Computer Vision

Jana Kosecka
http://cs.gmu.edu/~kosecka/cs482/
kosecka@cs.gmu.edu

Some slides thanks to S. Seitz, S. Lazebnik and others

Logistics

- **Grading:** Homeworks (about every 2 weeks) 40%, Exam 30%, Project 30%
- **Prerequisites:** basic statistical concepts, geometry, linear algebra, calculus
- **Recommended text:**
- **From Images to Geometric Models:** Y. Ma, S. Soatto, J.Kosecka and S. Sastry, Springer Verlag 2003 (I will provide relevant chapters online)
- **Computer Vision: Algorithms and Applications**. R. Szeliski, 2010 (available online)
- Computer Vision a Modern Approach (D. Forsyth, J. Ponce, Prentice Hall 2002)
- Image Processing, Analysis, and Machine Vision. Sonka, Hlavac, and Boyle. Thomson.
- **Computer Vision**. Ballard and Brown (available on-line)

- Required Software MATLAB (with Image Processing toolbox)
- Open CV library

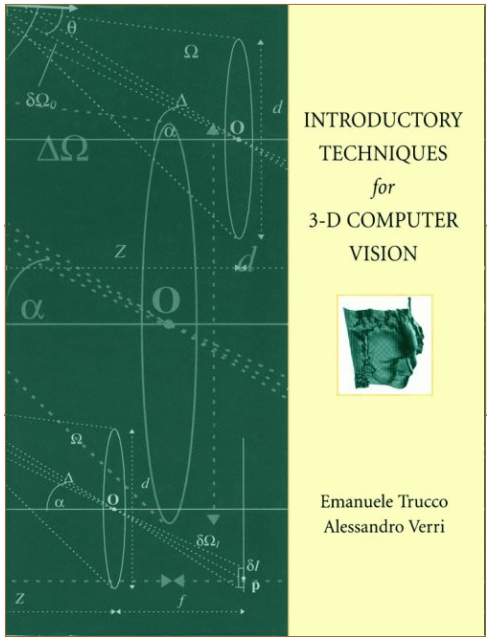
OpenCV Resources

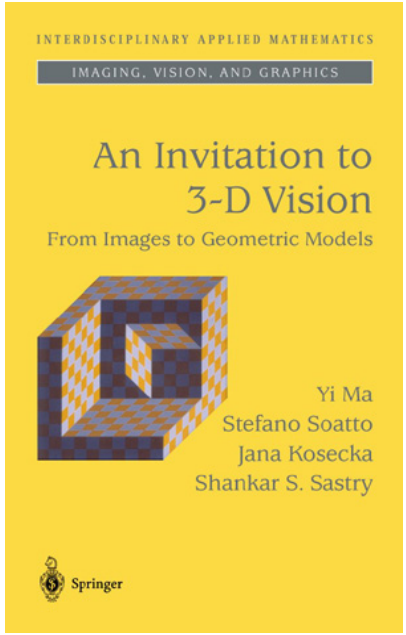
- OpenCV documentation etc (opencv.org)
- OpenCV Safari Books Free for GMU students
- OpenCV Computer Vision with Python A good source for installation in various OS, code examples, etc.
- Programming Computer Vision with Python Not OpenCV, but a lot of examples

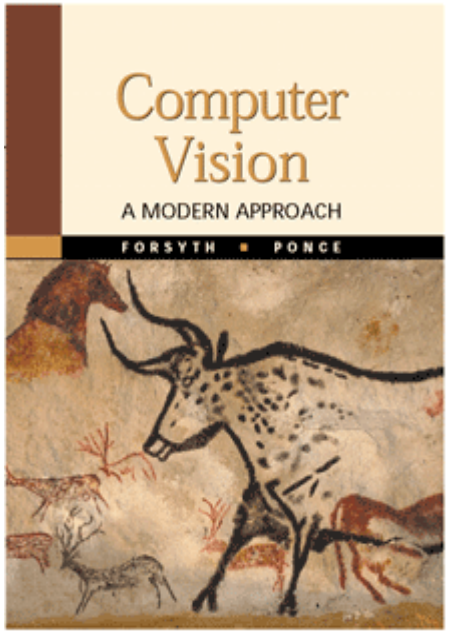
Late Policy

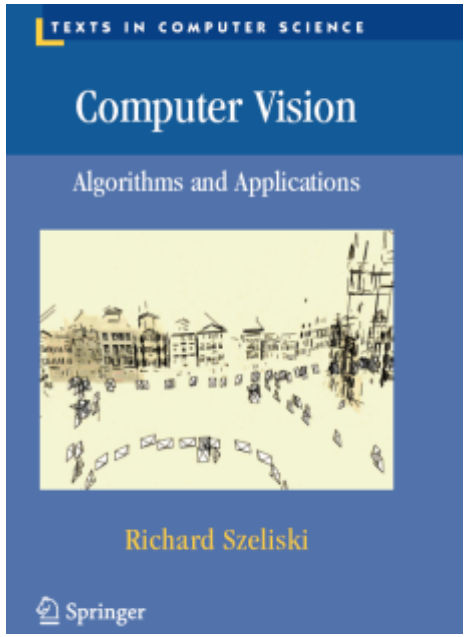
- Every student has a 3 late day credit which can be used towards handing in the homeworks and project proposal.

Textbooks


INTRODUCTORY
TECHNIQUES
for
3-D COMPUTER
VISION
Emanuele Trucco
Alessandro Verri


INTERDISCIPLINARY APPLIED MATHEMATICS
IMAGING, VISION, AND GRAPHICS
An Invitation to
3-D Vision
From Images to Geometric Models
Yi Ma
Stefano Soatto
Jana Kosecka
Shankar S. Sastry
Springer


Computer
Vision
A MODERN APPROACH
FORSYTH ■ PONCE


TEXTS IN COMPUTER SCIENCE
Computer Vision
Algorithms and Applications
Richard Szeliski
Springer

Project Deadlines

- Check Web site for proposals, or develop your own
- Teams 2-3 people
- Dates
  - November 2, project proposals due
  - December week of finals final report due
  - Project presentations

To define your own project…

- Generate project description for the Class Web site
- Find a mentor
- Gather data, process data
- Write suitable project proposal

Examples:

- Learn to find sports videos on youtube.com
- Match images of same location at flickr.com
- Fly autonomous helicopter with camera
- Reconstruct 3D scene from a moving camera
- Detect pedestrians from moving camera on the car
- HCI projects – use your hand as remote

- <your idea here>

Today's Goals

- Get Excited about Computer Vision
- Learn about image formation (Part 1)

What is vision?

- From the 3-D world to 2-D images: image formation (physics).
  - Domain of artistic reproduction (synthesis): painting, graphics.
- From 2-D images to the 3-D world: image analysis (mathematical modeling, inference).
  - Domain of vision: biological (eye+brain), computational

IMAGE SYNTHESIS: image-formation process

- Pinhole (perspective) imaging in most ancient civilizations.
- Euclid, perspective projection, 4th century B.C., Alexandria (Egypt)
- Pompeii frescos, 1st century A.D. (ubiquitous).
- Geometry understood very early on, then forgotten.

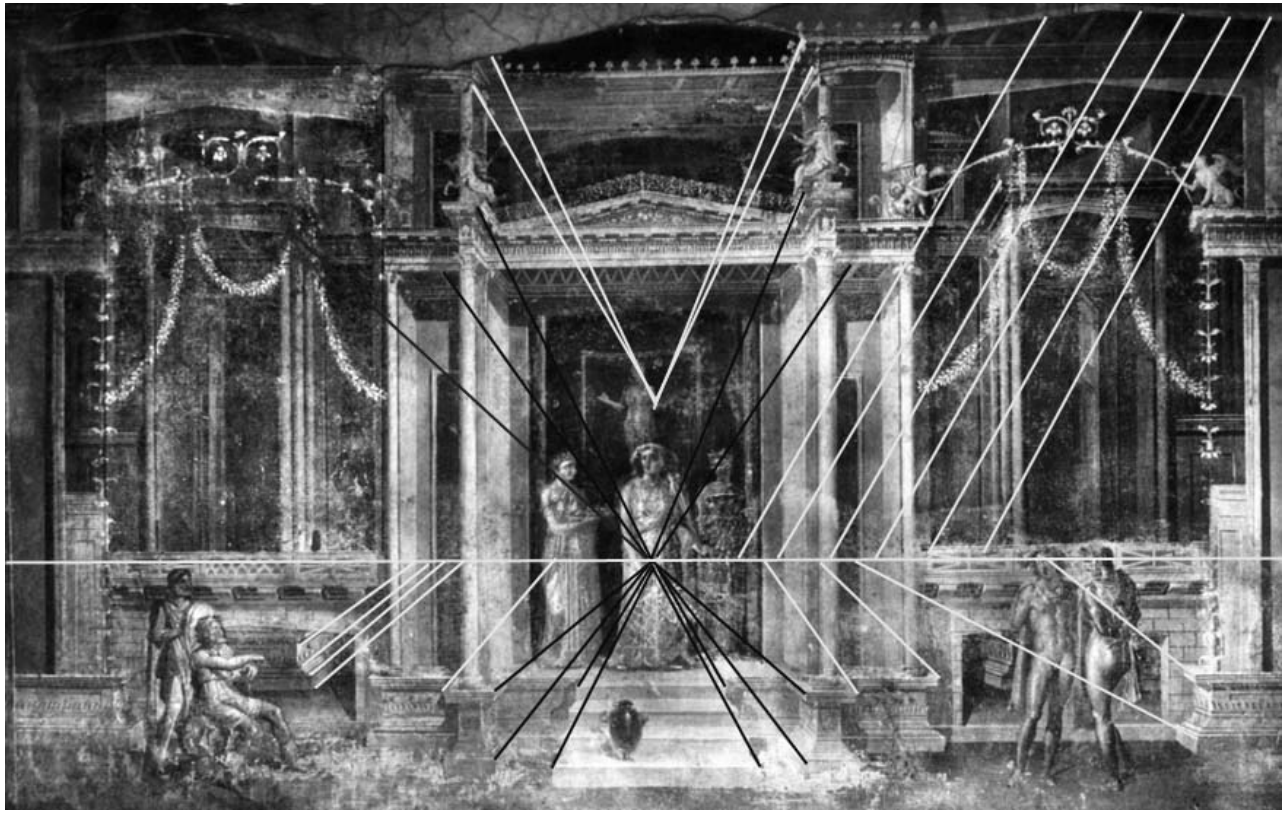


Image courtesy of C. Taylor

PERSPECTIVE IMAGING (geometry)

- Re-discovered and formalized in the Renaissance:
- Fillippo Brunelleschi, first Renaissance artist to paint with correct perspective,1413
- "Della Pictura", Leon Battista Alberti, 1435, first treatise
- Leonardo Da Vinci, stereopsis, shading, color, 1500s
- Raphael, 1518

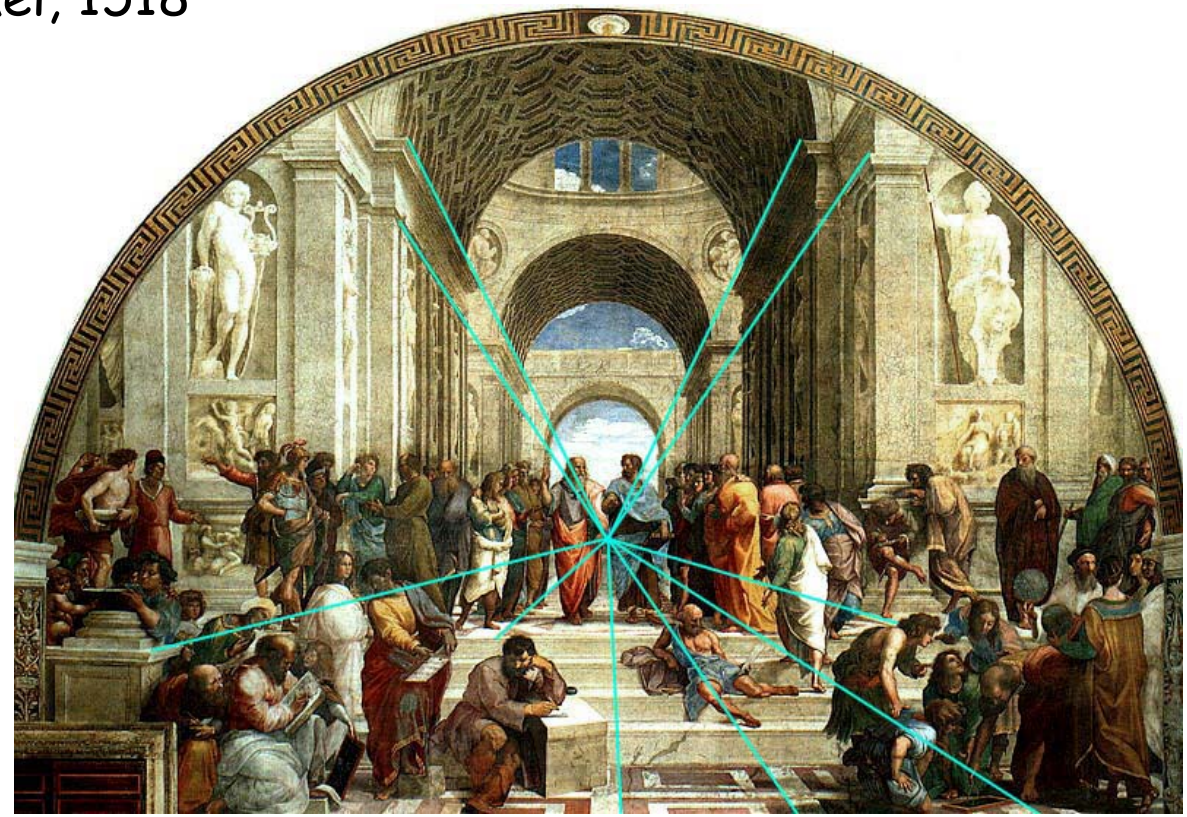


Image courtesy of C. Taylor

Goals of Computer Vision

- Build machines and develop algorithms which can automatically replicate some functionalities of biological visual system

- Systems which navigate in cluttered environments
- Systems which can recognize objects, activities
- Systems which can interact with humans/world

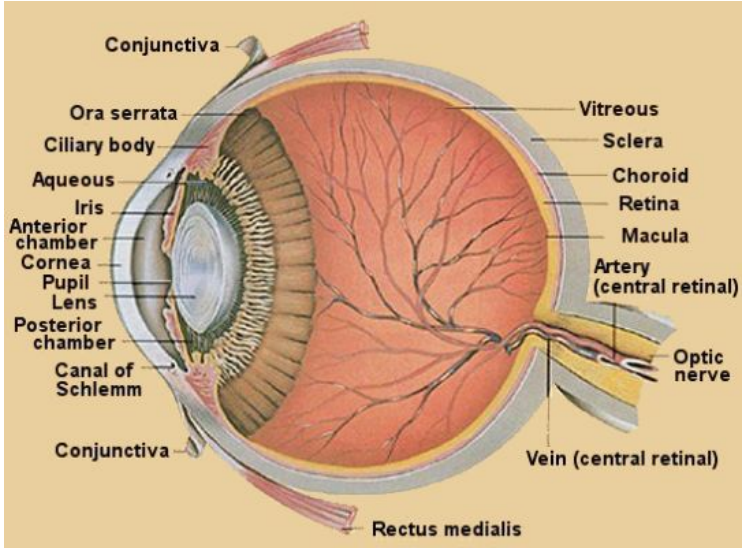
- Synergies with other disciplines and various applications
  Artificial Intelligence - robotics, natural language understanding

- Vision as a sensor - medical imaging, Geospatial Imaging, robotics, visual surveillance, inspection

Computer Vision

Visual Sensing

Images I(x,y) – brightness patterns



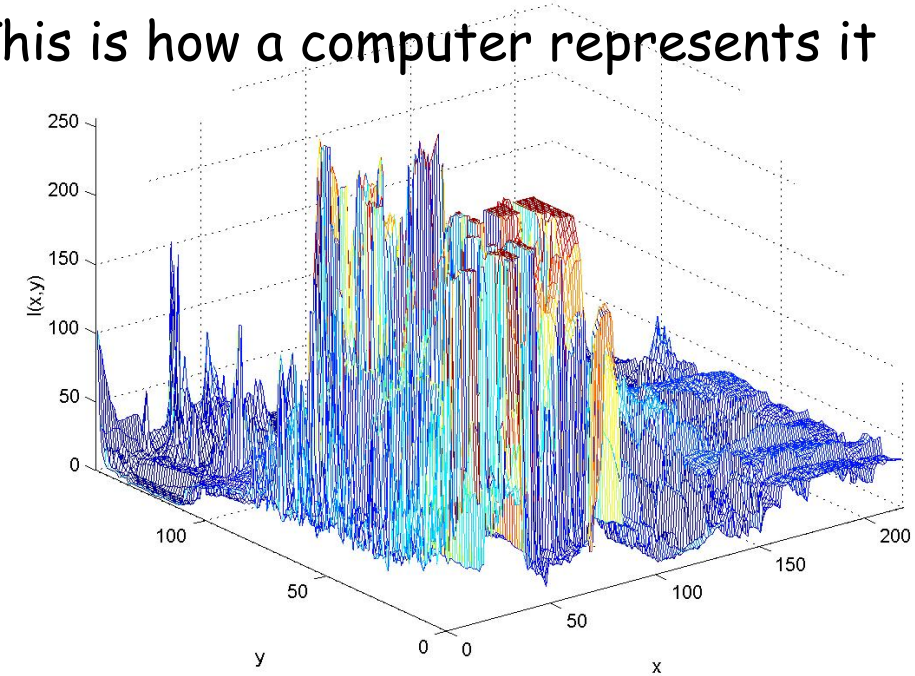
- image appearance depends on structure of the scene
- material and reflectance properties of the objects
- position and strength of light sources

Visual Information Processing

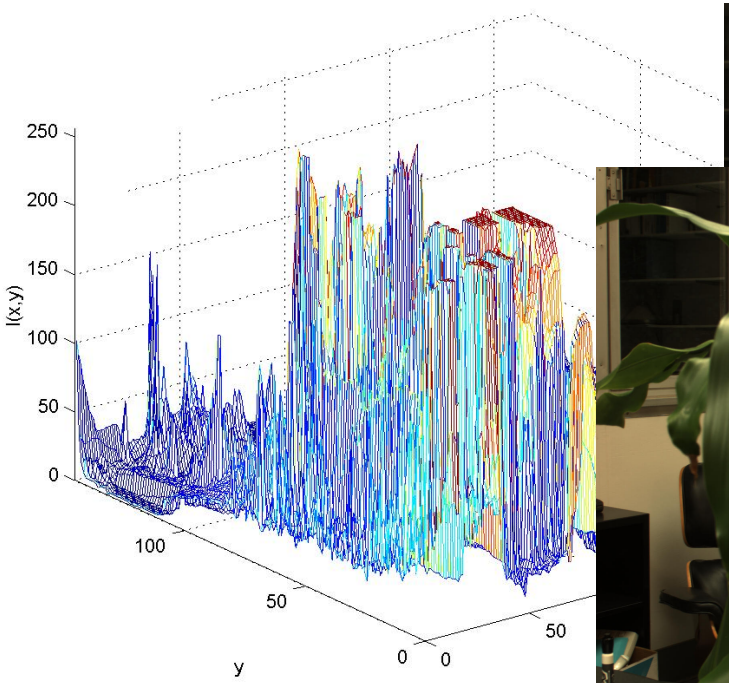
This is the part of your
brain that processes visual
information

[Felleman & Van Essen, 1991]

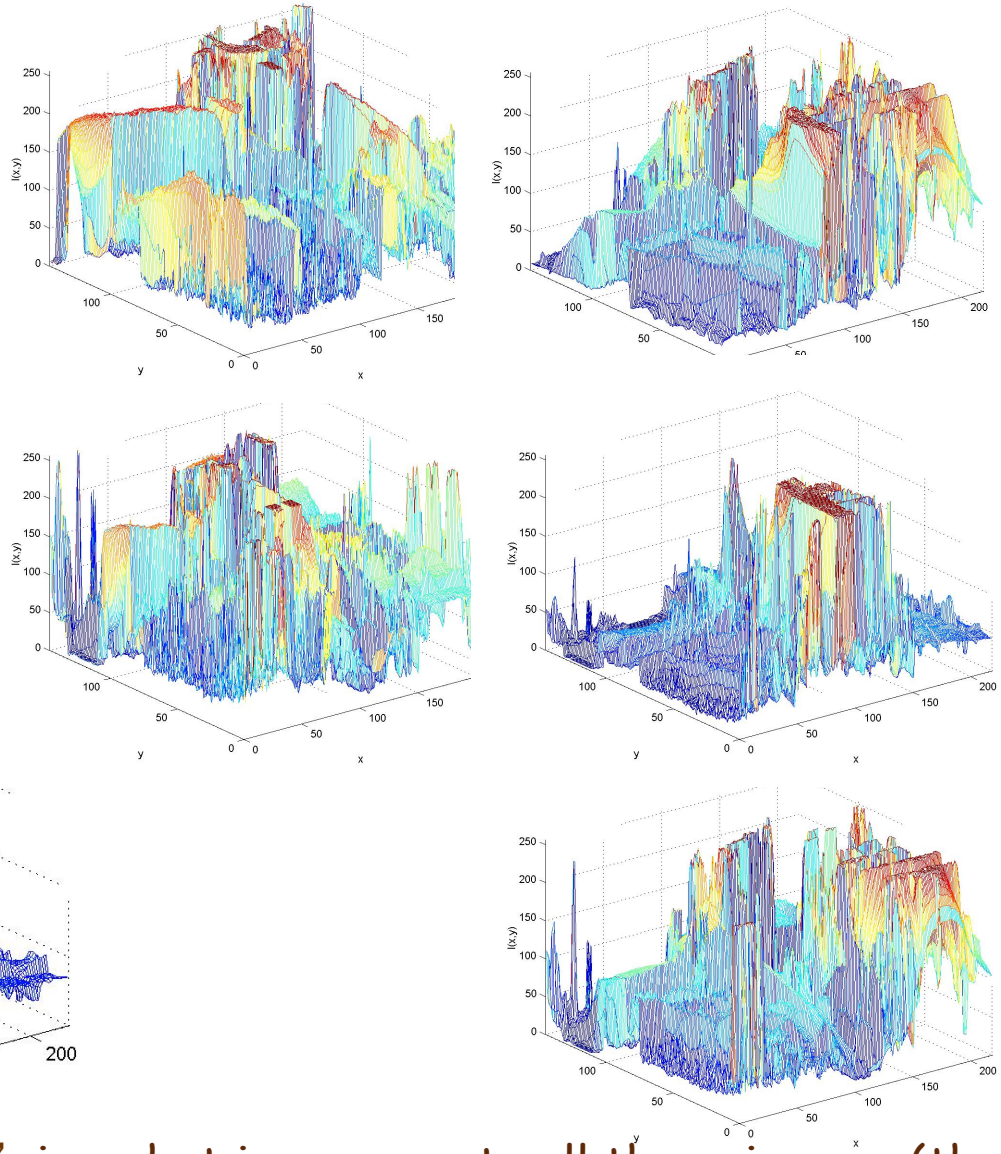
This is how a computer represents it

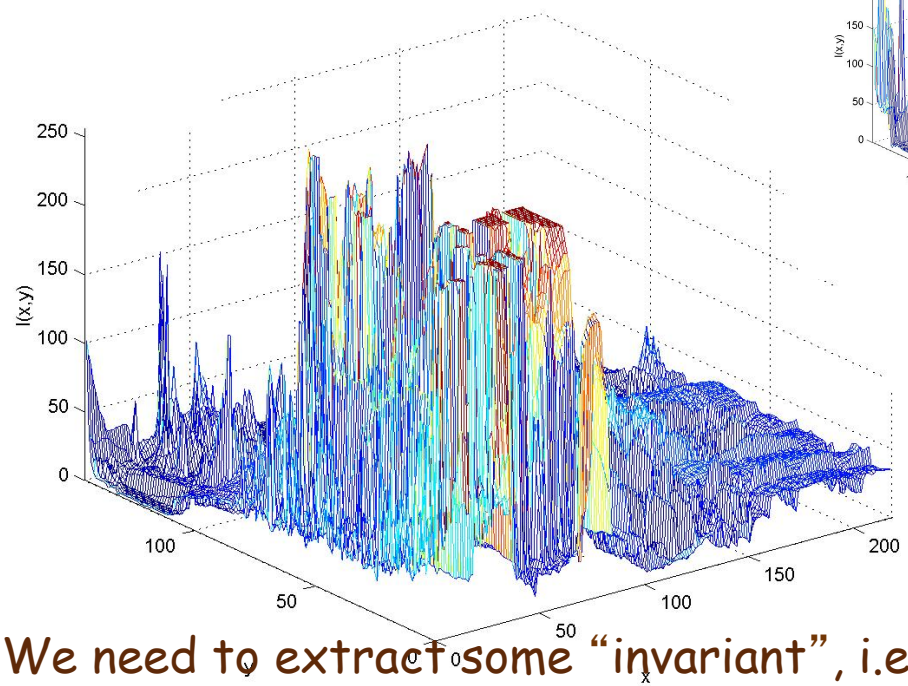



And so is this …


250
200
150
100
50
0
I(x,y)
100
50
0
y
0
50

And so are these!





We need to extract some "invariant", i.e. what is common to all these images (they are all images of an office)

☹ truly invariant (photometric and geometric) representations do not exist

BUMMER! THIS IS IMPOSSIBLE!

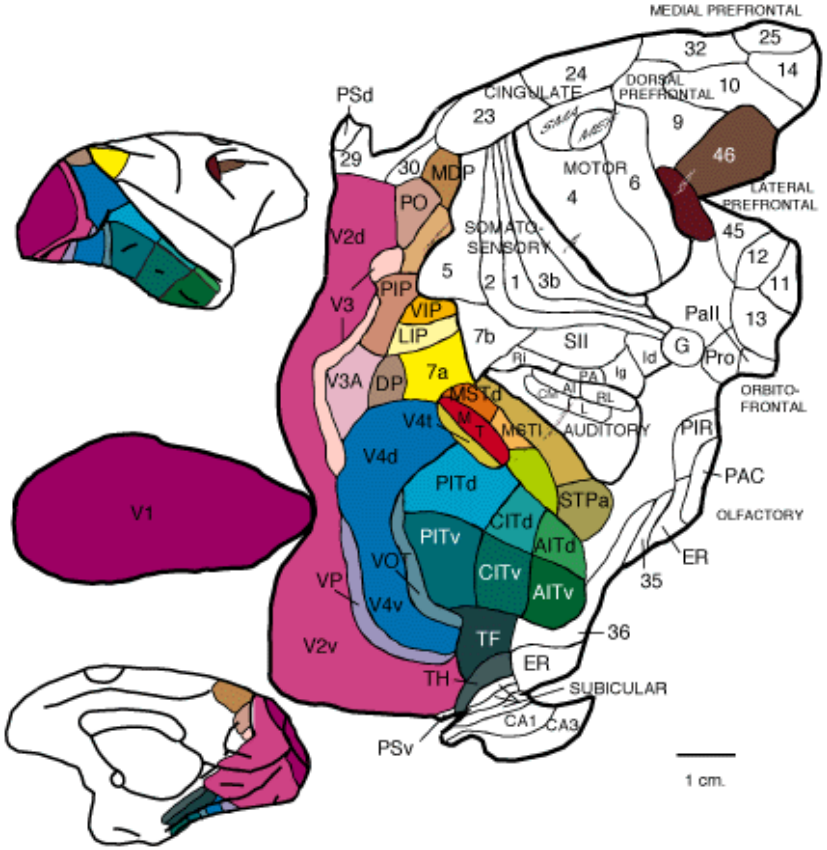
- THM: [Weiss, 1991]: There exists NO generic viewpoint invariant!
- THM: [Chen et al., 2003]: There exists NO photometric invariant!!

- So, how do we (primates) solve the problem?

Challenges/Issues

- About 40% of our brain is devoted to vision
- We see immediately and can form and understand images instantly

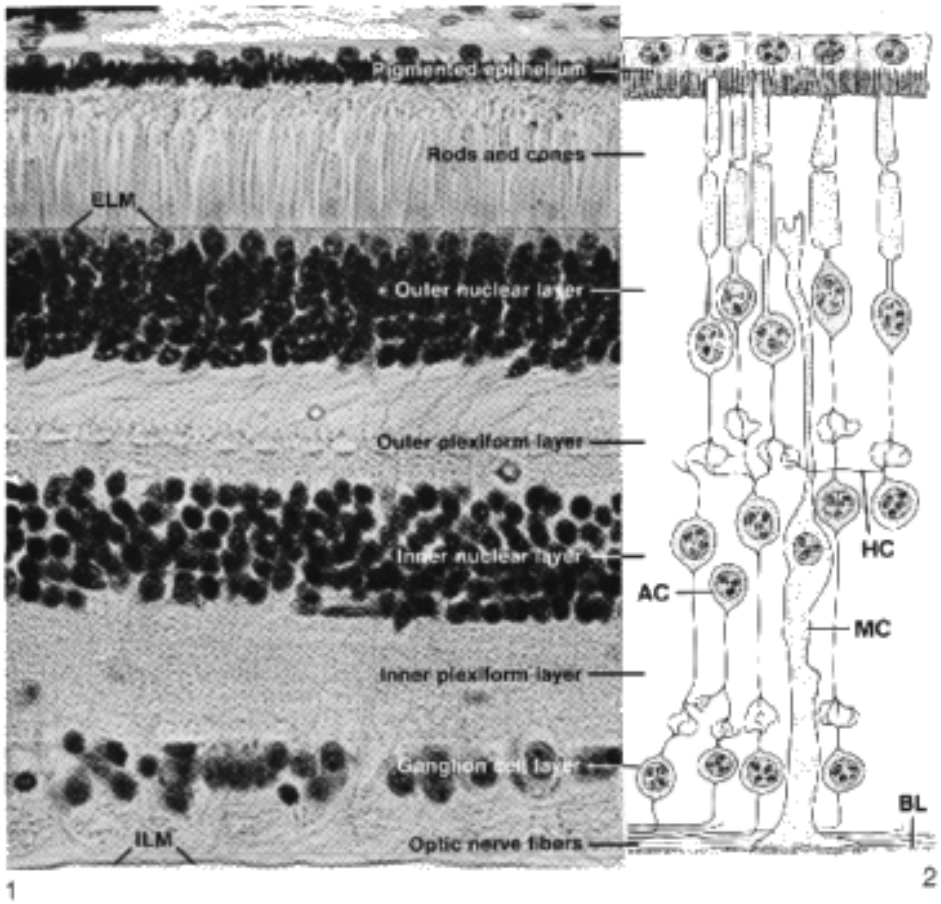
- Applications and examples



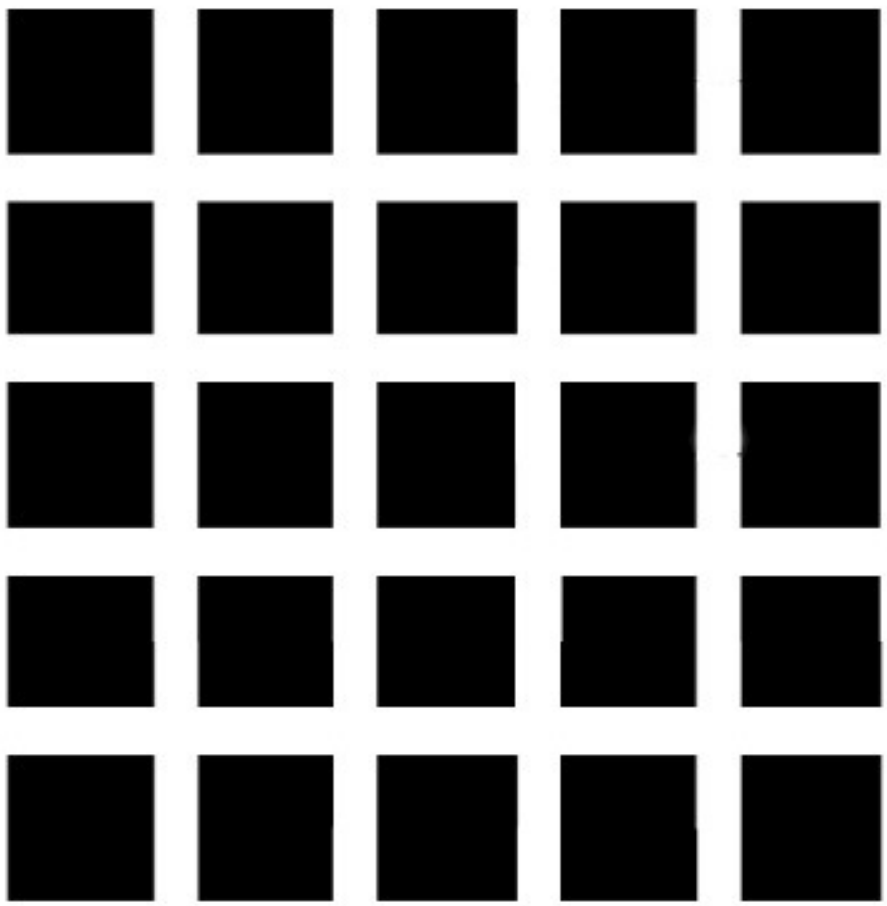
EXAMPLE OF A (VERY COMMON) SENSOR NETWORK

Retina performs distributed computation:

- Contrast adaptation (lateral inhibition)
- Enhancement/edgels (e.g. Mach bands)
- Motion detection (leap frog)

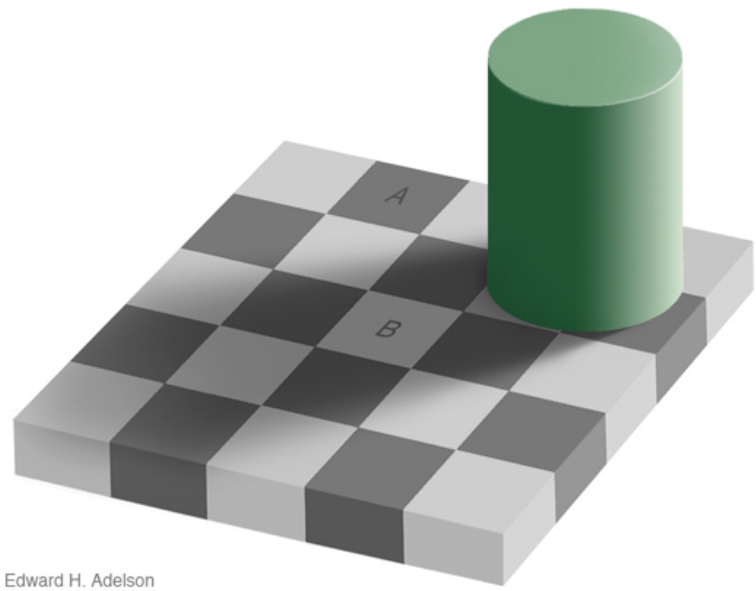


Optical Illusion



Look at the crosses they appear to be gray

Optical Illusion



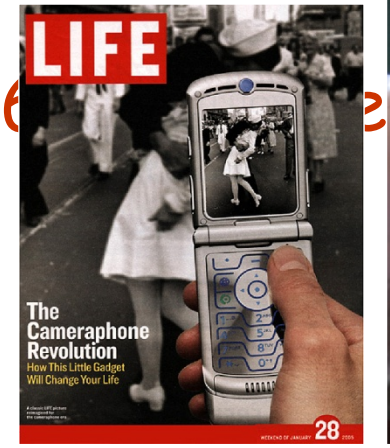
Checker A and B are of the same gray-level value

http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

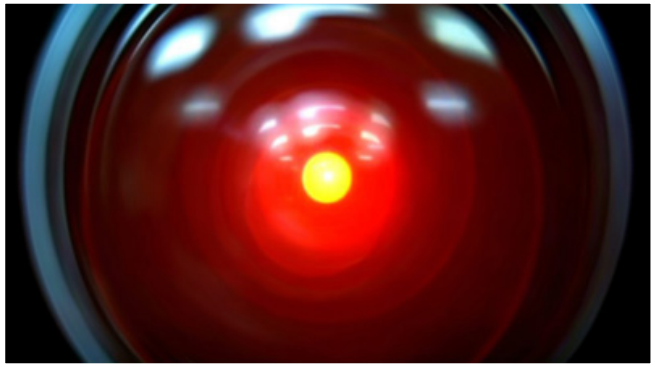
http://www.psy.rittsumei.ac.jp/~akitaoka/rotsnakee.html

- THESE ARE NOT SPIRALS
- AND THEY ARE NOT MOVING!

Cause Peripheral Drift


LIFE
The Cameraphone Revolution
How This Little Gadget Will Change Your Life
28

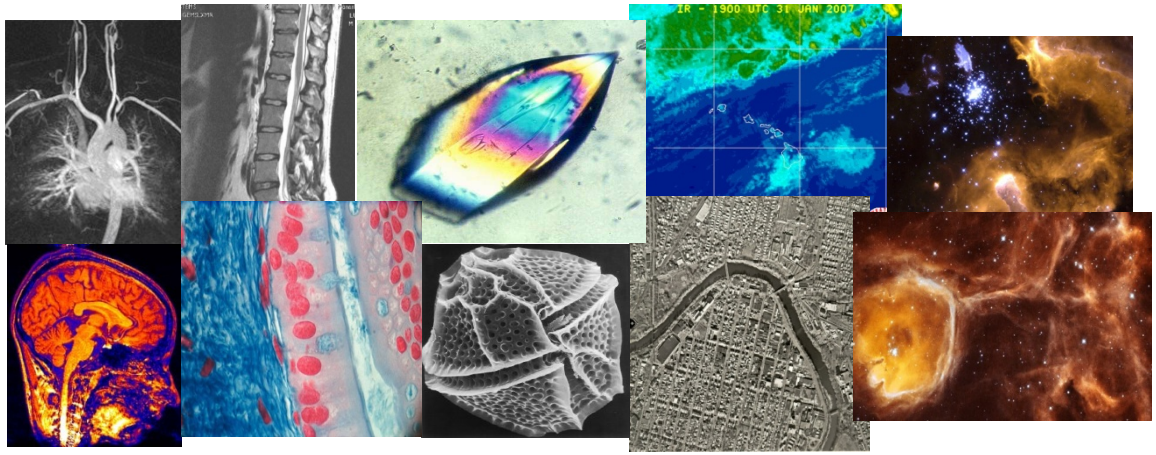

detector camera



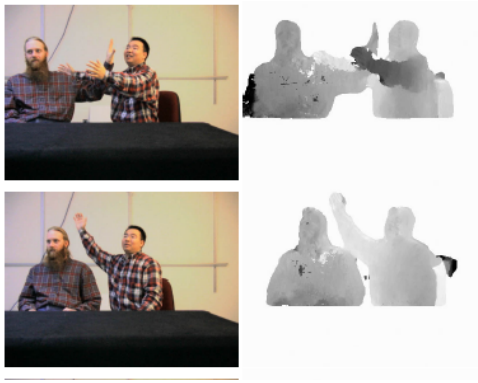
Why study computer vision?

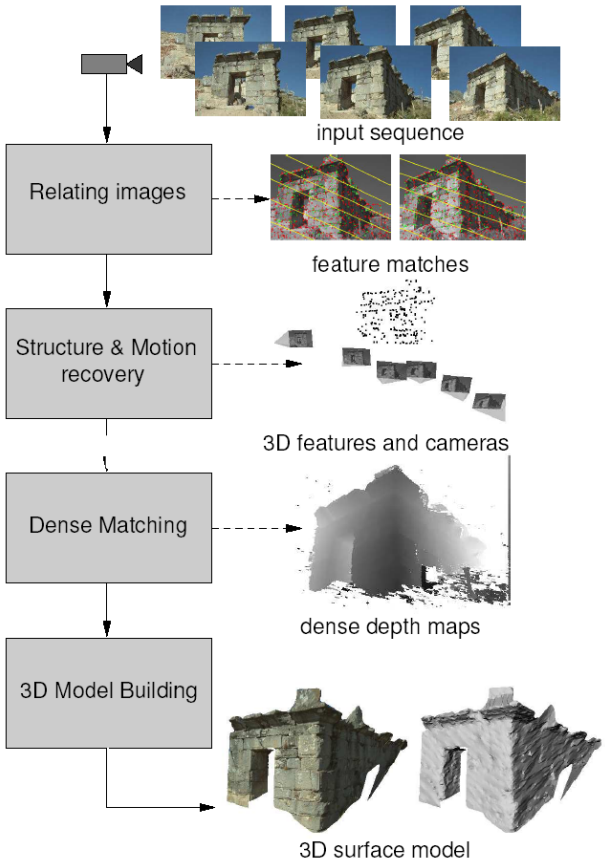
- Vision is useful: Images and video are everywhere!

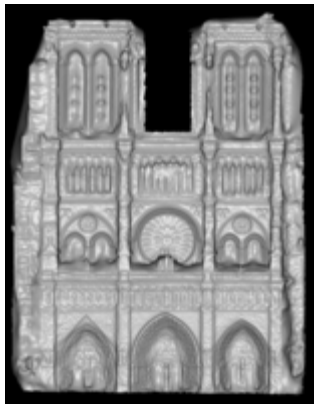
Google Image Search Picasa flickr webshots picsearch YouTube Broadcast Yourself



Vision as measurement device







Vision as a source of semantic information

slide credit: Fei-Fei, Fergus & Torralba

Object categorization

slide credit: Fei-Fei, Fergus & Torralba

Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

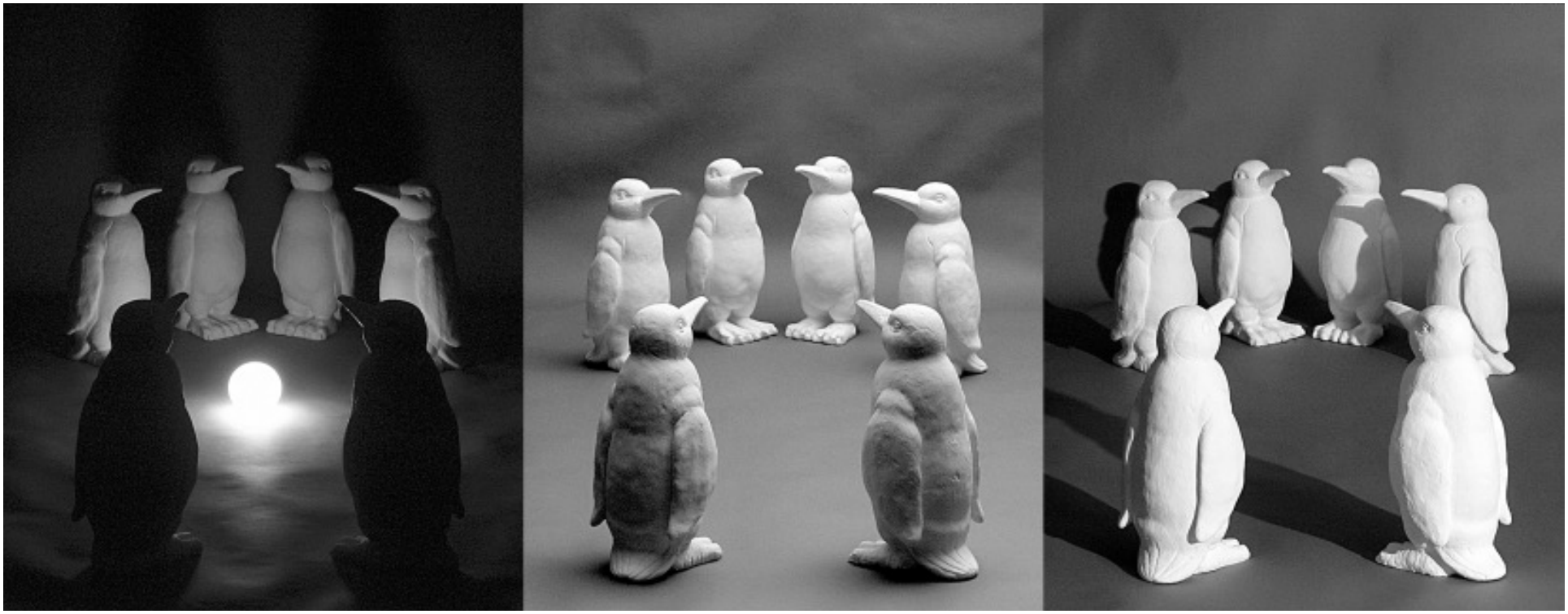
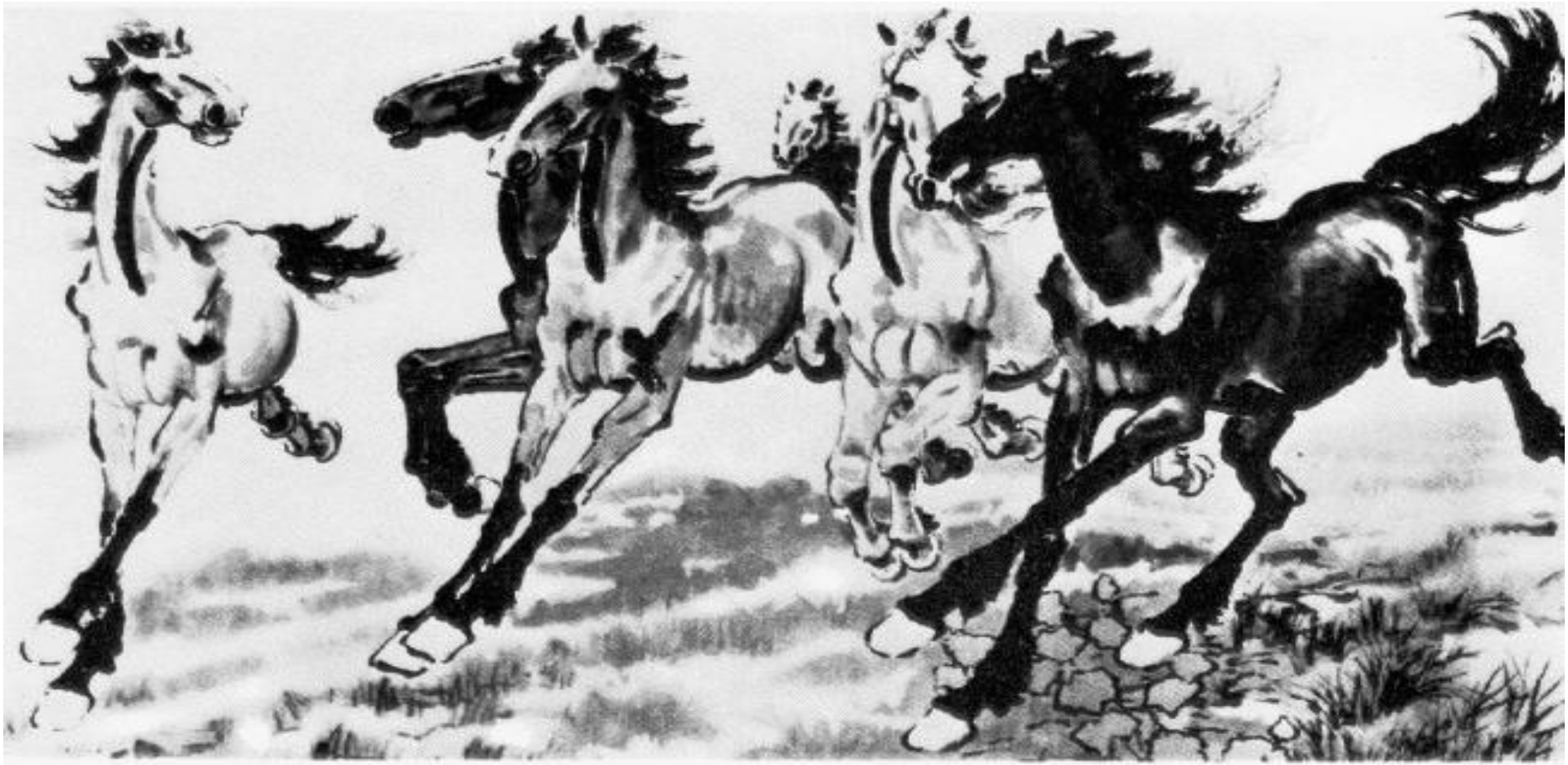


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

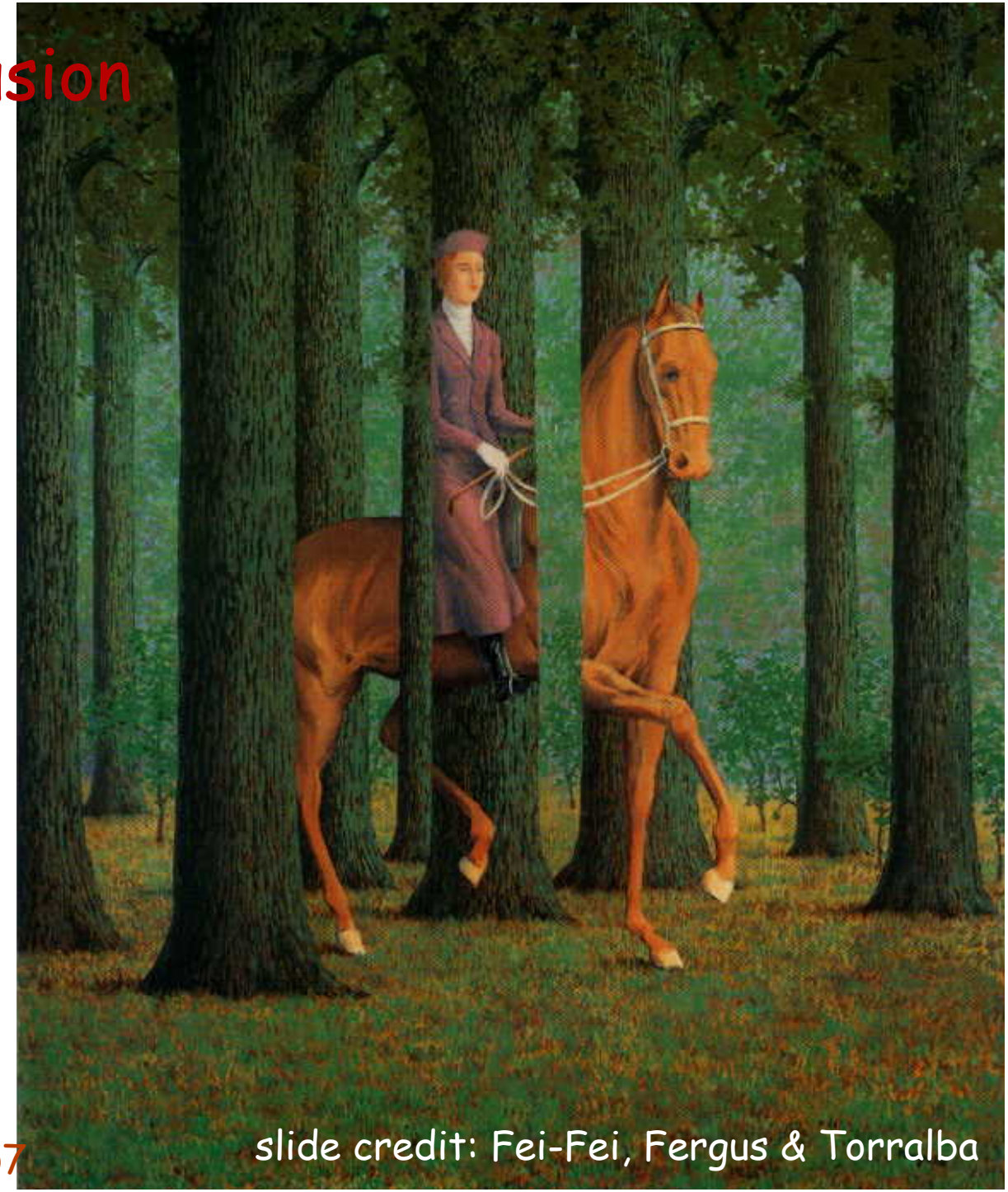
Challenges: deformation



Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenges: occlusion



Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

Challenges: background clutter

Challenges: Motion

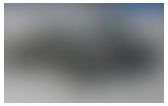
Challenges: object intra-class variation

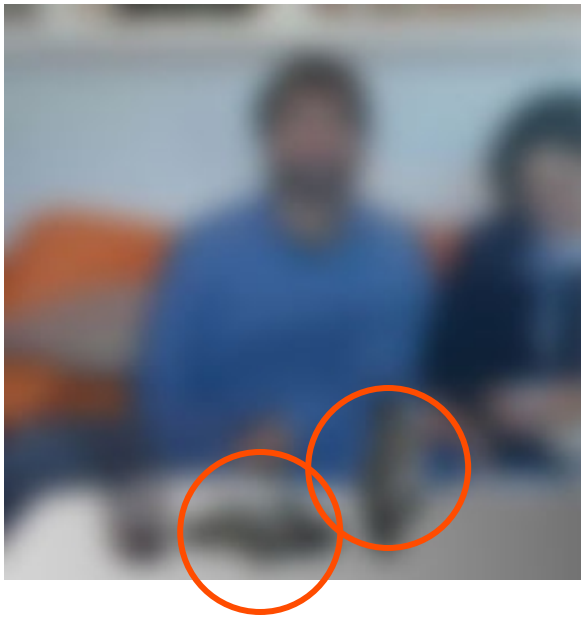


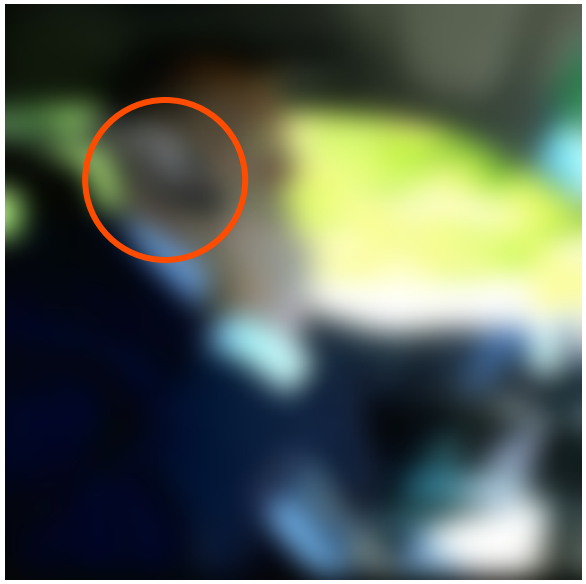


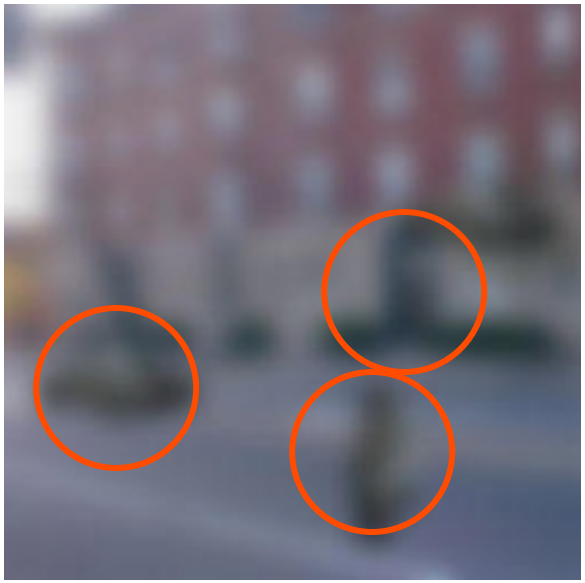

slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity









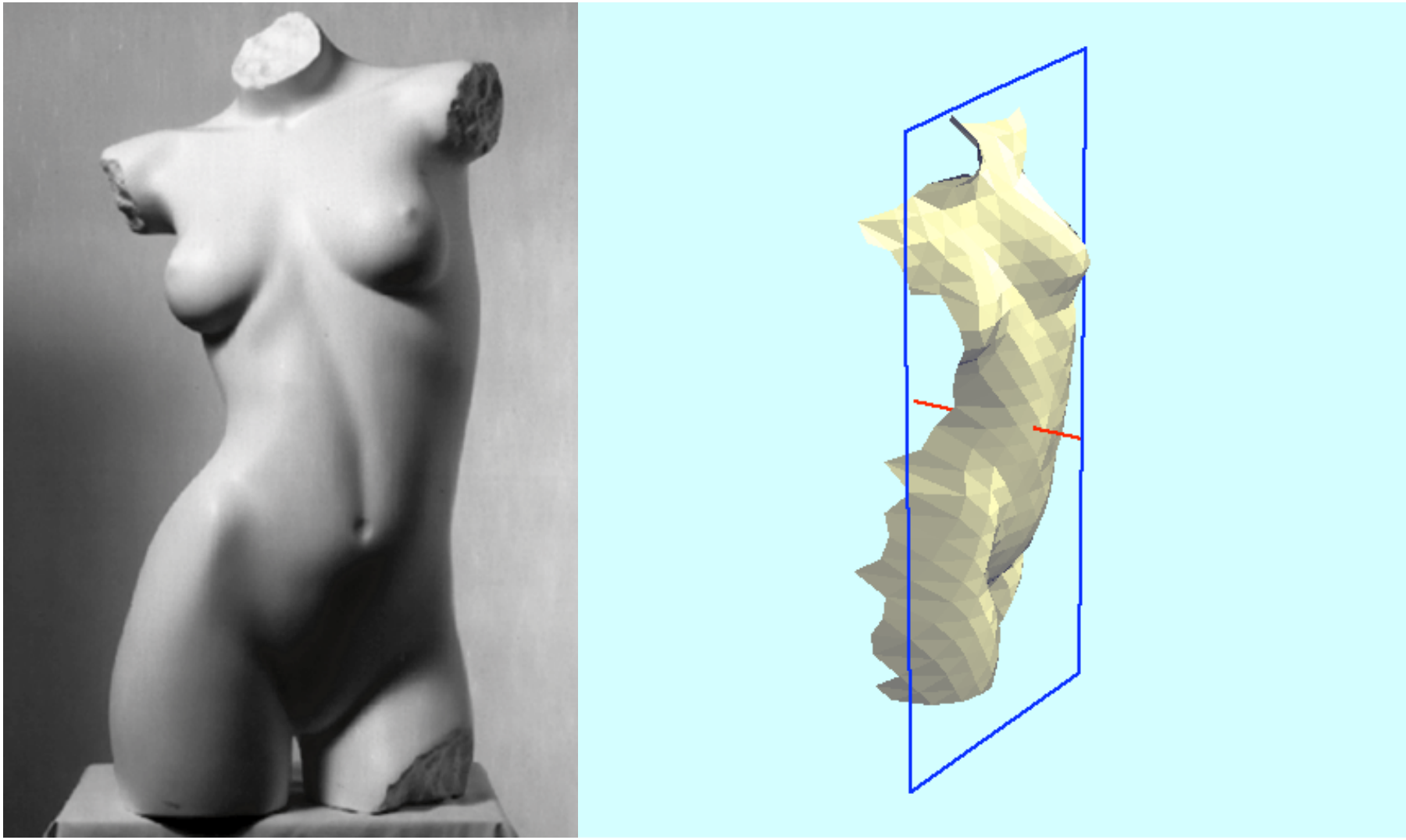
slide credit: Fei-Fei, Fergus & Torralba

Challenges or opportunities?

- Images are confusing, but they also reveal the structure of the world through numerous cues
- Our job is to interpret the cues!

Image source: J. Koenderink

NATIONALGEOGRAPHIC.COM



Source: J. Koenderink

Grouping cues: Similarity (color, texture, proximity)

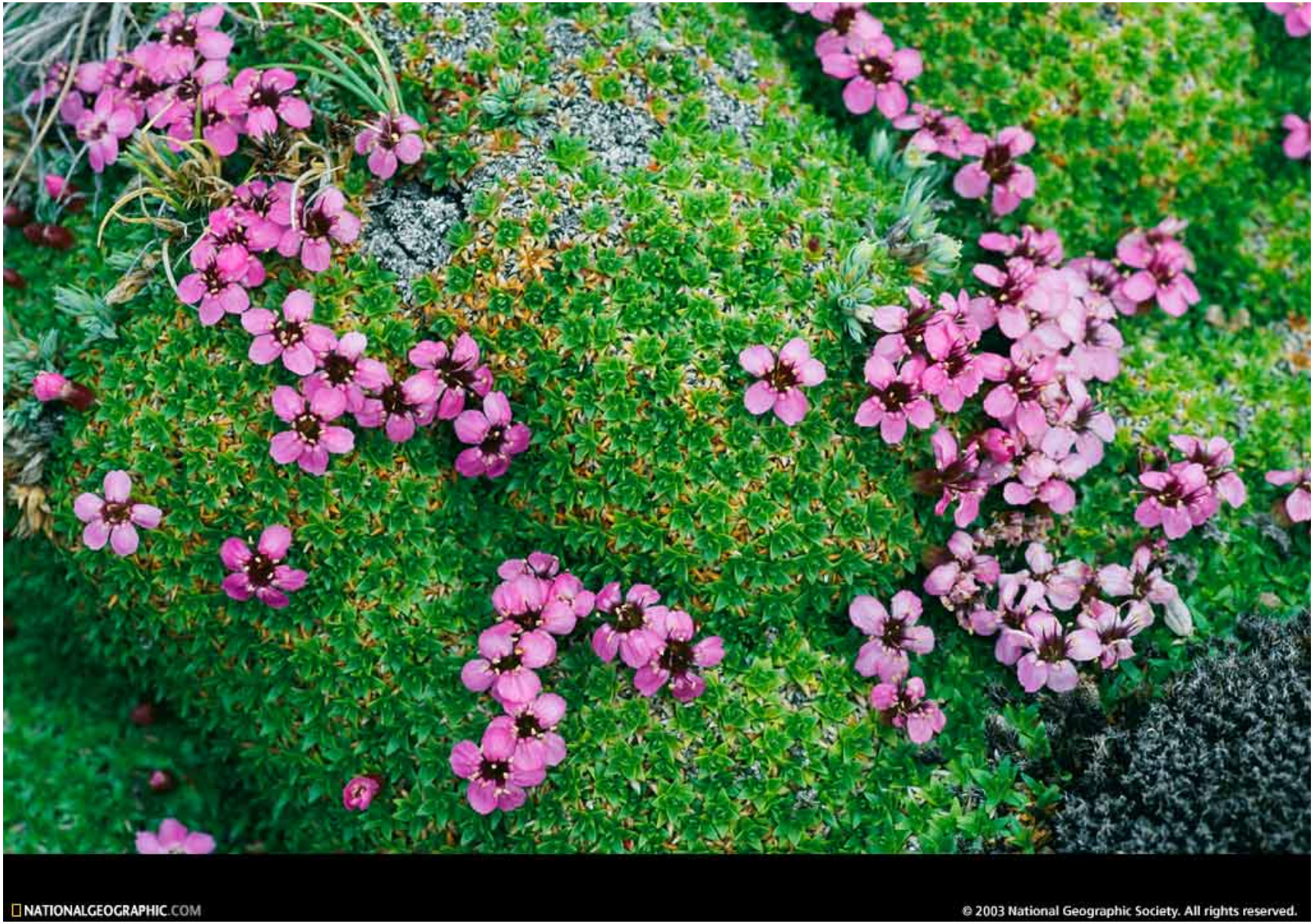
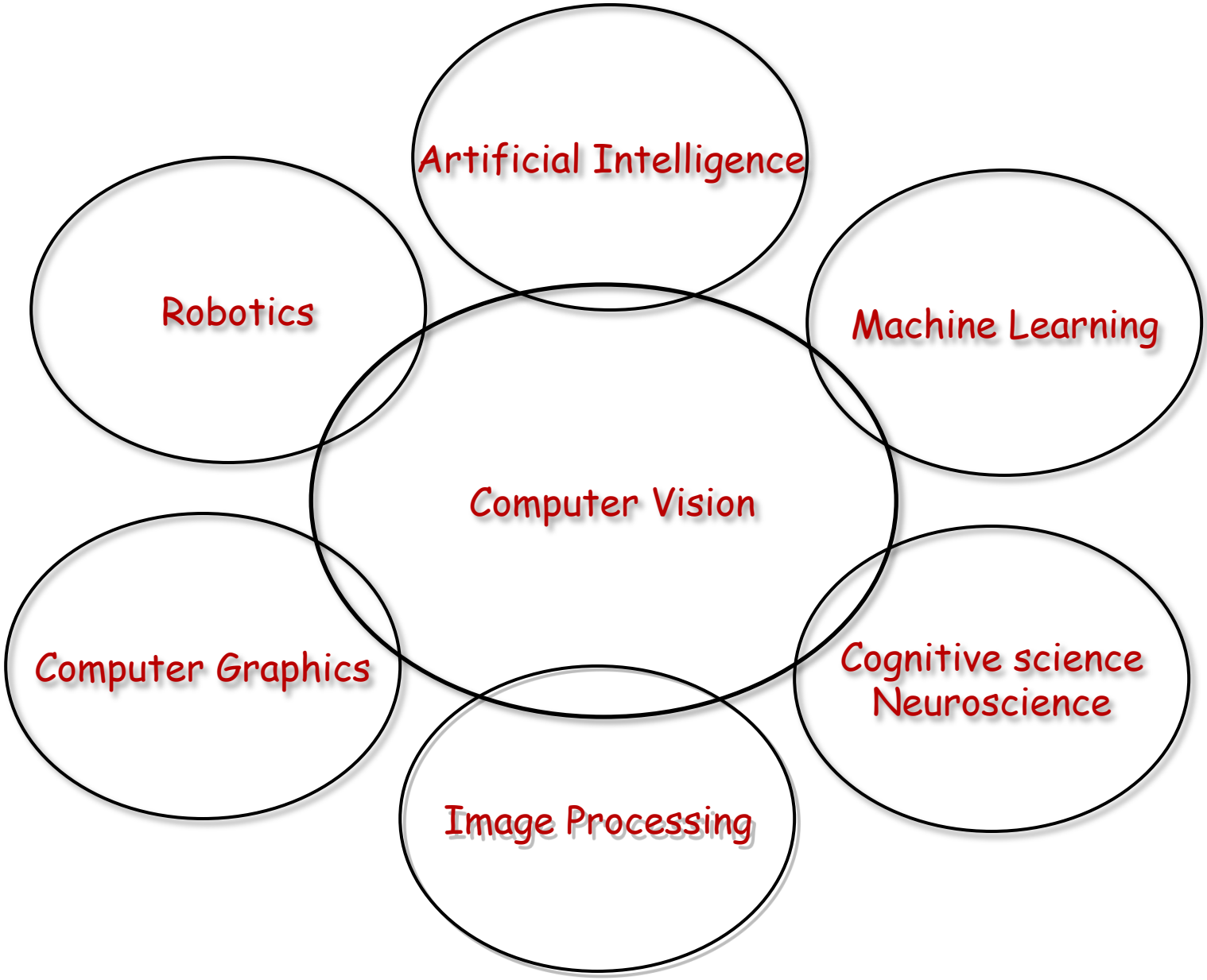


Image credit: Arthus-Bertrand (via F. Durand)

Figure/Ground Segmentation

Connections to other disciplines



Stereo

Example of stereo pipeline, from raw data, preprocessing, meshes, texture maps

See http://schwehr.org/photoRealVR/example.html

Structure From Motion

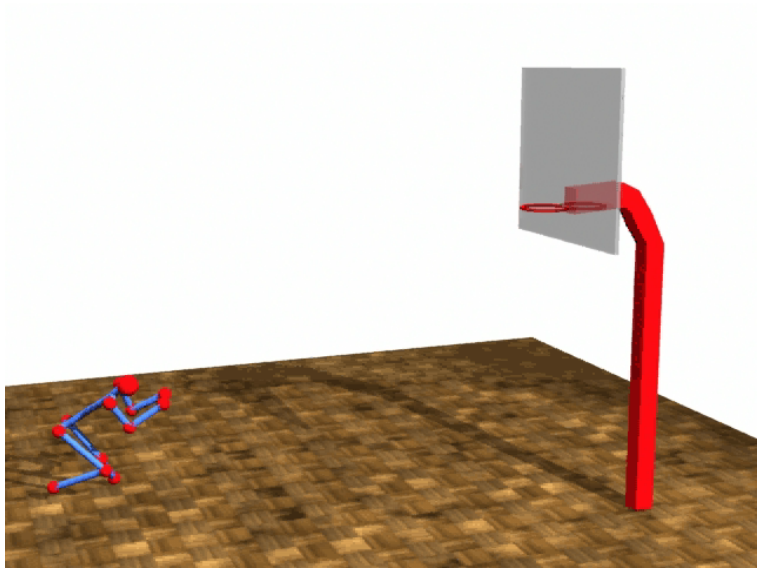
http://www.cs.unc.edu/Research/urbanscape

Structure From Motion

http://www.cs.unc.edu/Research/urbanscape

http://www.cs.unc.edu/Research/urbanscape

3D Modeling



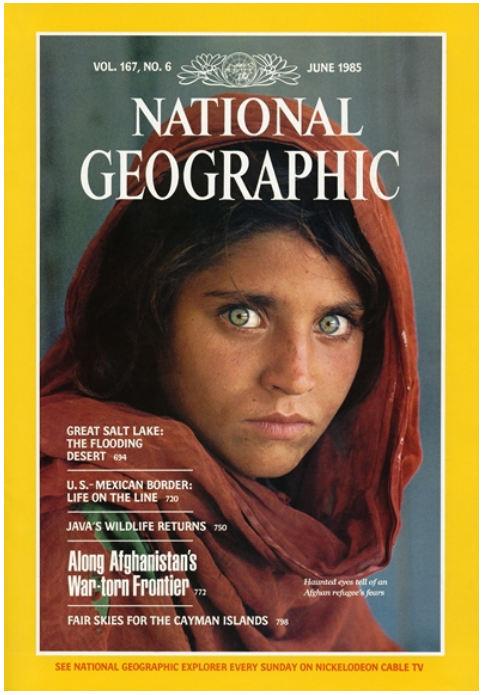
http://www.photogrammetry.ethz.ch/research/cause/3dreconstruction3.html

Special effects: shape and motion capture

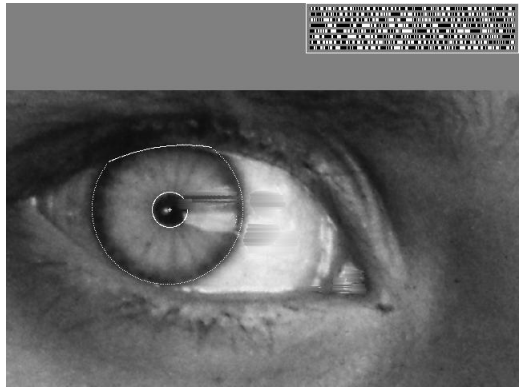
Source: S. Seitz

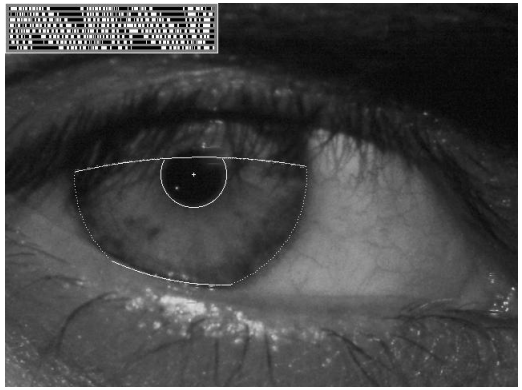
Face recognition: Apple iPhoto software

http://www.apple.com/ilife/iphoto/



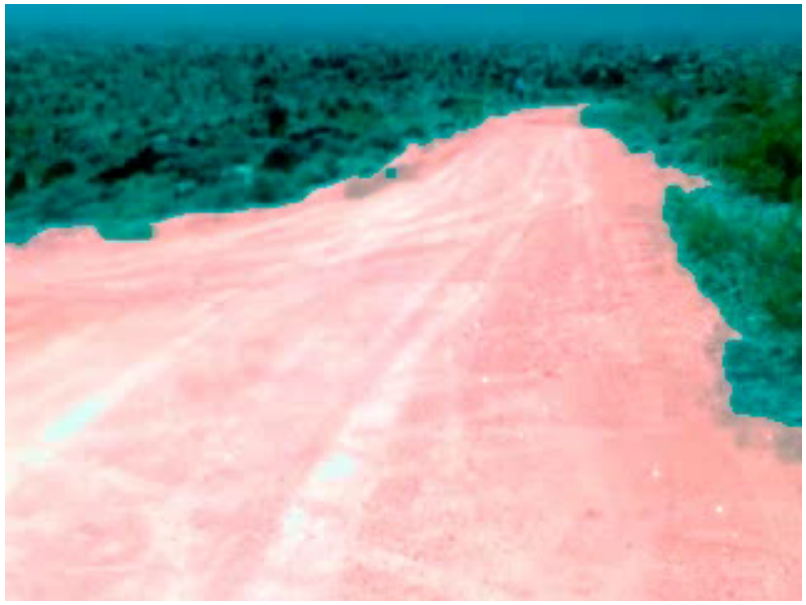
How the Afghan Girl was Identified by Her Iris Patterns





Source: S. Seitz

Classification



- Given video from a car camera recognize where drivable surfaces

Biometrics

Fingerprint scanners on many new laptops, other devices

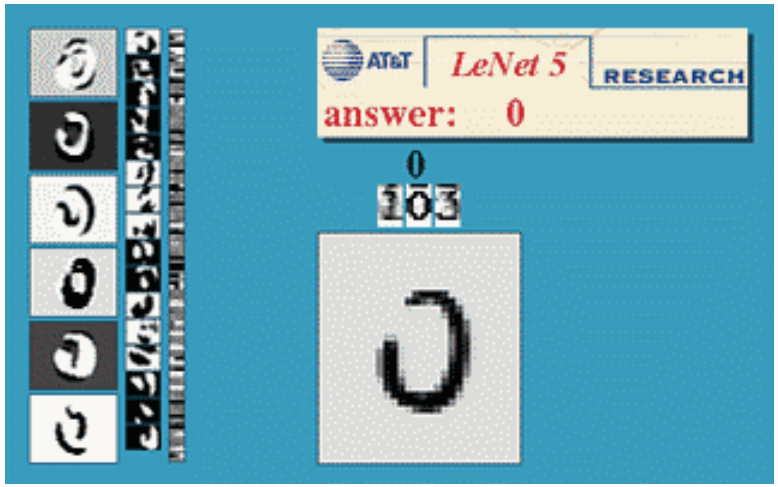
Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Source: S. Seitz

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



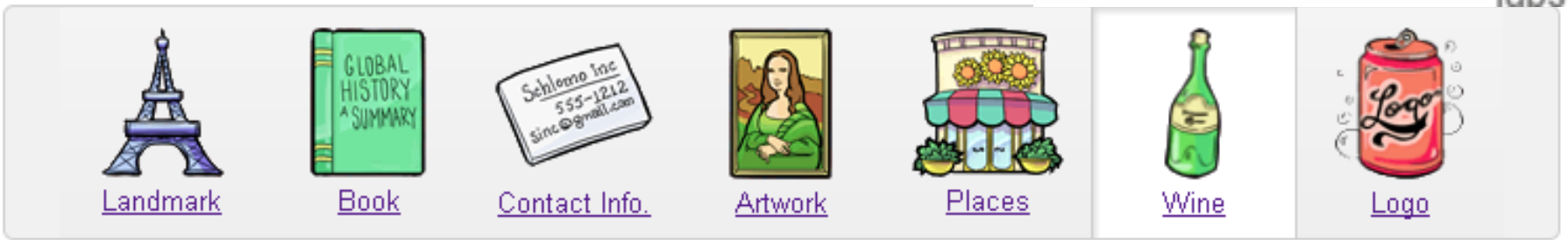
Digit recognition, AT&T labs

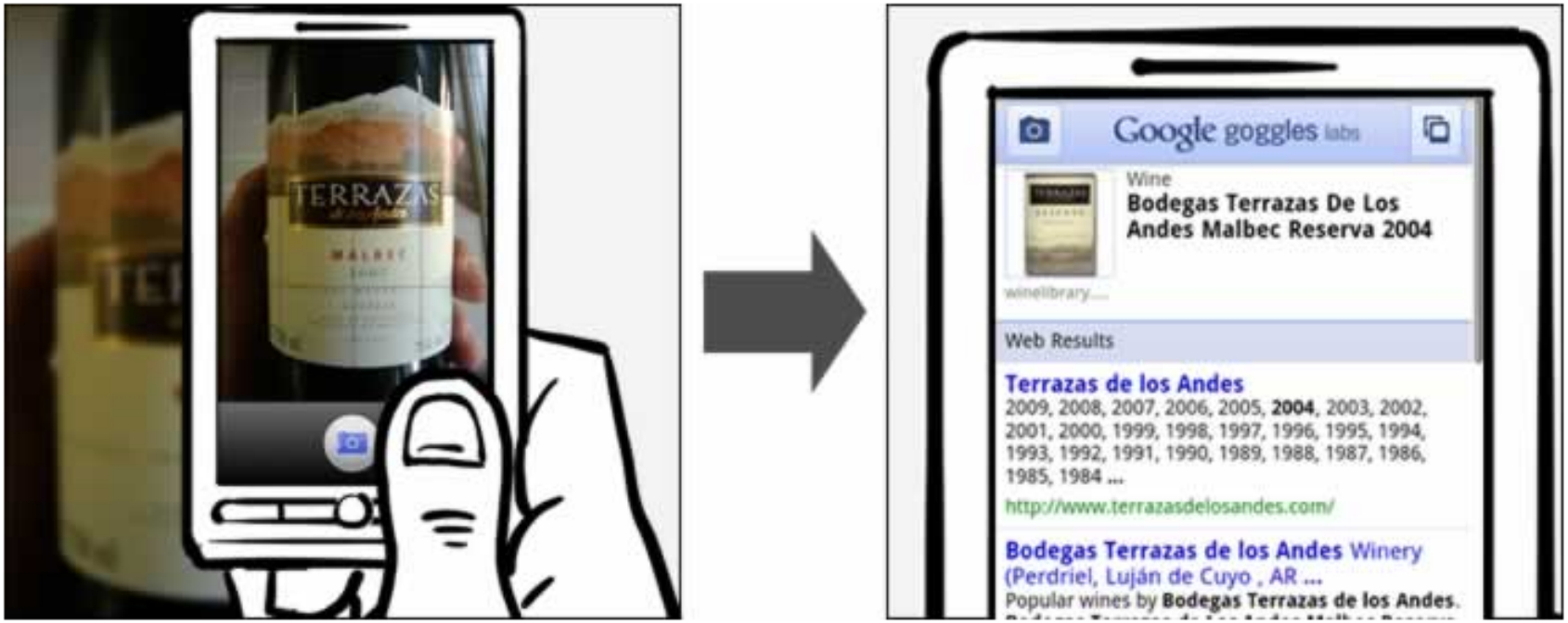
License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Source: S. Seitz

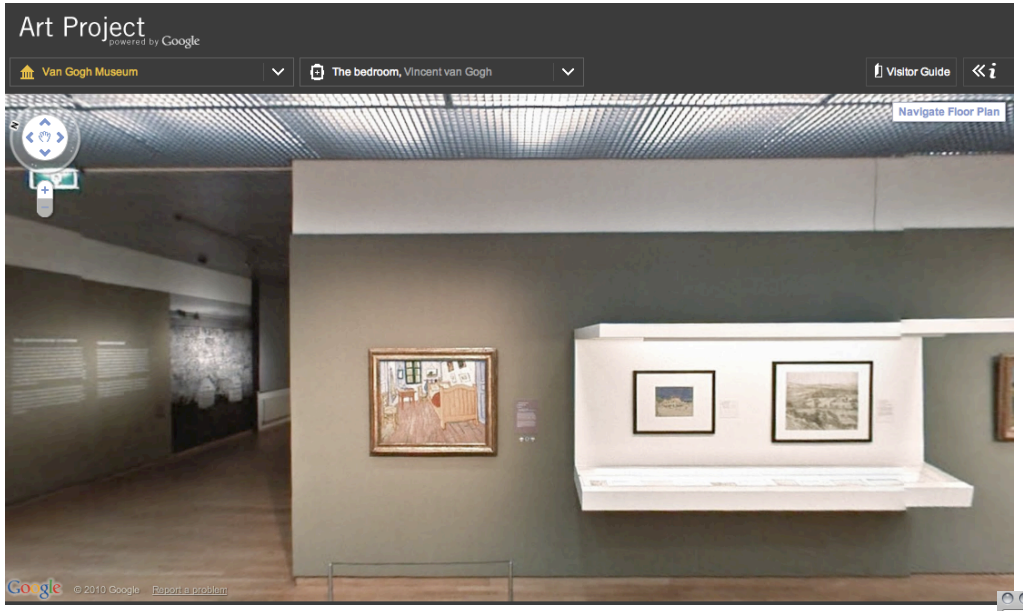
Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

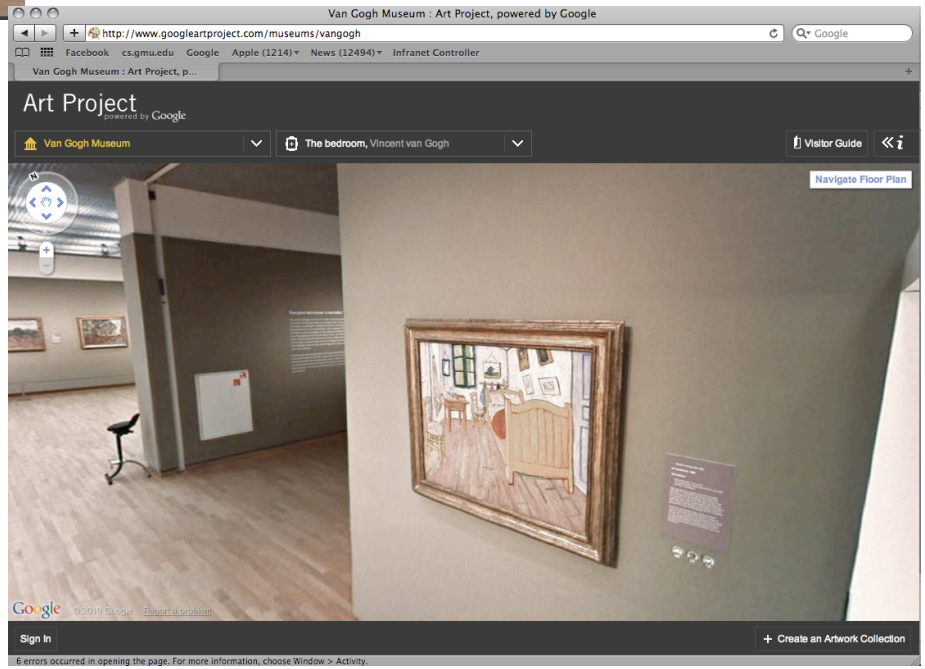




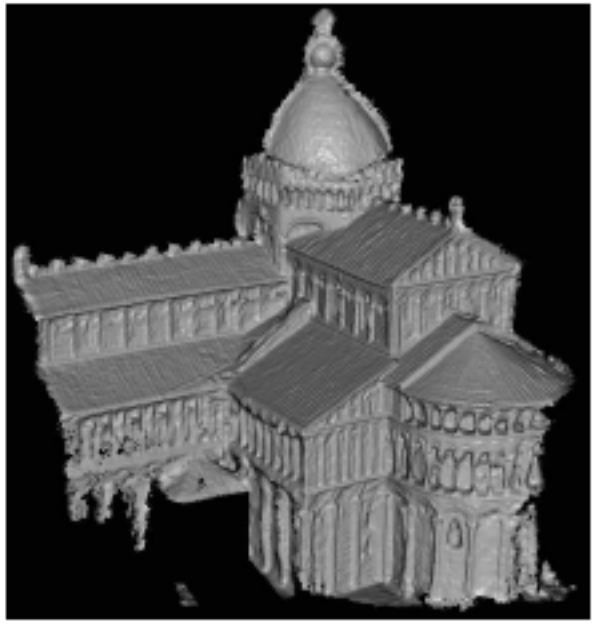
Google Art Museum Project



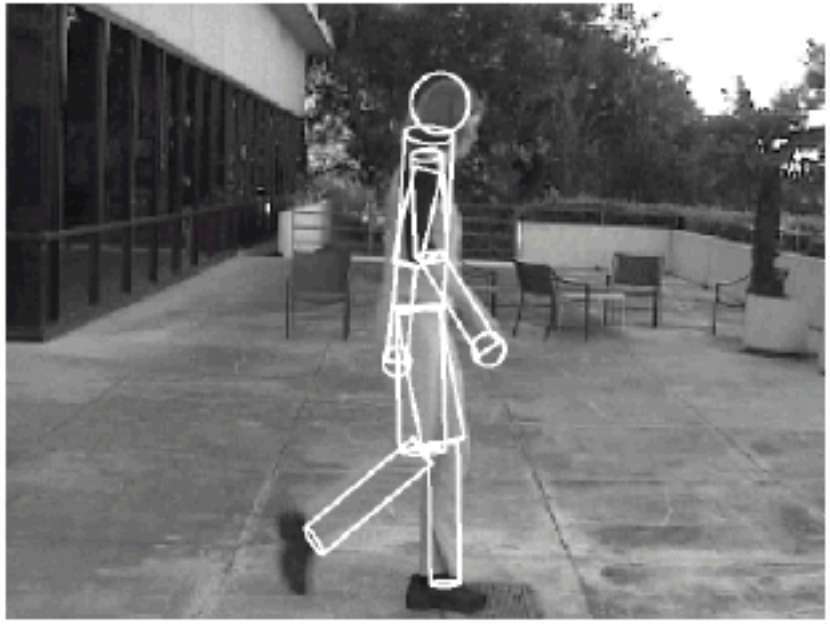
Navigate museums of the world



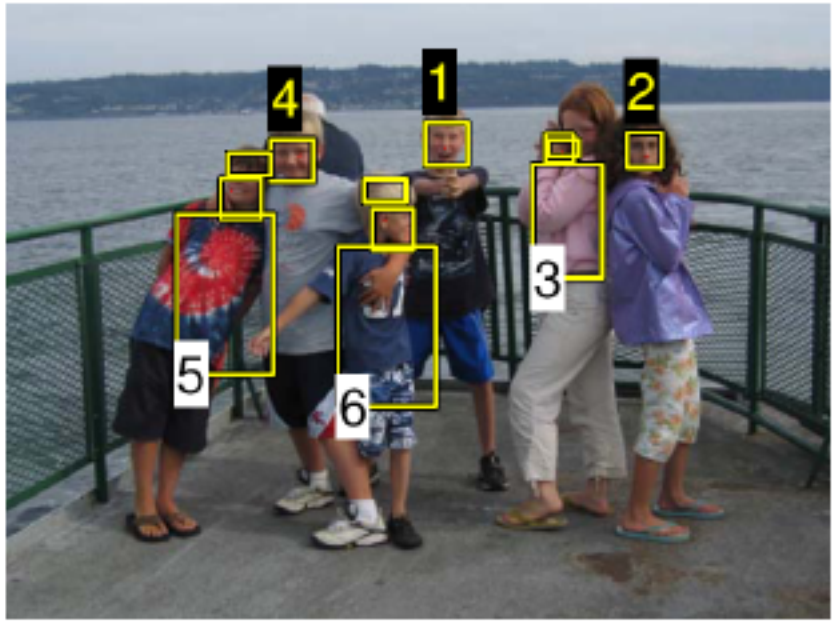
(a)



(b)



(c)



(d)

Automotive safety

- Mobileye: Vision systems in high-end BMW, GM, Volvo models
  - "In mid 2010 Mobileye will launch a world's first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians."

Source: A. Shashua, S. Seitz

Vision in supermarkets

LaneHawk by EvolutionRobotics

"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… "

Source: S. Seitz

Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

Sony EyeToy

Assistive technologies

Xbox and Kinect sensor

Source: S. Seitz

Computer Vision Stages

- Early vision – local operations, compute maps, or statistics of individual pixels (edges, motion fields, depth maps)
- Midlevel vision – assembly of local information (segmentation, contour completions, grouping)
- Scene analysis – recognition of objects, scenes
- Active vision – how to control and use the resources to adjust the sensor to gather additional information
- Goal directed vision – control behaviors based on visual information
- Deep learning Techniques – categorization tasks, biologically inspired methods

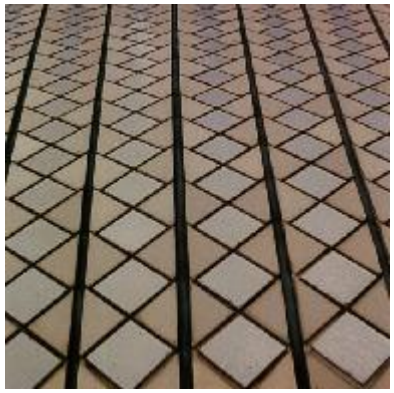
Course Overview

1. Image formation and Image Processing
   (low level vision)
2. Features, matching, correspondences
3. 3D reconstruction
4. Grouping, segmentation
5. Object Detection and Recognition

1. Geometric aspects

Shape and Motion Recovery, Matching, Alignment Problems
Reconstruction (from 2D to 3D)

Computation of Pictorial cues – shading, texture, blur, contour, stereo, motion cues



2. Object Detection and Recognition

Object representation, detection in cluttered scenes
Recognition of object categories
Scene recognition and understanding

1. Geometry of Single and Multiple Views, Video

How to reliably recover and represent the geometric model from single image or video and camera motion/pose

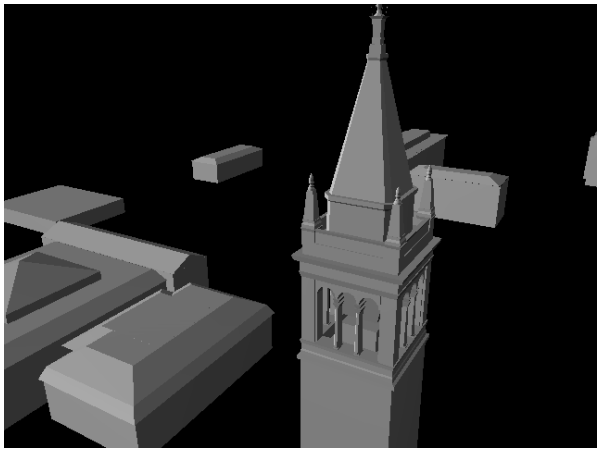
Representation issues depends on the task/applications

- Image-based rendering, Computer Graphics
- Virtual and Augmented Reality
- Vision based control, surveillance, target tracking
- Human computer interaction

- Medical imaging (alignment, monitoring of change)
- Video Analysis

Vision and Computer Graphics

- image based rendering techniques
- 3D reconstruction from multiple views or video
- single view modeling
- view morphing (static and dynamic case)



Modeling with Multiple Images

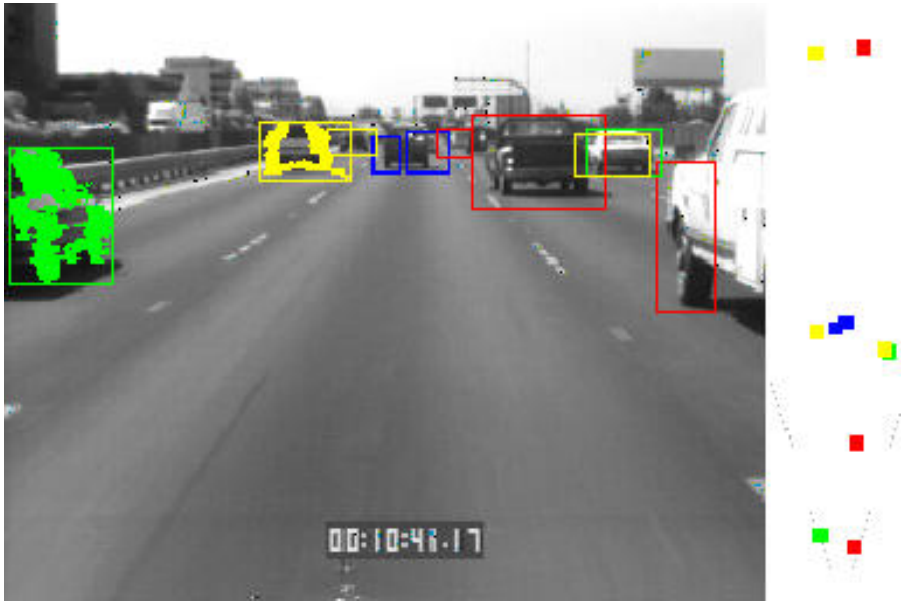
University High School, Urbana, Illinois
Three of Twelve Images, courtesy Paul Debevec

Final Model

Visual surveillance

wide area surveillance, traffic monitoring

Interpretation of different activities



Virtual and Augmented Reality, Human computer Interaction

Virtual object insertion
various gesture based interfaces
Interpretation of human activities
Enabling technologies of intelligent homes, smart spaces

Kinect Fusion https://www.youtube.com/watch?v=q8jfgTilFmo

Surveillance

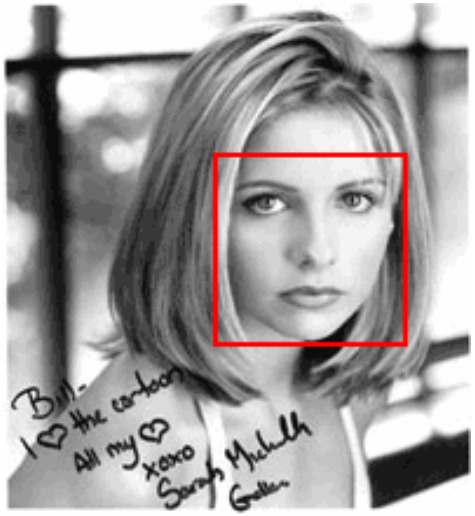
Image Morphing, Mosaicing, Alignment

Images of CSL, UIUC

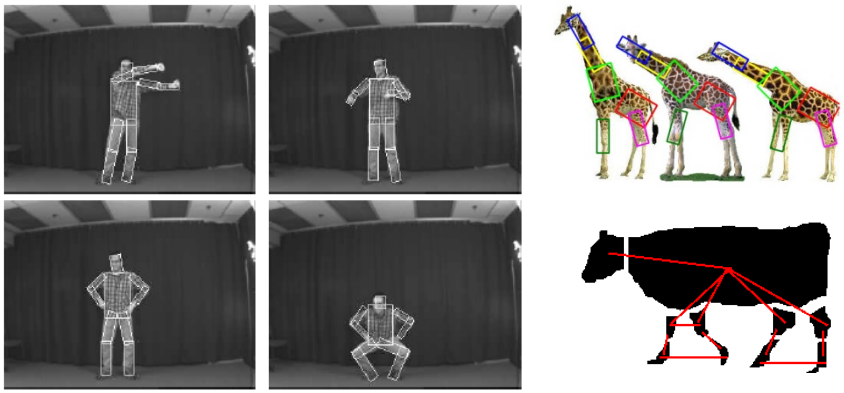
V. Advanced Topics

- Time permitting…

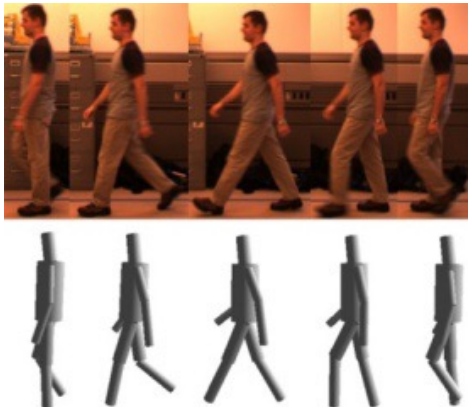
Segmentation



Face detection



Articulated models



Motion and tracking

3D RECONSTRUCTION FROM MULTIPLE VIEWS: GEOMETRY AND PHOTOMETRY

jin-soatto-yezzi; image courtesy: j-y bouguet - intel

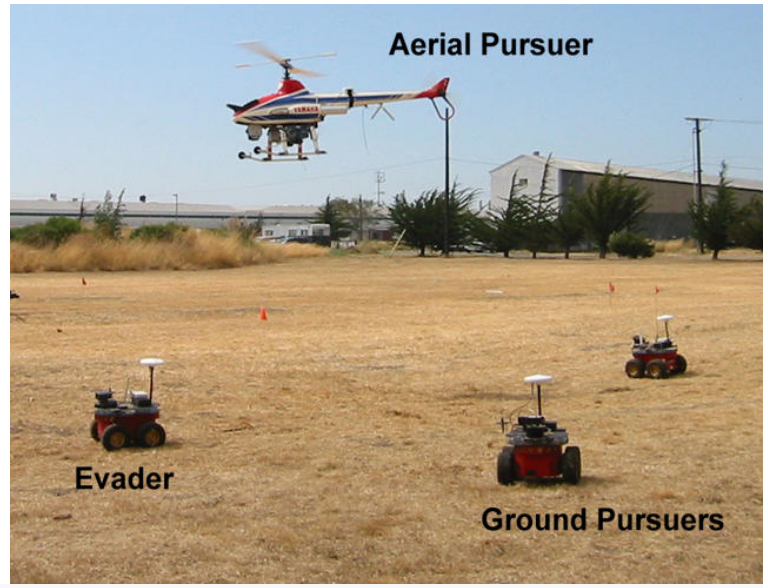
estimated shape

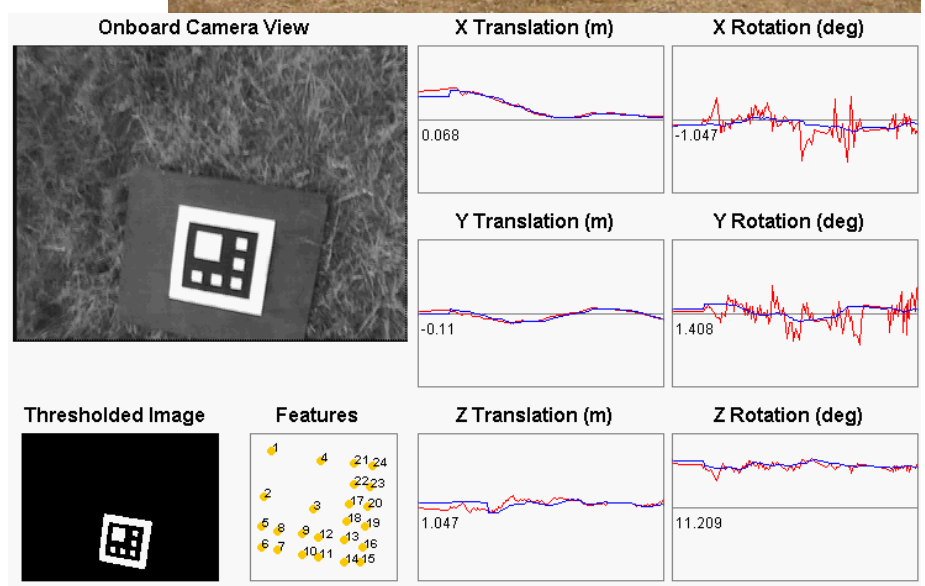
laser-scanned,
manually polished

jin-soatto-yezzi


Vincent
van Gogh

Unmanned Aerial Vehicles (UAVs)

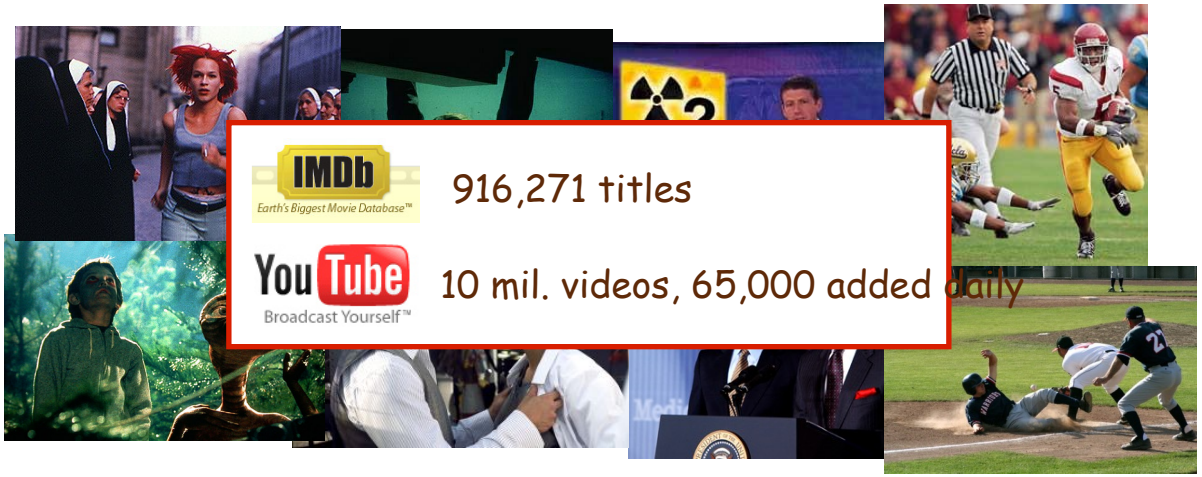


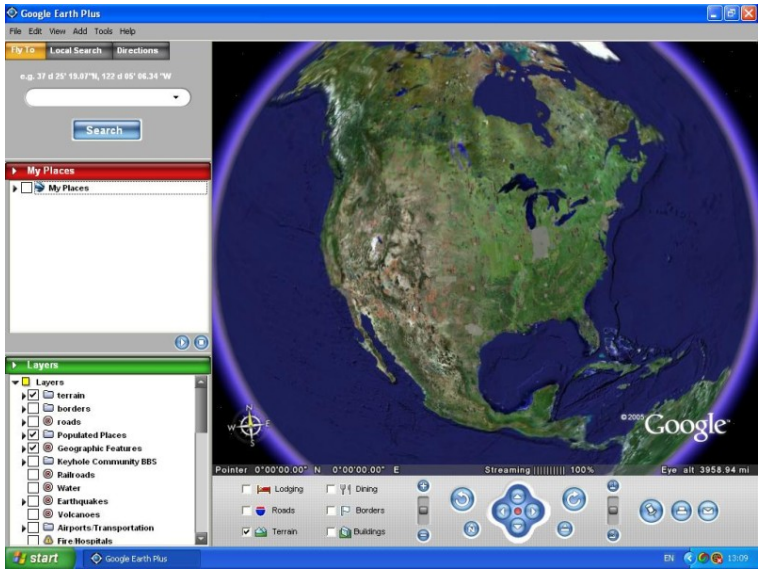


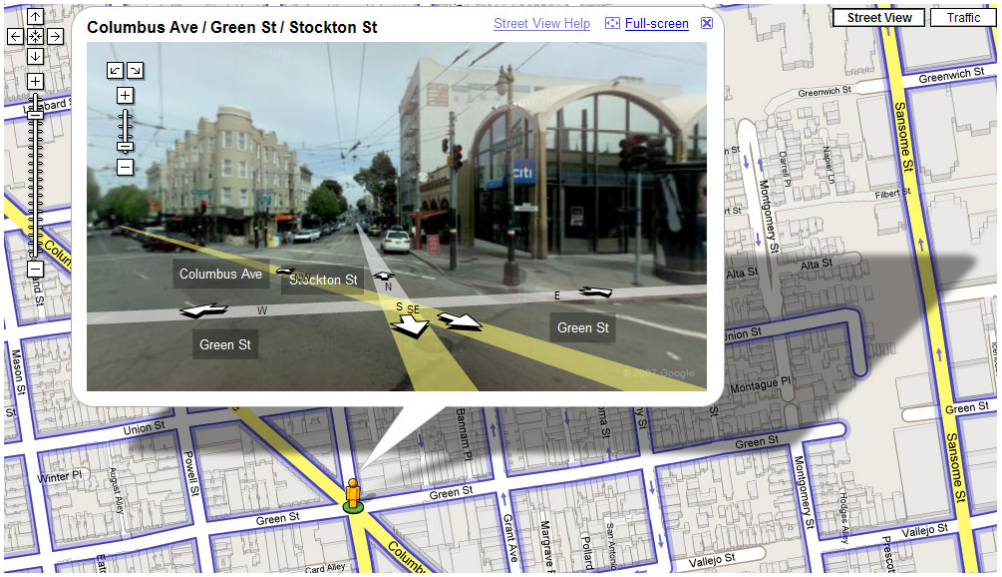
Rate: 10Hz
Accuracy: 5cm, 4°

Berkeley Aerial Robot (BEAR) Project

Computer Vision and the Web

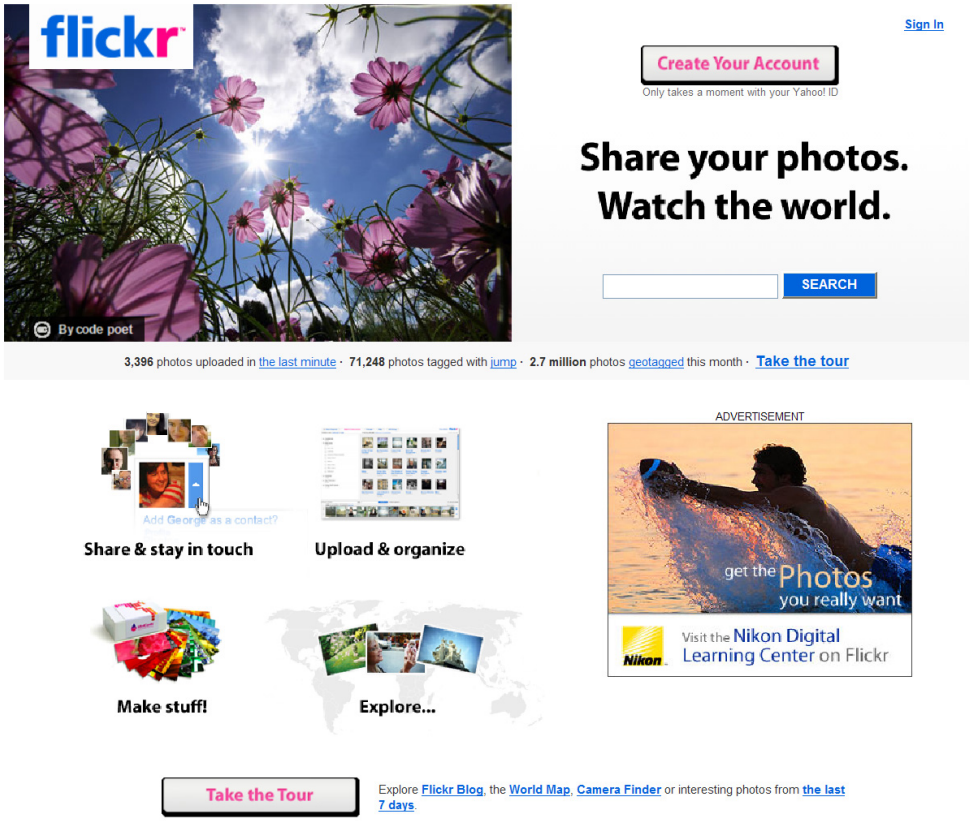






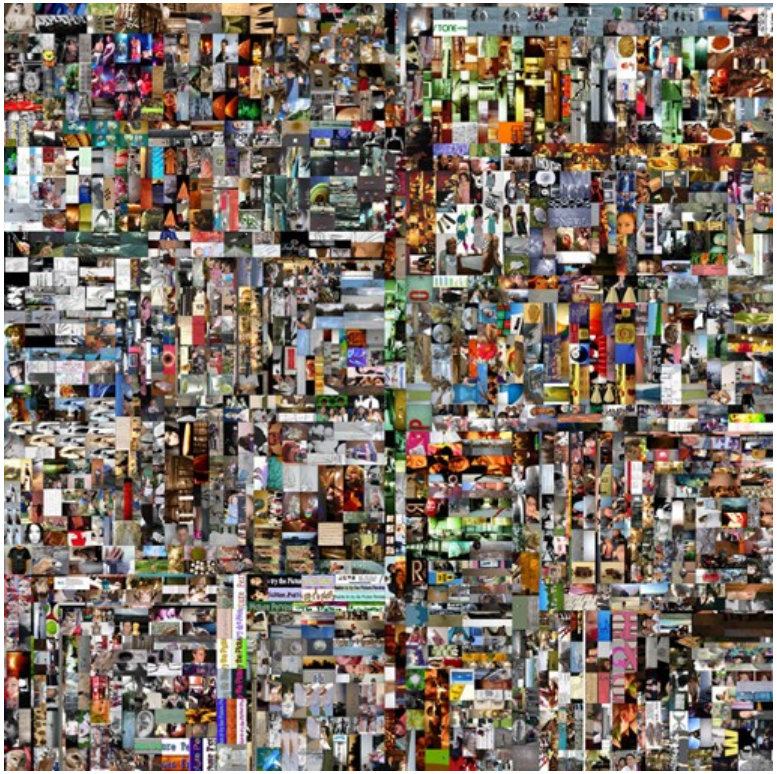
Exploring online photo collections: Flickr.com

- Flickrvision
- Flickr tag browser
- Flickr time graphs



Artistic Visualization

- Jason Salavon
  - 100 Special Moments
  - Homes for Sale
  - Every Playboy centerfold



- To find anything on the Web, you must search with words!

- But what if the question is a picture?

Visual sentiment analysis

- Attractiveness: Hot or Not (research project)

- Cuteness: Cute overload

- Interestingness: Flickr

Exploring photo collections in 3D

- Photo tourism
  Snavely, Seitz and Szeliski (SIGGRAPH 2006)

- Automatic photo pop-up
  Hoiem and Efros (SIGGRAPH 2005)