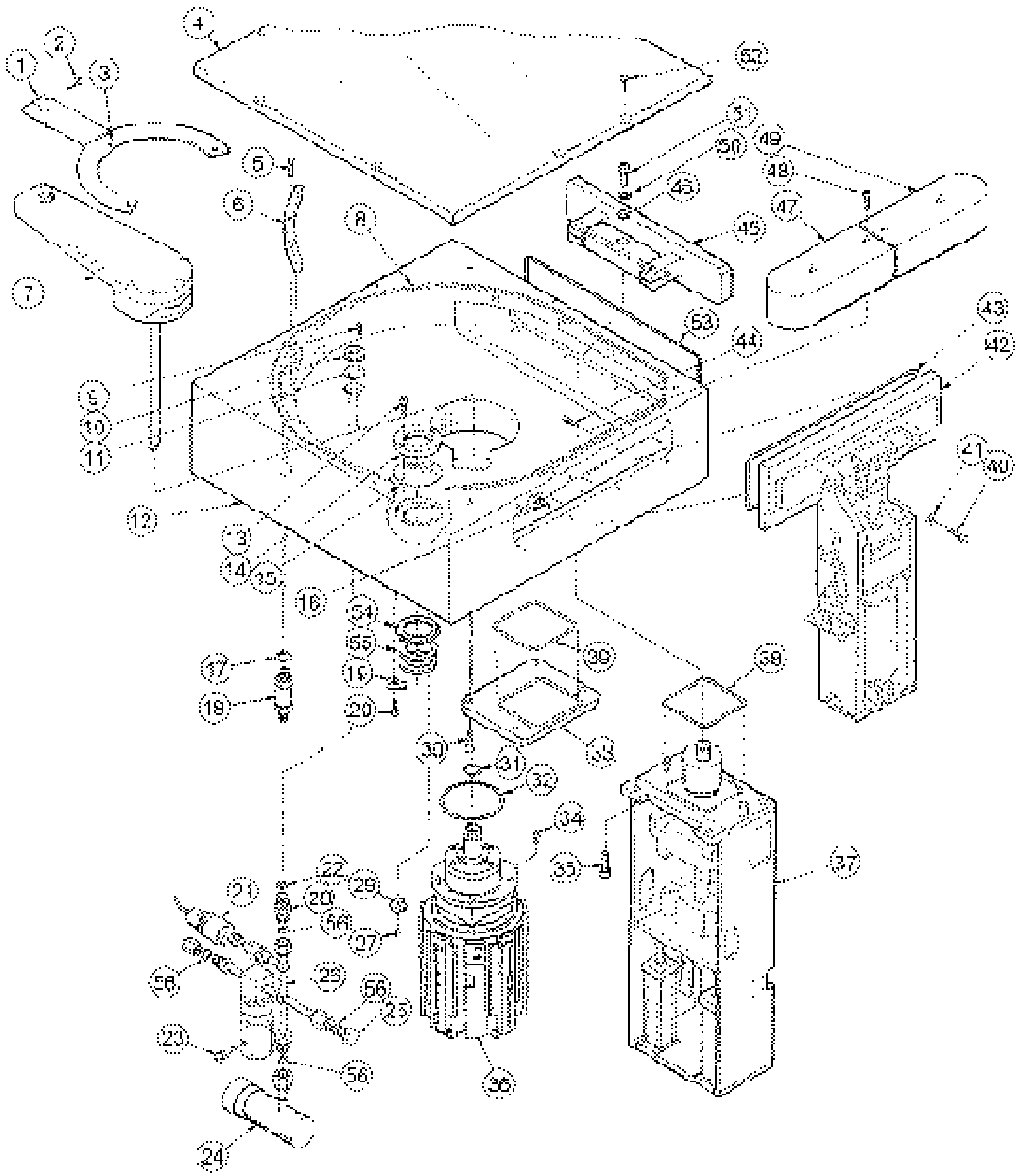


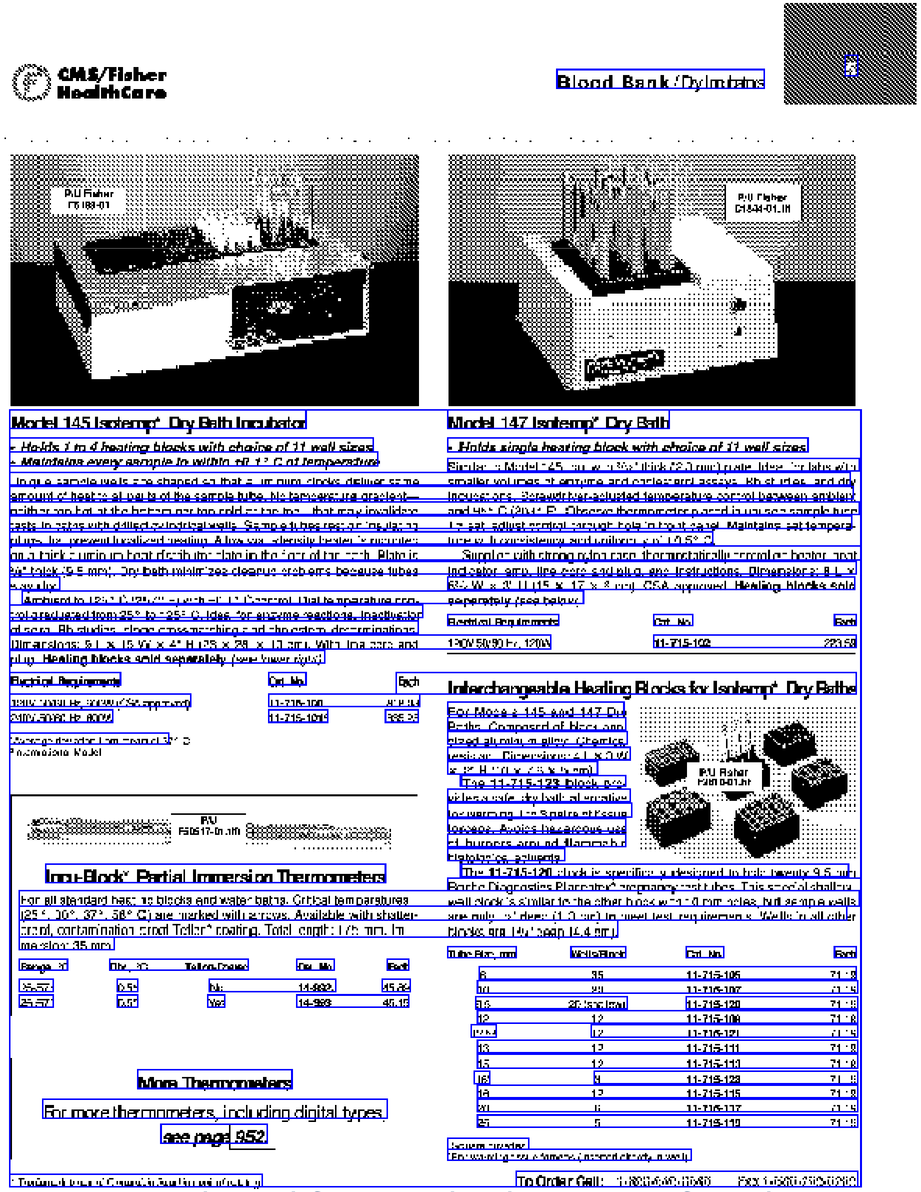
I looked as hard as I could see,
beyond 100 plus infinity
an object of bright intensity
– it was the back of me!

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 Incubators

Model 145 Isotemp* Dry Bath Incubator

Model 147 Isotemp* Dry Bath

Interchangeable Heating Blocks for Isotemp* Dry Baths

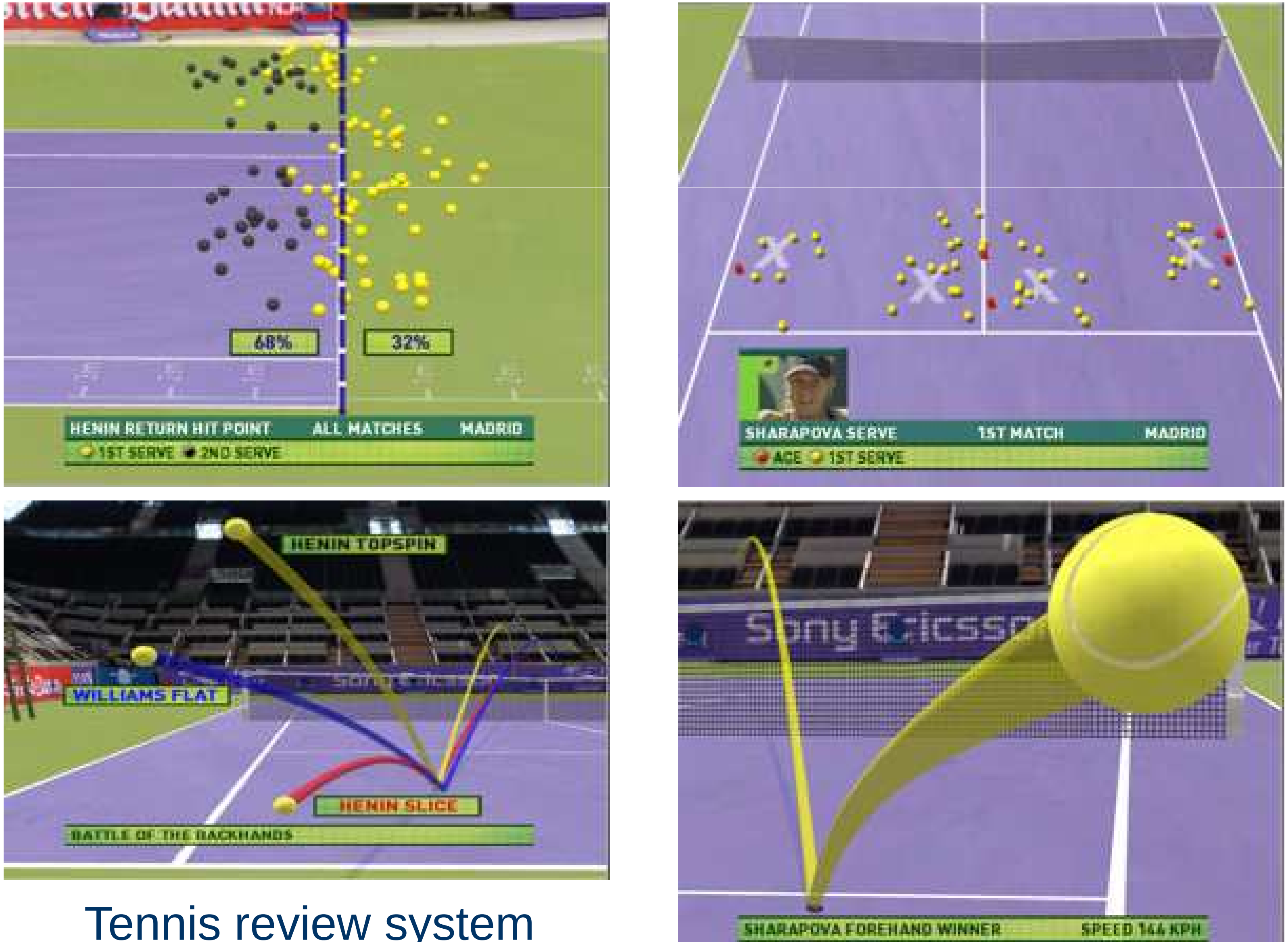
Incu-Block* Partial Immersion Thermometers

More Thermometers

For more thermometers, including digital types, see page 952

Adapted from Linda Shapiro, U of Washington

Sports video analysis

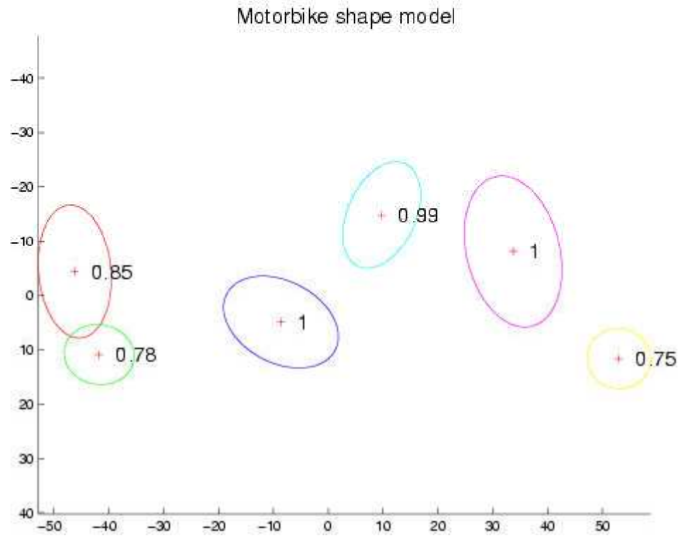


Tennis review system

http://www.hawkeyeinnovations.co.uk

Scene classification

Object recognition

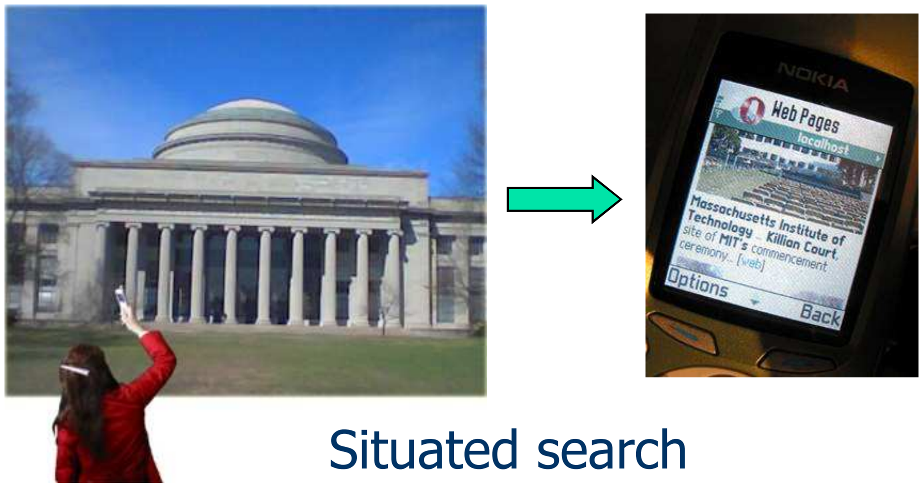


Adapted from Rob Fergus, MIT

Object recognition

Lincoln, Microsoft Research

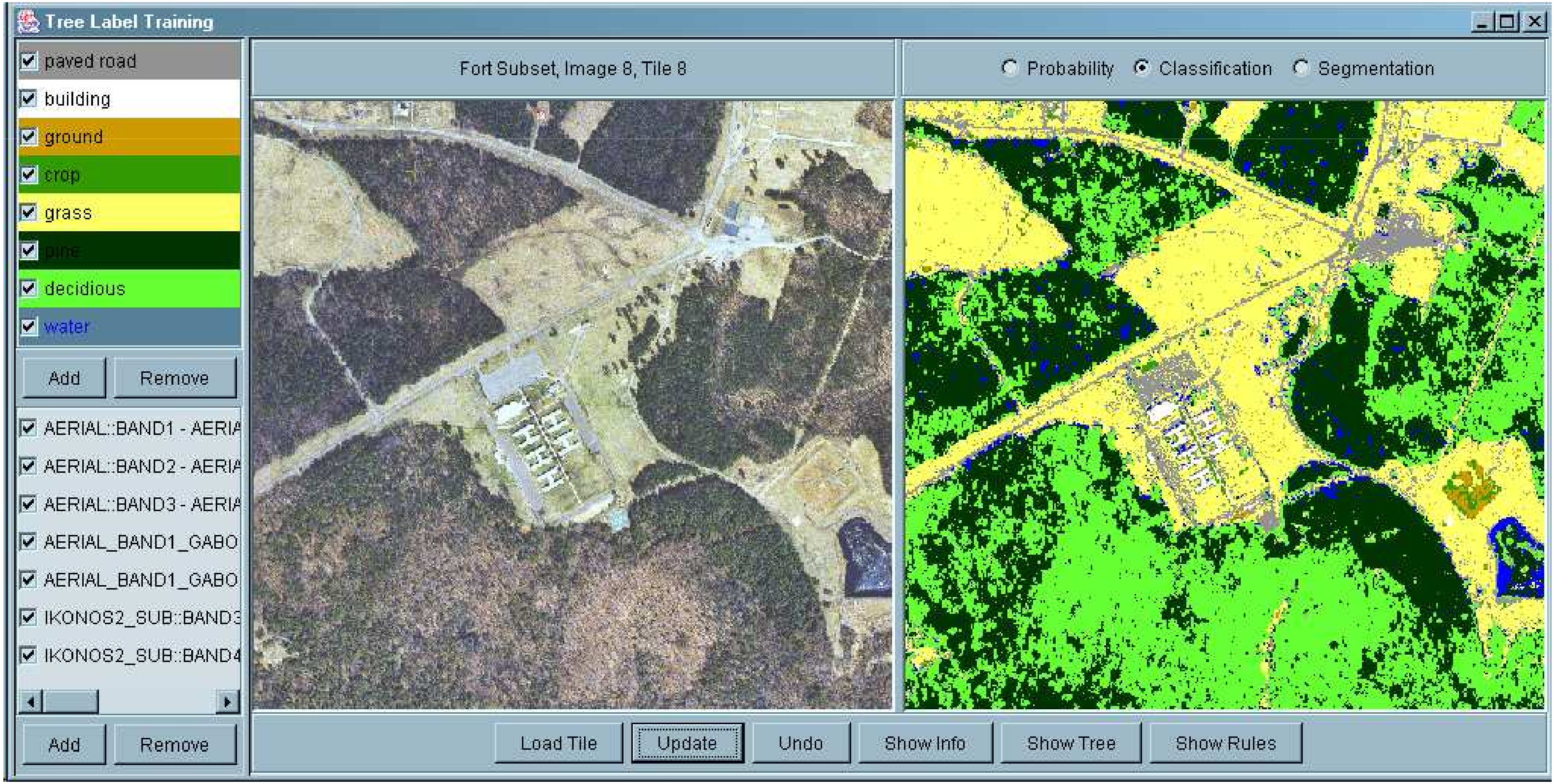
kooaba



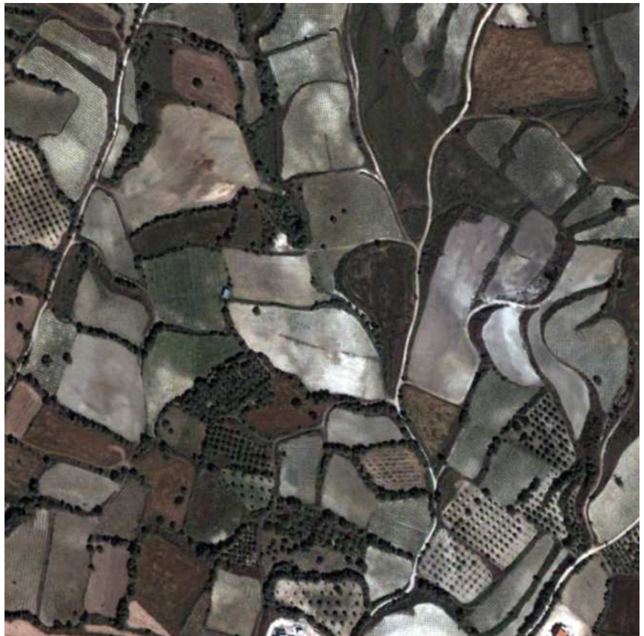
Situated search
Yeh et al., MIT

Google Goggles
Bing Vision

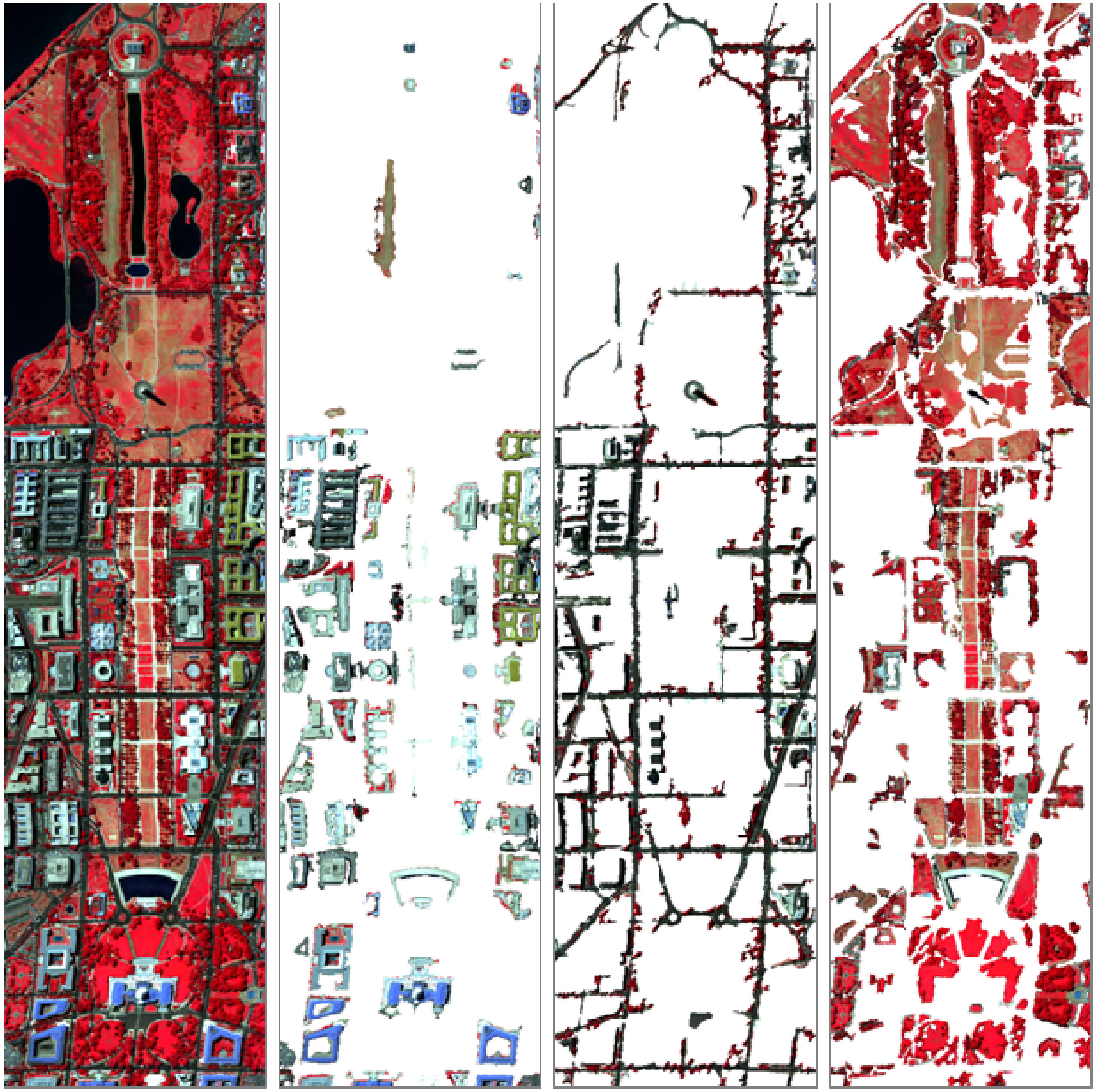
Land cover classification



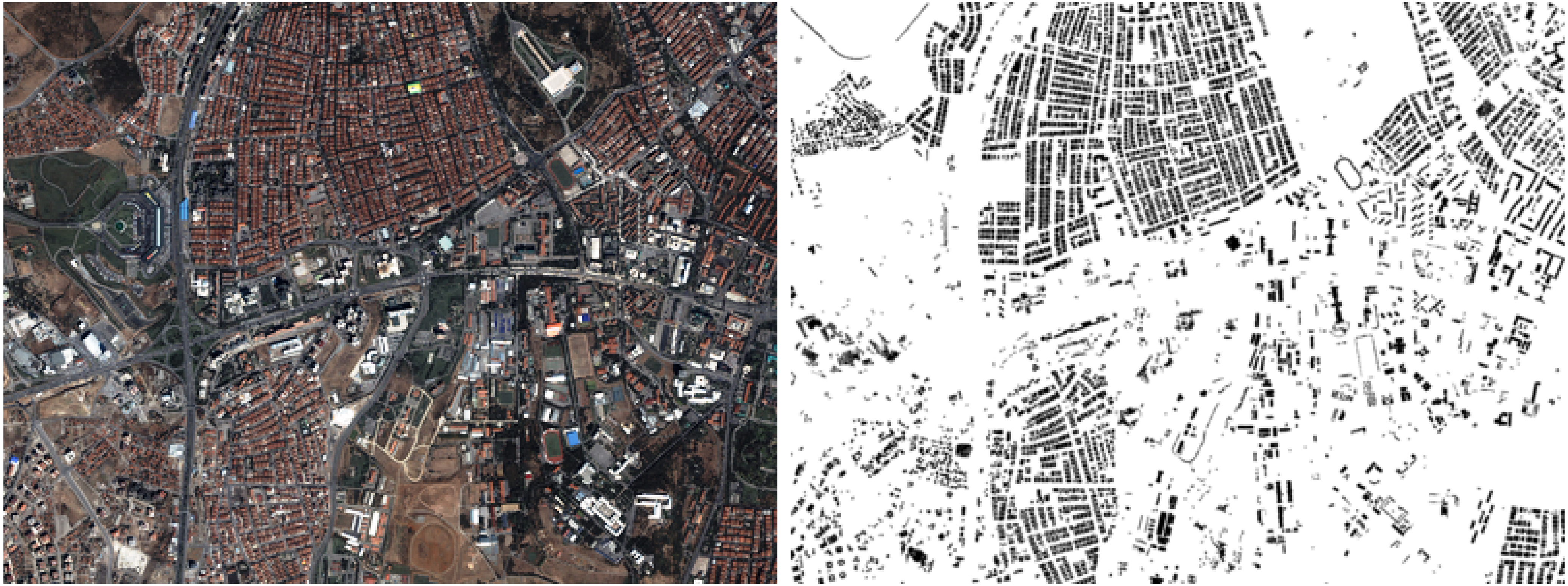
Land cover classification



Object recognition

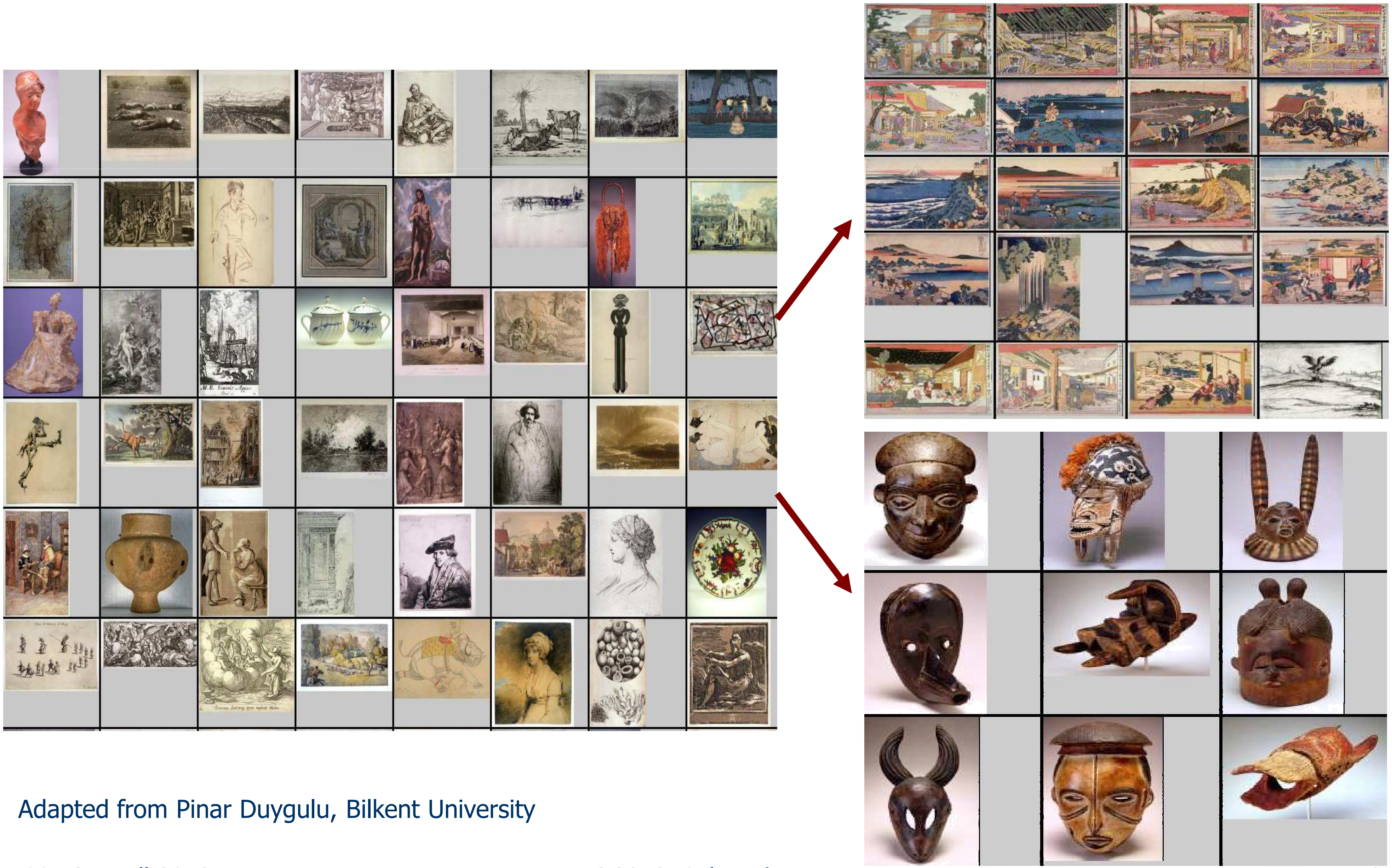


Object recognition



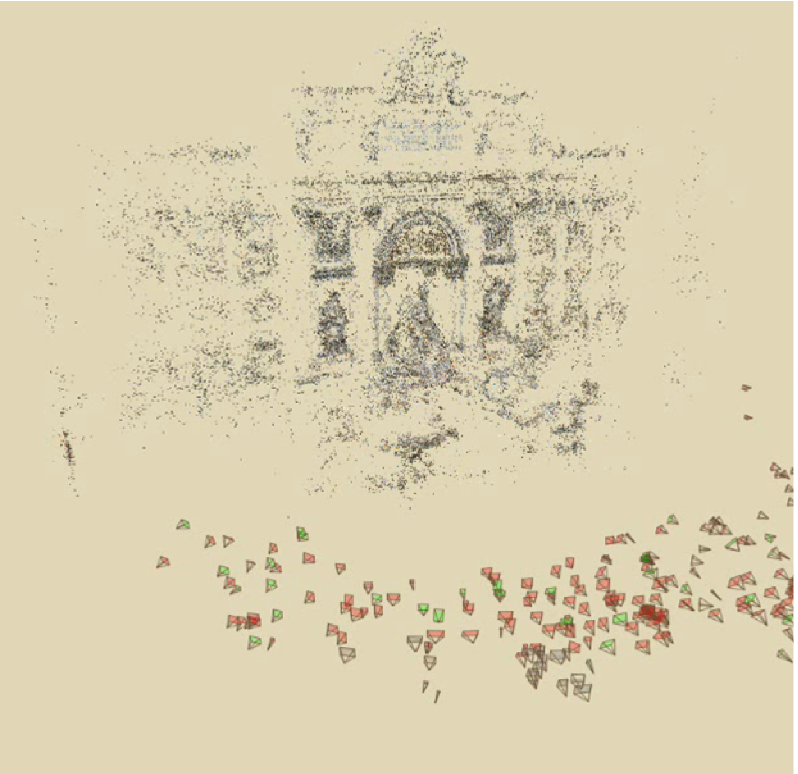
Recognition of buildings and building groups

Organizing image archives



Adapted from Pinar Duygulu, Bilkent University

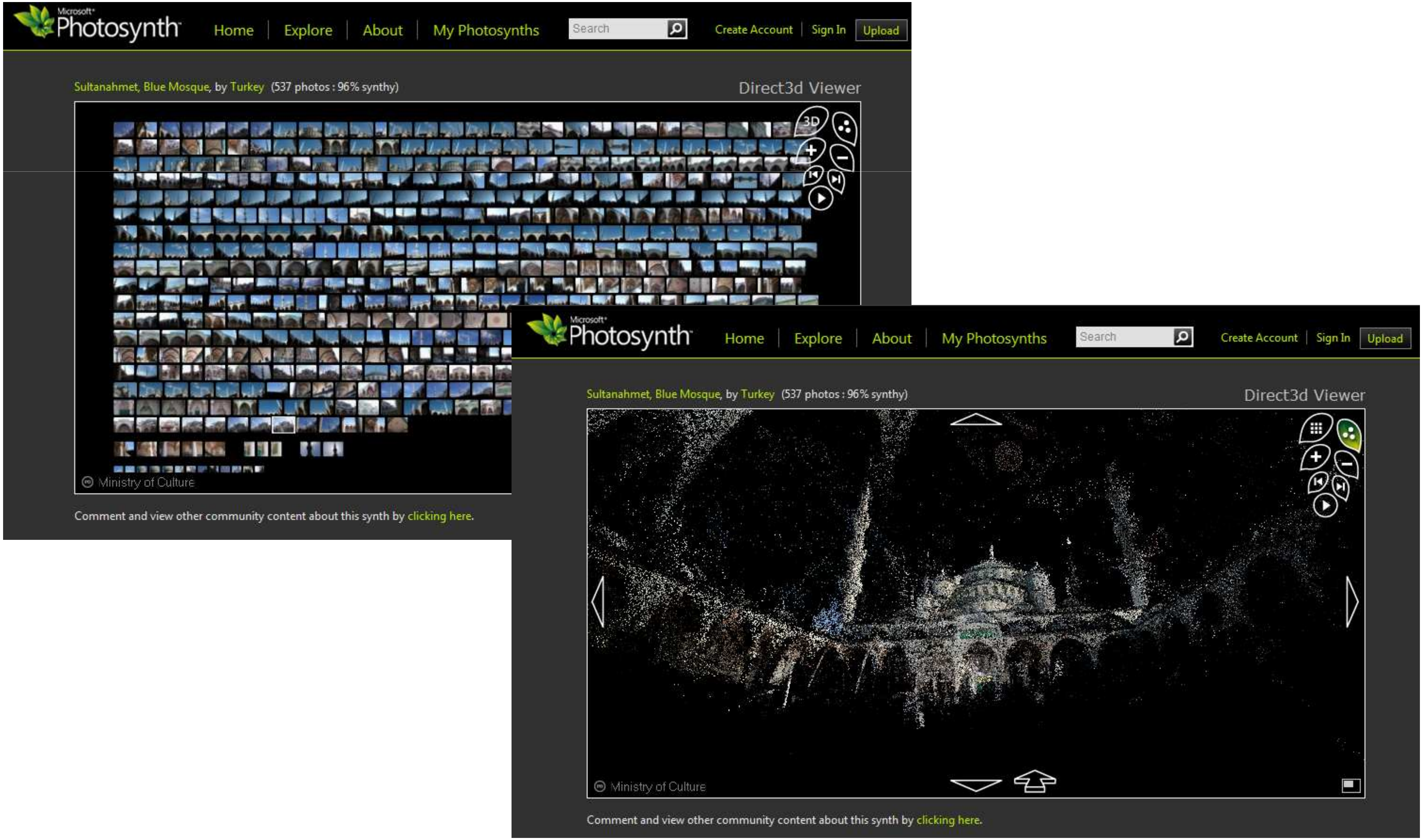
Photo tourism: exploring photo collections



Building 3D scene models from individual photos

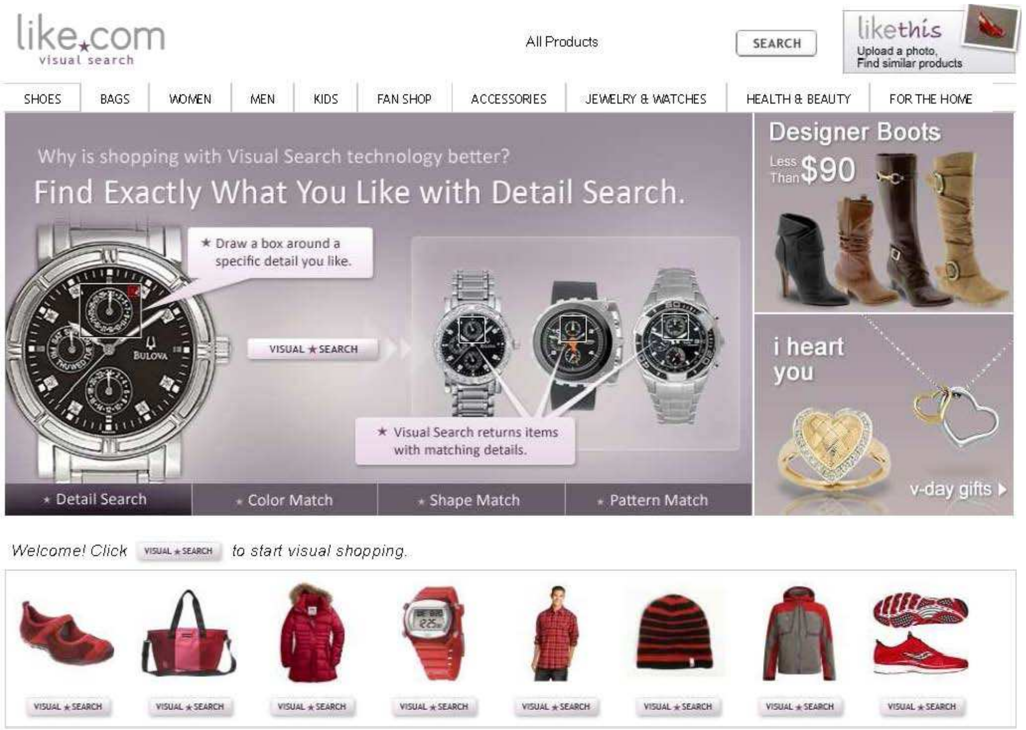
Adapted from Steven Seitz, U of Washington

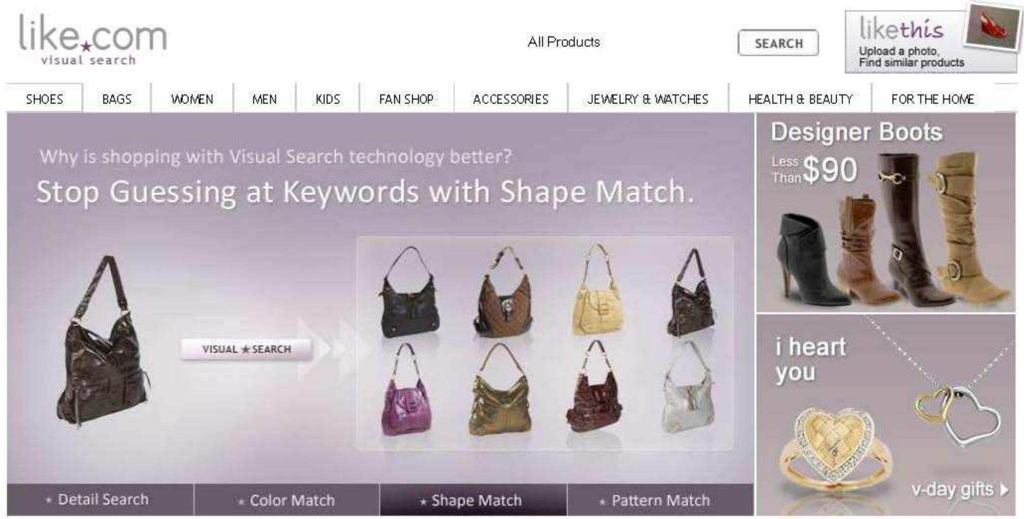
Photosynth

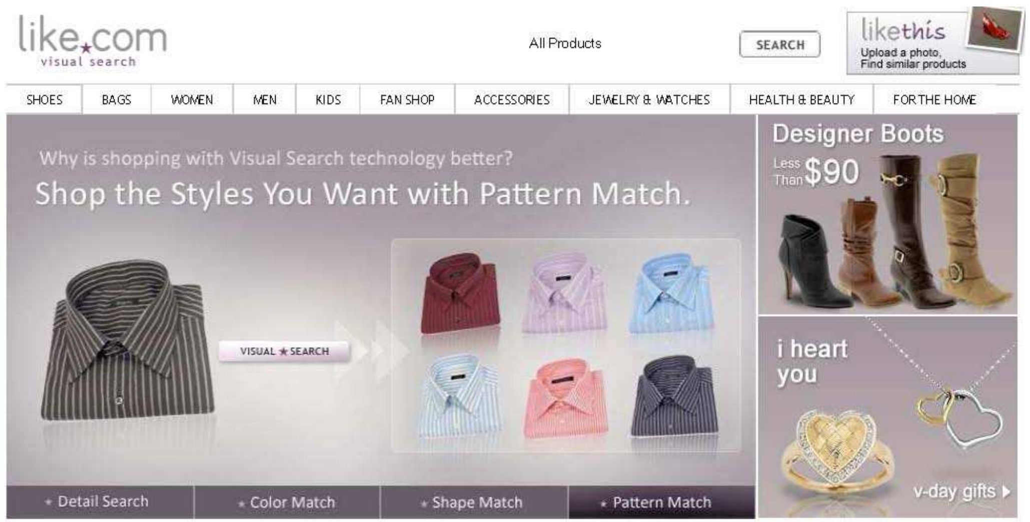


Microsoft® Photosynth
Home
Explore
About
My Photosynths
Search
Create Account
Sign In
Upload
Sultanahmet, Blue Mosque, by Turkey (537 photos : 96% synthy)
Direct3d Viewer
3D
Ministry of Culture
Comment and view other community content about this synth by clicking here.

Content-based retrieval



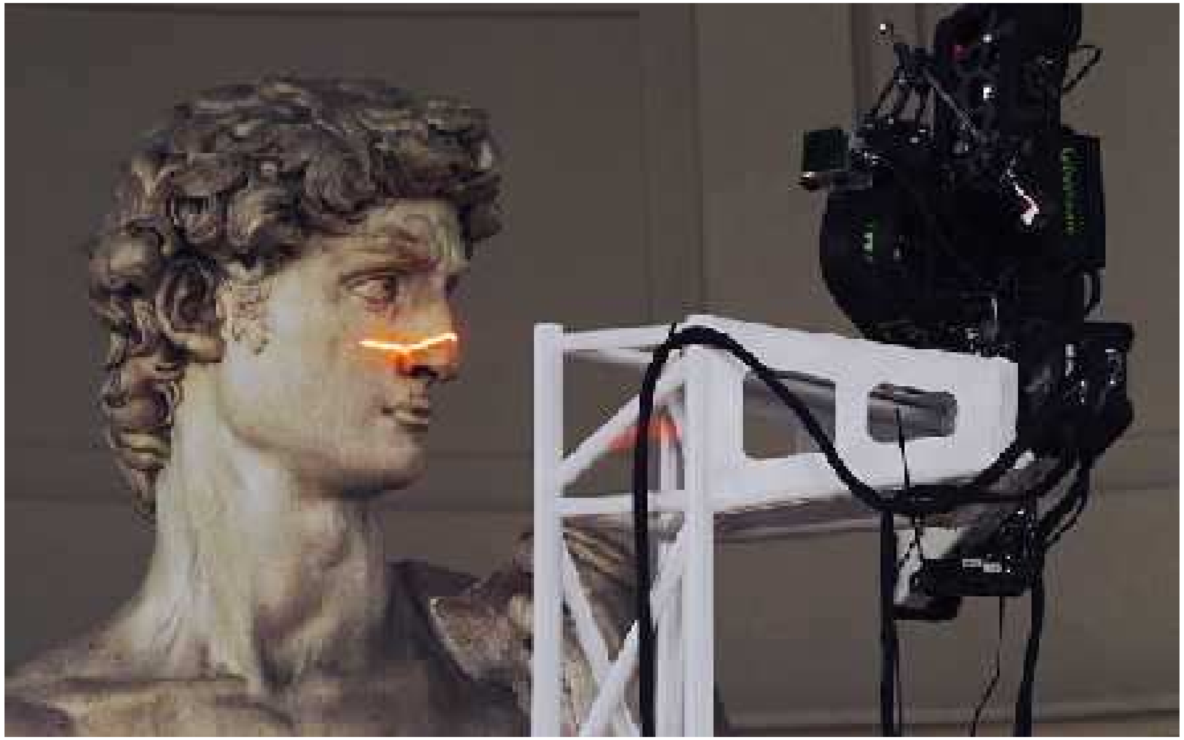




Online shopping catalog search

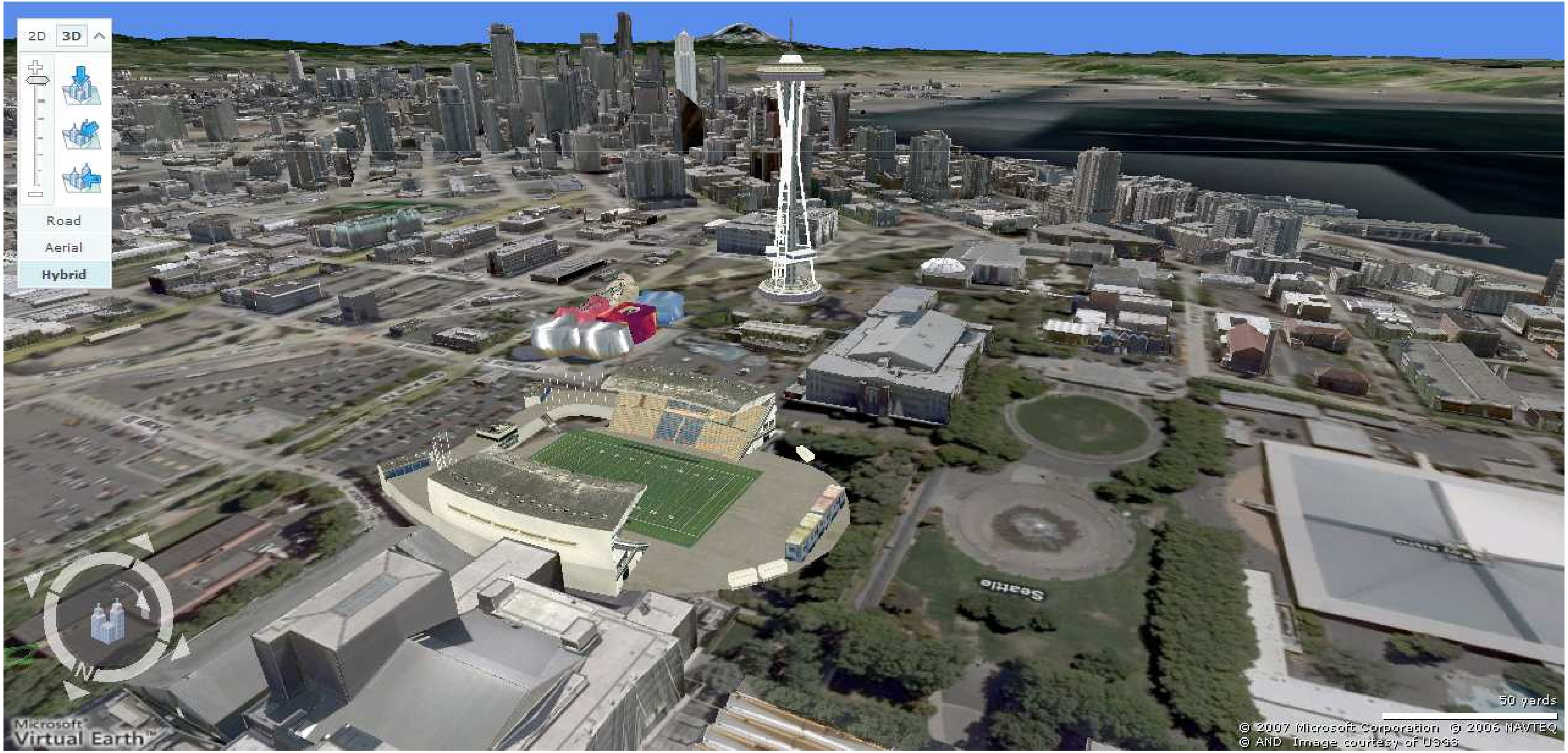
http://www.like.com

3D scanning and reconstruction



Adapted from Linda Shapiro, U of Washington

Earth viewers (3D modeling)



Motion capture

Adapted from Linda Shapiro, U of Washington

Visual effects

Adapted from CSE 455, U of Washington

Motion capture

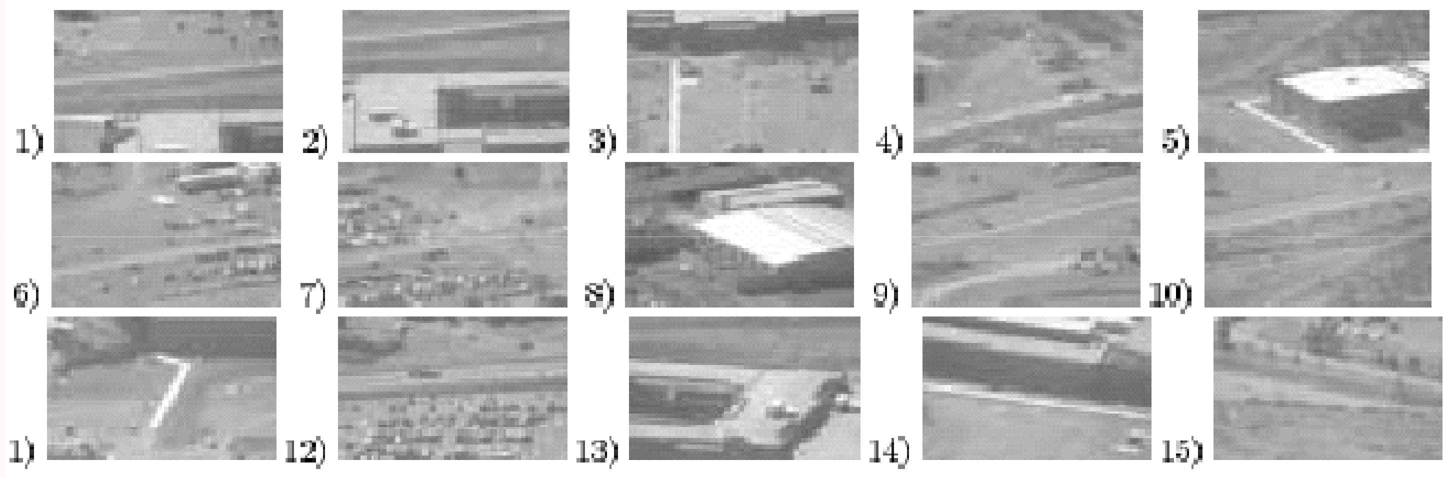
Microsoft's XBox Kinect

Adapted from CSE 455, U of Washington

Mozaic

Adapted from David Forsyth, UC Berkeley

Mozaic

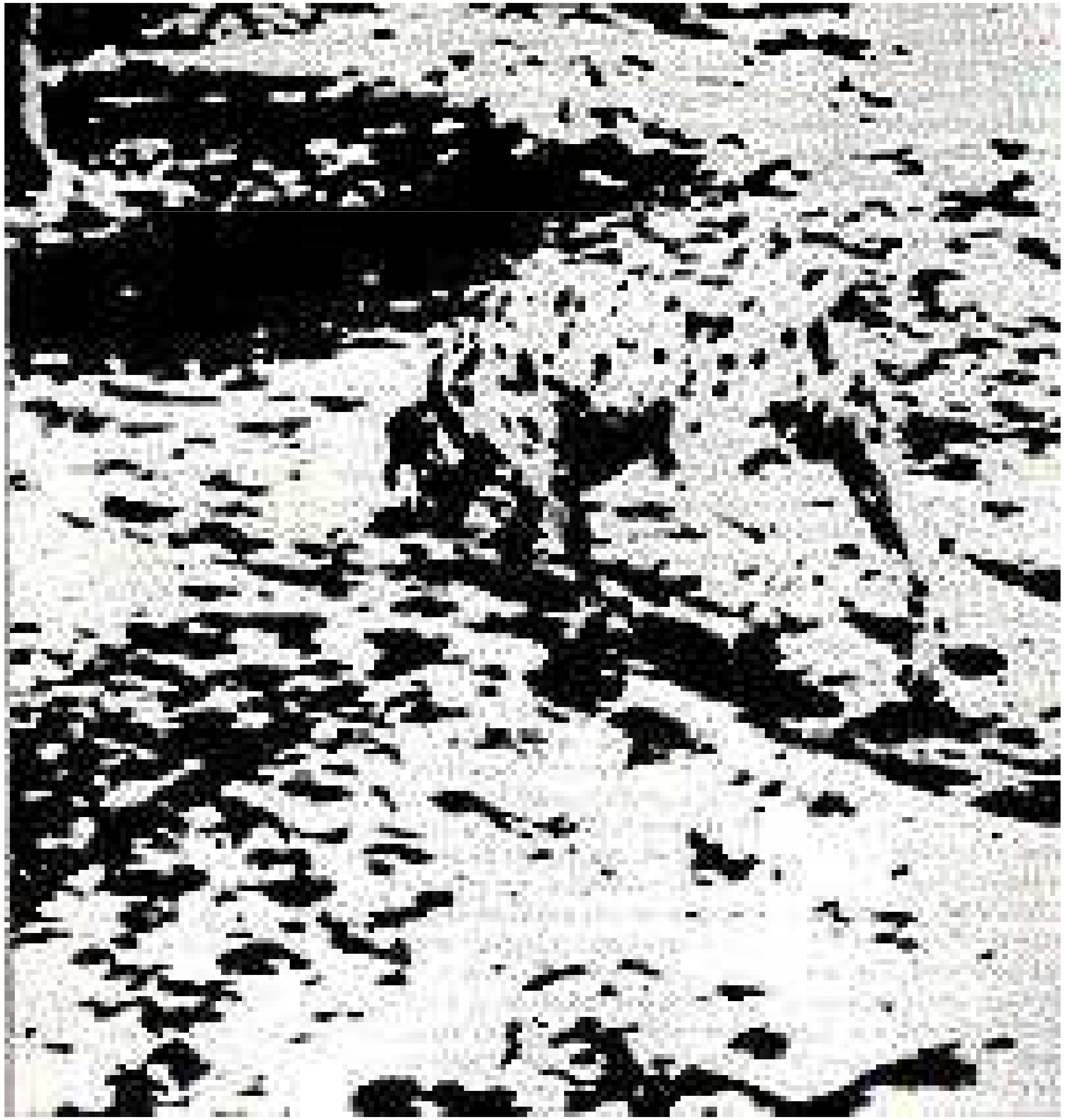


Adapted from David Forsyth, UC Berkeley

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

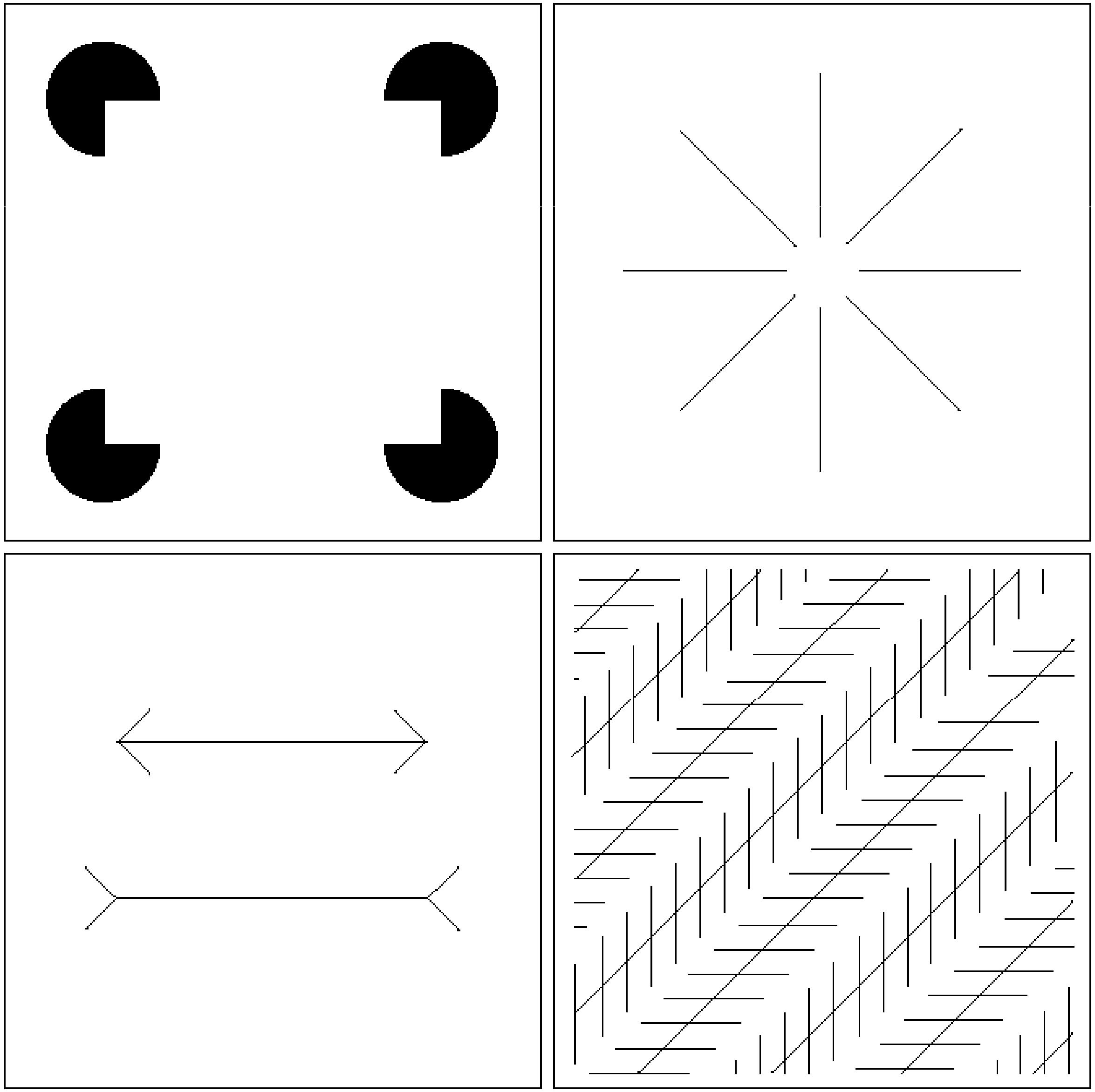


Subjective contours

Perception and grouping

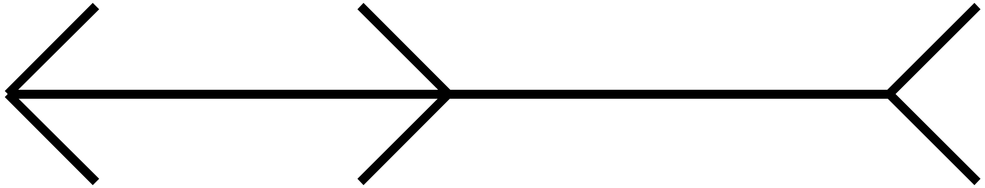
a b
c d

FIGURE 2.9 Some well-known optical illusions.

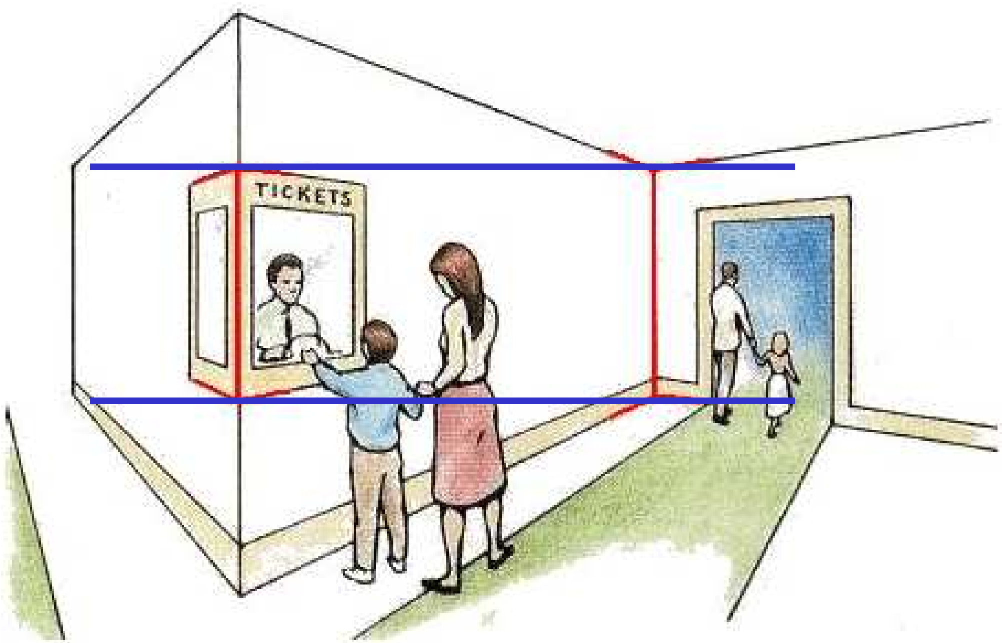


Adapted from Gonzales and Woods

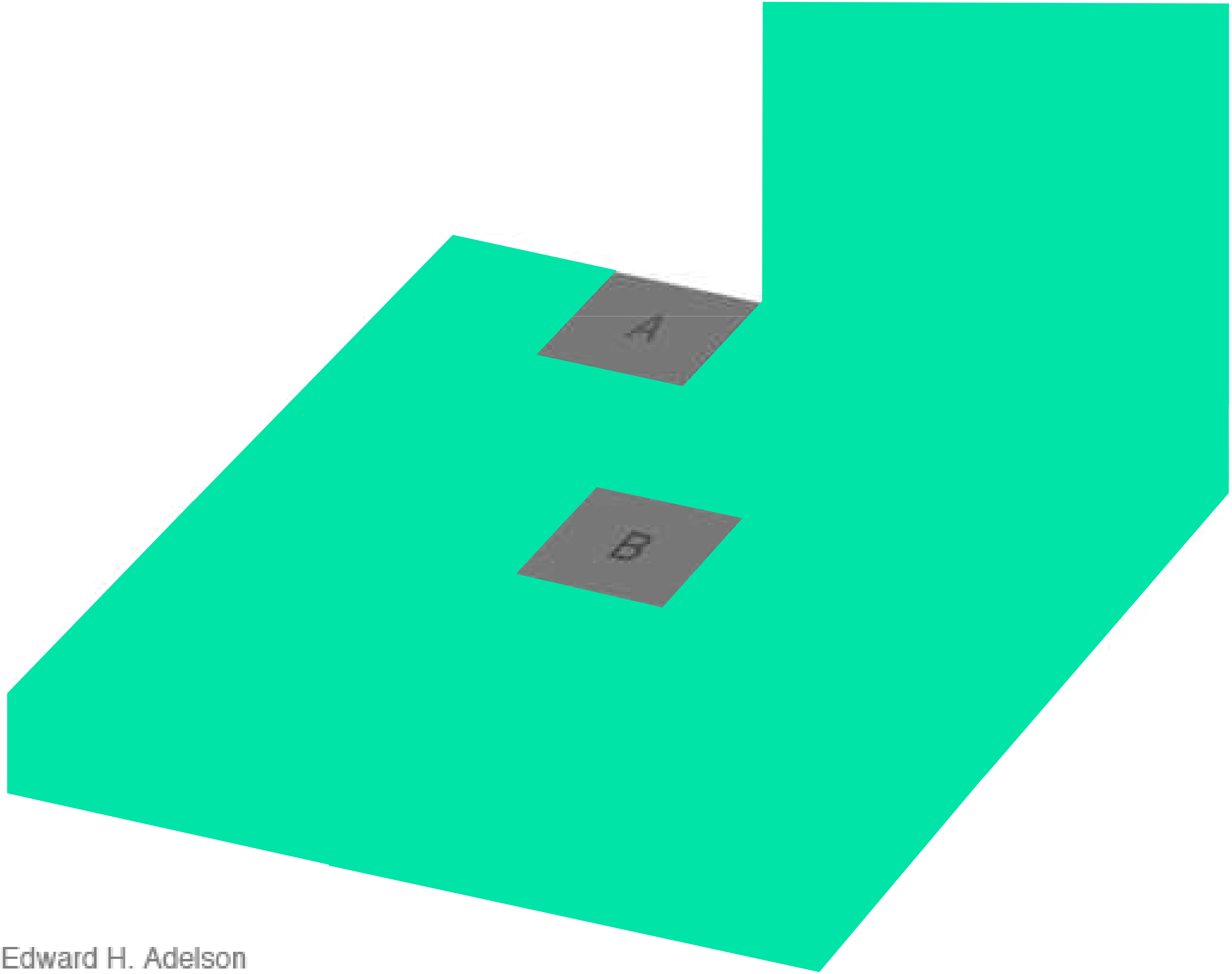
Perception and grouping



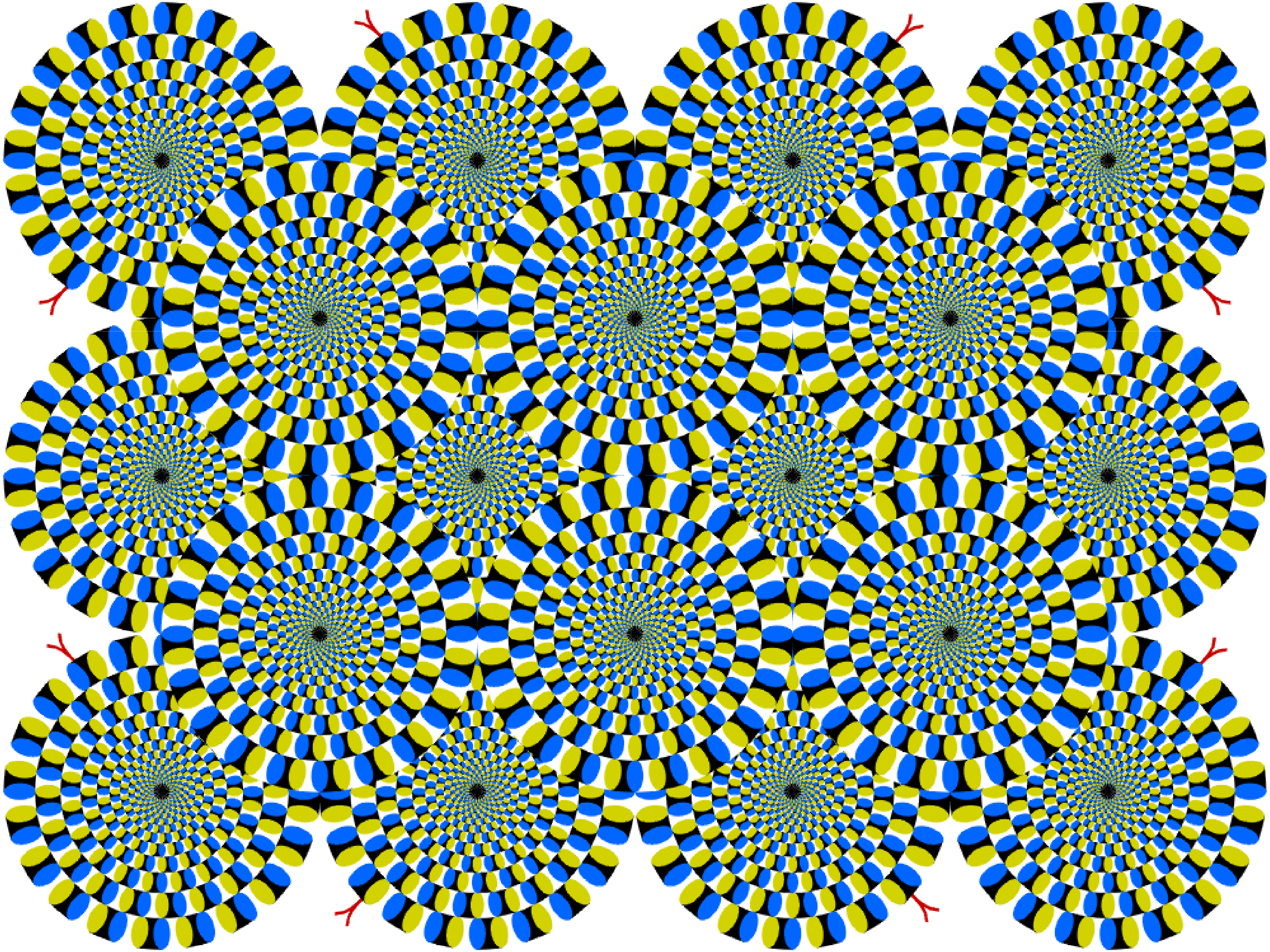
Müller-Lyer Illusion



Adapted from Alyosha Efros, Carnegie Mellon



Edward H. Adelson



Copyright A.Kitaoka 2003

Perception and grouping

I. Rock, The Logic of Perception, 1983.

Occlusion

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

Adapted from Michael Black, Brown University

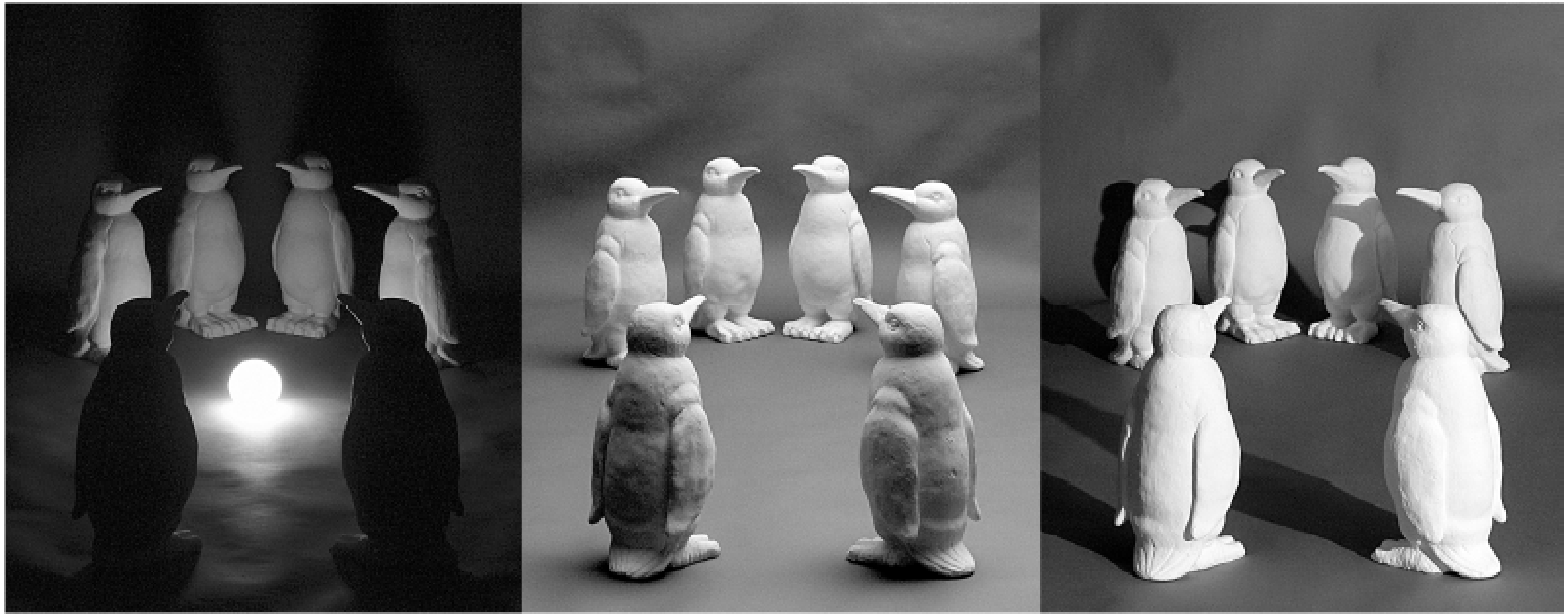
What the computer gets

Challenges 1: view point variation

Michelangelo 1475-1564

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

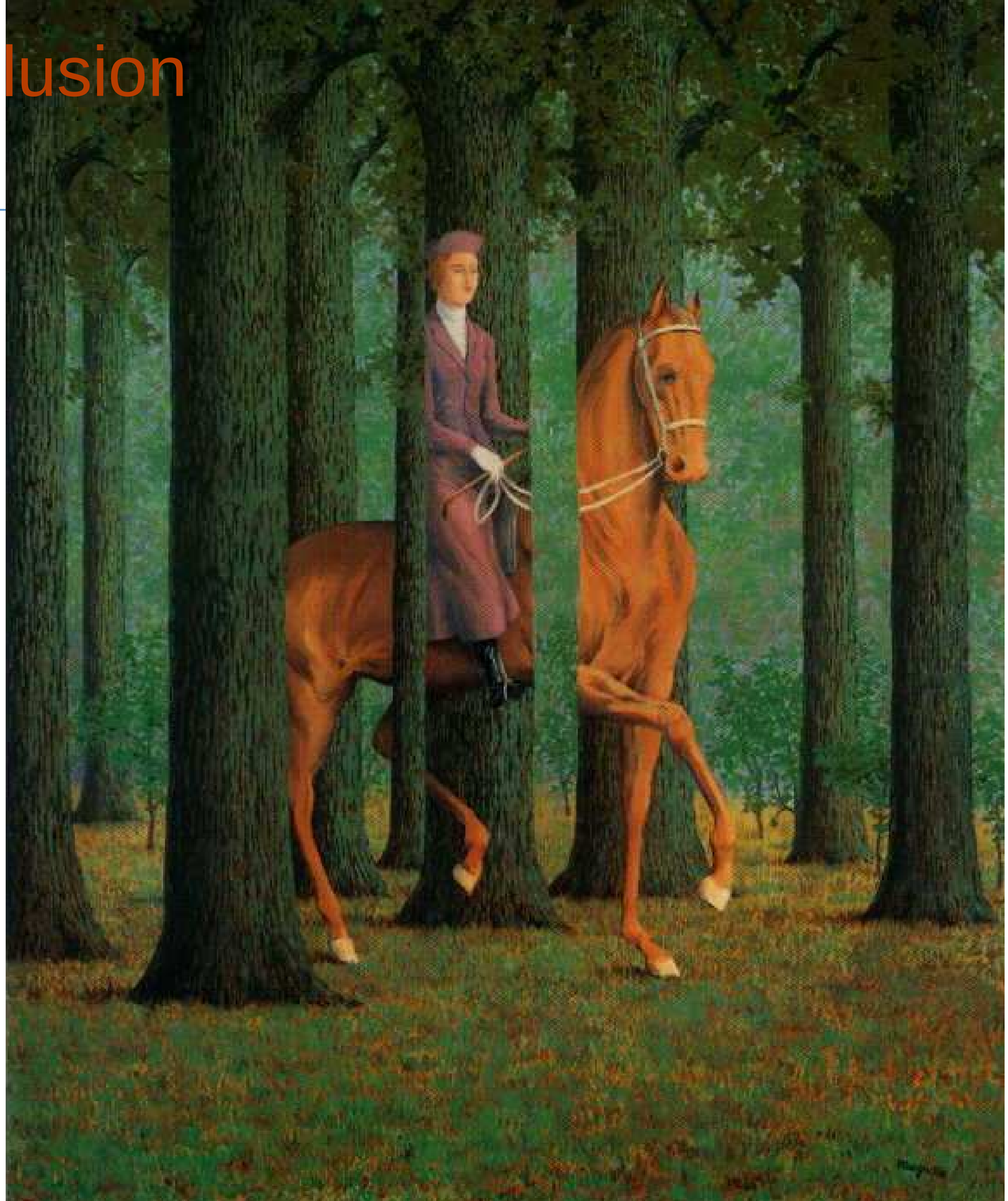
Challenges 2: illumination



Adapted from Fei-Fei Li

Challenges 3: occlusion

Magritte, 1957

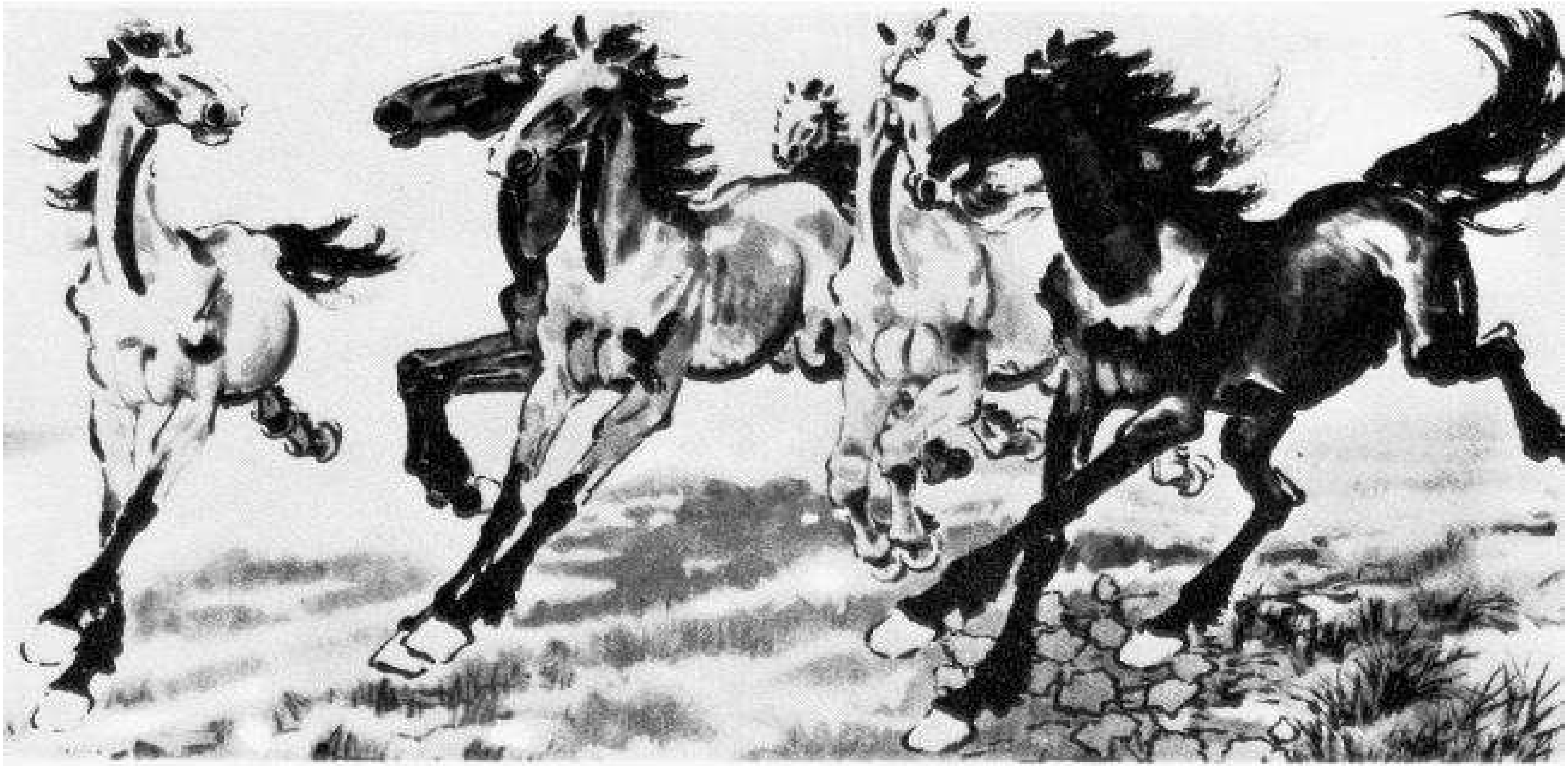


Adapted from L. Fei-Fei, R. Fergus, A. Torralba

Challenges 4: scale

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

Challenges 5: deformation



Xu, Beihong 1943

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

Challenges 6: background clutter

Adapted from Fei-Fei Li

Challenges 7: intra-class variation



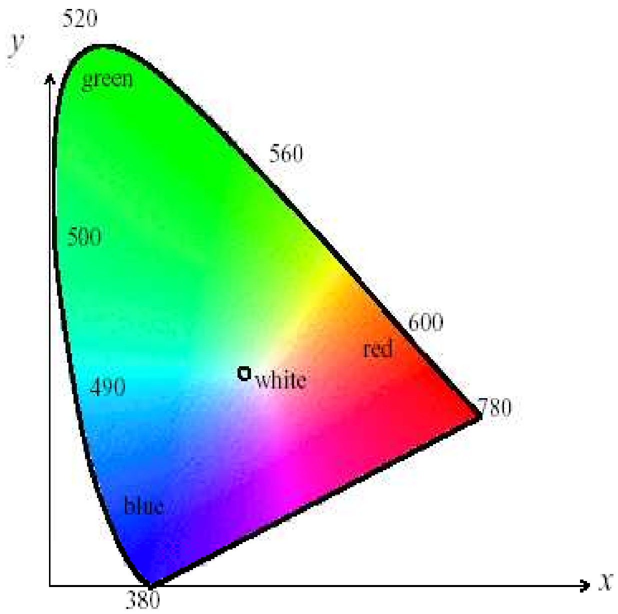



Adapted from L. Fei-Fei, R. Fergus, A. Torralba

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

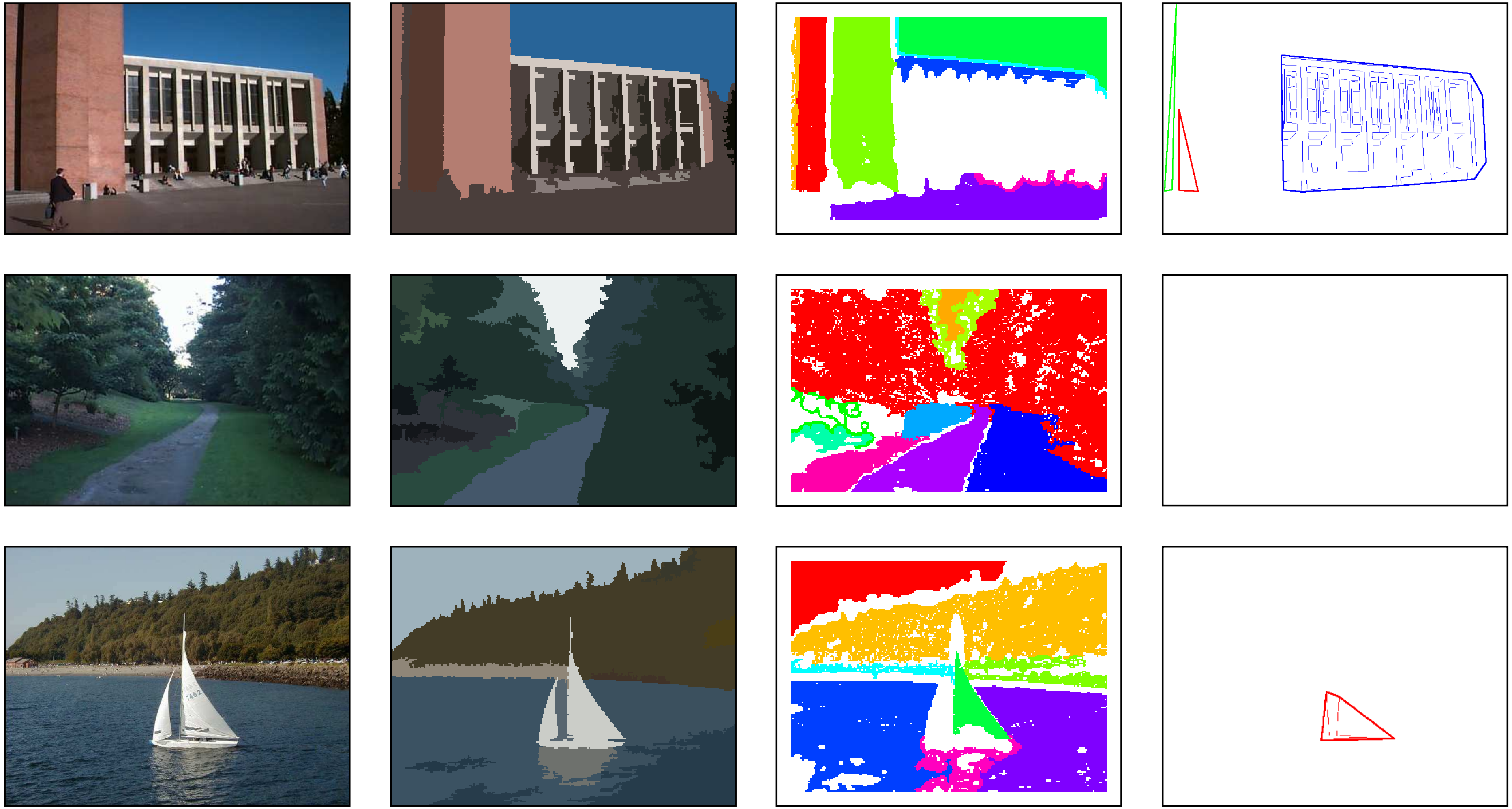
Adapted from David Forsyth, UC Berkeley

Color, texture, and proximity

Adapted from Fei-Fei Li

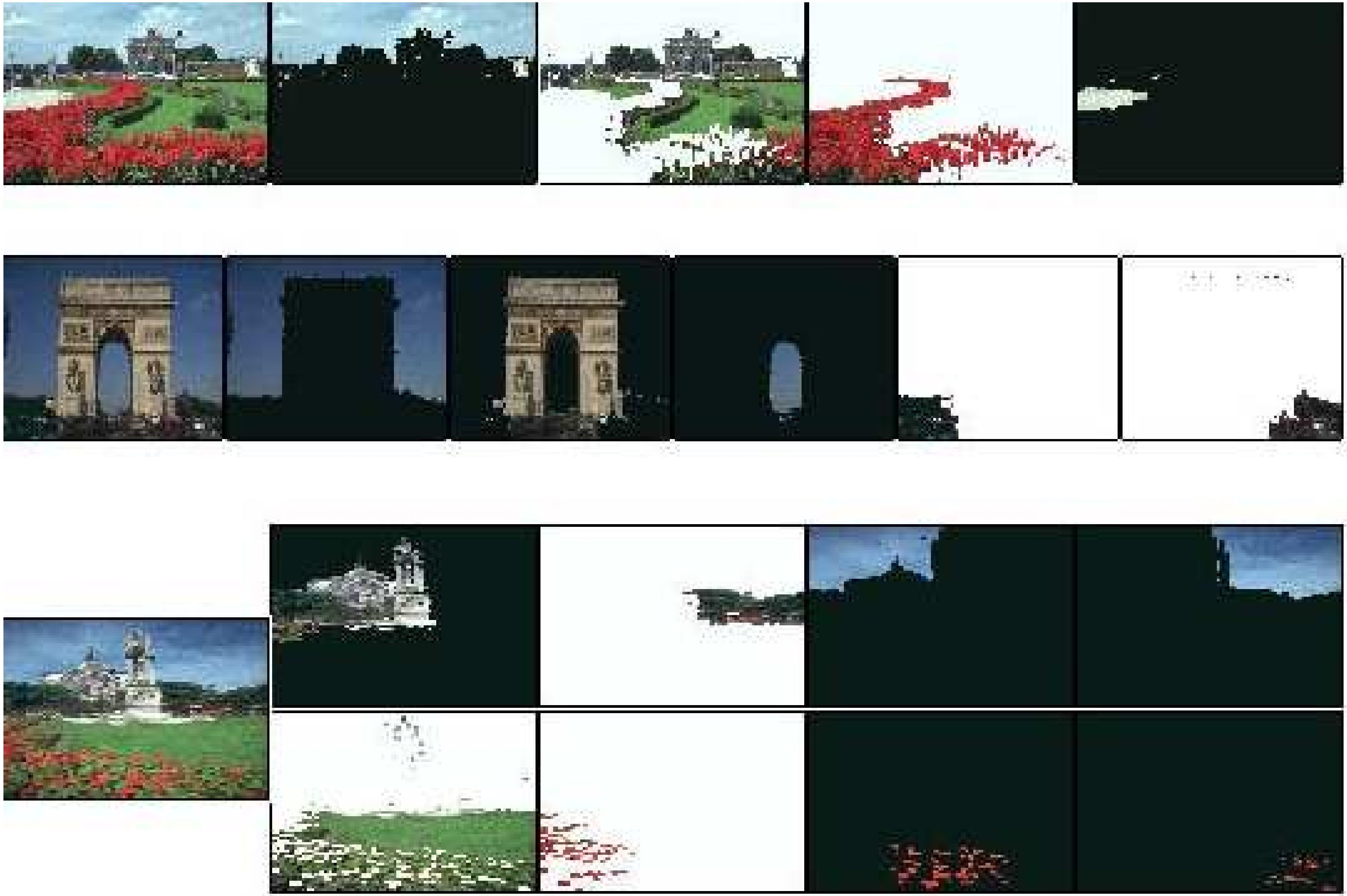
Segmentation

Original Images | Color Regions | Texture Regions | Line Clusters



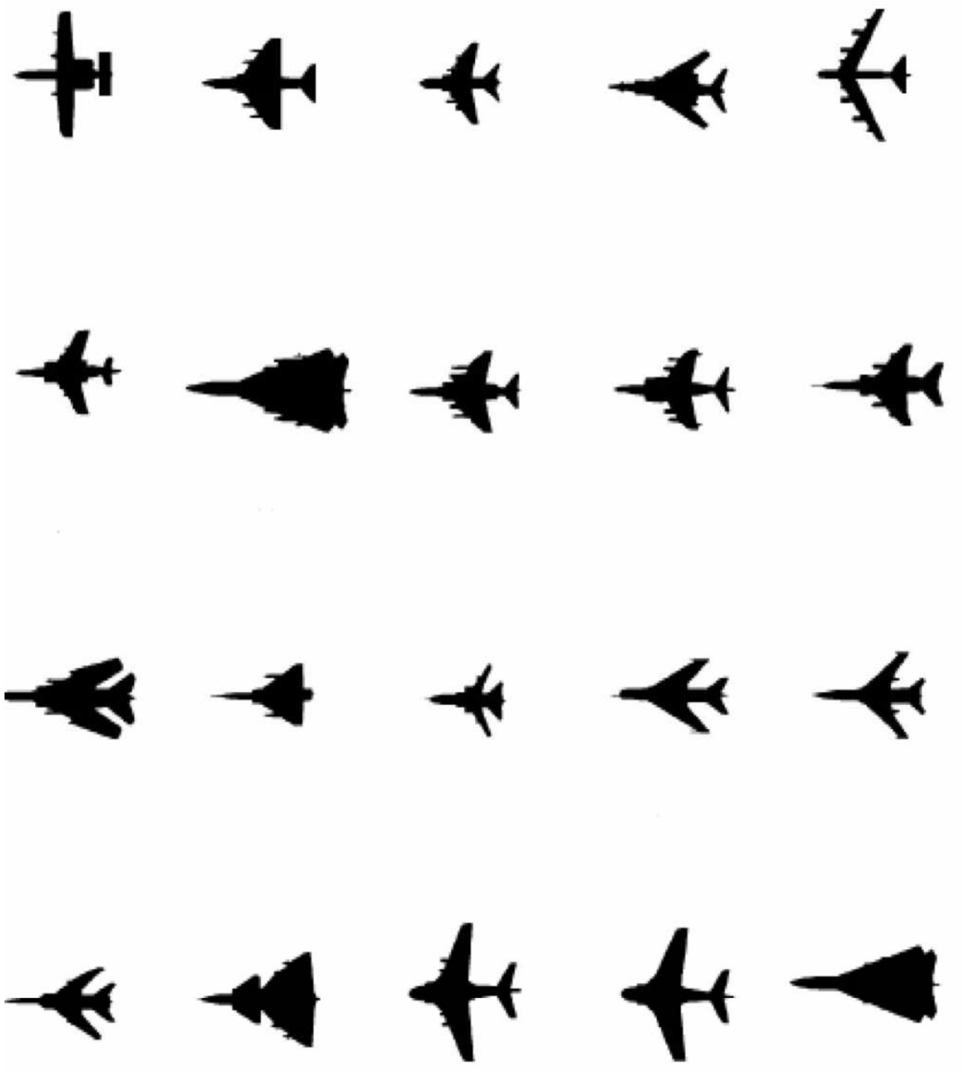
Adapted from Linda Shapiro, U of Washington

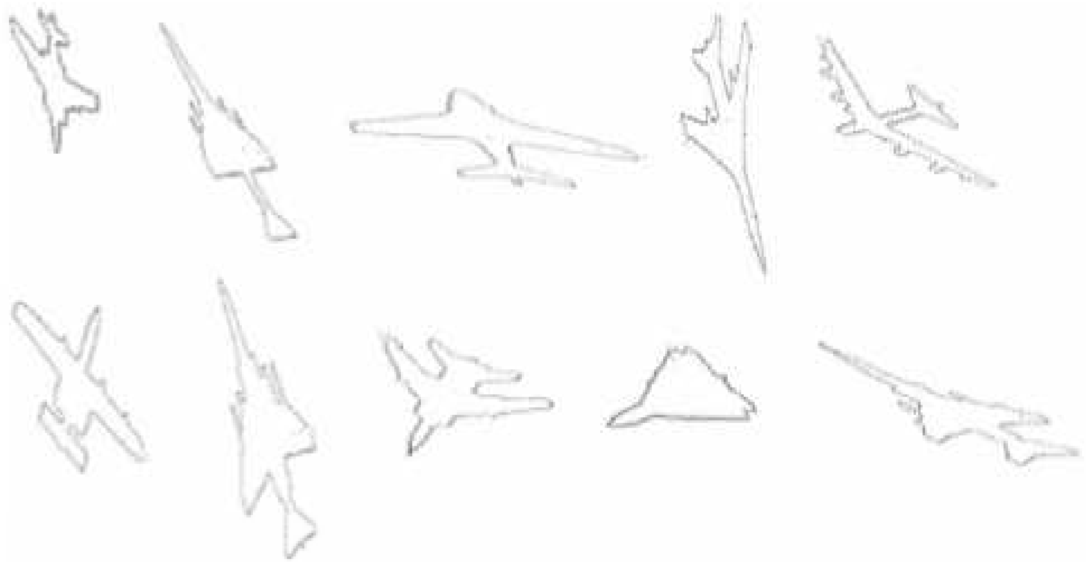
Segmentation



Adapted from Jianbo Shi, U Penn

Shape





Recognized objects

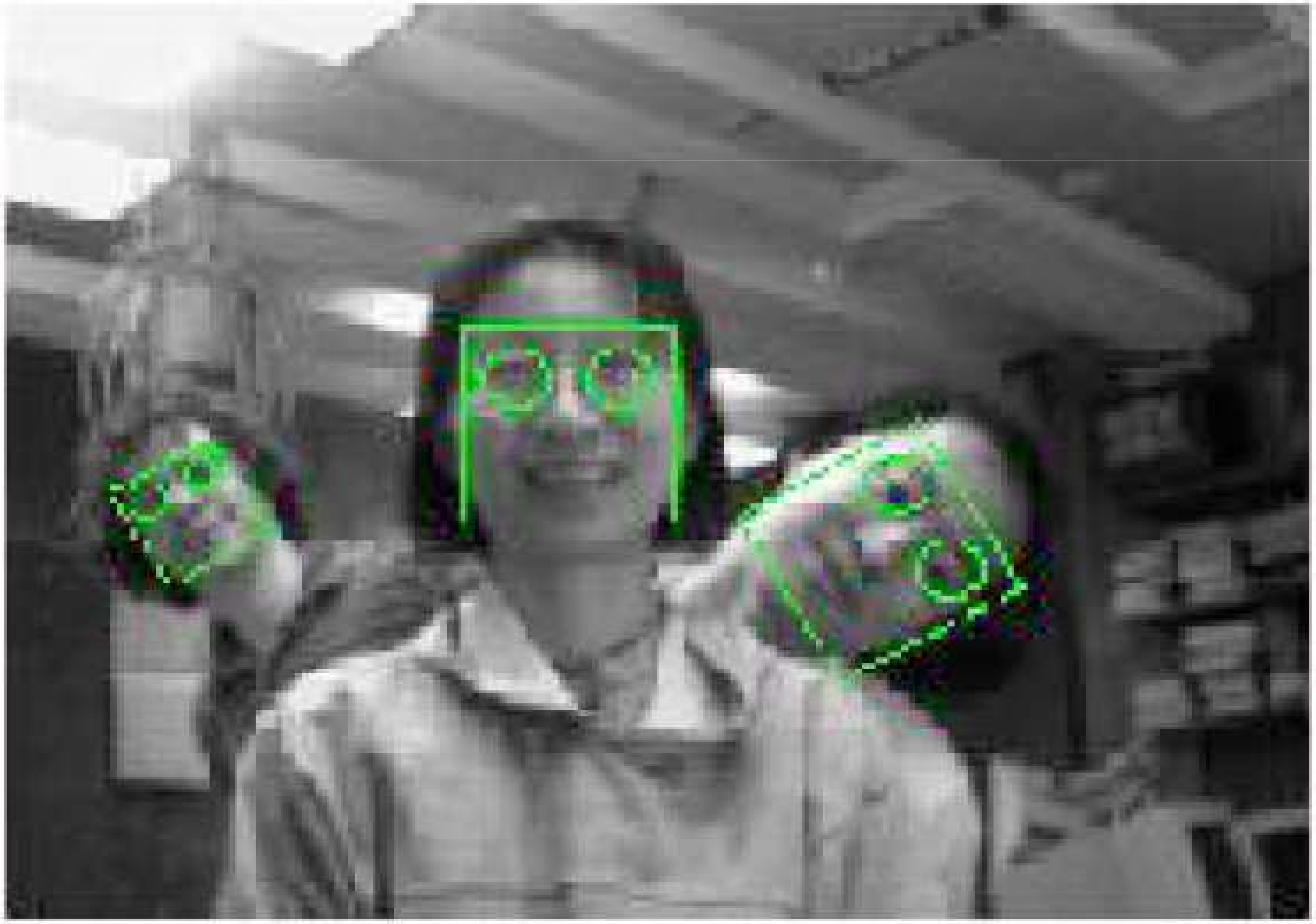
Model database

Adapted from Enis Cetin, Bilkent University

Motion

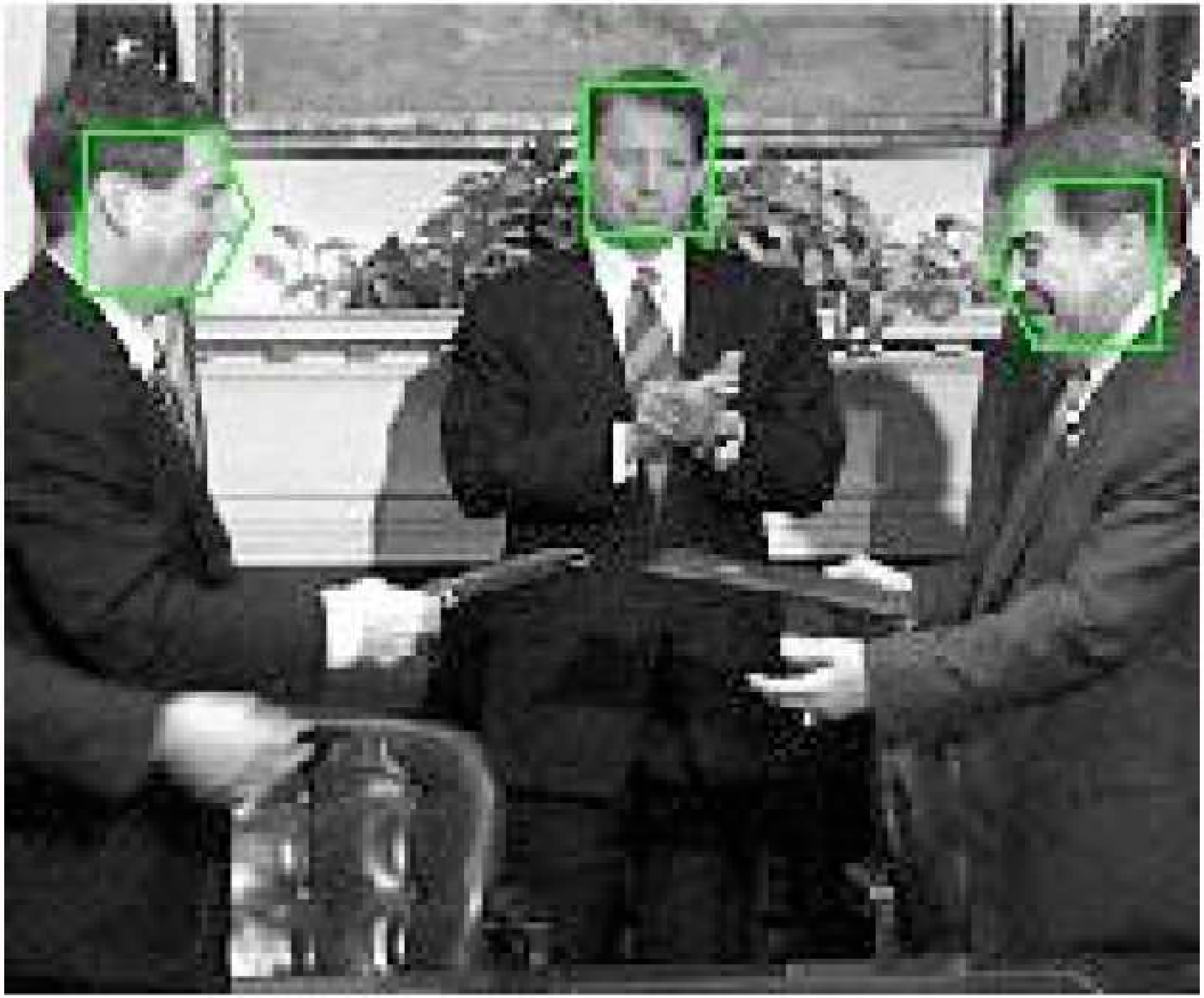
Adapted from Michael Black, Brown University

Detection



Adapted from David Forsyth, UC Berkeley

Detection

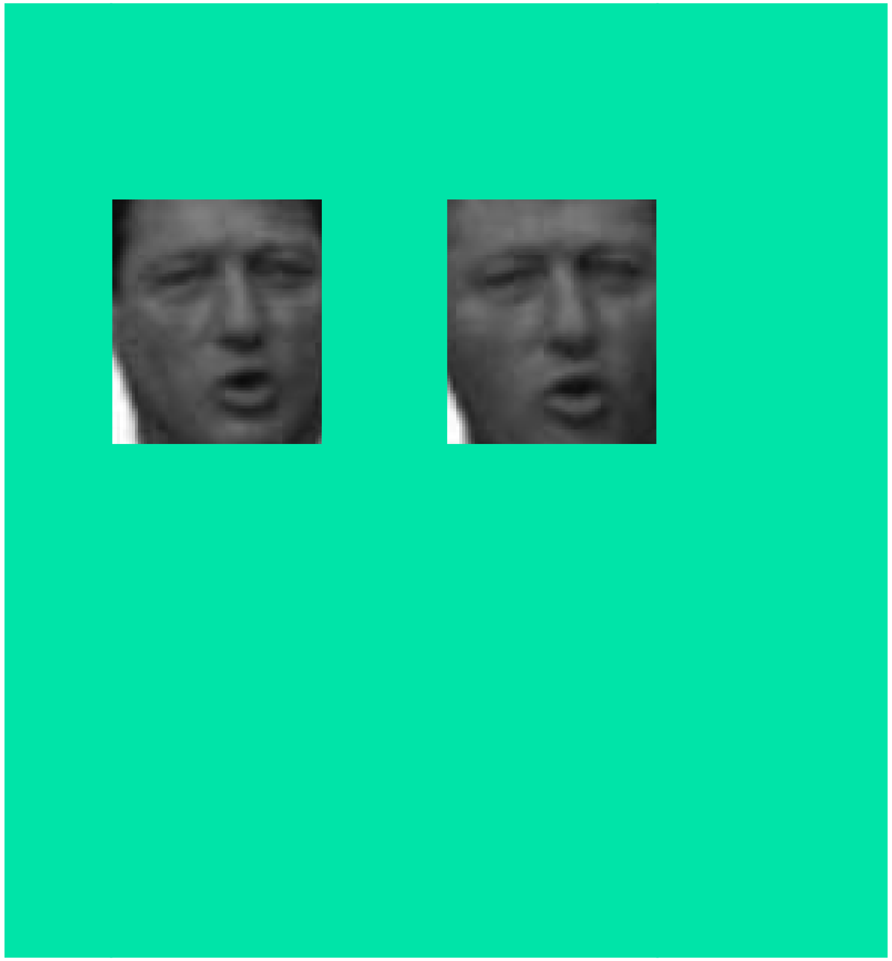


Adapted from David Forsyth, UC Berkeley

Detection

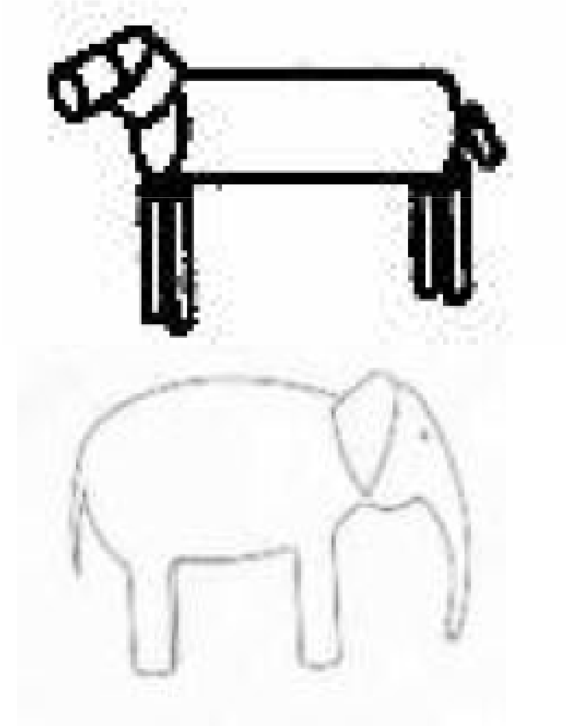
What are our "models"?

How good are they?

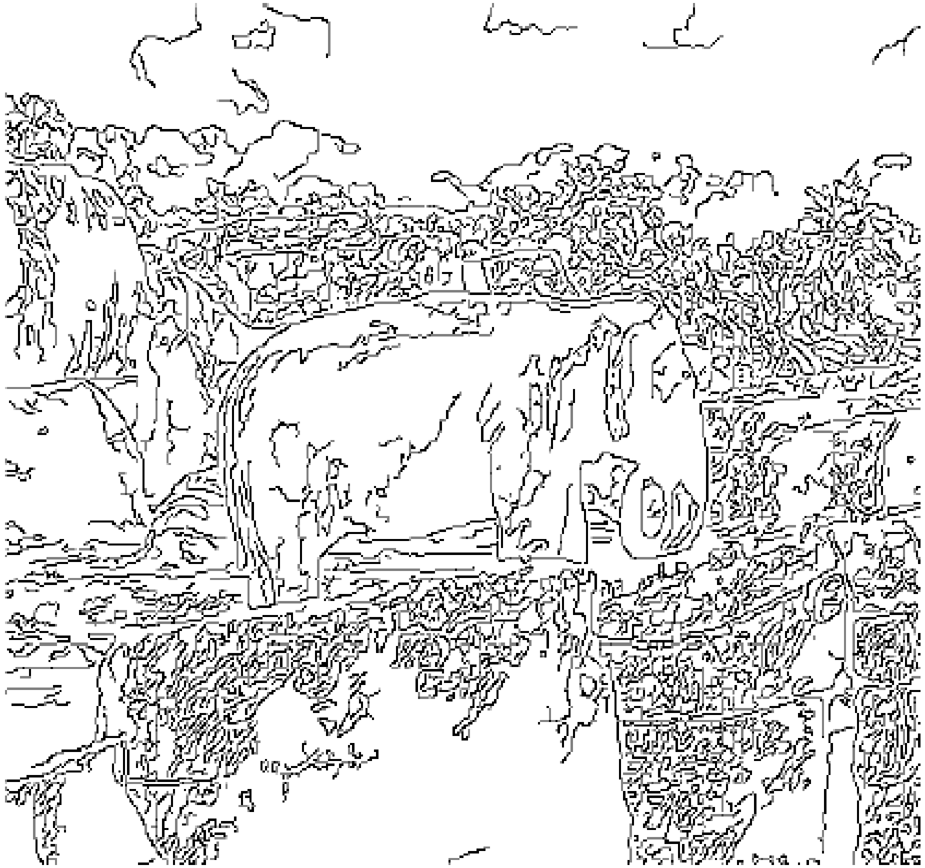


Adapted from Michael Black, Brown University

Recognition

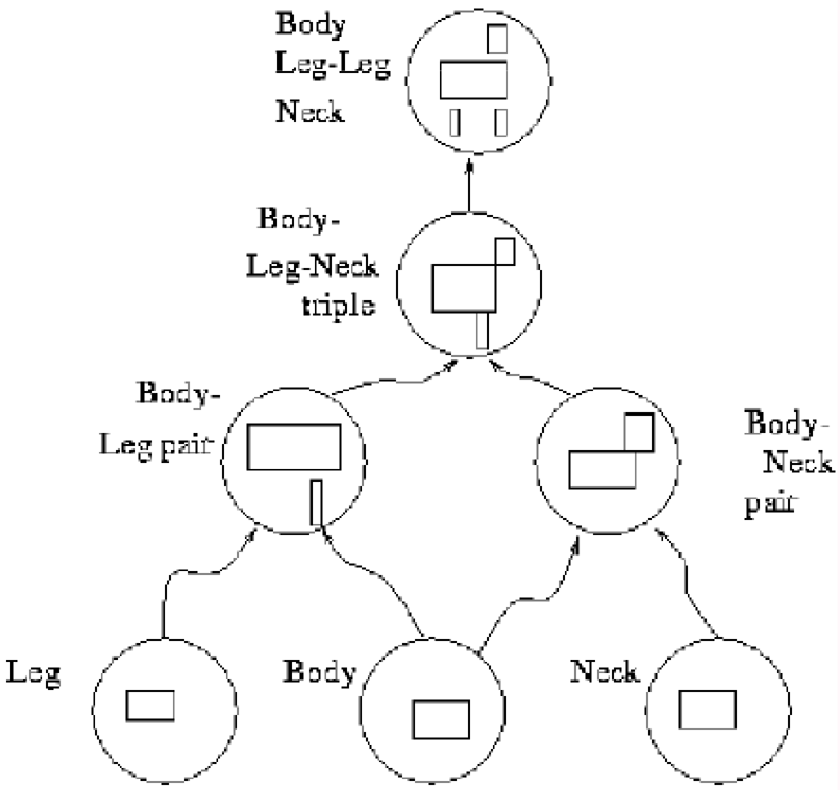


Match “model” to measurements?



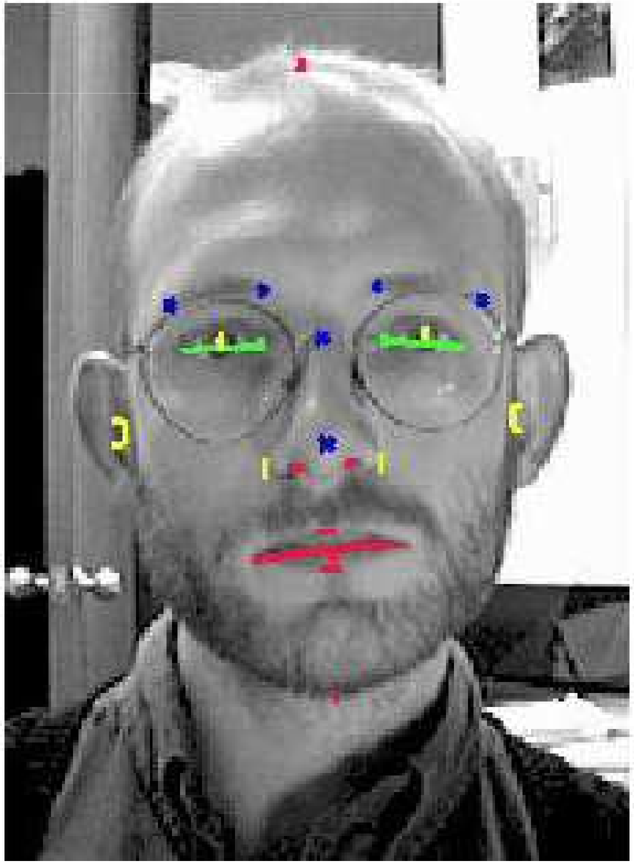
Adapted from Michael Black, Brown University

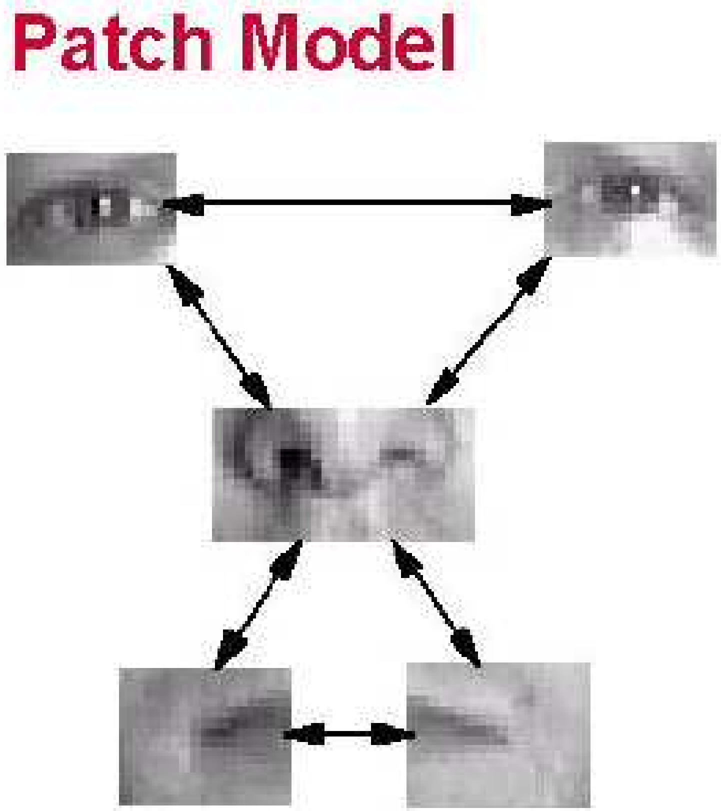
Recognition



Adapted from David Forsyth, UC Berkeley

Parts and relations





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

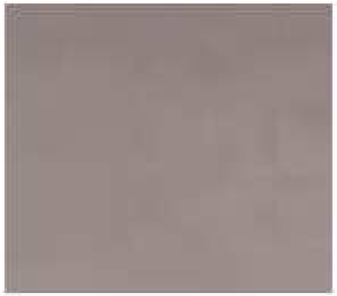
Context

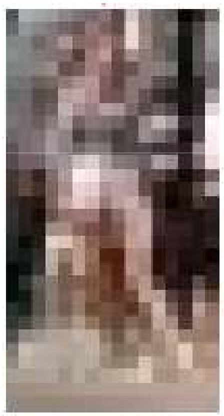
Adapted from Antonio Torralba, MIT

Context

Adapted from Antonio Torralba, MIT

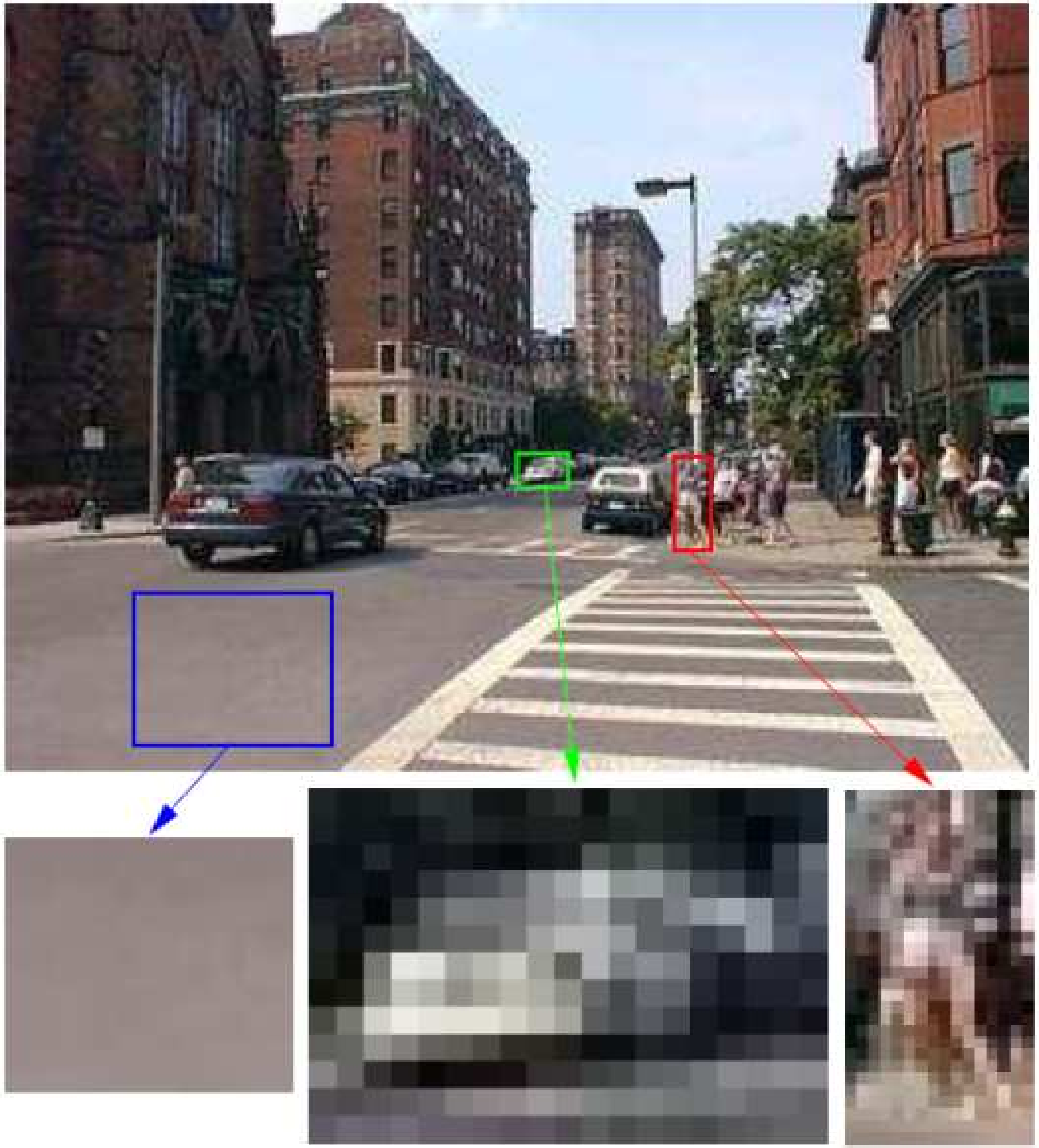
Context





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

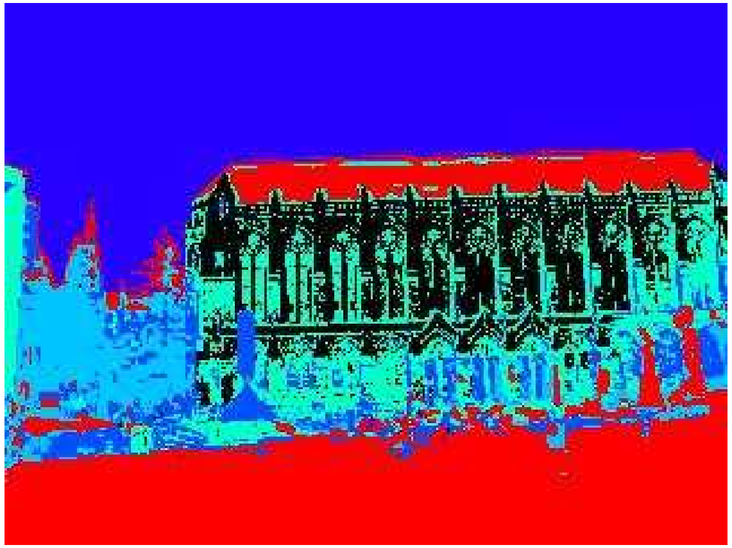
Adapted from Linda Shapiro, U of Washington

Mid-level

original color image

K-means clustering

(followed by connected component analysis)



regions of homogeneous color

data structure

Adapted from Linda Shapiro, U of Washington

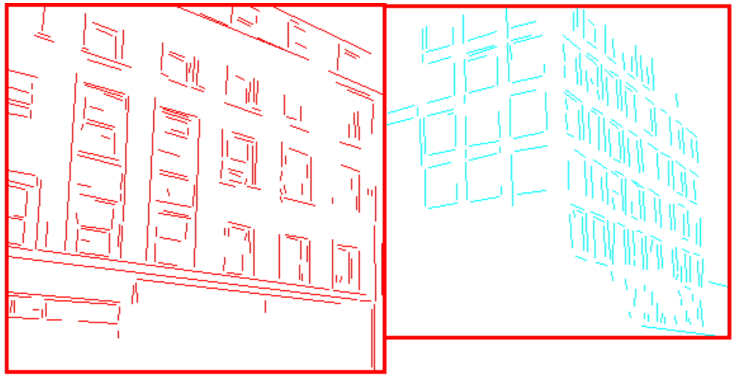
Low-level to high-level

low-level → edge image

mid-level ↓

consistent line clusters

high-level ←



Adapted from Linda Shapiro, U of Washington