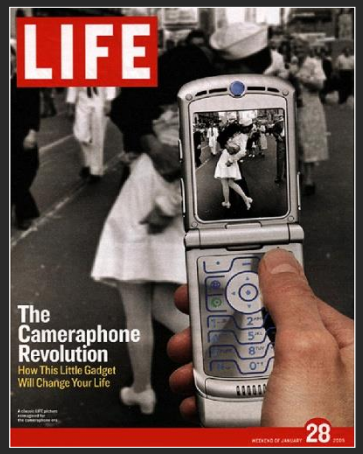
ELEC 345: Computer Vision

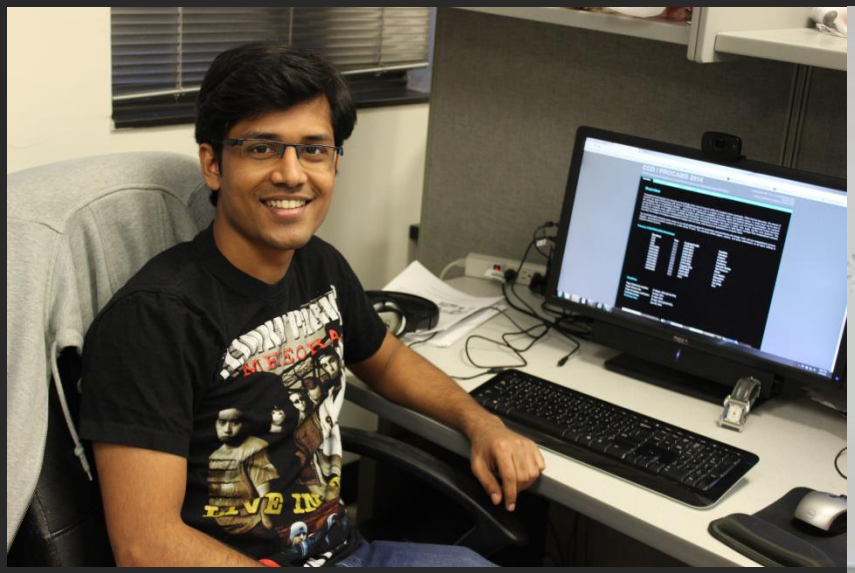


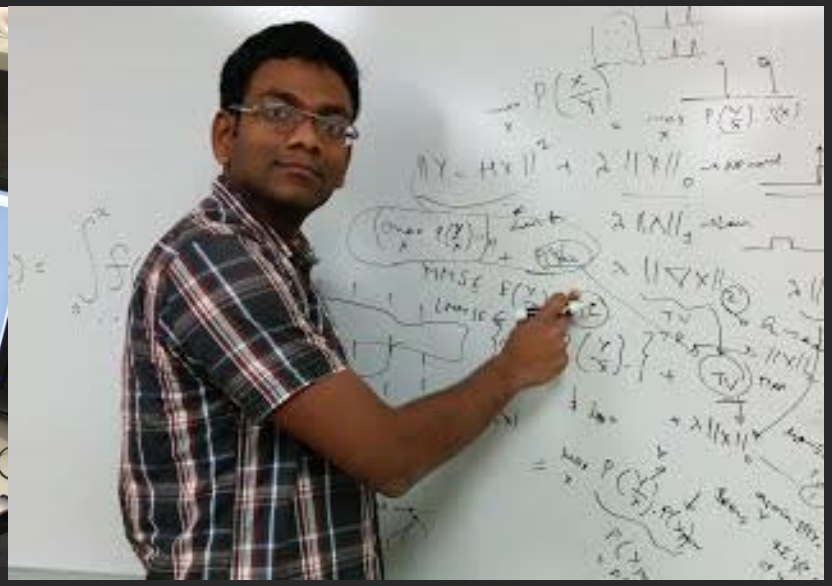
ELEC 345

Introduction to Computer Vision and its applications

- Course catalog number: ELEC 345 (Spring 2017)
- Instructor: Prof. Ashok Veeraraghavan (vashok (at) rice.edu)
  DH 3030
- Office Hours: Wednesday 9:00-10:00 AM
- Meeting time: Tuesday 4:00 – 5:15 PM
  Thursday 4:00 – 5:15 PM
- Meeting place: HRZ 210

Teaching Assistants





Vivek Boominathan (vb10@rice.edu)

Adithya Pediredla (akp4@rice.edu)

Recitation Sessions

Mondays or Wednesdays

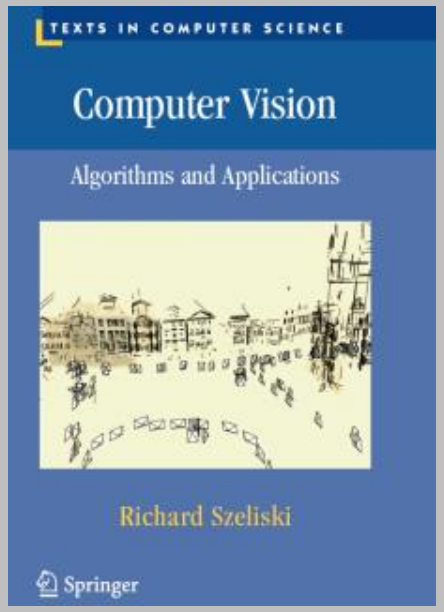
5:00-6:00 PM

Venue: TBD

Course Info

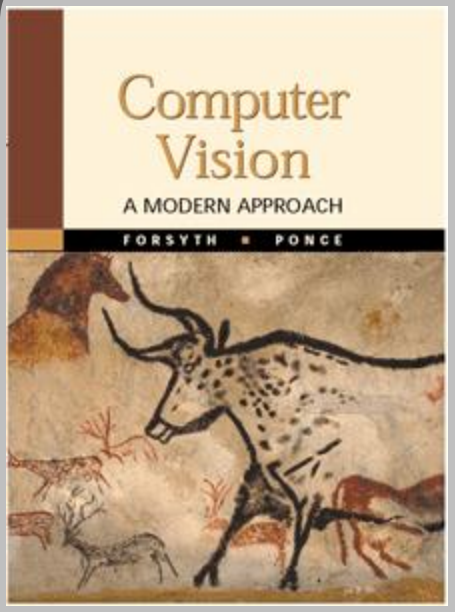
- Course assumes prior
  - Programming experience
  - Experience with basic Linear algebra
  - Introductory knowledge of Vector calculus
  - Creativity and enthusiasm
- Course also requires that you have some
  - Matlab experience
  - Interest in imaging or photography
  - Hard work ---- this course is not going to be easy – We are going to build some real working computational photography systems and algorithms as projects. Be prepared to work hard and have loads of fun
- Do you own a camera? Do you have access to a camera?

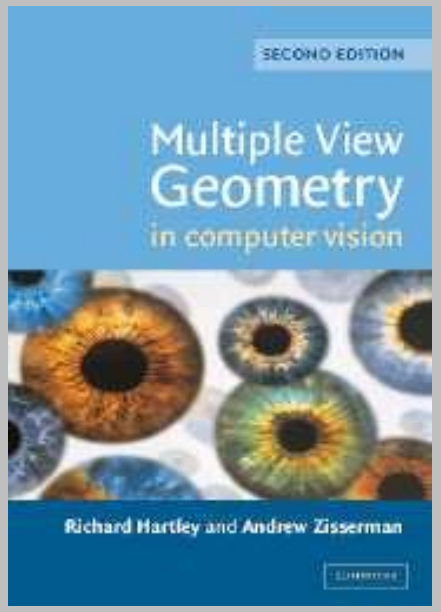
Books

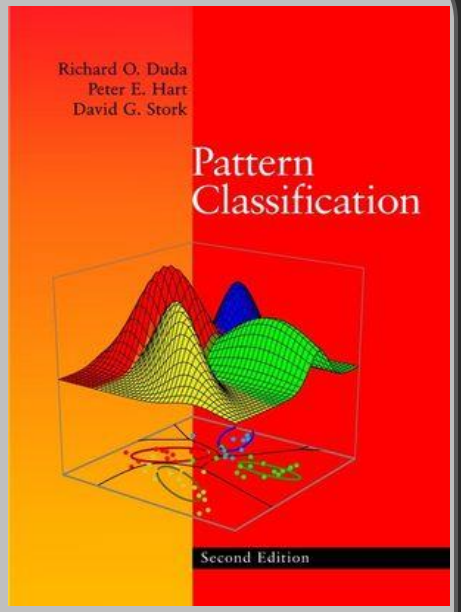


Textbook

Free to download







Reference Books

Selected chapters are recommended reading

Grading

Grading (subject to change)

Take Home Assignments	-	65%
Mid Term (take home)	-	15%
Final Exam (in class)	-	5%
Final Exam (take home)	-	15%

All Assignments and Exams will have 3 Parts

Part A - Straightforward from Class Notes

Part B - Creativity and Extensions to Concepts Reqd.

Part C – Gradaute Course Credit

Quick Round of Introductions

BRING YOUR OWN NAME TAG

Today

- Course Syllabus
- Assignments and Examinations
- Course Rules, Policies and Procedures
- Introduction to computer vision, animals and illusion

Course Syllabus

http://www.youtube.com/watch?v=Vxq9yj2pVWk

Course Syllabus

		Introduction		
01		Introduction to Computer Vision and Sample Applications. Animal Eyes, Perception and Illusions		
02		Linear filters and Edge detection Assignment0 release		
03		Feature extraction (Harris + SIFT)		
04		Feature Extraction 2		
05		Model Fitting and RANSAC Assignment1 release		
06		Alignment		
07		Photometric tools (Gradient domain processing, Laplacian etc)		
08		REVIEW AND DISCUSSION Assignment2 release		

Course Syllabus

		Image Formation		
09		Cameras, Projection		
10		Projective Geometry and meterology		
11		Computational Photography Assignment3 release		
12		REVIEW AND DISCUSSION		
13		Camera Calibration		
		Spring Break (No classes)		
15		Stereo and Multi-view stereo Assignment5 release		
16		Midterm Exam		
16		Structured Light and Kinect		
17		Structure from Motion		
18		Light, Shading and Color (Material Properties) Assignment6 release		
19		Photometric Stereo and Shape from Shading		

Course Syllabus

Recognition and Pattern Classification/Machine Learning Methods				
20		History and Overview		
		Midterm Recess (No Class)		
21		Recognition and Machine Learning Assignment7 release		
22		Bags of features and part based models		
23		Face + Review and Discussion		
24		Motion (Tracking, Optical Flow)		
25		Photometry: Segmentation		

- Assignments – 65%
  - There will be periodic assignments distributed
  - Assignments will be mix of
    - Reading papers
    - Implementing algorithms
    - Writing a summary of a paper etc.
    - Assignments will progressively get tougher: the first couple of assignments might be easy to get your feet wet.
    - The total grade of assignments will be graded for a total of 65% credit

- Examinations
  - One Midterm exam (Take Home)
  - Final Exam
    - One part taken in class
    - Another part taken at home

Class Rules

- Questions
  - All questions are welcome
  - Interrupt me if you do not understand something
  - I am available in class to help you all learn--- make use of that
- Talk to each other
  - Discuss papers that you like
- Shape the syllabus
  - If there are particular papers or topics you think would be interesting, let me know.
- Laptops, Ipads etc are not permitted.
  - If you want to take class notes on your laptop or IPAD let me know ahead of time and get permission.
  - Otherwise, I expect your concentration in class.

Class Rules

I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.
I do not like late comers in class.

Setting Course Expectations

- HARD
- DEFINITELY TIME CONSUMING
- YOU WILL LEARN CV IF YOU COMPLETE THE COURSE REQUIREMENTS
- I EXPECT AND WILL TREAT ALL OF YOU AS ADULTS
- DEFINITELY TIME CONSUMING COURSE

- Many of the slides used in the course have their origin in various computer vision and related courses in several universities. Thanks to many of the original slide creators and those who modified it subsequently.

Slide Credits:

Steve Seitz	Rick Szeliski
Srinivas Narasimhan	Fei-Fei-Li
Svetlana Lazebnik	Neel Joshi
Antonio Torralba	Rob Fergus
Frank Dellaert	Michael Black

Introduction to computer vision

The goal of computer vision

- To extract “meaning” from pixels

What we see

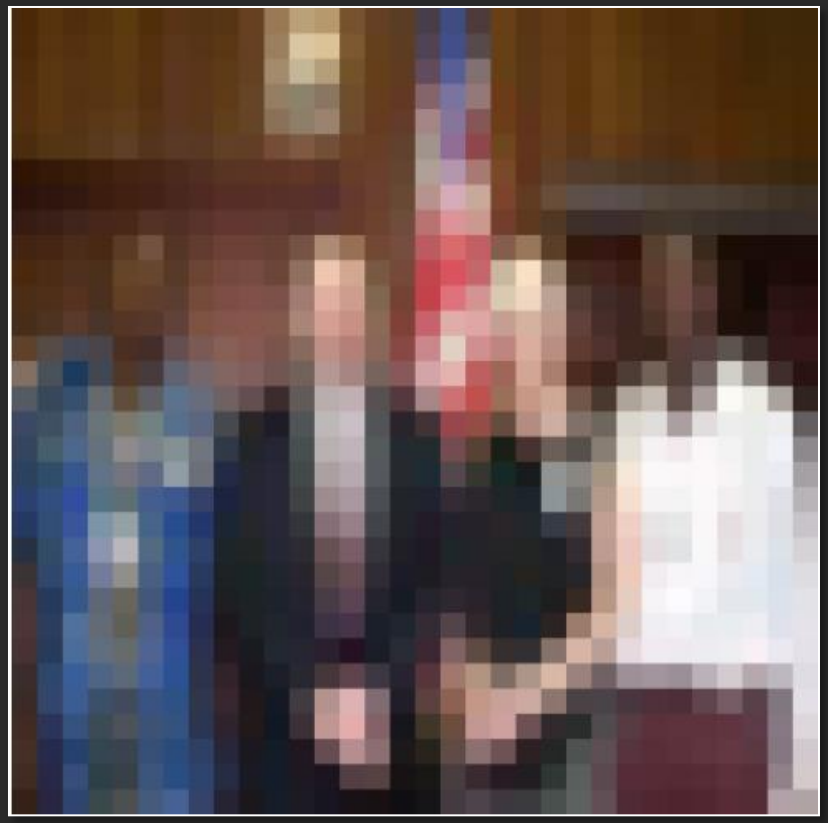
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

The goal of computer vision

- To extract “meaning” from pixels



Humans are remarkably good at this…

Source: “80 million tiny images” by Torralba et al.

The goal of computer vision

- To extract “meaning” from pixels

Humans are remarkably good at this…

Source: Rob Fergus and Antonio Torralba

What kind of information can be extracted from an image?

- Metric 3D information
- Semantic information

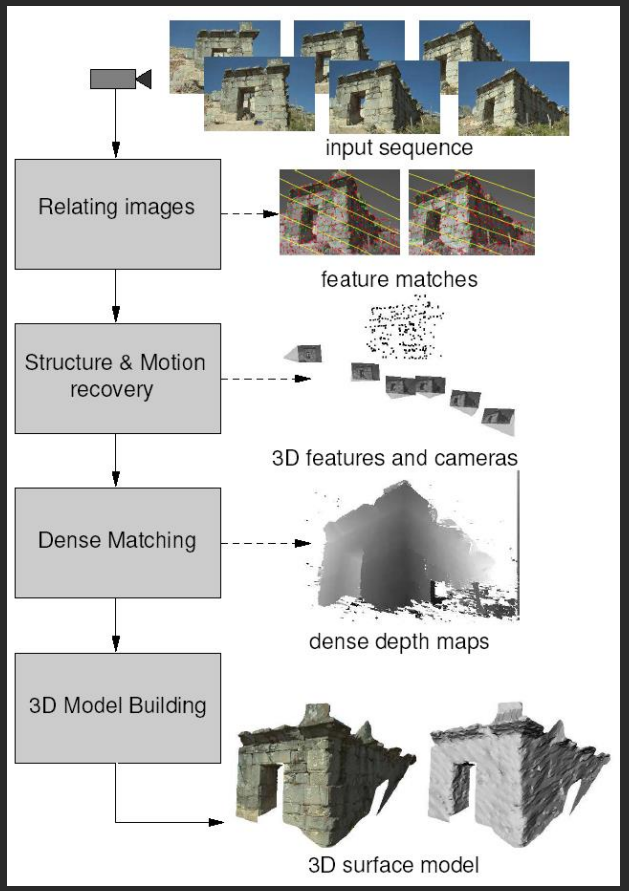
Vision as measurement device

Real-time stereo

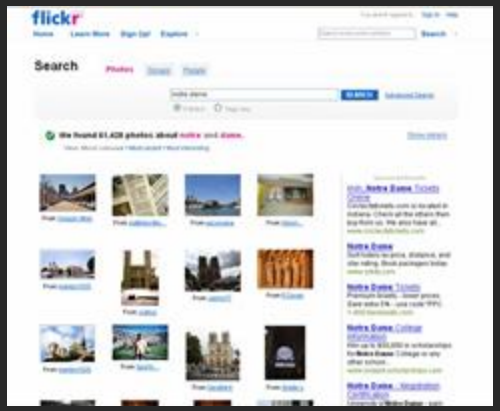
NASA Mars Rover

Pollefeys et al.

Structure from motion



Reconstruction from Internet photo collections



Goesele et al.

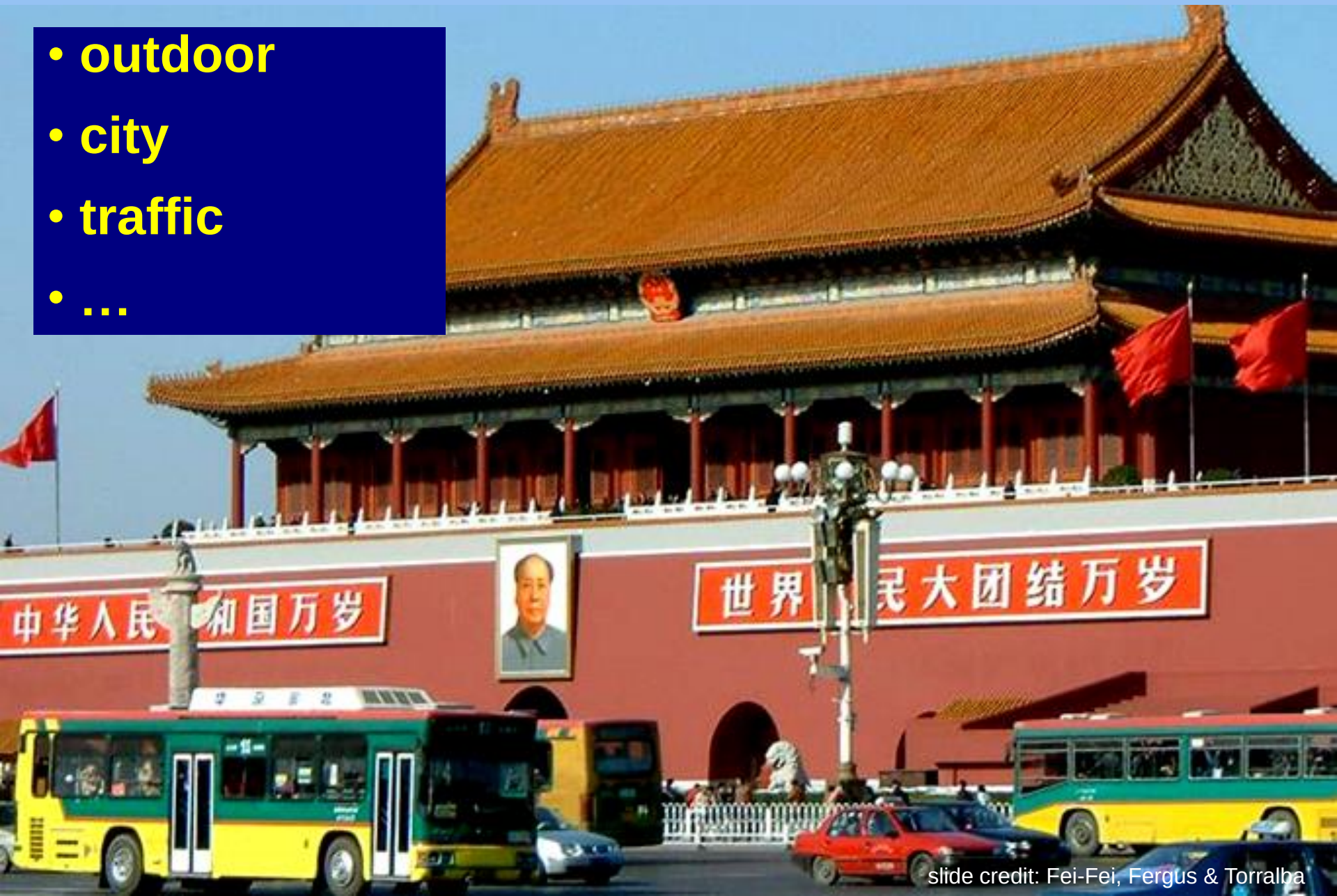
Vision as a source of semantic information

slide credit: Fei-Fei, Fergus & Torralba

Object categorization

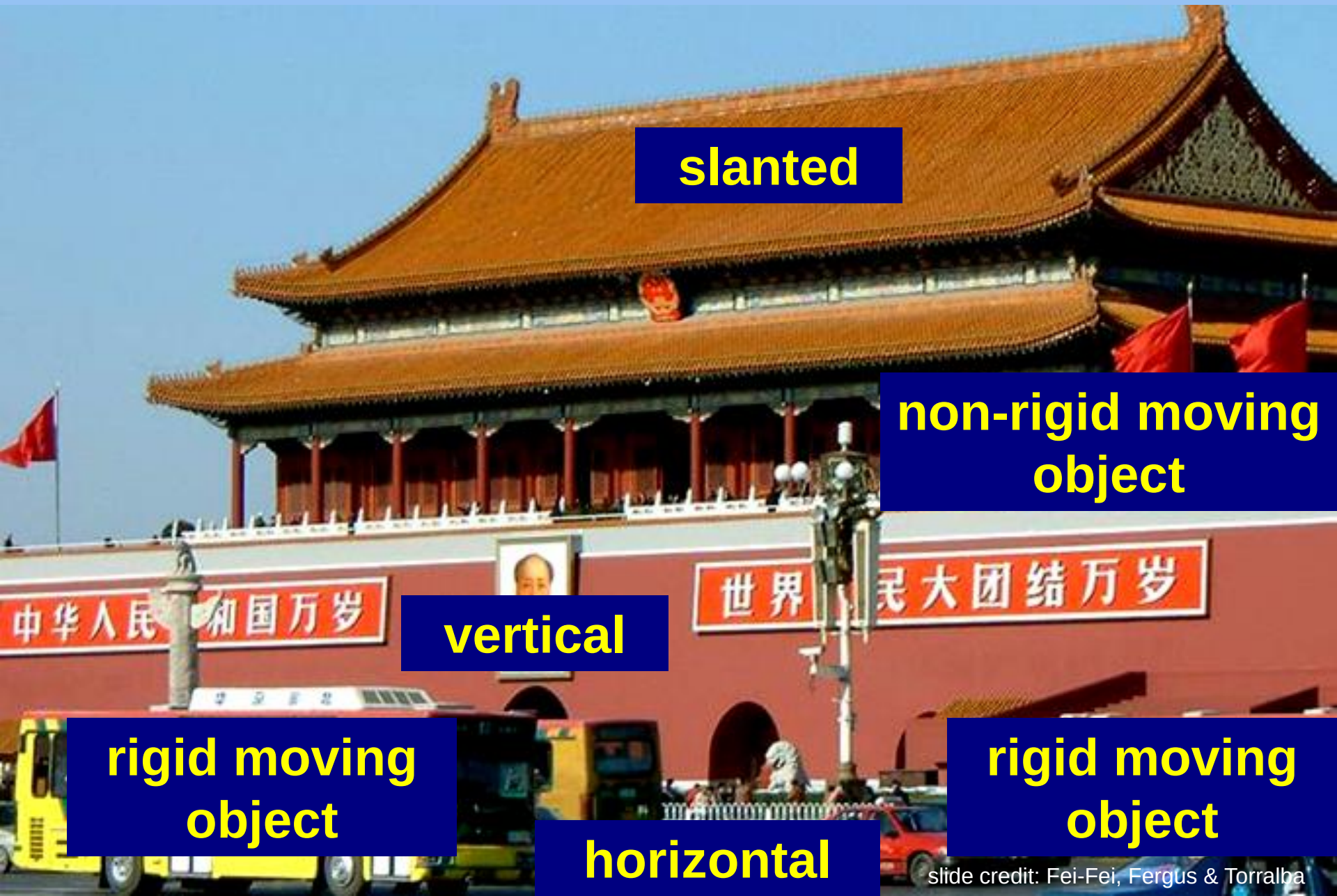
Scene and context categorization

- **outdoor**
- **city**
- **traffic**
- **…**



slide credit: Fei-Fei, Fergus & Torralba

Qualitative spatial information

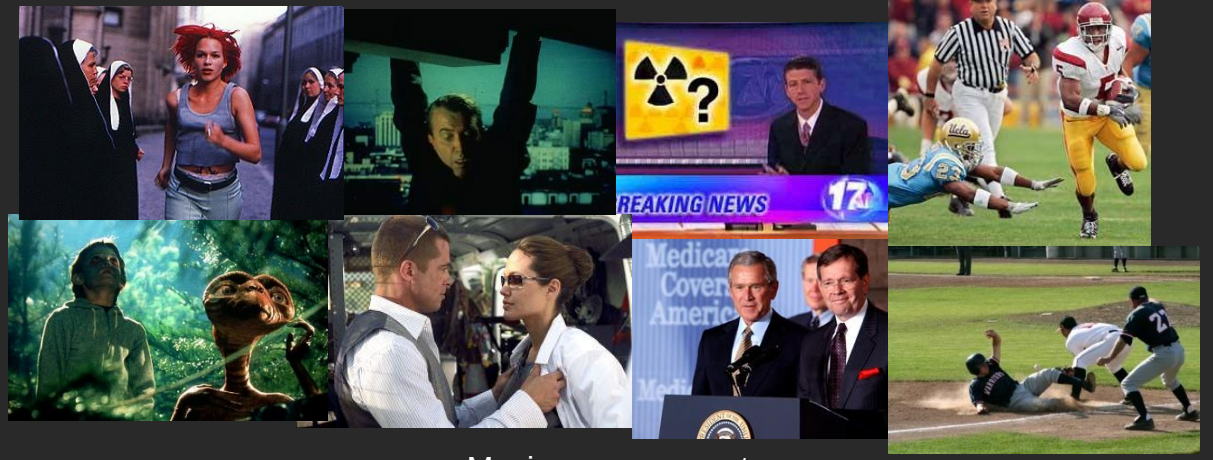


slide credit: Fei-Fei, Fergus & Torralba

Why study computer vision?

- Vision is useful: Images and video are everywhere!

Personal photo albums



Movies, news, sports

Google Image Search

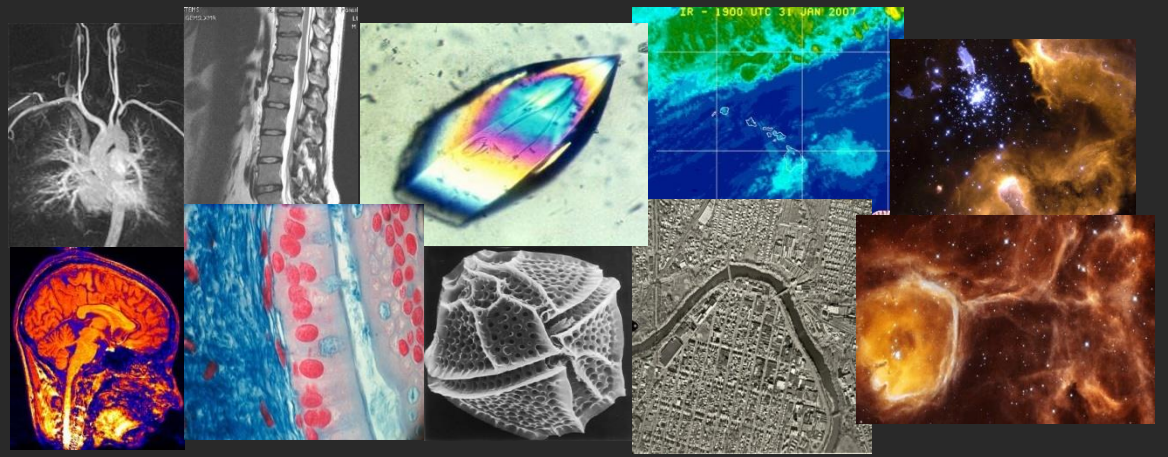
Picasa

flickr

facebook

YouTube Broadcast Yourself

Surveillance and security



Medical and scientific images

Why study computer vision?

- Vision is useful
- Vision is interesting
- Vision is difficult
  - Half of primate cerebral cortex is devoted to visual processing
  - Achieving human-level visual perception is probably “AI-complete”

Why is computer vision difficult?

Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

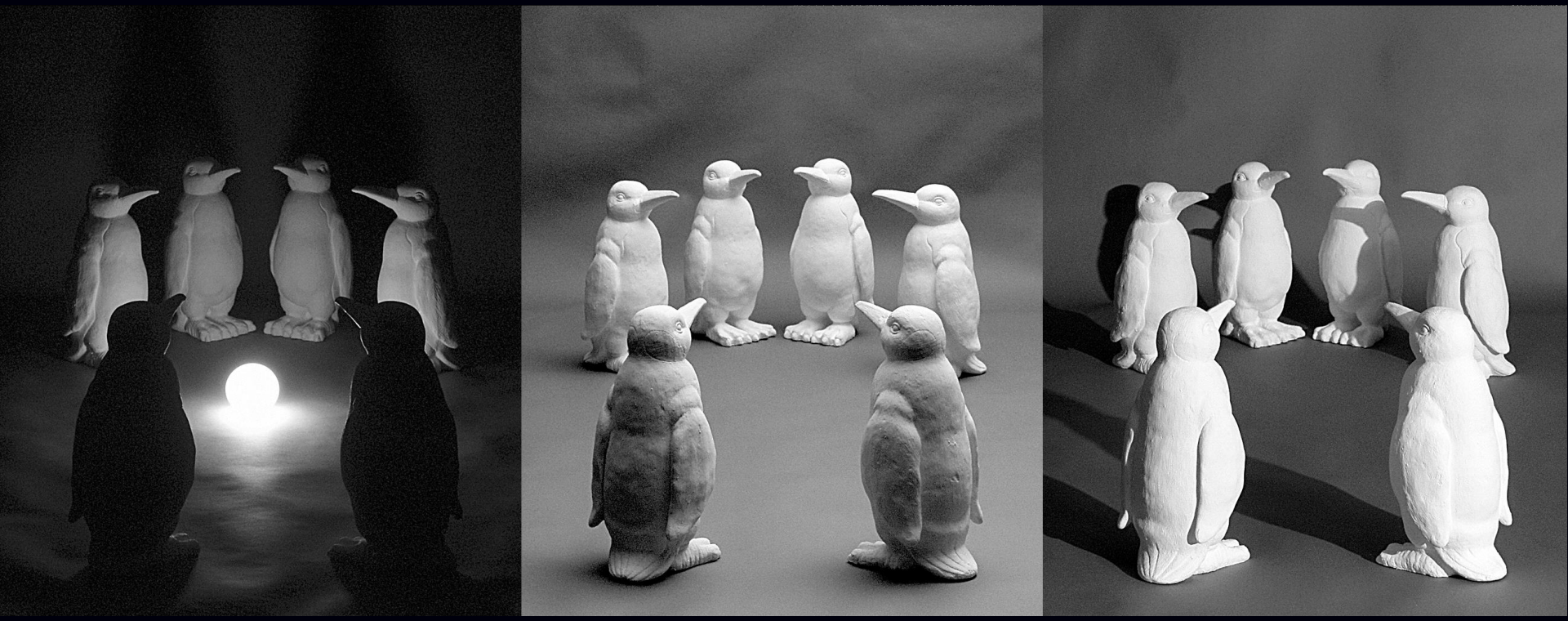
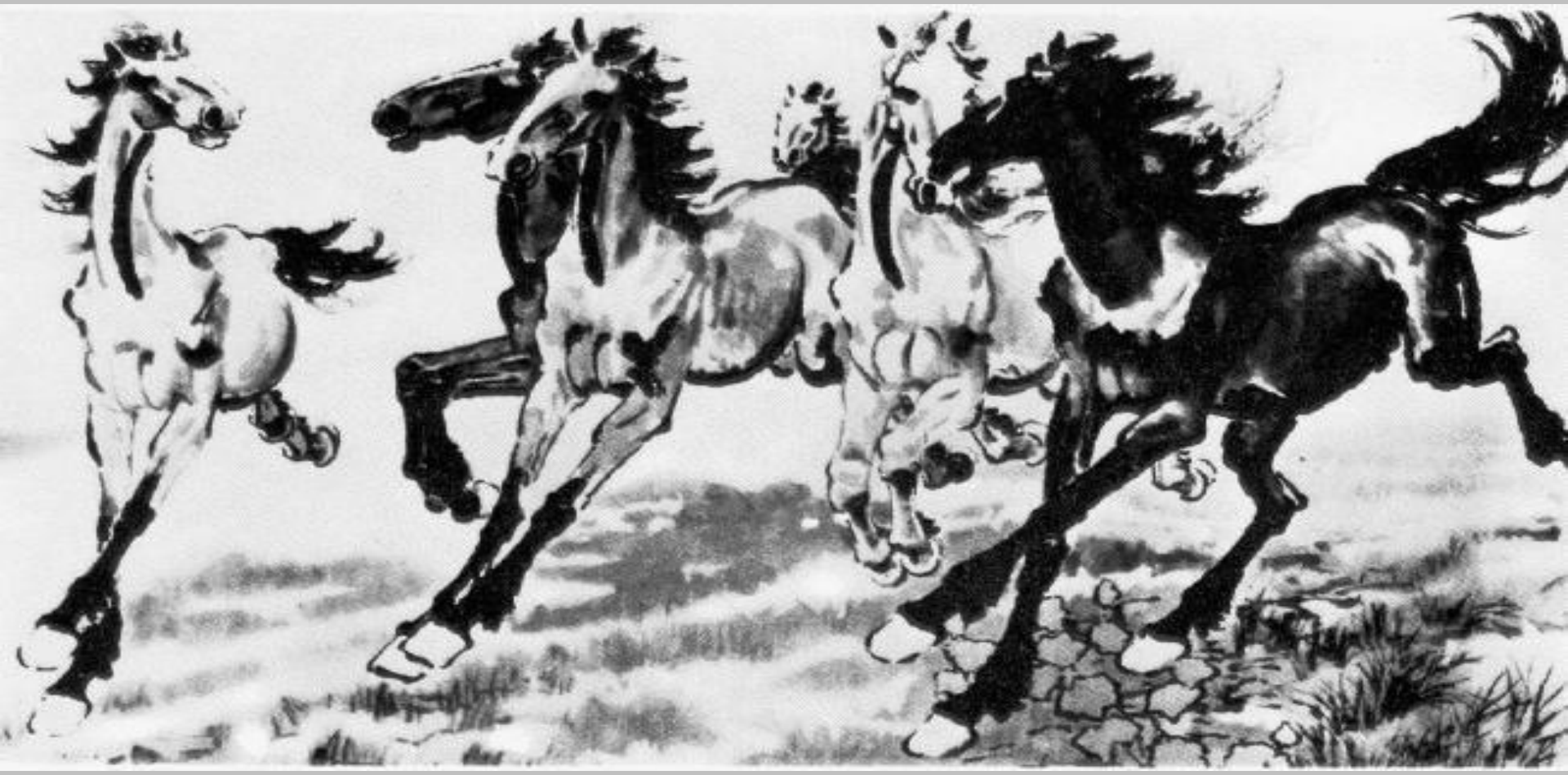


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

Challenges: deformation



Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenges: occlusion

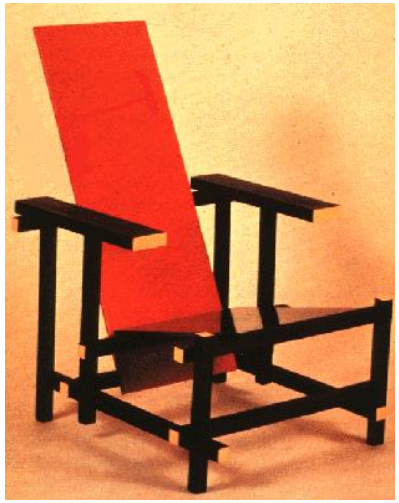
Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

Challenges: background clutter

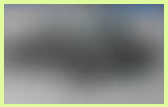
Challenges: Motion

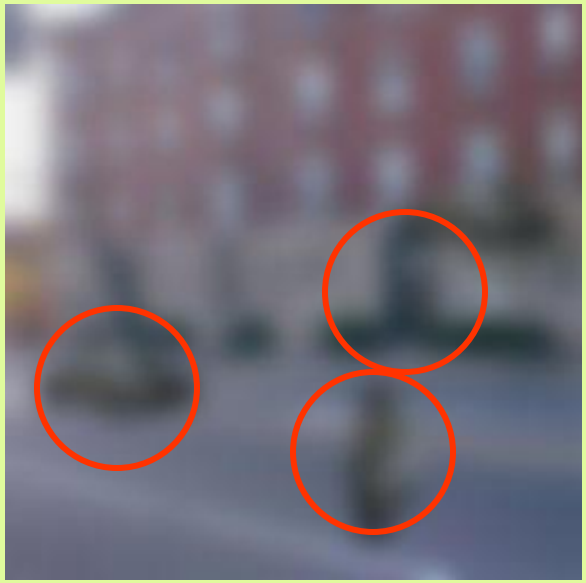
Challenges: object intra-class variation



slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity





slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity

Source: Rob Fergus and Antonio Torralba

Challenges: local ambiguity

Source: Rob Fergus and Antonio 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

Depth cues: Linear perspective

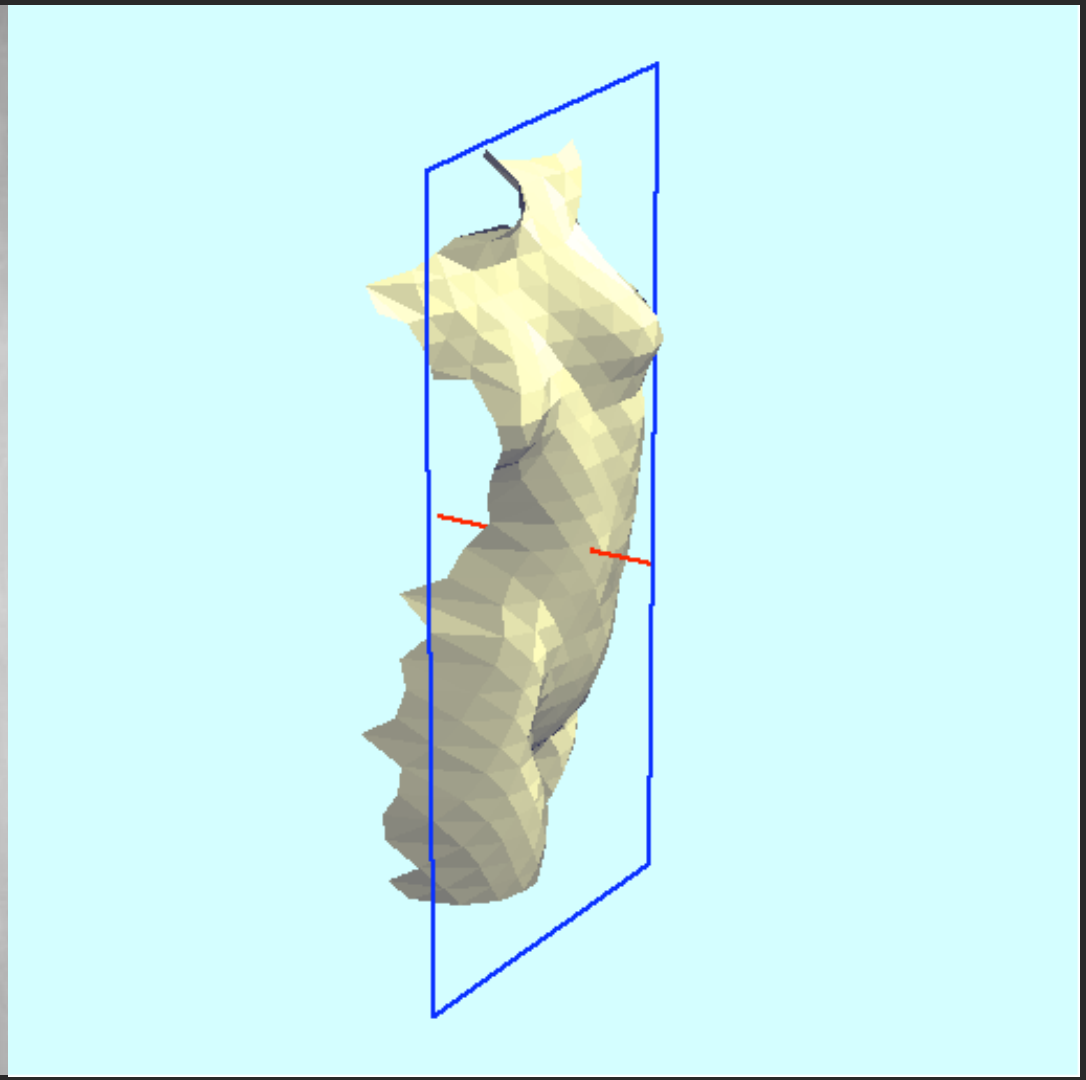
Depth cues: Aerial perspective

Depth ordering cues: Occlusion

Source: J. Koenderink

Shape cues: Texture gradient

Shape and lighting cues: Shading



Source: J. Koenderink

Position and lighting cues: Cast shadows

Source: J. Koenderink

Grouping cues: Similarity (color, texture, proximity)

Grouping cues: “Common fate”

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

Inherent ambiguity of the problem

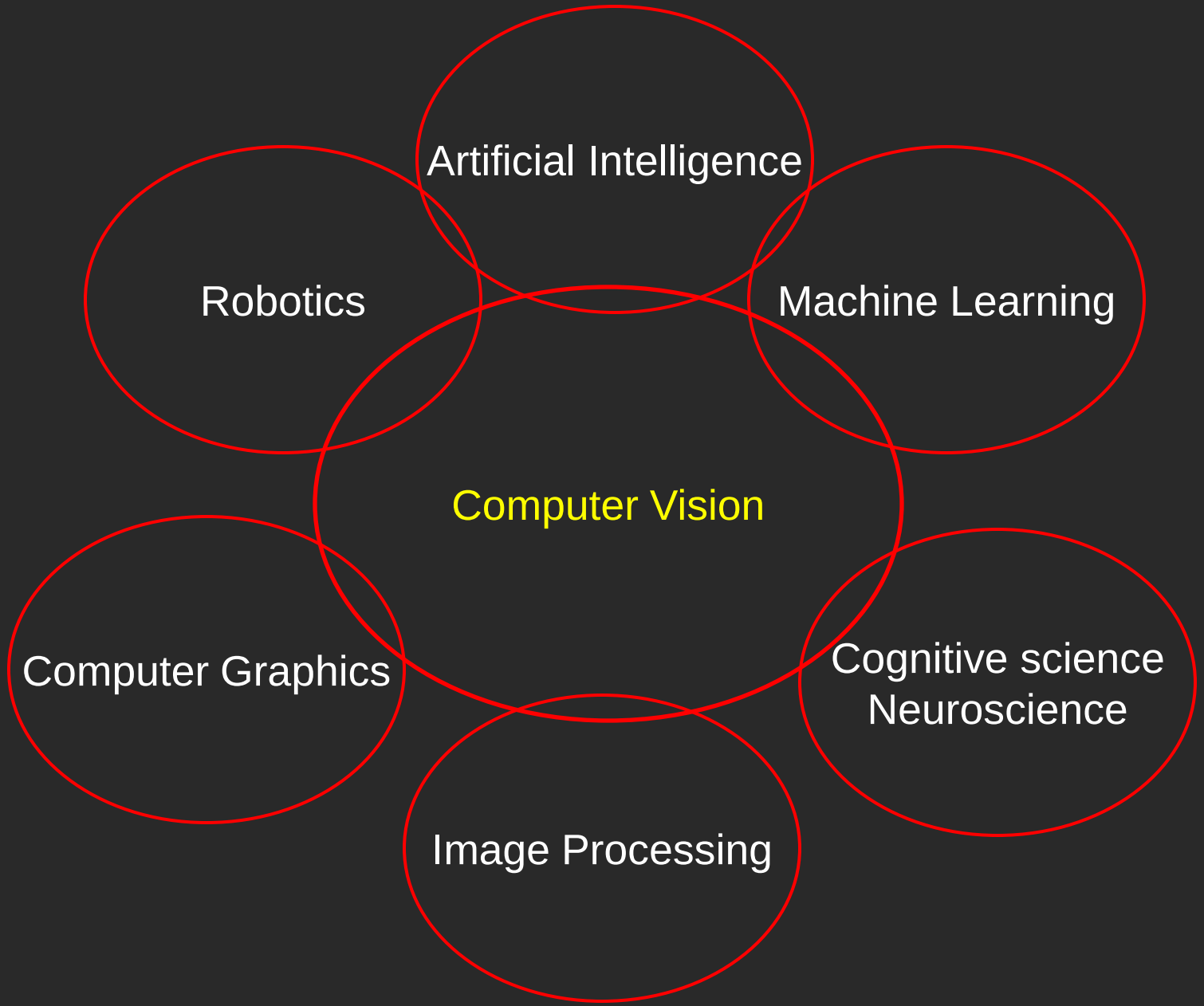
- Many different 3D scenes could have given rise to a particular 2D picture

Inherent ambiguity of the problem

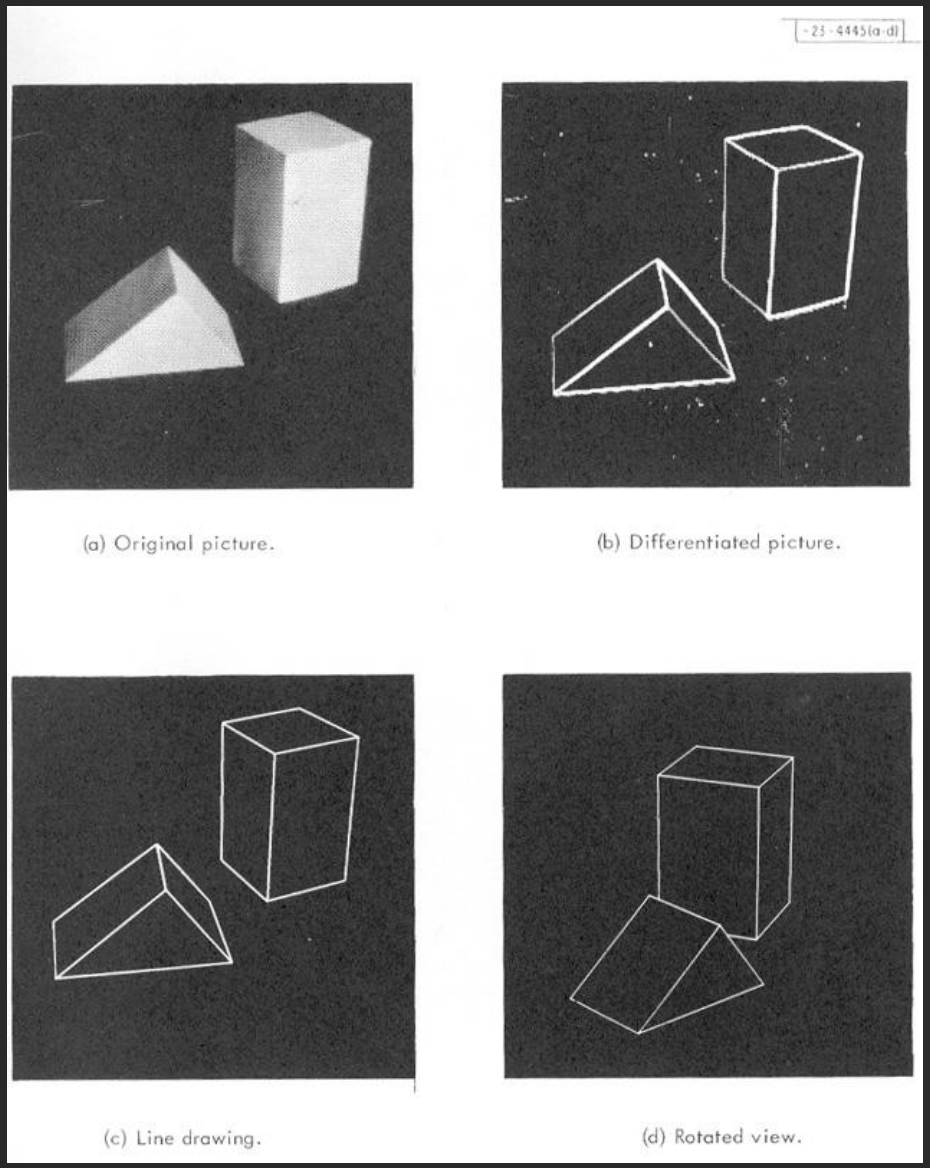
- Many different 3D scenes could have given rise to a particular 2D picture

- Possible solutions
  - Bring in more constraints (more images)
  - Use prior knowledge about the structure of the world
- Need a combination of geometric and statistical methods

Connections to other disciplines



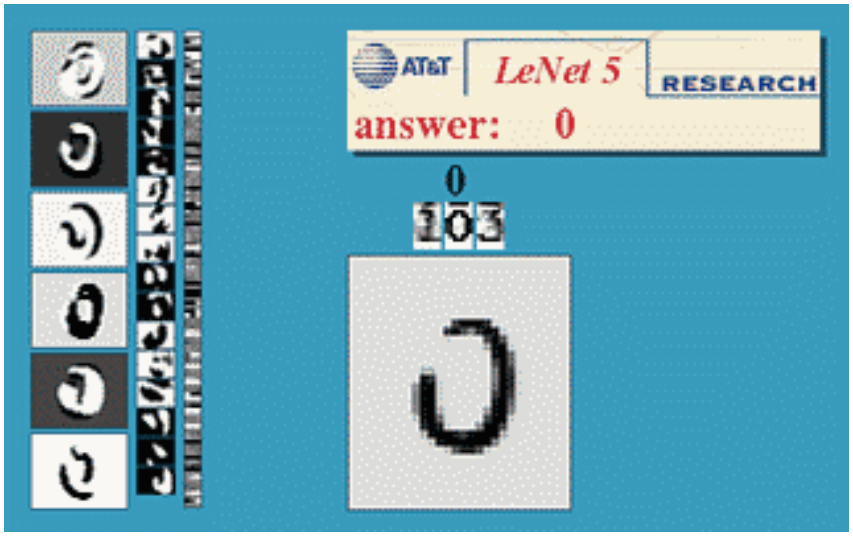
Origins of computer vision



L. G. Roberts, *Machine Perception of Three Dimensional Solids,* Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

Successes of computer vision to date

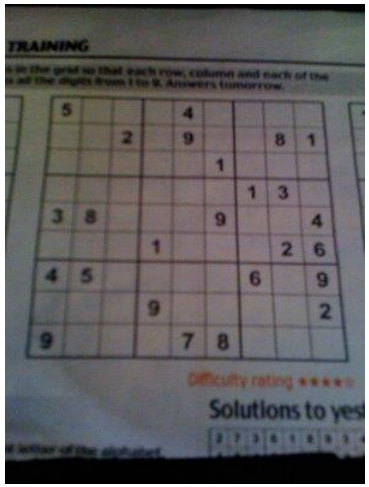
Optical character recognition (OCR)



Digit recognition
yann.lecun.com

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

Automatic check processing



Sudoku grabber

http://sudokugrab.blogspot.com/

Source: S. Seitz, N. Snavely

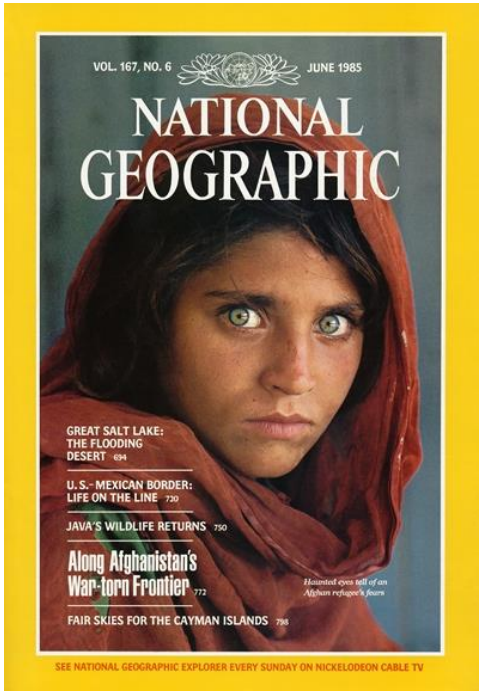
Biometrics

Fingerprint scanners on many new laptops, other devices

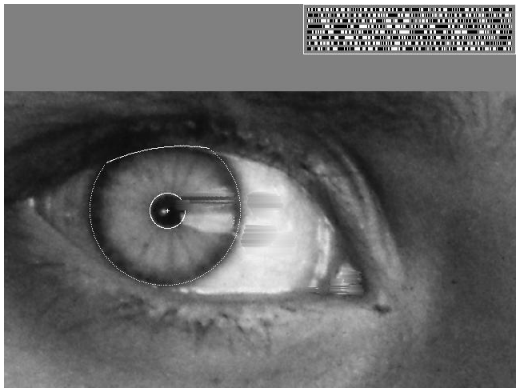
Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

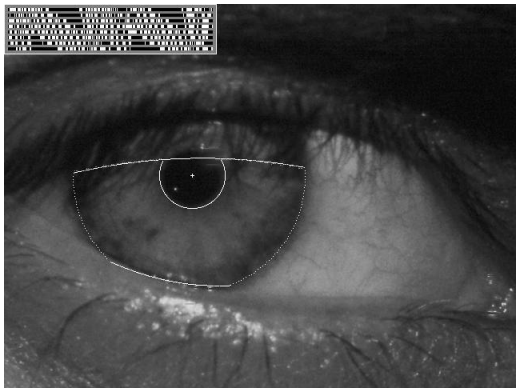
Source: S. Seitz

Biometrics



[How the Afghan Girl was Identified by Her Iris Patterns]





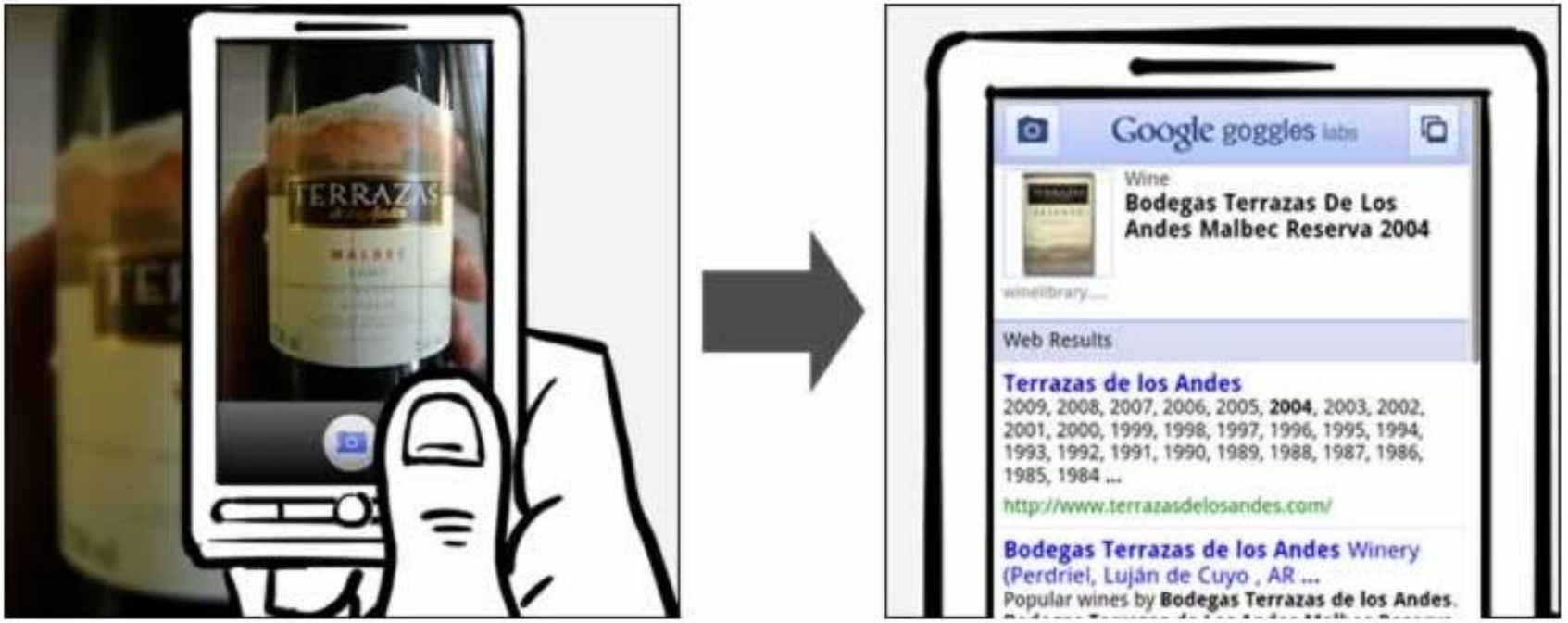
Source: S. Seitz

Mobile visual search: Google Goggles

Google Goggles in Action

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

Google goggles labs



Object recognition (in mobile phones)

This is becoming real:

- Lincoln Microsoft Research
- Point & Find, Nokia

Source: Neel Joshi

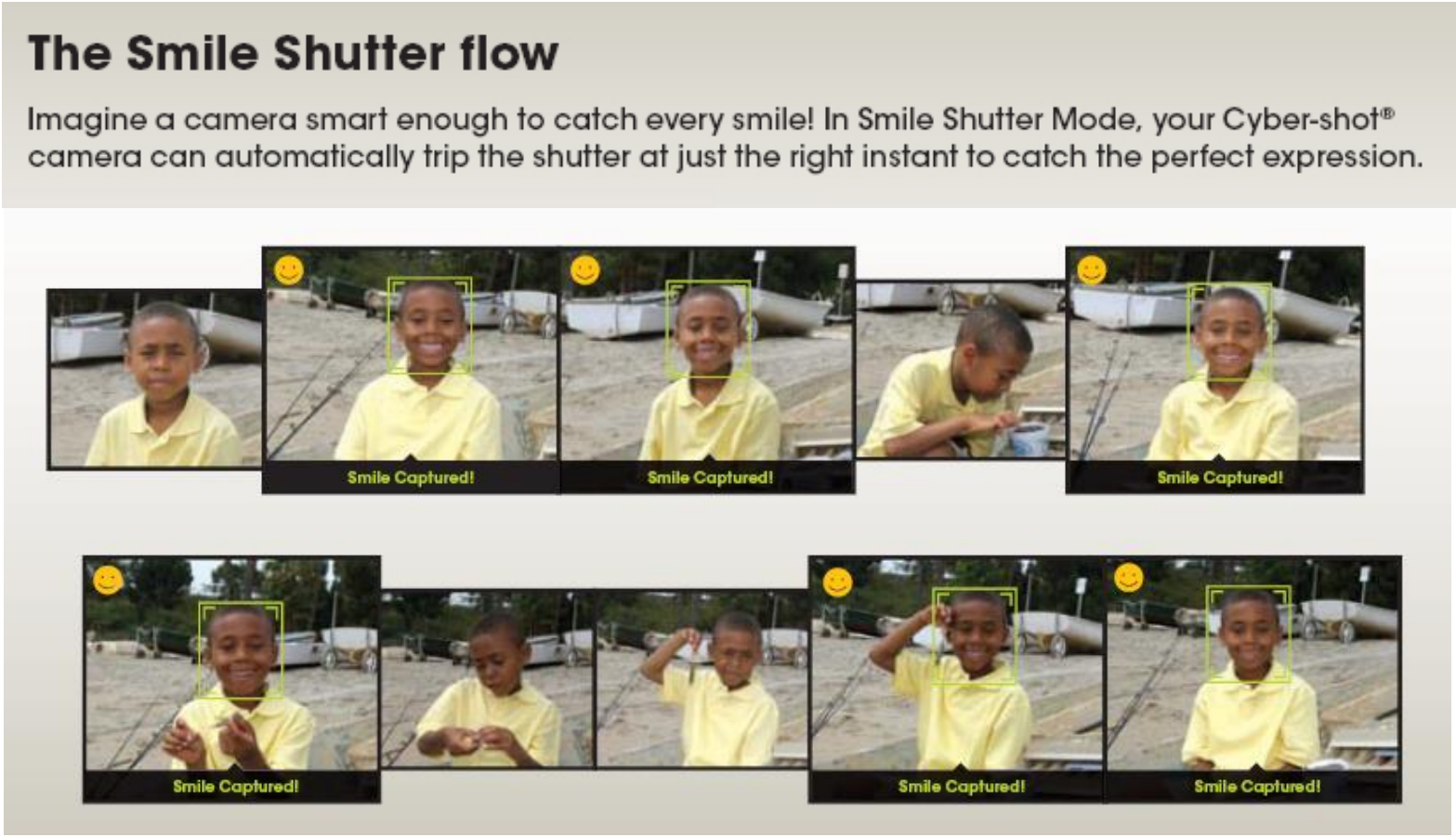
Face detection

Many new digital cameras now detect faces

- Canon, Sony, Fuji, …

Source: S. Seitz

Smile detection



[Sony Cyber-shot® T70 Digital Still Camera]

Source: S. Seitz

Face recognition: Apple iPhoto software

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

Object recognition (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: Steve Seitz and Neel Joshi

Sports

***Sportvision* first down line**

Nice [explanation](#) on www.howstuffworks.com

Source: Steve Seitz and Neel Joshi

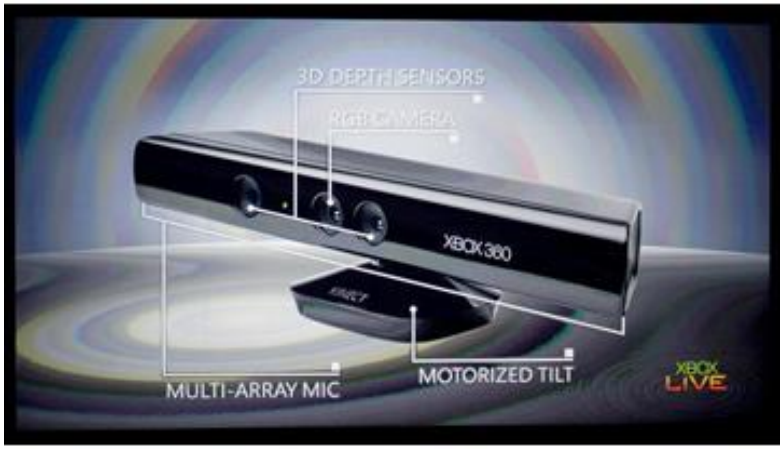
Automotive safety

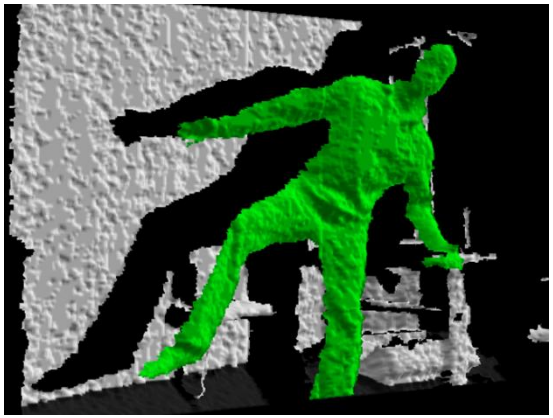
Mobileye: Vision systems in high-end BMW, GM, Volvo models

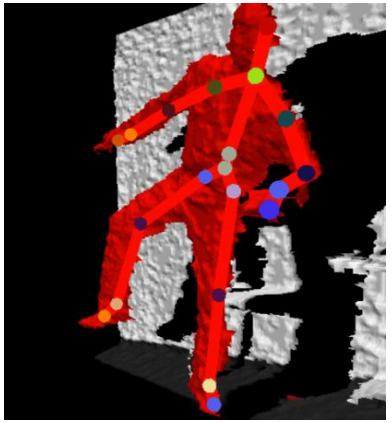
- Pedestrian collision warning
- Forward collision warning
- Lane departure warning
- Headway monitoring and warning

Source: A. Shashua, S. Seitz

Vision-based interaction: Xbox Kinect







http://blogs.howstuffworks.com/2010/11/05/how-microsoft-kinect-works-an-amazing-use-of-infrared-light/

http://electronics.howstuffworks.com/microsoft-kinect.htm

http://www.xbox.com/en-US/Live/EngineeringBlog/122910-HowYouBecometheController

http://www.ismashphone.com/2010/12/kinect-hacks-more-interesting-than-the-devices-original-intention.html

Special effects: shape and motion capture

Source: S. Seitz

Earth viewers (3D modeling)

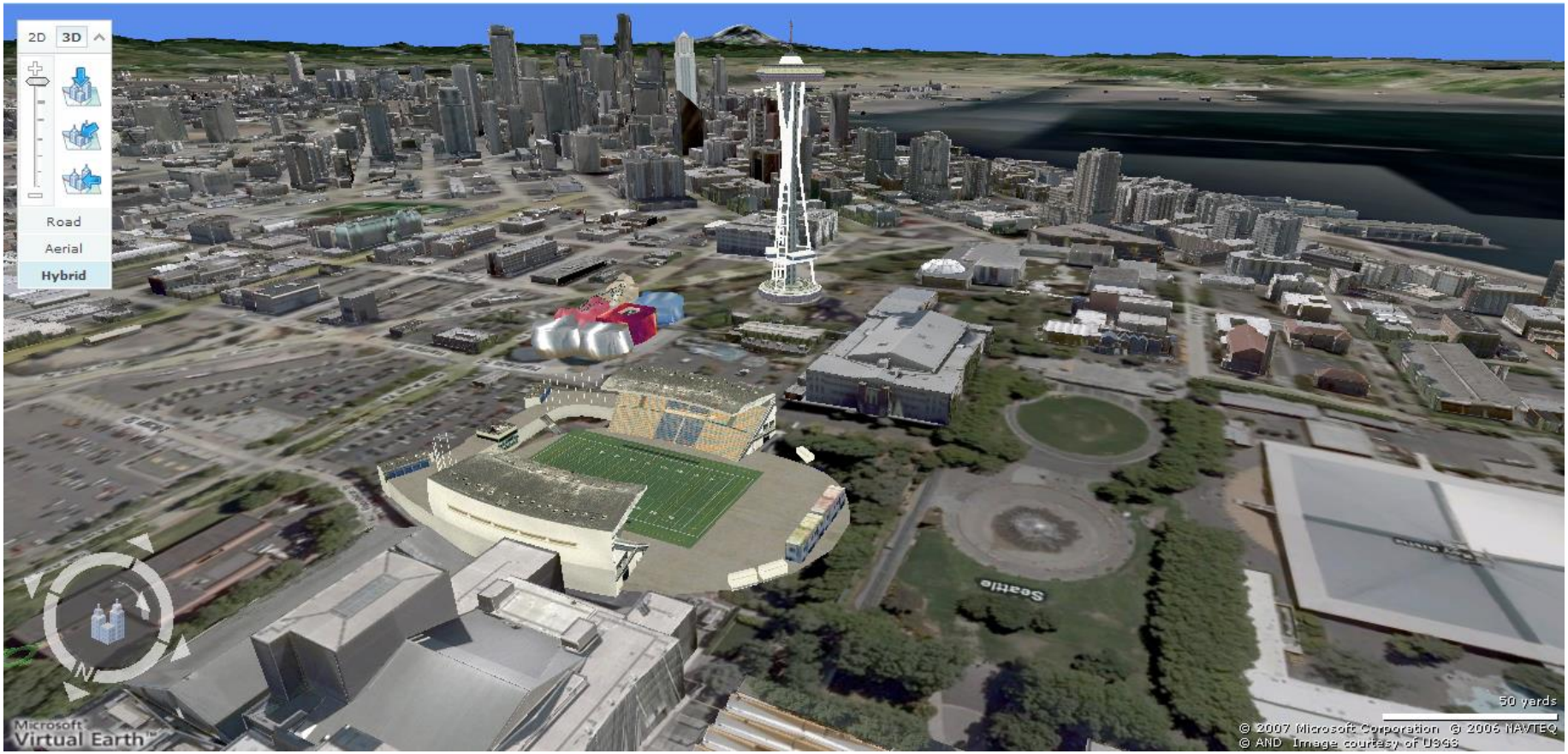


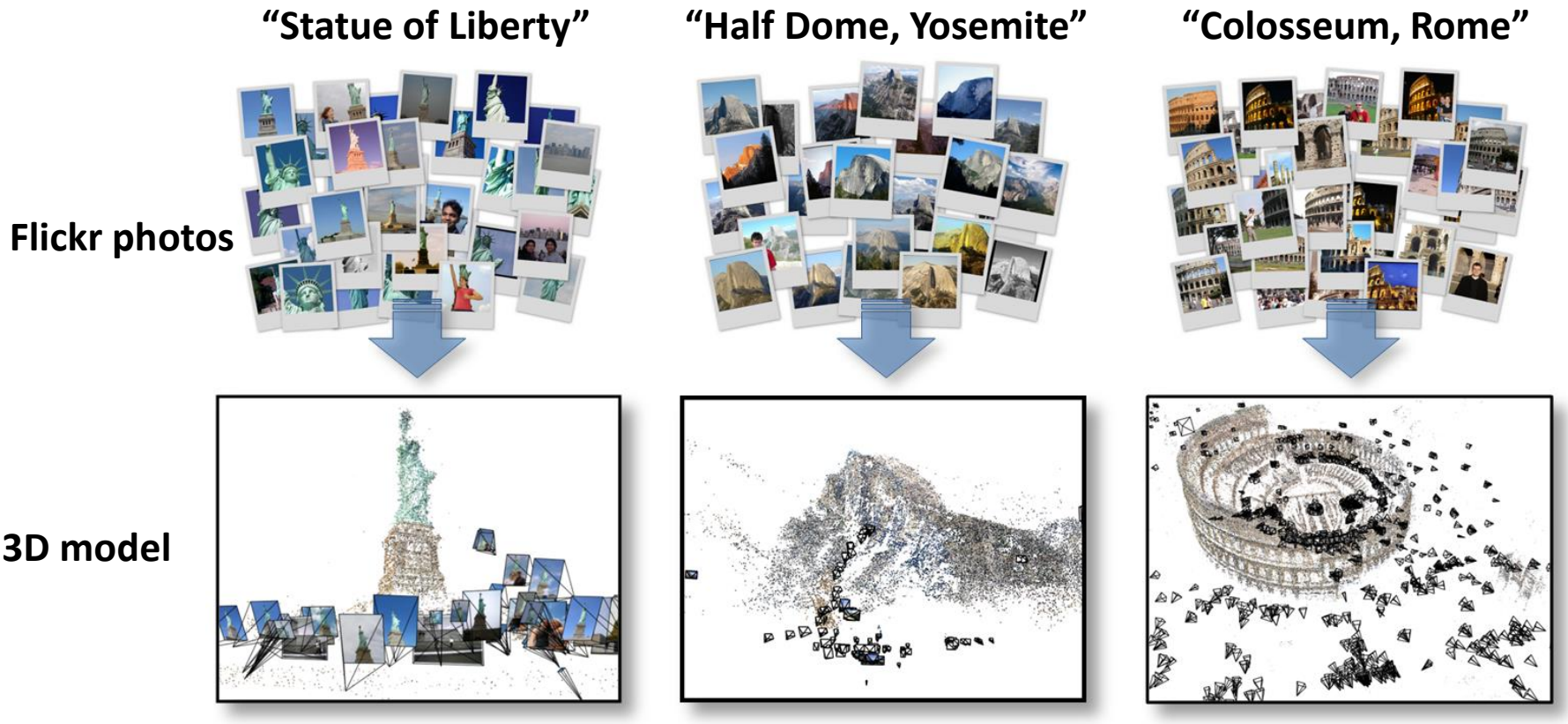
Image from Microsoft's Virtual Earth

(see also: Google Earth)

Source: Steve Seitz and Neel Joshi

Phototourism

Automatic 3D reconstruction from Internet photo collections



Source: Neel Joshi

3D visualization: Microsoft Photosynth

http://photosynth.net

Source: S. Seitz

Vision for robotics, space exploration

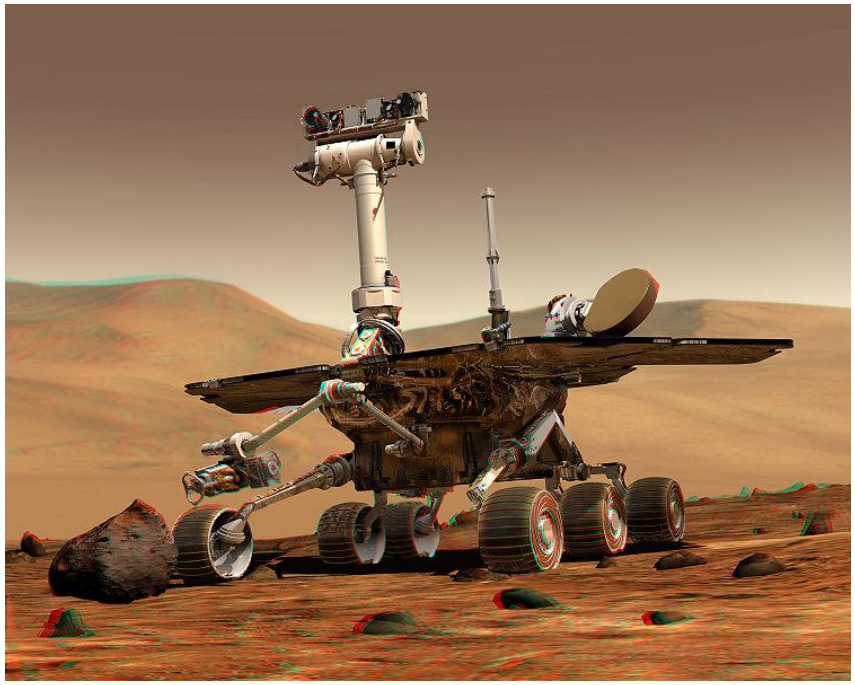
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

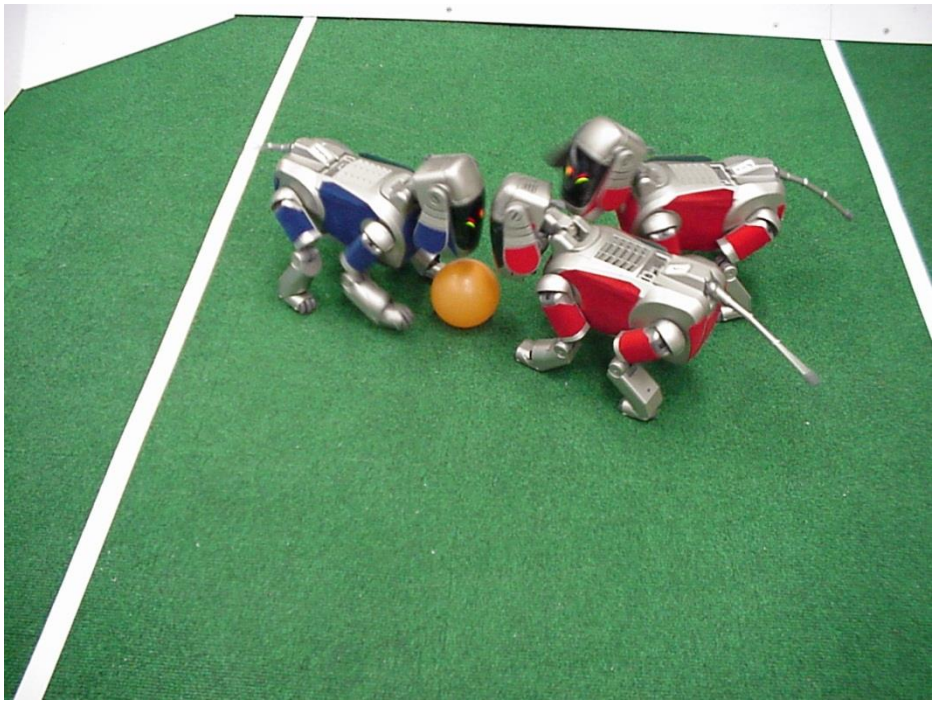
Source: S. Seitz

Robotics



NASA's Mars Spirit Rover

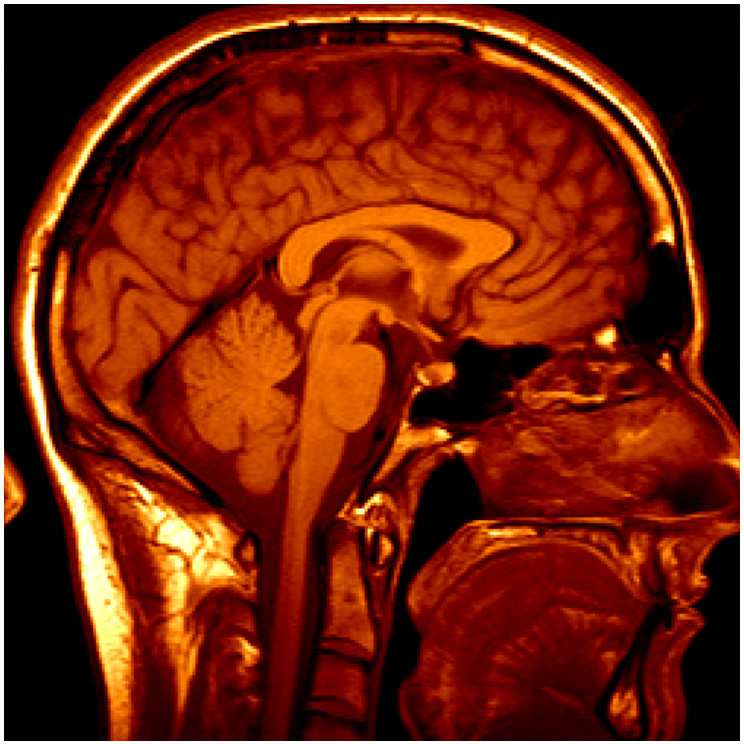
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Source: Steve Seitz and Neel Joshi

Medical imaging



3D imaging

MRI, CT

Image guided surgery

Grimson et al., MIT

Source: Steve Seitz and Neel Joshi

The computer vision industry

- A list of companies here:

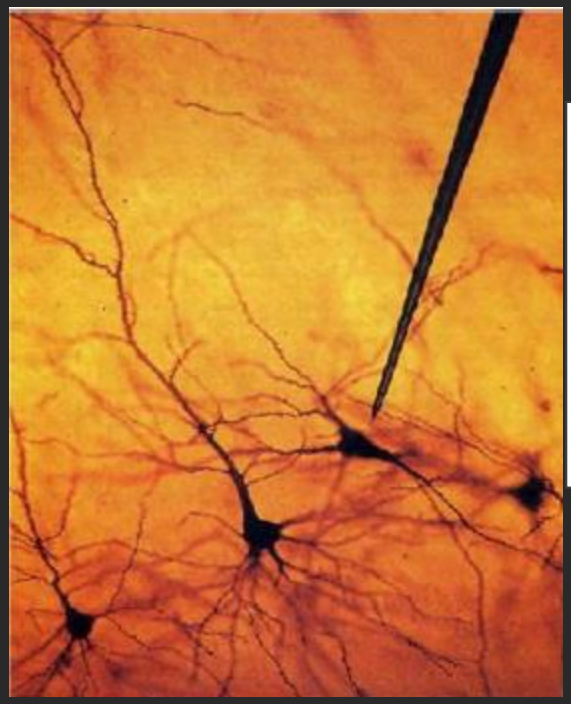
  http://www.cs.ubc.ca/spider/lowe/vision.html

Next Class

Animal Eyes

Compound eye and its evolutionary advantages?

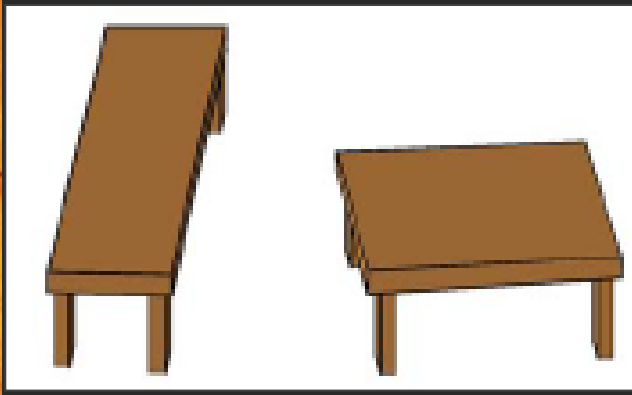
Perception



1981 Nobel Prize in Medicine

Hubel and Wiesel

Illusions



Compare the length and breadth of these two tables. Why did I ask this question?