Chapter 1

Introduction to Computer Vision

James Hays, Brown University

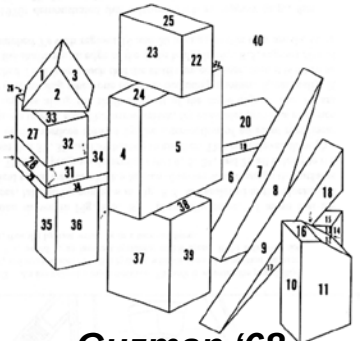
Department of Mechatronics

Contents

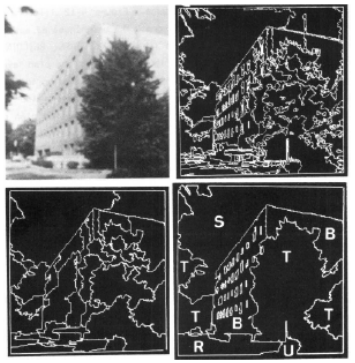
- Introductions
- What is Computer Vision?
- Why study computer vision?
- Why computer vision matters?
- How vision is used now?

Brief history of computer vision

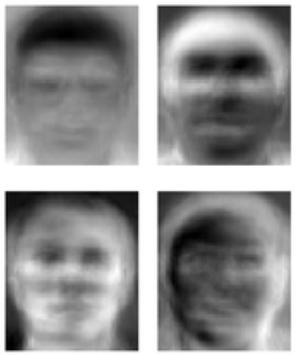
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2030's: robot uprising?



Guzman '68



Ohta Kanade '78



Turk and Pentland '91

Computer Vision

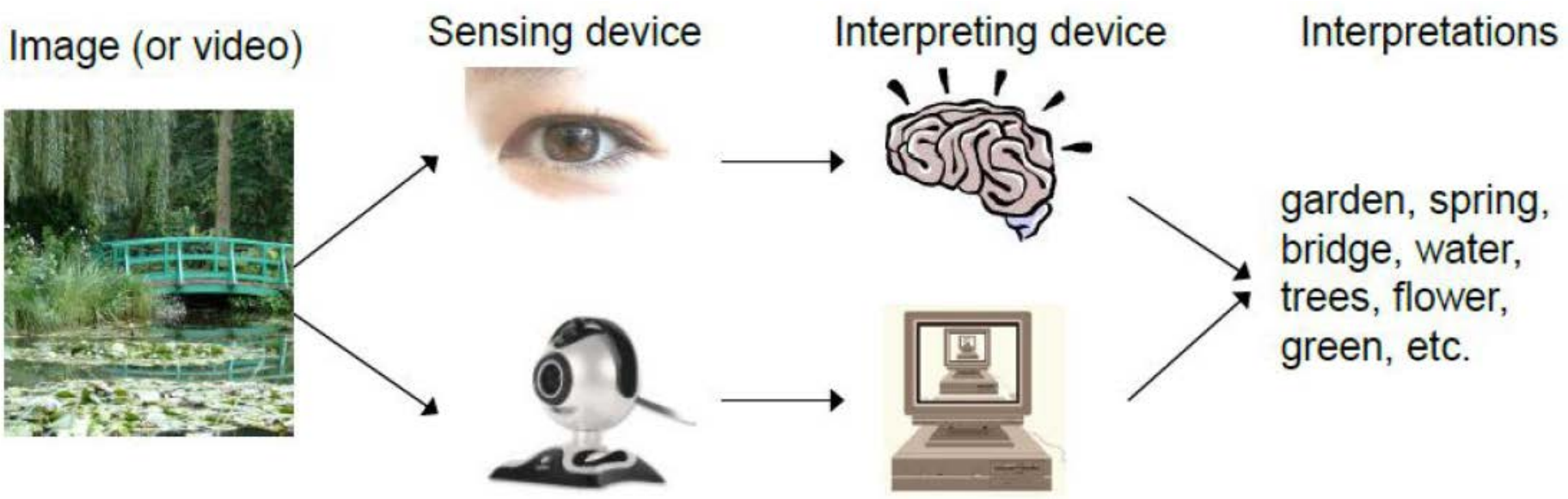
- Make computers understand images and video.

What kind of scene?

Where are the cars?

How far is the building?

What is (computer) vision?



Human brain:
~100 ÷ 1.000 trillion synapses

(Artificial) Neural network:
~1 ÷ 10 billion synapses

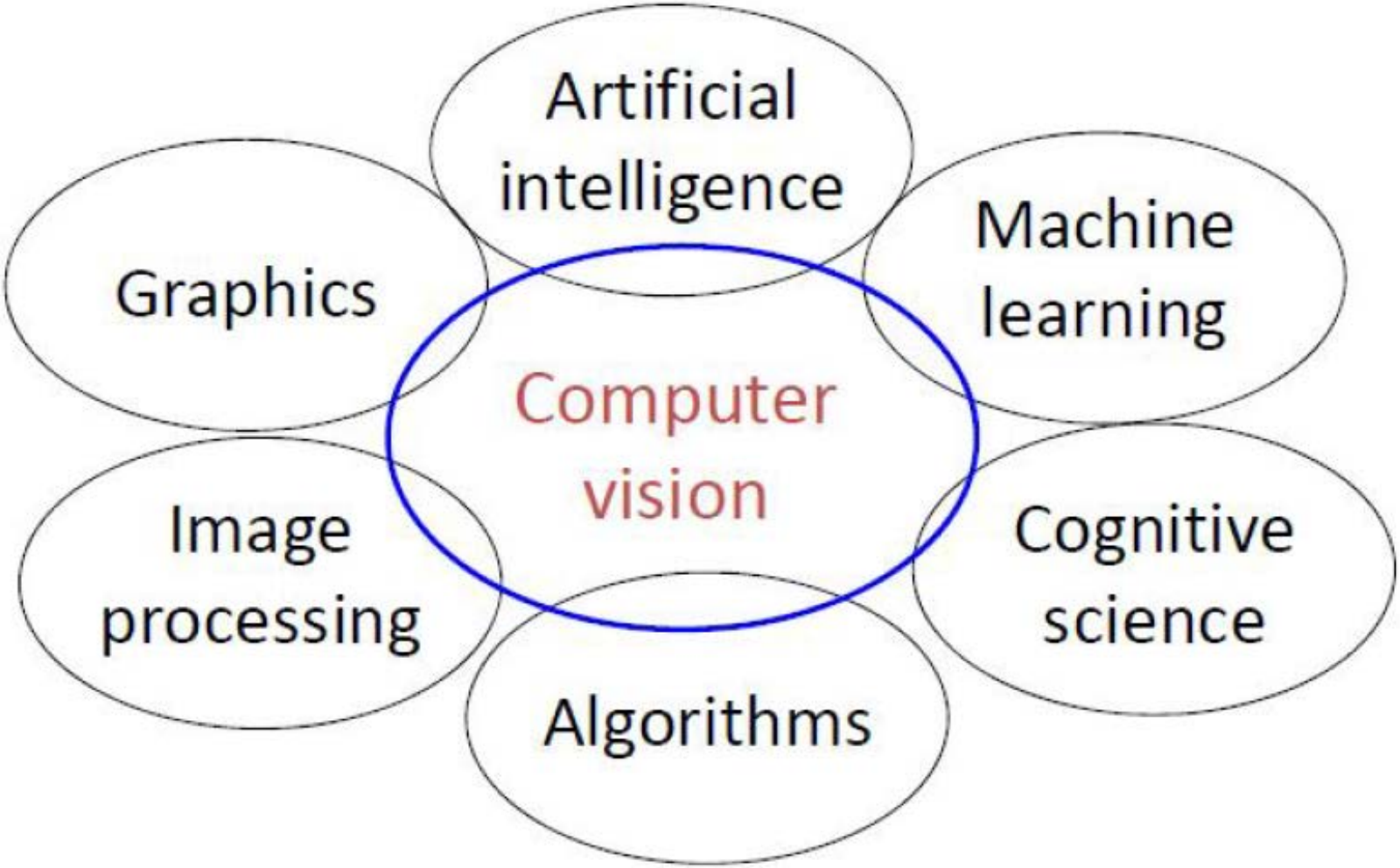
Human brains have ~10.000 computational power than computer brains

Lex Fridman, MIT

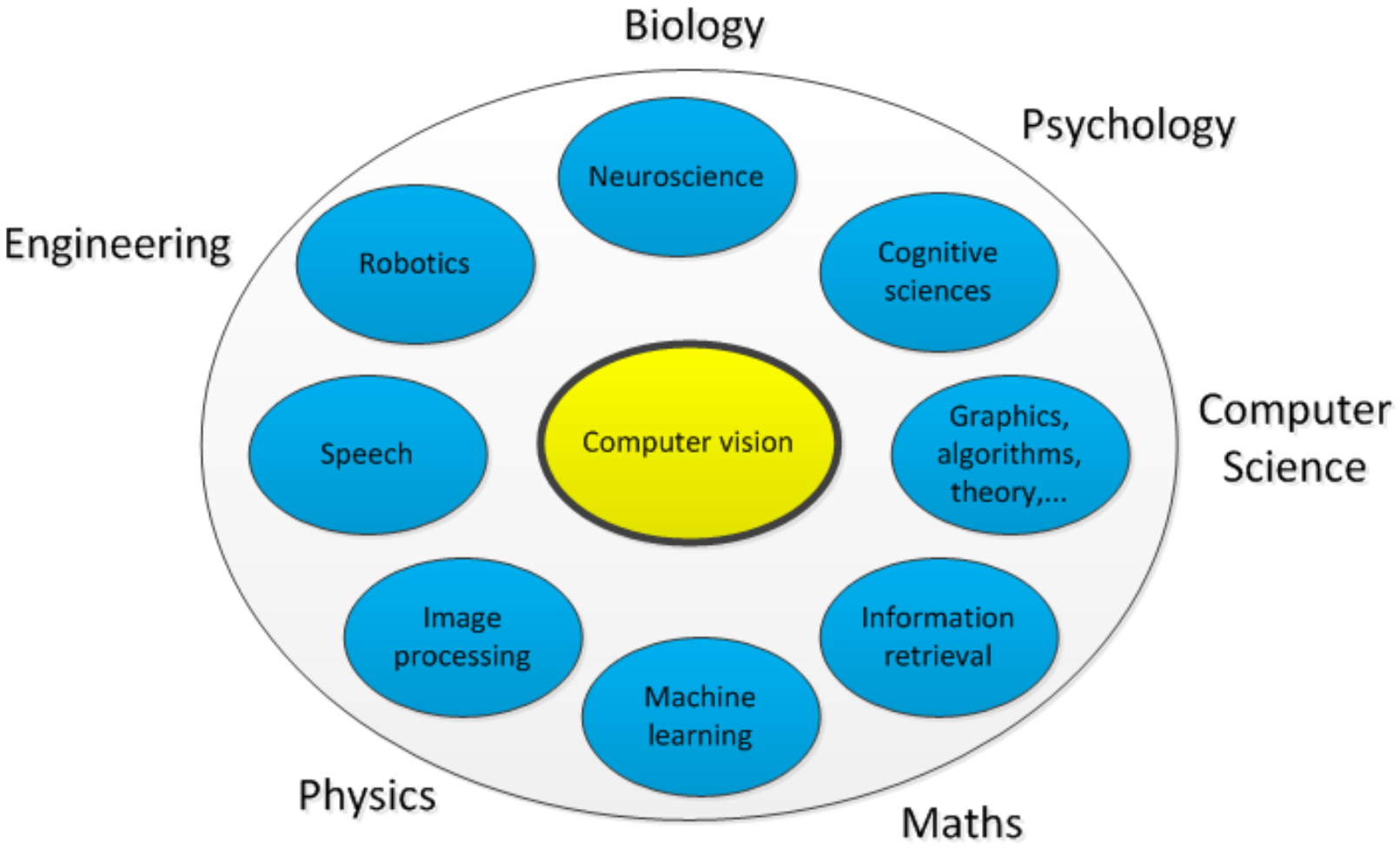
Computer Vision and Nearby Fields

- Computer Graphics: Models to Images
- Computer Photography: Images to Images
- Computer Vision: Images to Models

Related disciplines



What is it related to?



The goal of computer vision

- **To bridge the gap between pixels and “meaning”.**

What we see

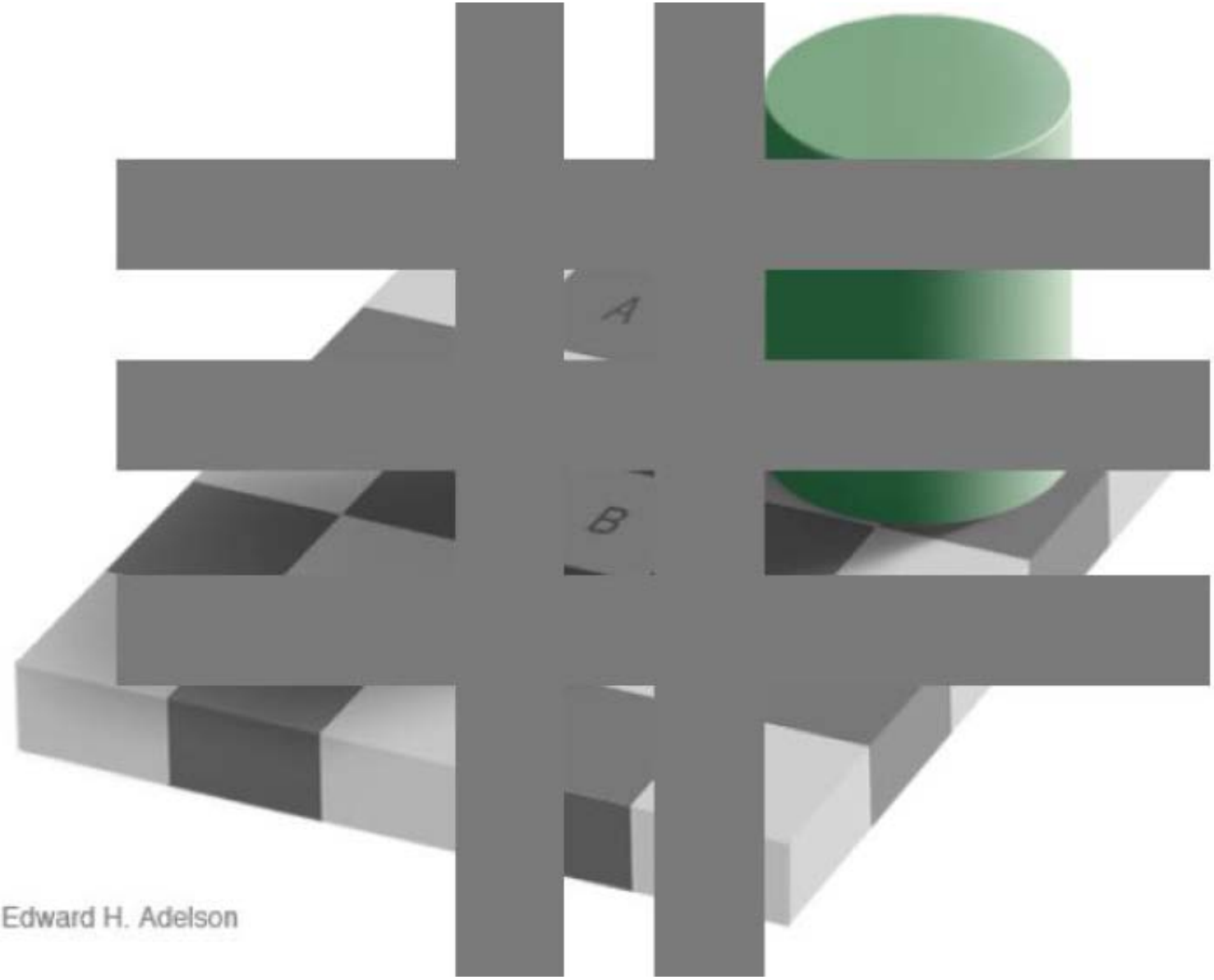
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

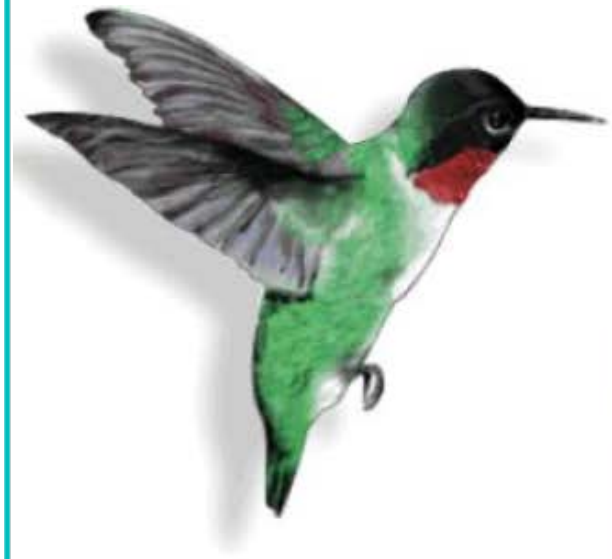
What a computer sees

Thorpe, et al. *Nature*, 1996

Segmentation

Perception

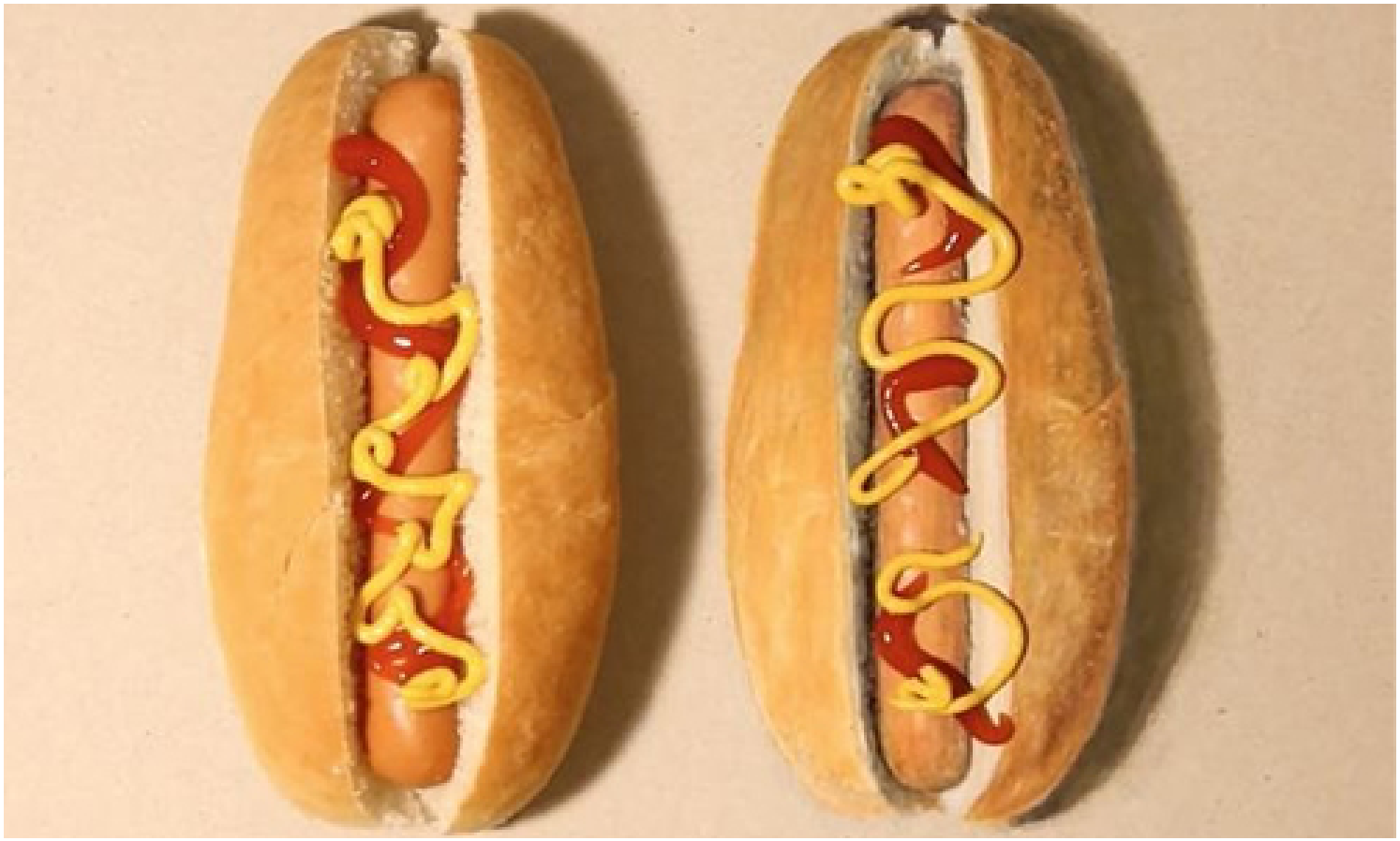

A
B
Edward H. Adelson



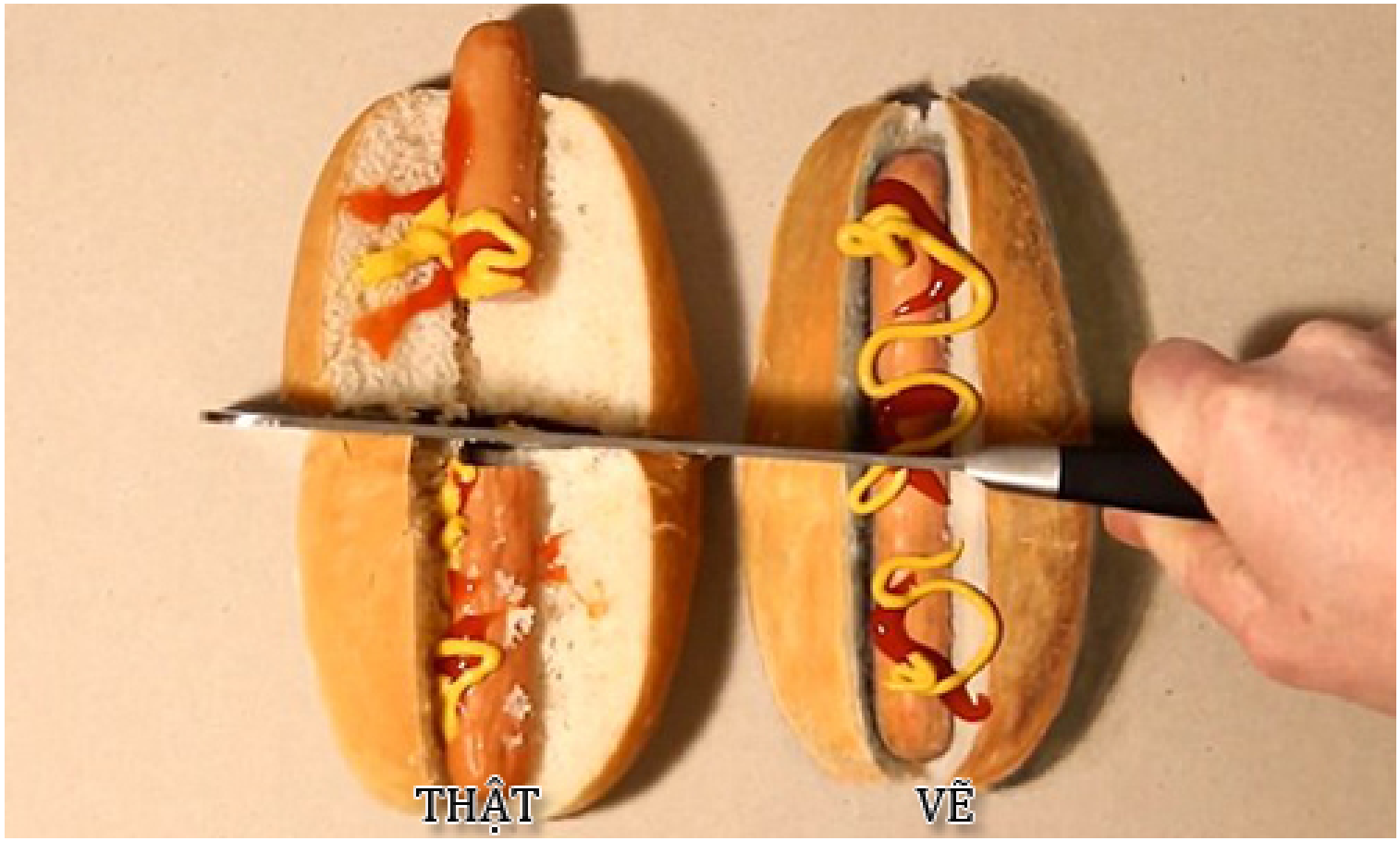
What kind of information can we extract from an image?

- **Metric 3D information**
- **Sematic information**

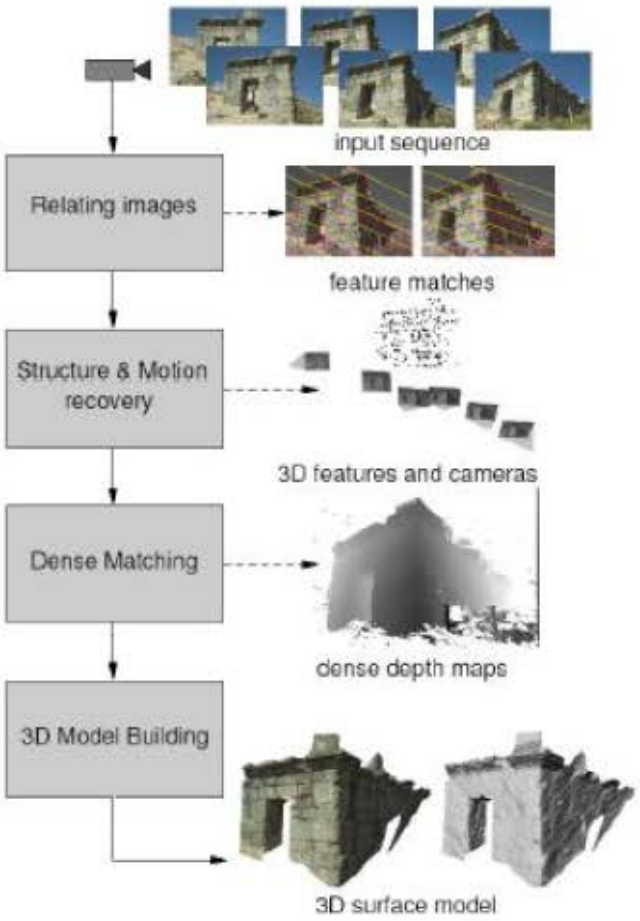
Guess which bread is real?



Left one is real



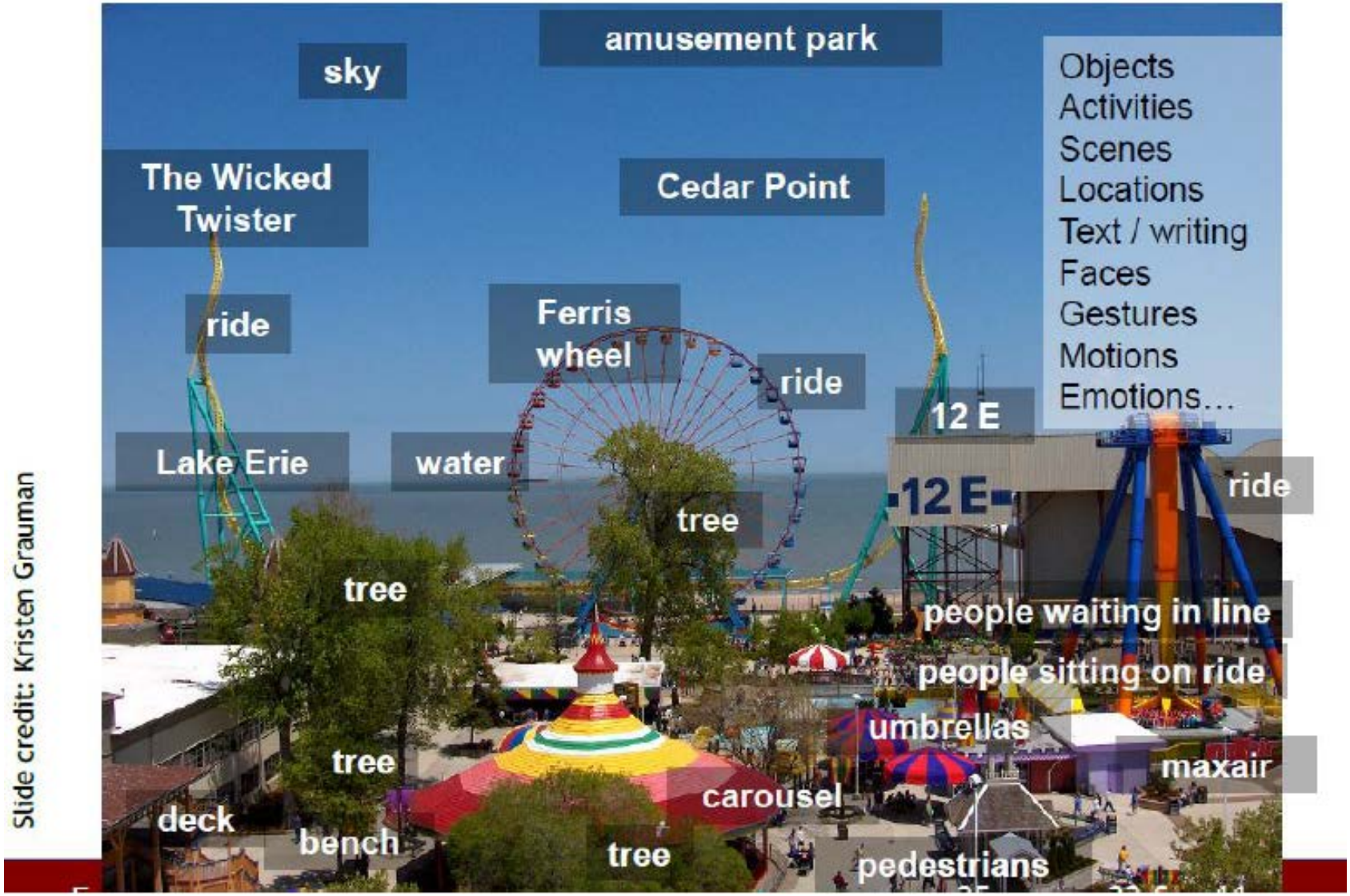
Vision as measurement device



Pollefeys et al.

Goesele et al.

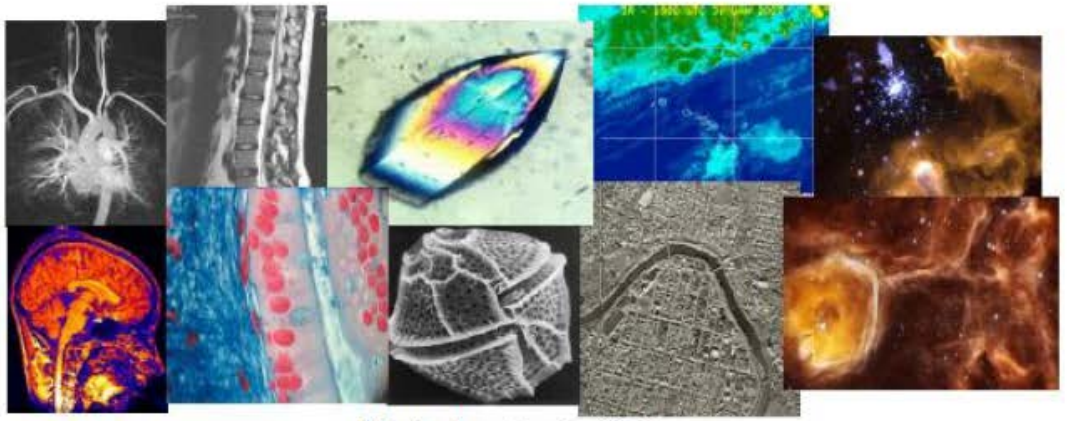
Vision as a source of sematic information



Why study computer vision?

- **Vision is useful: images and video are everywhere.**

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".**

*In the field of artificial intelligence, the most difficult problems are informally known as **AI-complete** or **AI-hard**, implying that the difficulty of these computational problems is equivalent to solving the central artificial intelligence problem—making computers as intelligent as people, or strong AI.*

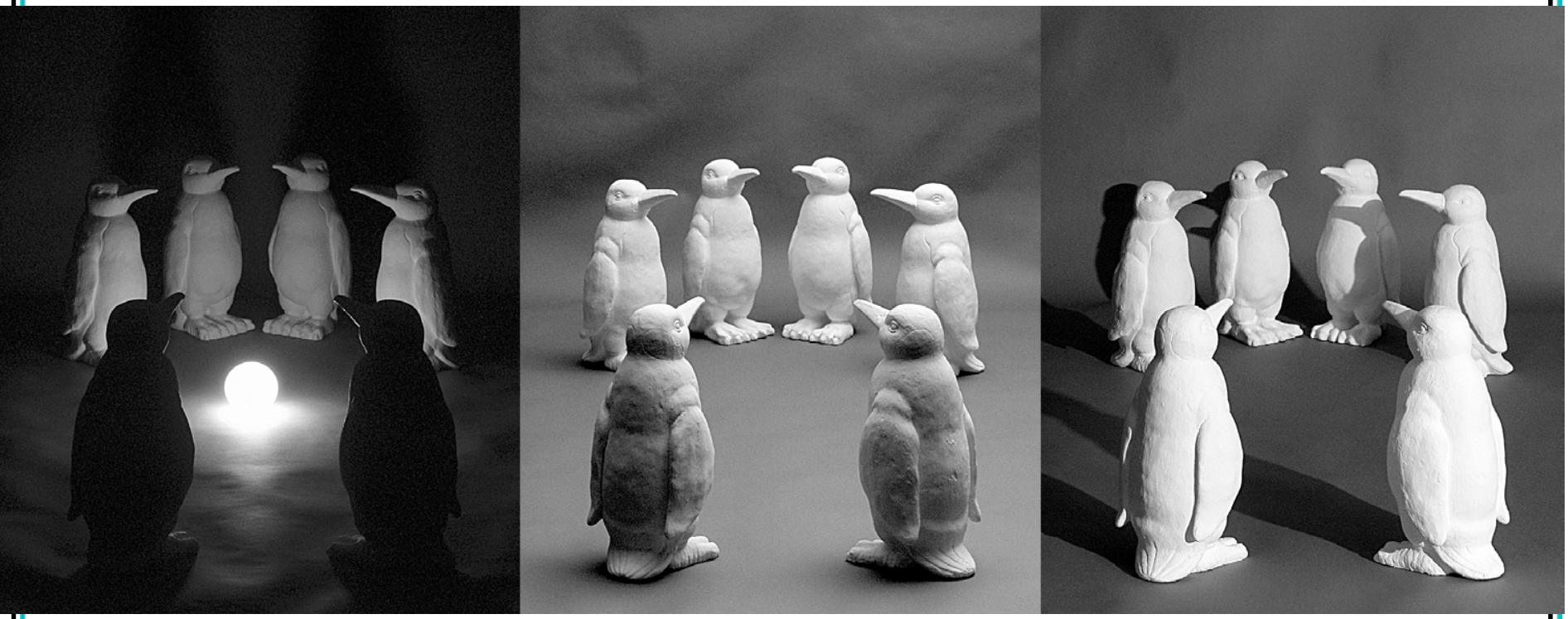
AI-complete problems are hypothesised to include computer vision, natural language understanding, and dealing with unexpected circumstances while solving any real world problem.

Vision is really hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of Macaque brain
  - More human brain devoted to vision than anything else

Challenges: view variation

Challenges: illumination

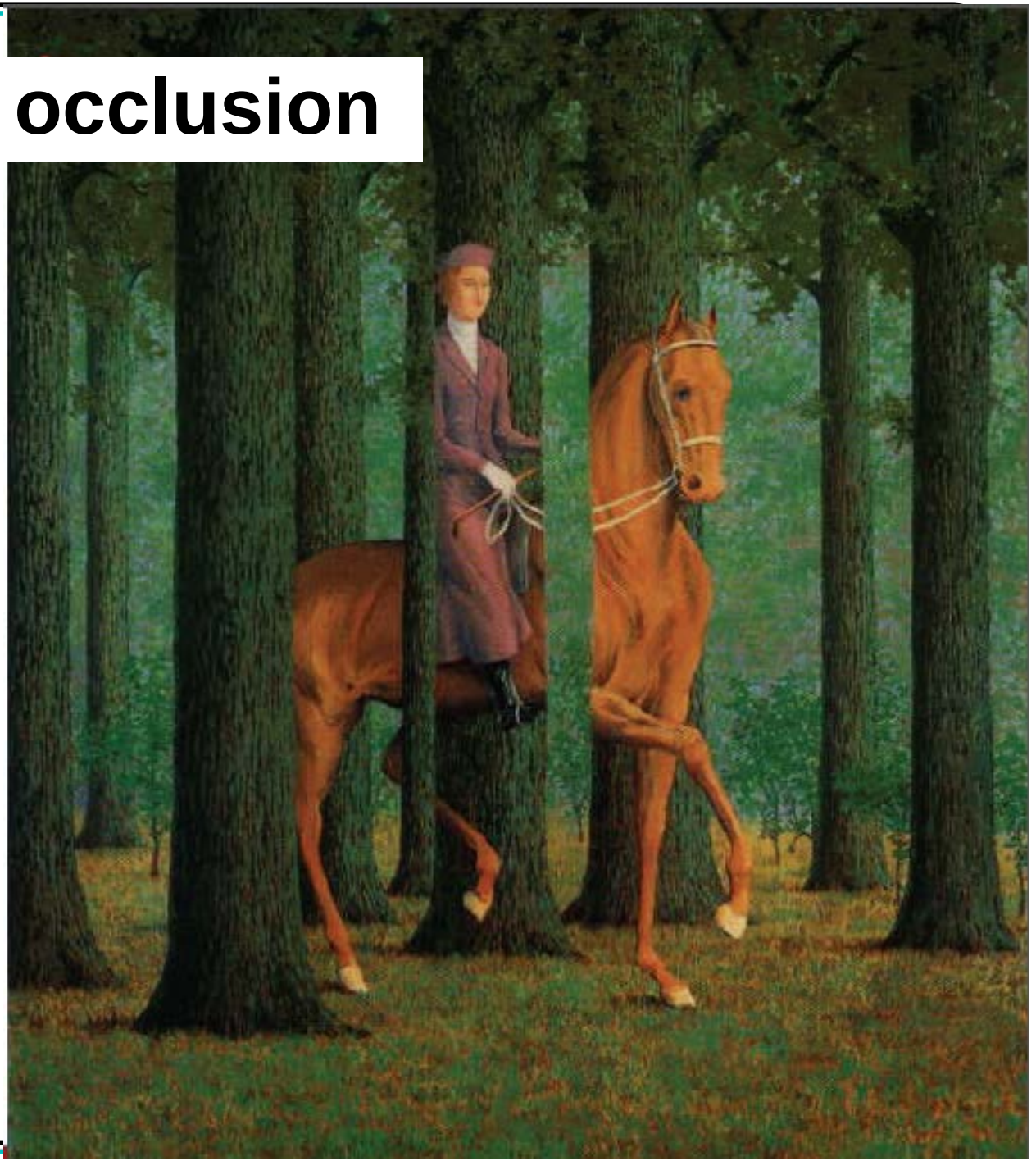


Challenges: scale

Challenges: deformation

Xu, Beihong 1943

Challenges: occlusion



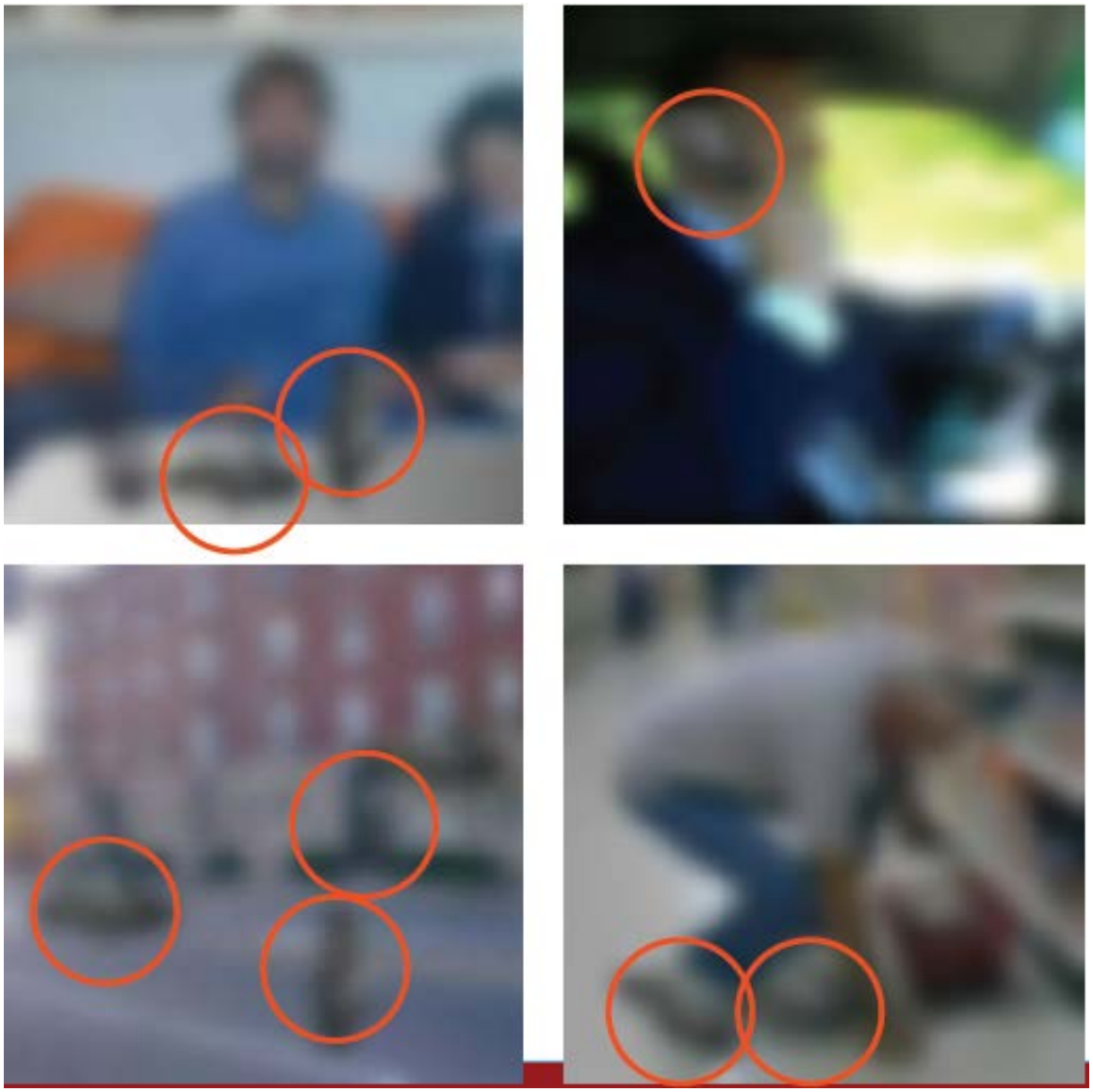
Magritte, 1957

Challenges: background clutter

Challenges: motion

Challenges: object intra-class variation

Challenges: local ambiguity

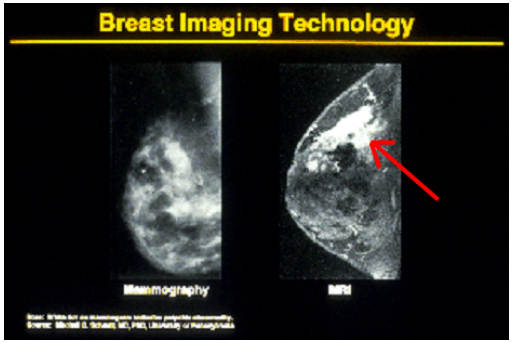


Challenges or Opportunities?

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

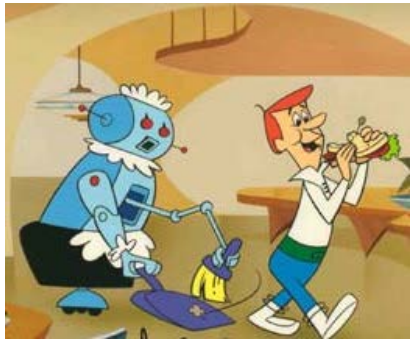
Why computer vision matters

Safety



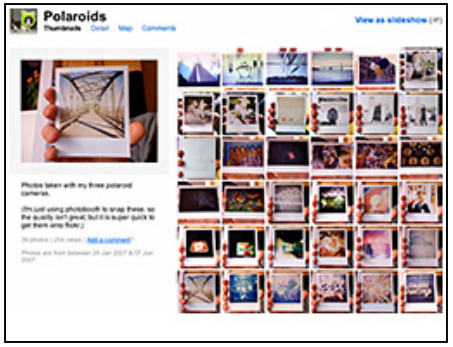
Health

Security



Comfort

Fun



Access

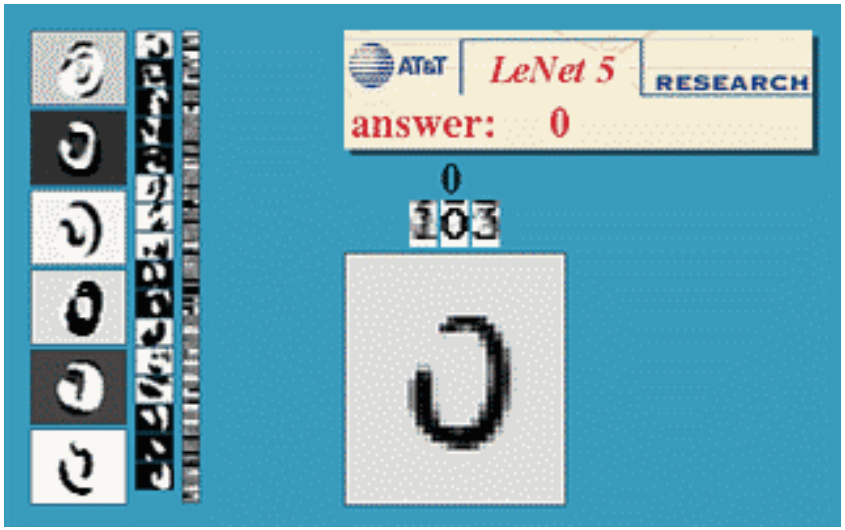
How vision is used now

- **Examples of state-of-the-art**

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
http://www.research.att.com/~yann/

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

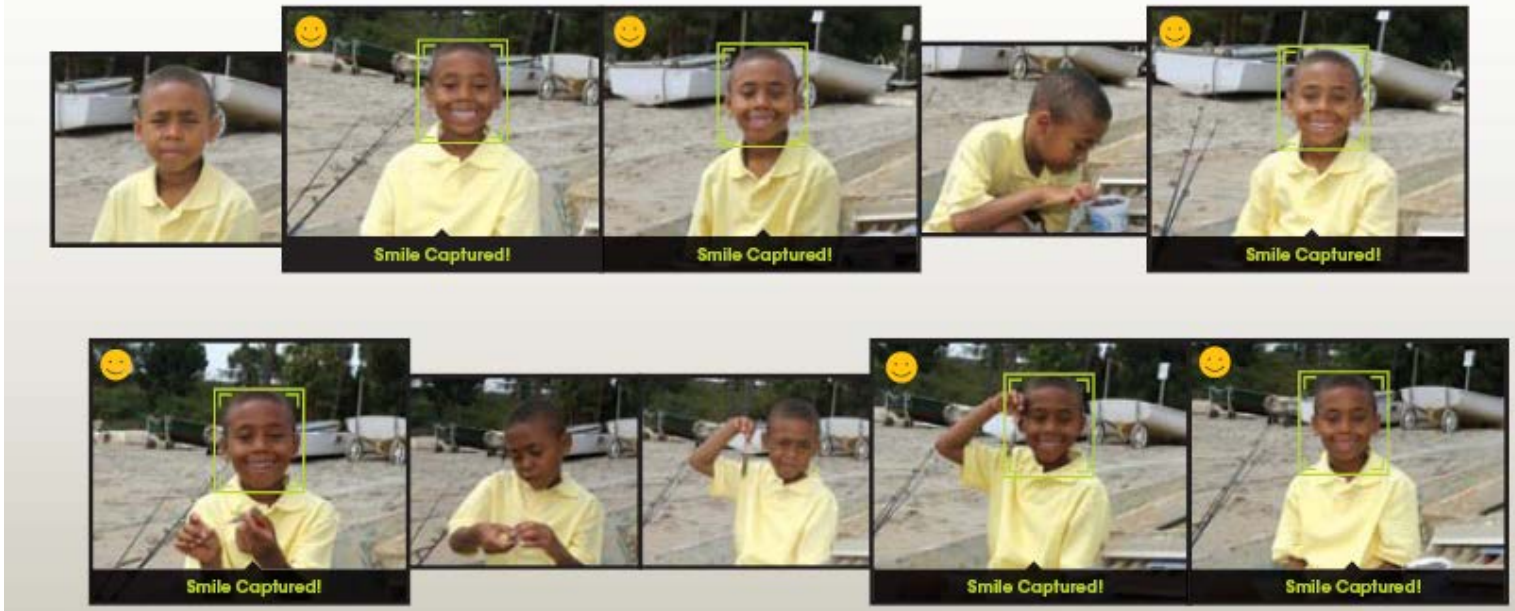
Face detection

- **Many new digital cameras now detect faces**
  - **Canon, Sony, Fuji, …**

Smile detection

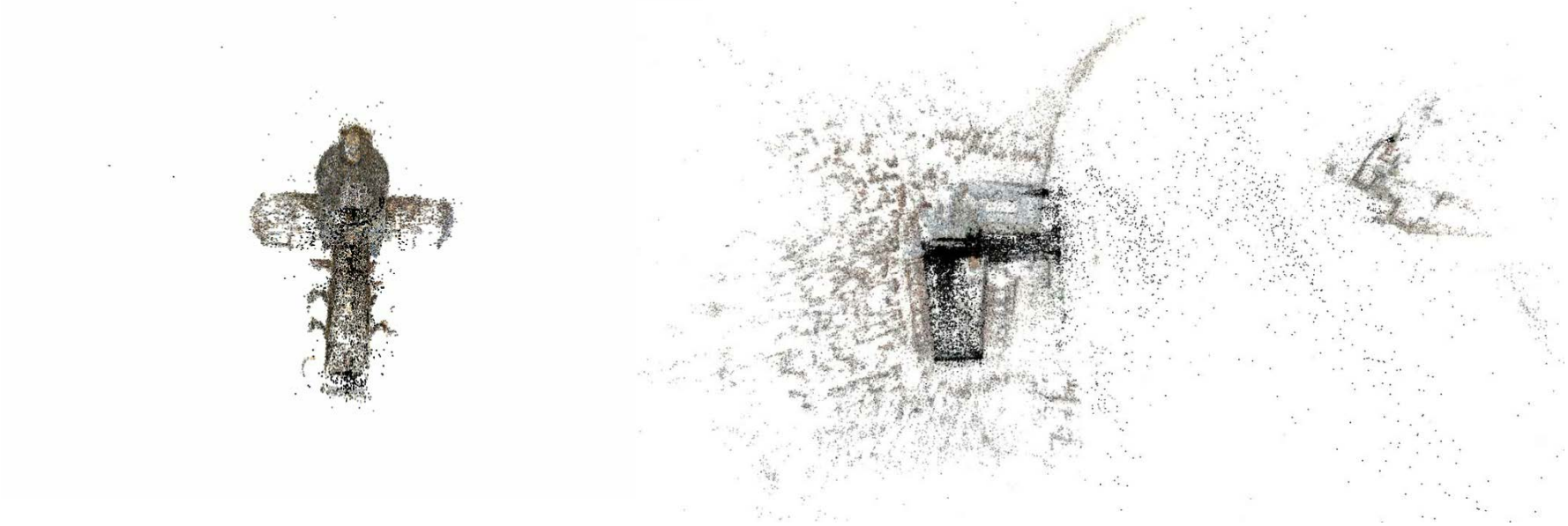
The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Sony Cyber-shot® T70 Digital Still Camera

3D from thousands of images



Building Rome in a Day: Agarwal et al. 2009

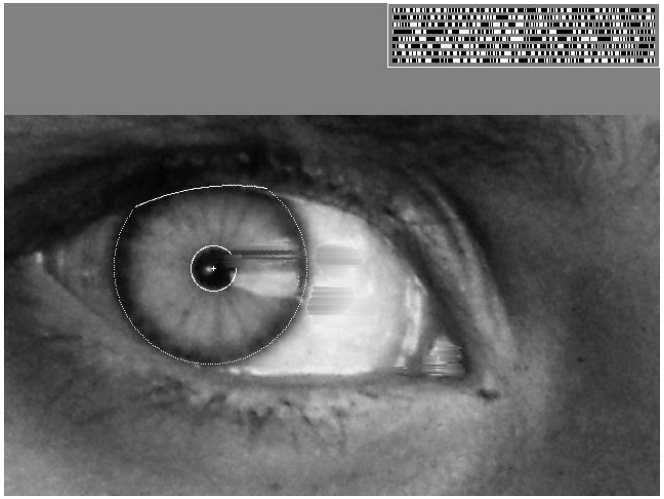
Object recognition (in supermarkets)

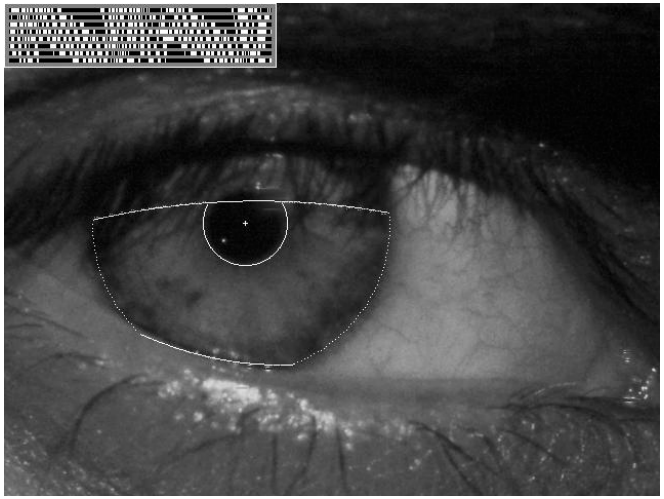
LaneHawk by EvolutionRobotics

"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… "

Vision-based biometrics

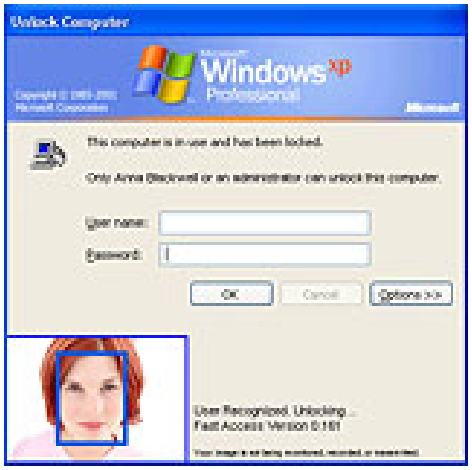
"How the Afghan Girl was Identified by Her Iris Patterns" Read the story wikipedia





Login without a password…

Fingerprint scanners on many new laptops, other devices



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

Object recognition (in mobile phones)

Point & Find, Nokia

Google

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Sports

Sportvision first downline

Nice explanation on www.howstuffworks.com

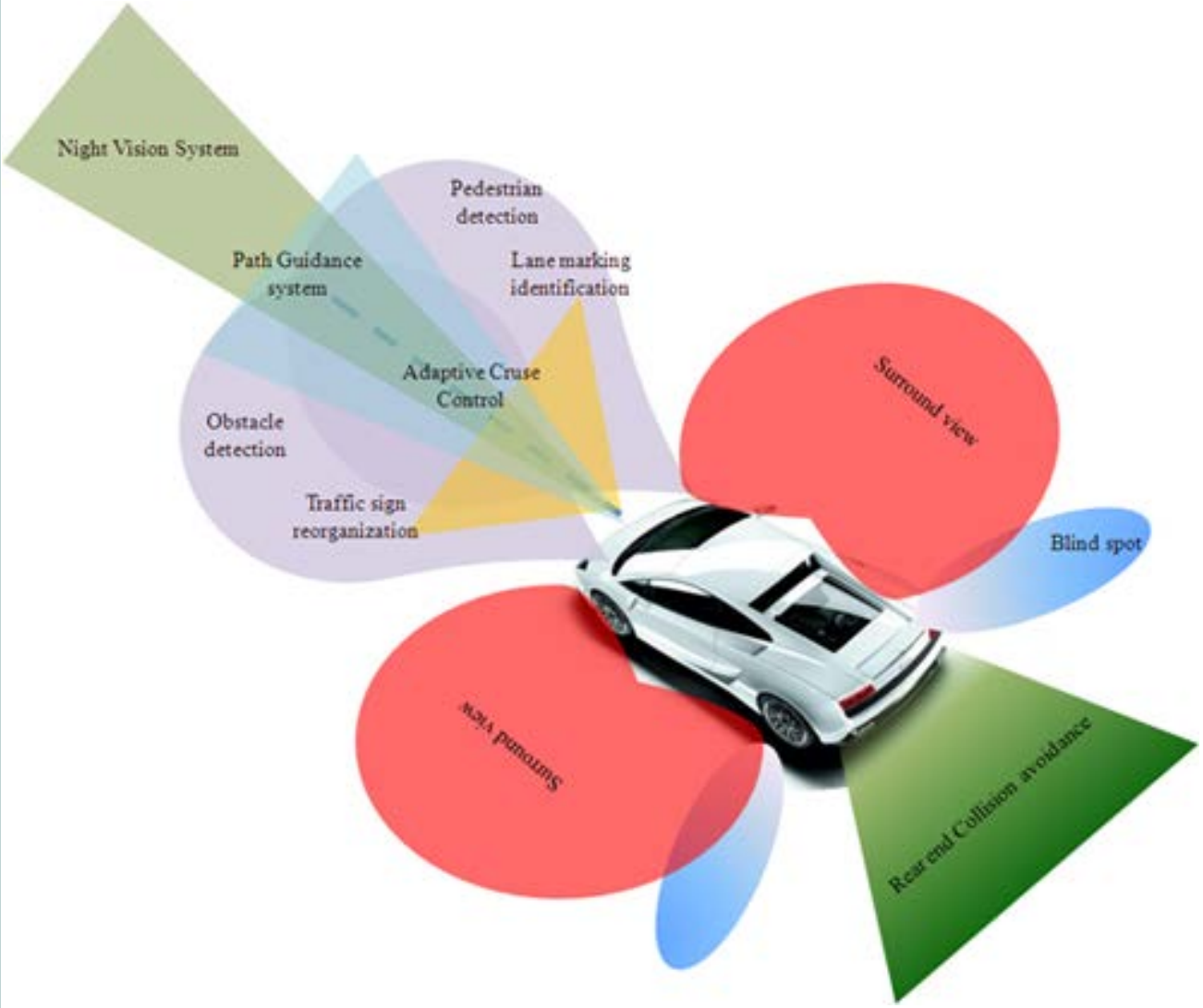
http://www.sportvision.com/video.html

Smart cars

- **Mobileye**
  - **Vision systems currently in high-end BMW, GM, Volvo, Volkswagen models**
  - **By 2010: 70% of car manufacturers.**

Slide content courtesy of Amnon Shashua

Advanced Driver Assistance Systems



ADAS applications

- Pedestrian Detection
- Blind Spot Detection
- Traffic Sign Detection
- Collision Avoidance
- Lane Detection
- Backup Obstacle Detection
- Occupant Sensing

Traffic Sign Recognition

Google cars

Oct 9, 2010. "Google Cars Drive Themselves, in Traffic". The New York Times. John Markoff

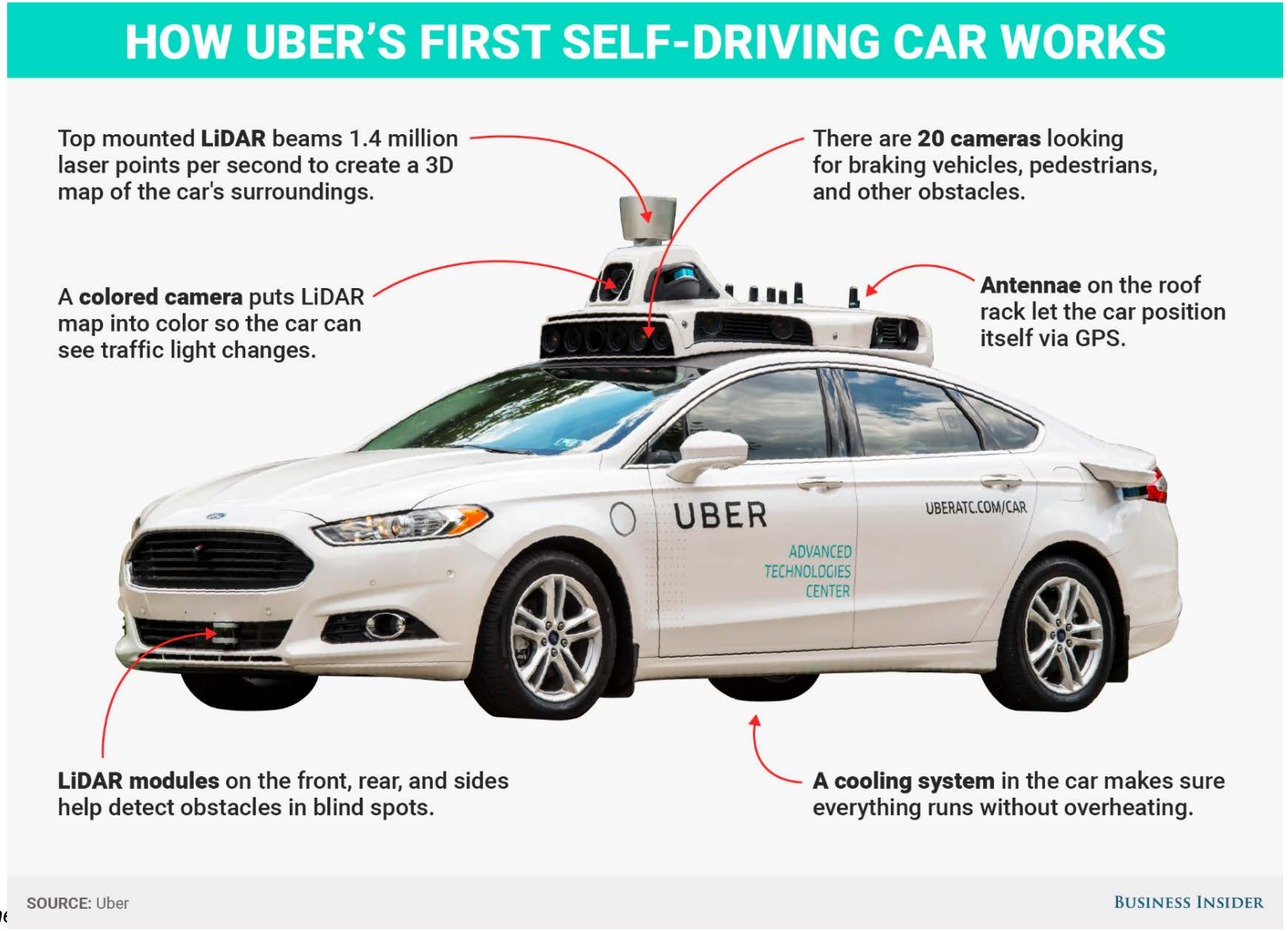
June 24, 2011. "Nevada state law paves the way for driverless cars". Financial Post. Christine Dobby

Aug 9, 2011, "Human error blamed after Google's driverless car sparks five-vehicle crash". The Star (Toronto)

Google cars

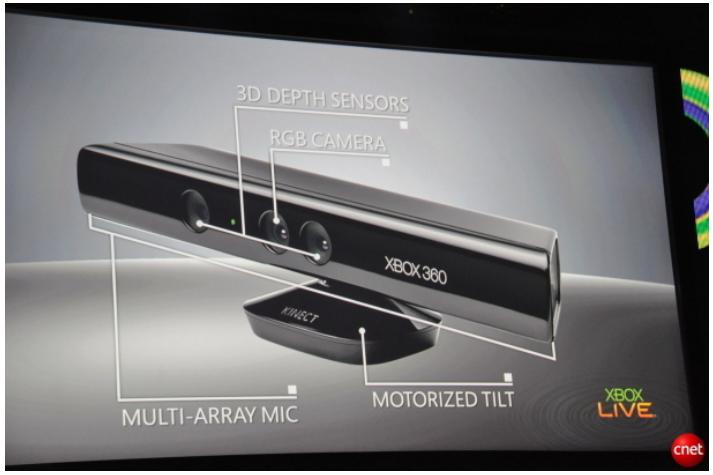
- **WAYMO – Self-driving car.**

Uber Self-Driving Car



Interactive Games: Kinect

- **Object Recognition:** http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- **Mario:** http://www.youtube.com/watch?v=8CTJL5lUjHg
- **3D:** http://www.youtube.com/watch?v=7QrnwoO1-8A
- **Robot:** http://www.youtube.com/watch?v=w8BmgtMKFbY



Vision in space

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

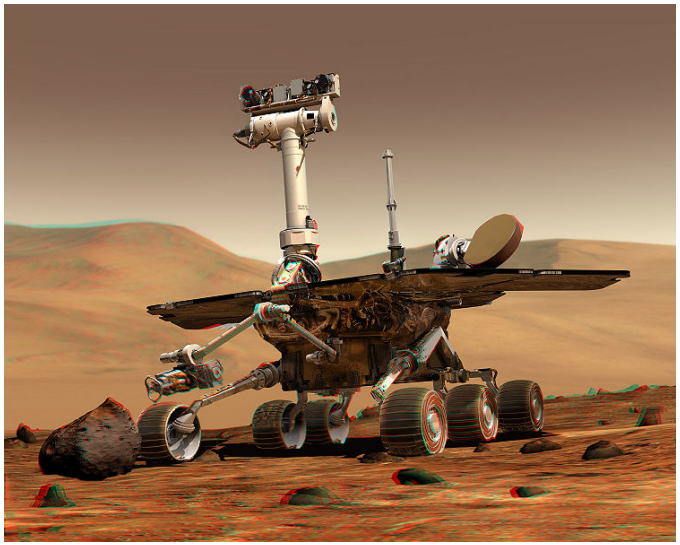
Vision systems (JPL) used for several tasks

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

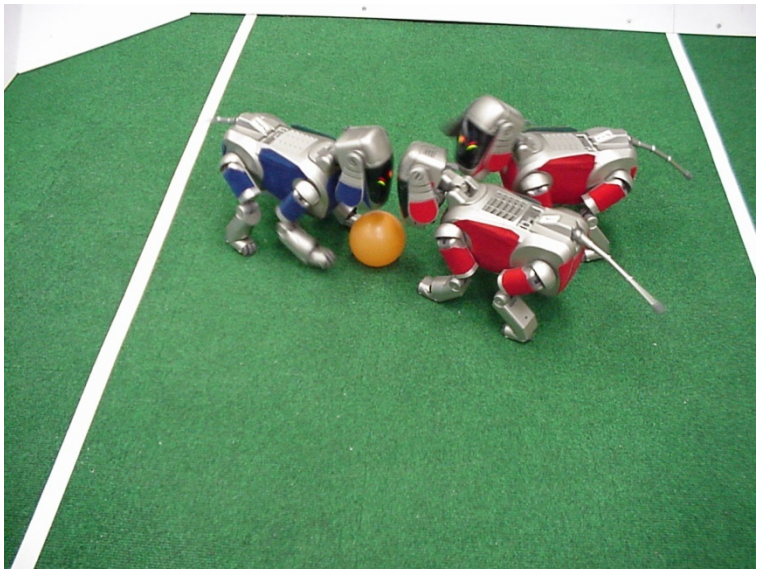
Industrial robots

Vision-guided robots position nut runners on wheels

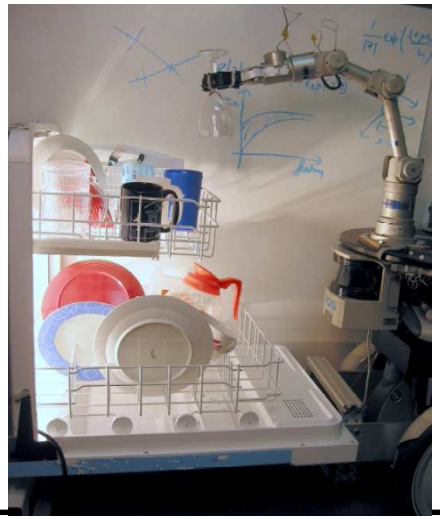
Mobile robots



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

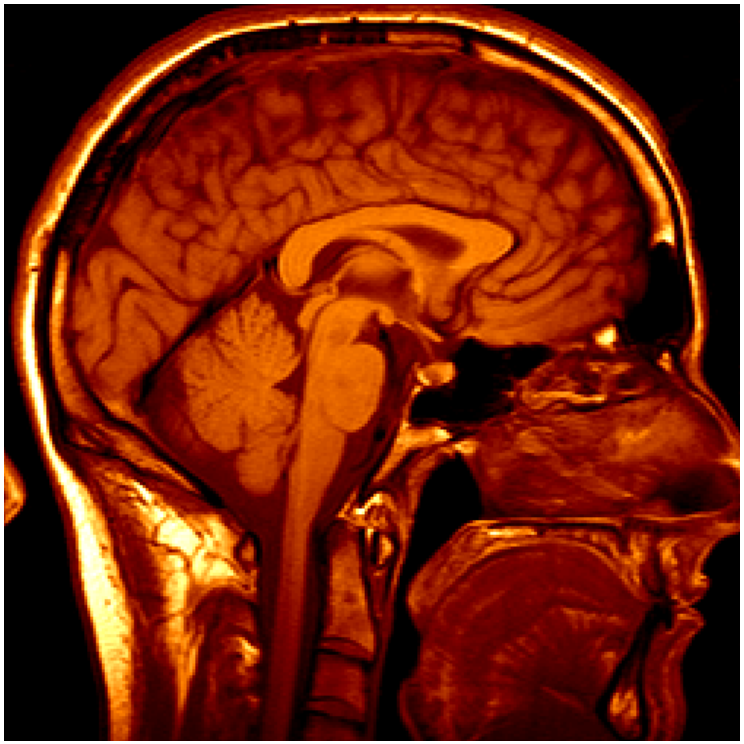


http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT