

CSC 391

Introduction to Computer

Vision

V. Paul Pauca and Rongzhong Li

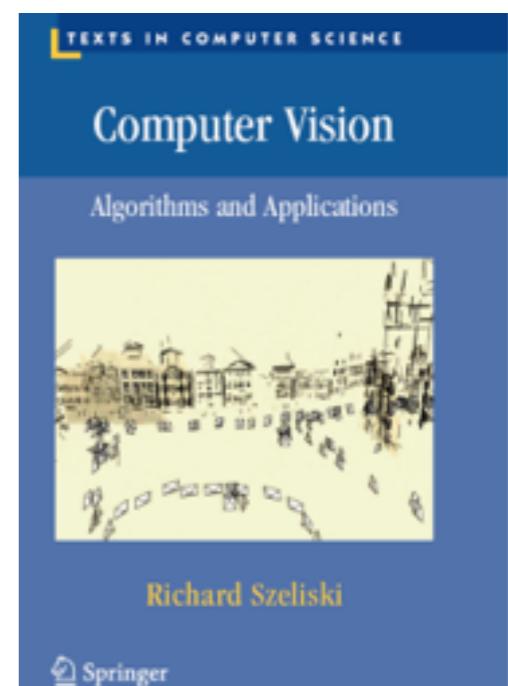
Fall 2016

Wake Forest University

Basic Info

- Instructors
 - V. Paul Pauca
 - Email: paucavp@wfu.edu
 - Office: Manchester Hall 235
 - Office hours: T-Th 9:00 - 11am
- Course webpage:
 - csweb.cs.wfu.edu/~pauca/CSC391-CV.html
- Textbook:

Computer vision: algorithms and applications by R. Szeliski



Course Details

- Grading
 - Programming projects: 40%
 - Final project: 20%
 - Midterm exam: 15%
 - Final exam: 20%
 - Attendance & participation: 5%
- Programming projects
 - OpenCV and Python
 - Weekly programming assignments
 - Turn in through Sakai



Course Details

- Final project
 - Individuals or up to groups of 2
 - Explore computer vision topics in robotics
 - Due on last week of class
- Midterm quiz and final
 - Test material comprehension as well as application
- Attendance and participation
 - Regular attendance required, participation helps a lot

Course Details

- Academic integrity
 - Can work with other students
 - But programming assignments are individual
 - Can use online resources as long as sources are credited

What is computer vision?

- *Automatic* understanding of images and video
 - Compute 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)

Human vision

- Easily perceive 3-d structure of the world
 - Shape, translucency, segmentation, recognition, etc.
- *How does it work?*
- Can computers match or beat human vision?

Humans vs computers



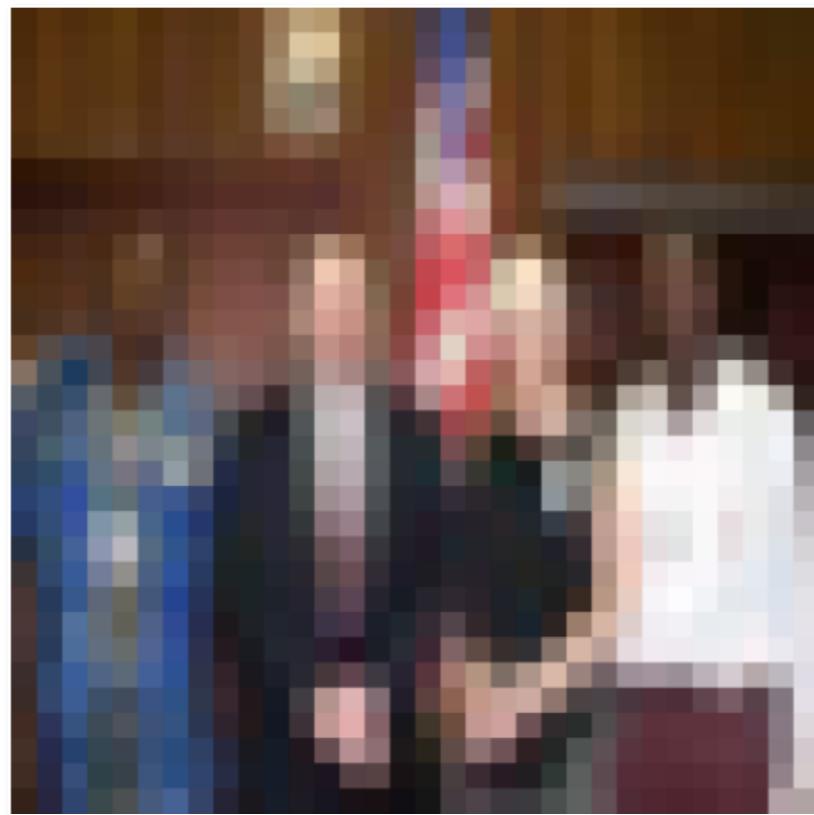
What we see

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

- The goal in computer vision is to extract *meaning* from pixels

Humans vs computers



Humans are remarkably good at extracting meaning from pixels

Human vs. computers

- But human perception has its shortcomings...

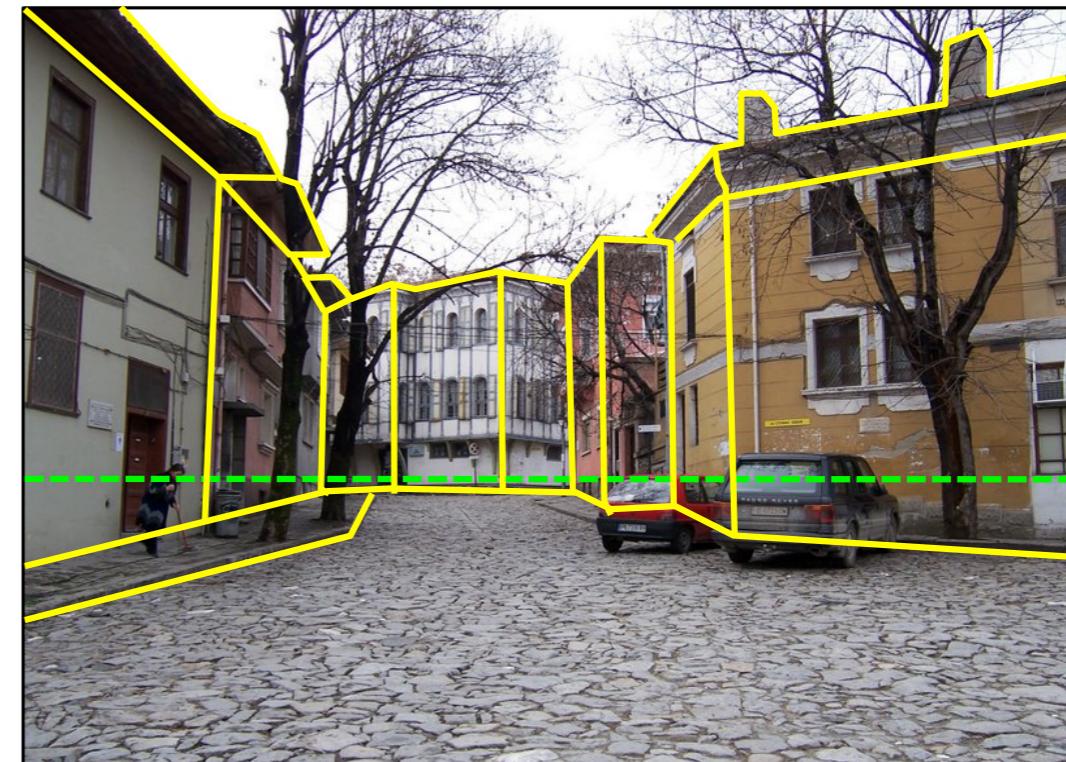


Sinha and Poggio, *Nature*, 1996

What kind of information can be extracted from an image?

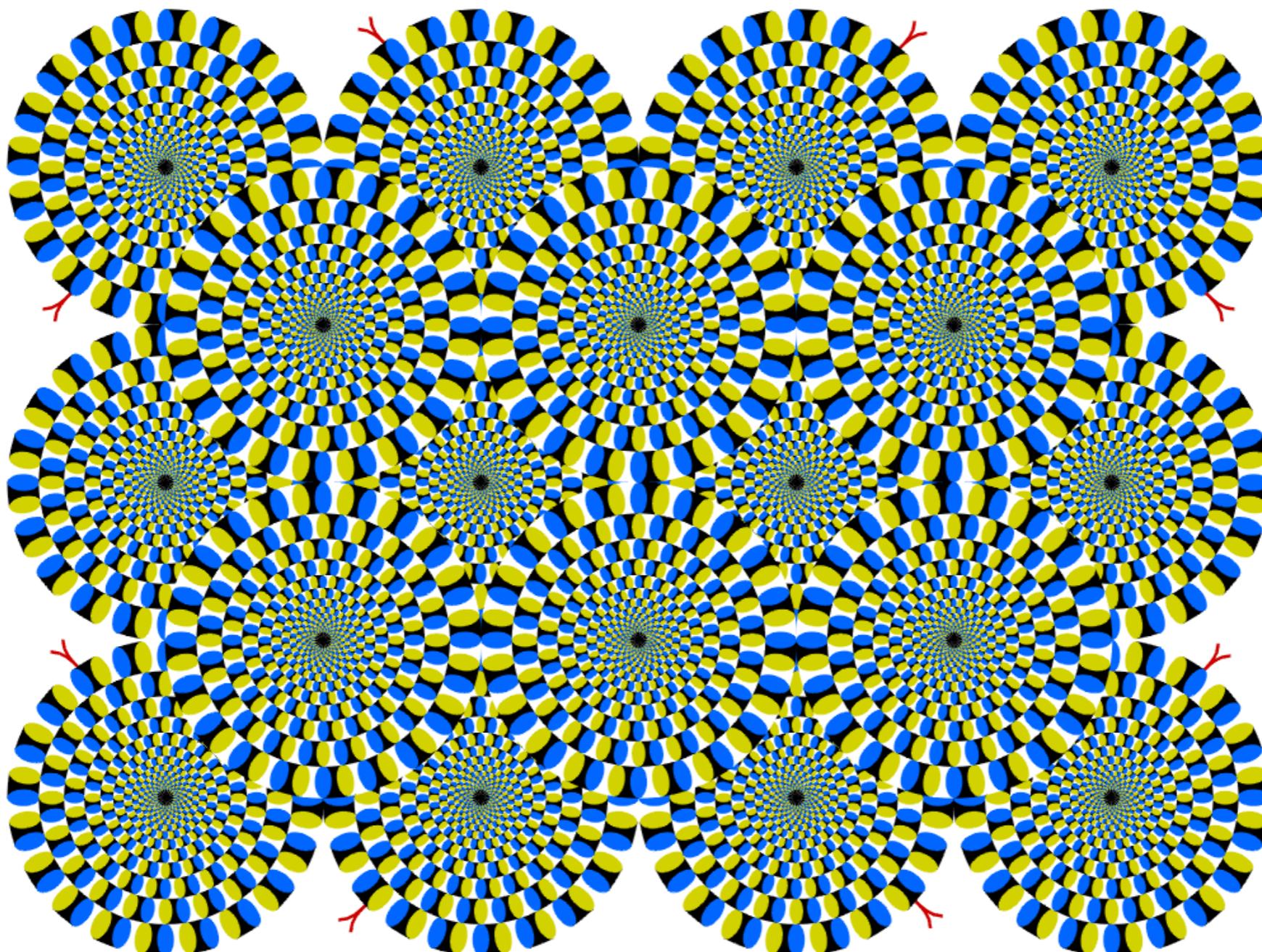


Semantic information



Geometric information

Humans vs. computers

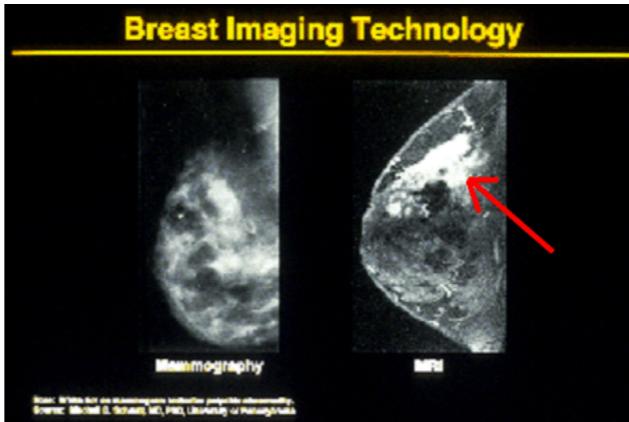


Copyright A. Kitaoka 2003

Why computer vision matters



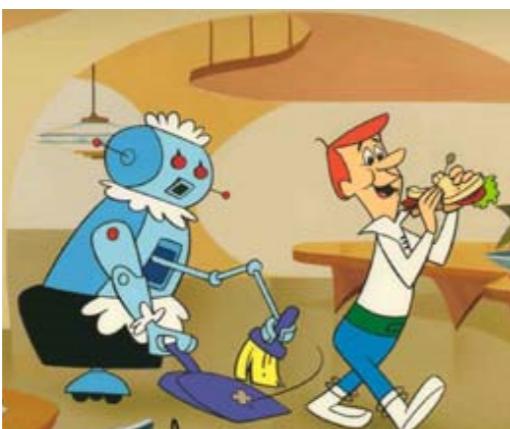
Safety



Health



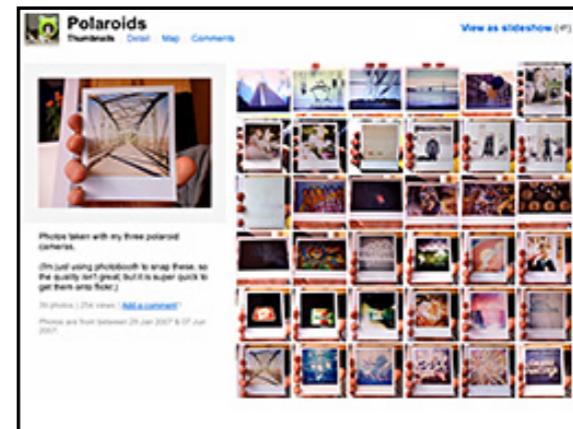
Security



Comfort



Fun

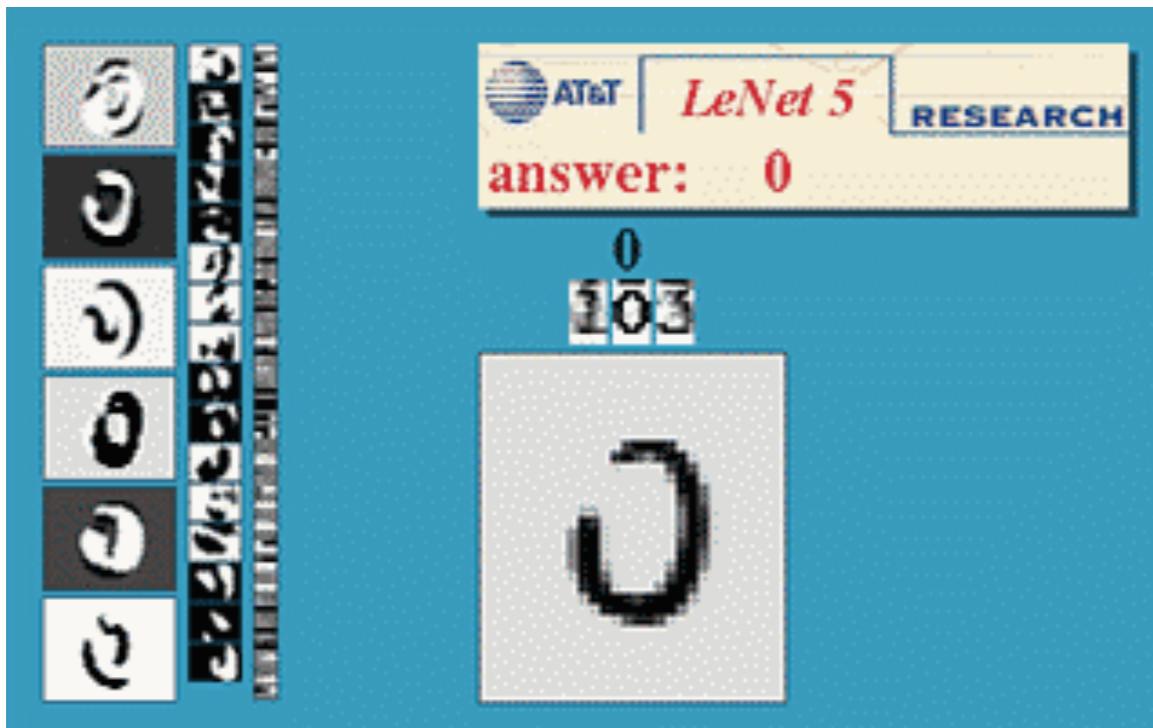


Access

Credit: J. Hays

Success of CV to date

Optical character recognition (OCR)



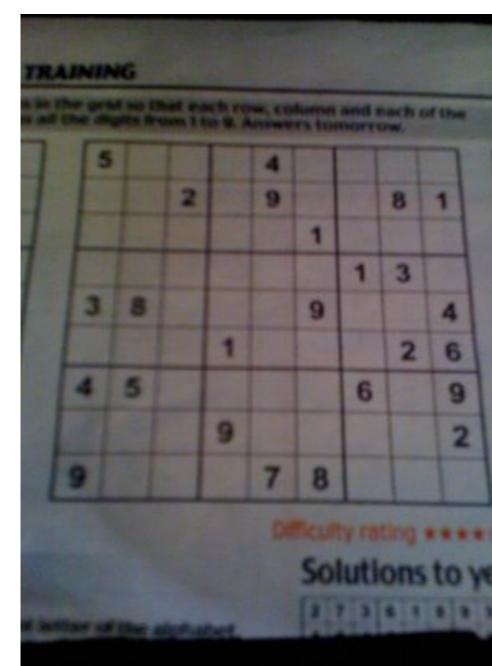
Digit recognition
yann.lecun.com



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

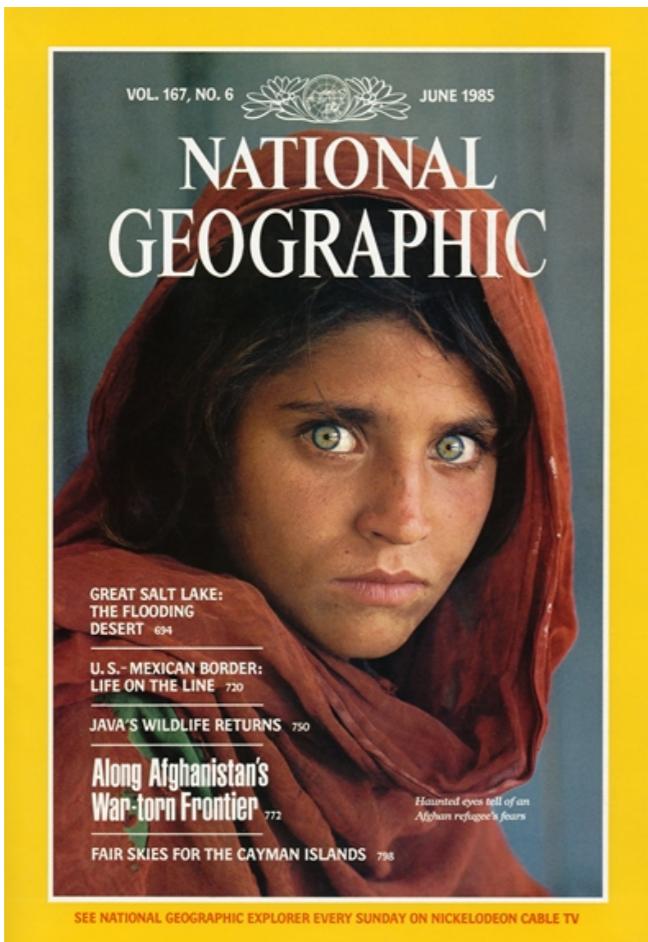


Fingerprint scanners on many new laptops, other devices

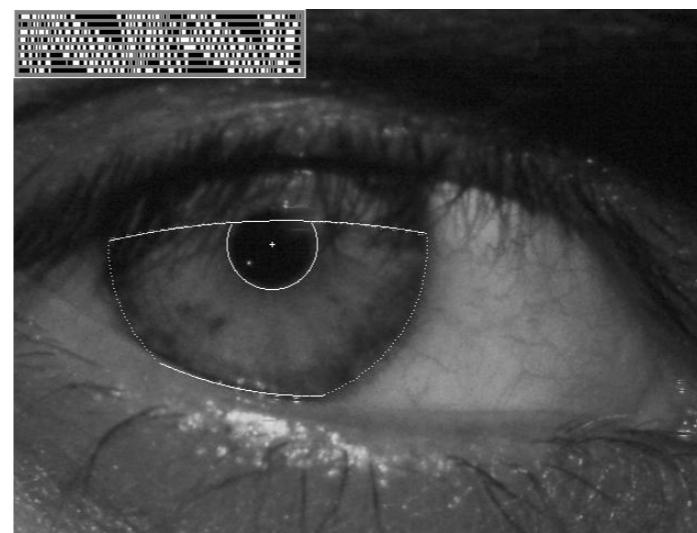
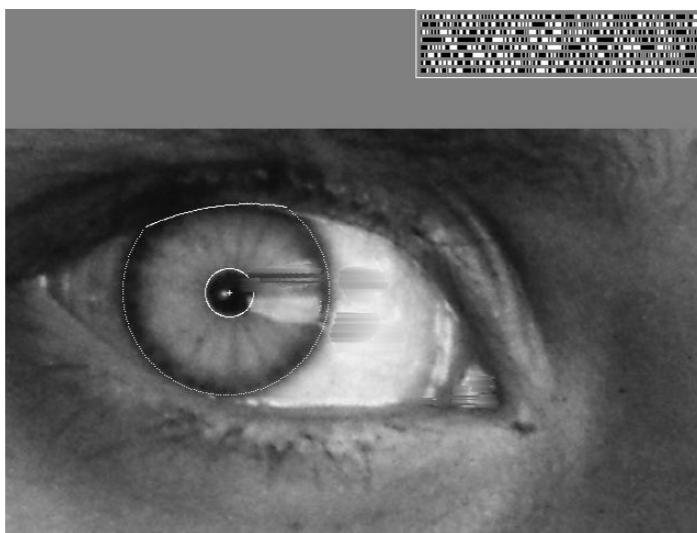


Face recognition systems now beginning to appear more widely
<http://www.sensiblevision.com/>

Biometrics



How the Afghan Girl was Identified by Her Iris Patterns



Source: S. Seitz

Face detection

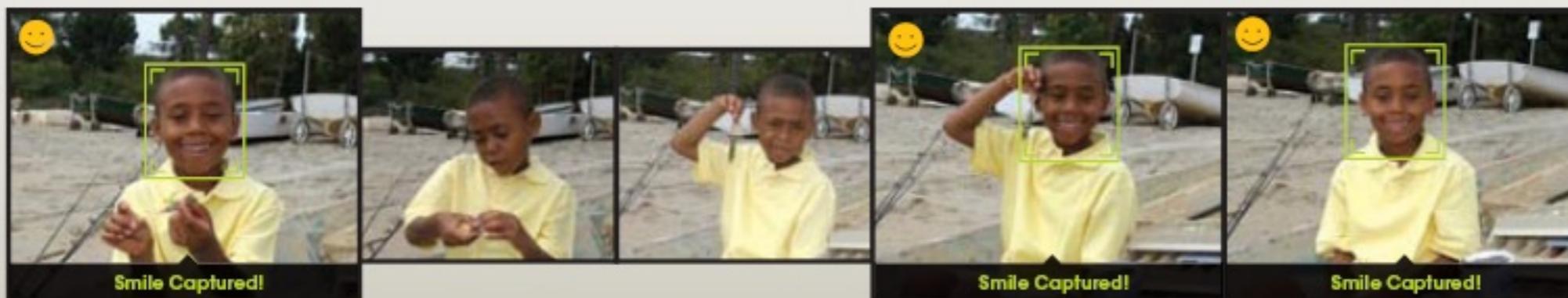
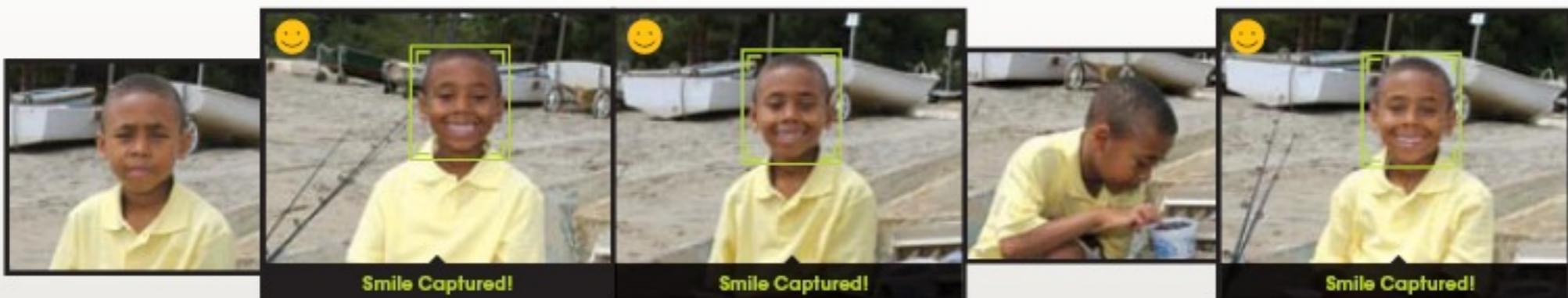


Many consumer digital cameras now detect faces

Smile detection

The Smile Shutter flow

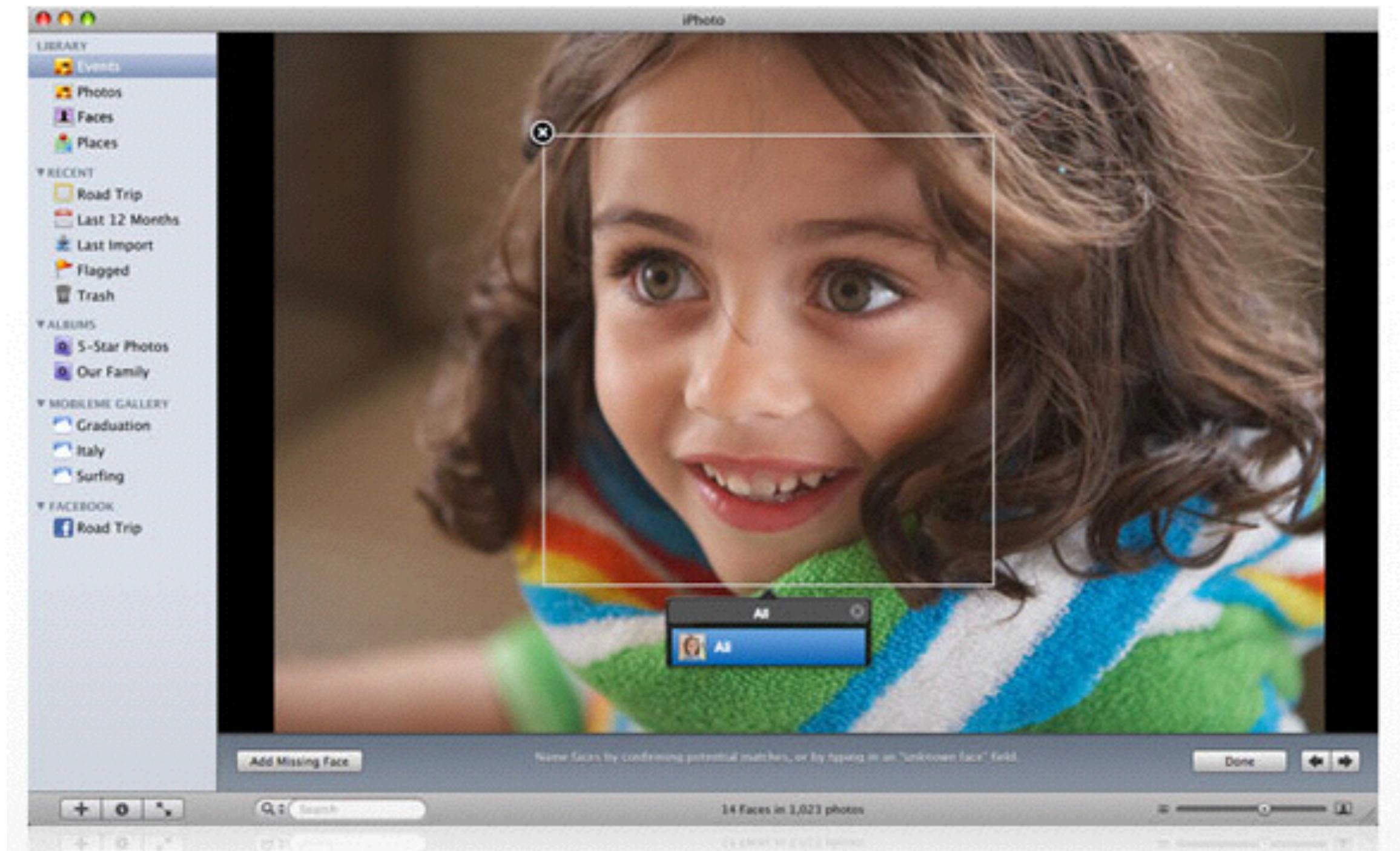
Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Privacy protection



Face recognition: Apple iPhoto software



Visual search: Google search by image



Four ways to search by image



Drag and drop

Drag and drop an image from the web or your computer into the search box on images.google.com.



Upload an image

On images.google.com, click the camera icon, then select "Upload an image." Select the image you want to use to start your search.



Copy and paste the URL for an image

Found an image on the web you're curious about? Right-click the image to copy the URL. On images.google.com, click the camera icon, and "Paste image URL".

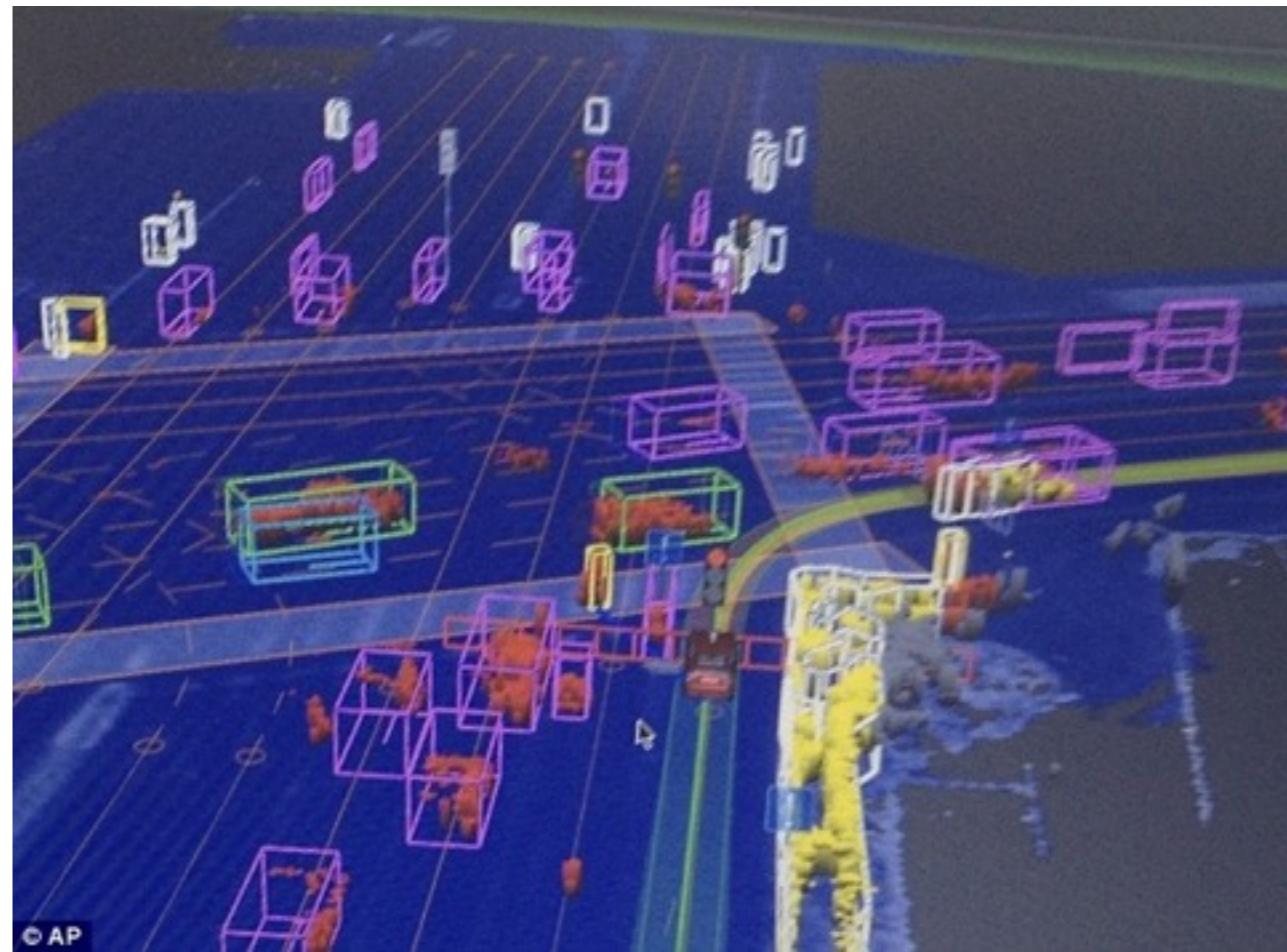


Right-click an image on the web

To search by image even faster, download the [Chrome extension](#) or the [Firefox extension](#). With the extension installed, simply right-click an image on the web to search Google with that image.

Source: S. Lazebnik

Google self-driving cars



- Google's self-driving car passes 1.5 million miles
- Nissan pledges affordable self-driving car models by 2020
(CNET, 8/27/2013)

Automotive safety

The screenshot shows the Mobileye website interface. At the top, there are navigation tabs: 'manufacturer products' (with arrows) and 'consumer products' (with arrows). Below this is a main heading 'Our Vision. Your Safety.' with a central image of a car from above, showing three cameras: 'rear looking camera' (viewing the rear), 'forward looking camera' (viewing the front), and 'side looking camera' (viewing the sides). To the right is a 'News' sidebar with two articles: 'Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System' and 'Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end'. Below the main heading are three product cards: 'EyeQ Vision on a Chip' (showing a chip image), 'Vision Applications' (showing a pedestrian crossing scene), and 'AWS Advance Warning System' (showing a screen with a car icon and '0.8'). Each card has a 'read more' link.

▷ manufacturer products consumer products ◀◀

Our Vision. Your Safety.

rear looking camera forward looking camera side looking camera

▶ EyeQ Vision on a Chip



▶ Vision Applications



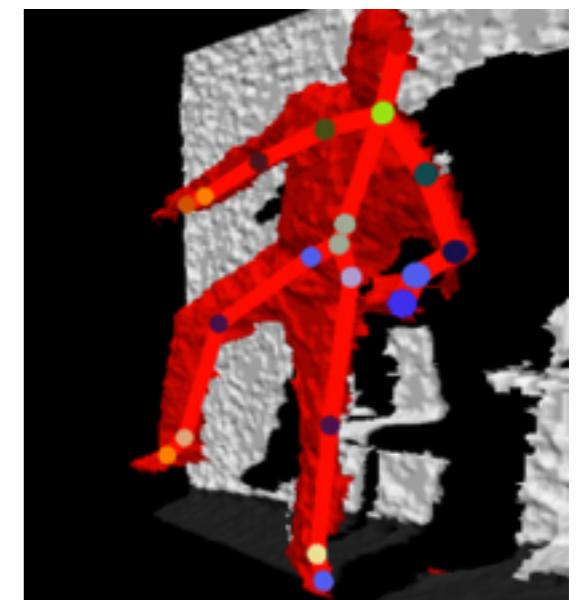
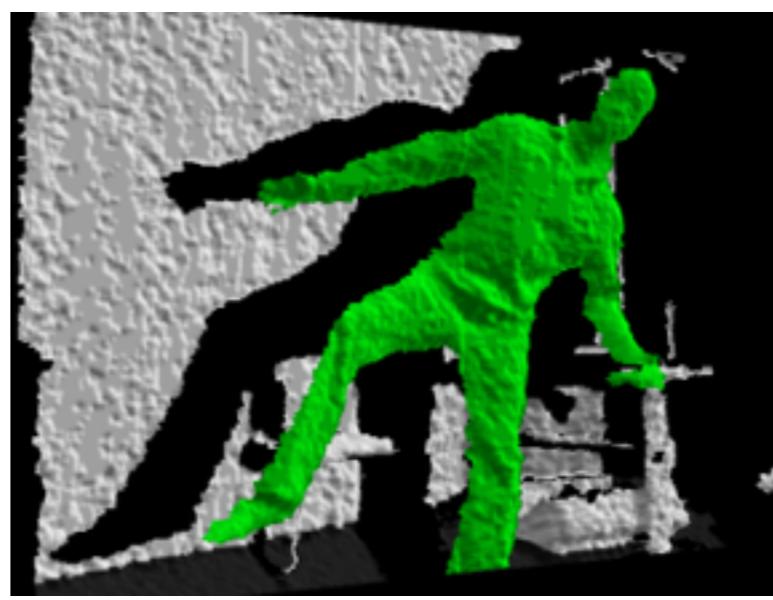
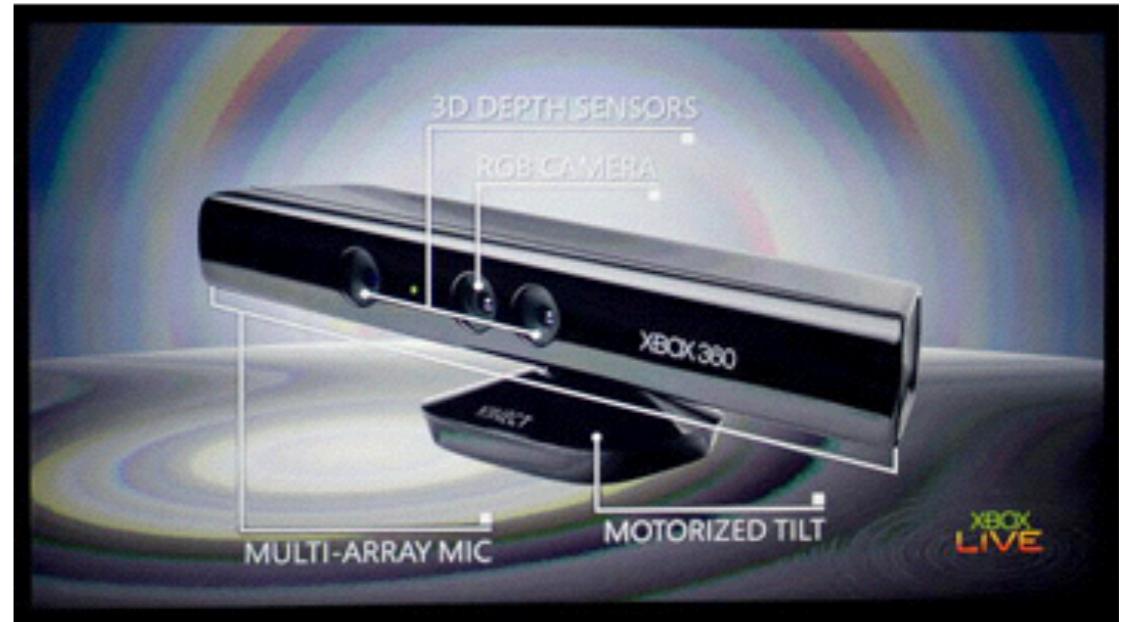
▶ AWS Advance Warning System



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

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

Vision-based interaction: Xbox Kinect



3D Reconstruction: Kinect Fusion



[YouTube Video](#)

Source: S. Lazebnik

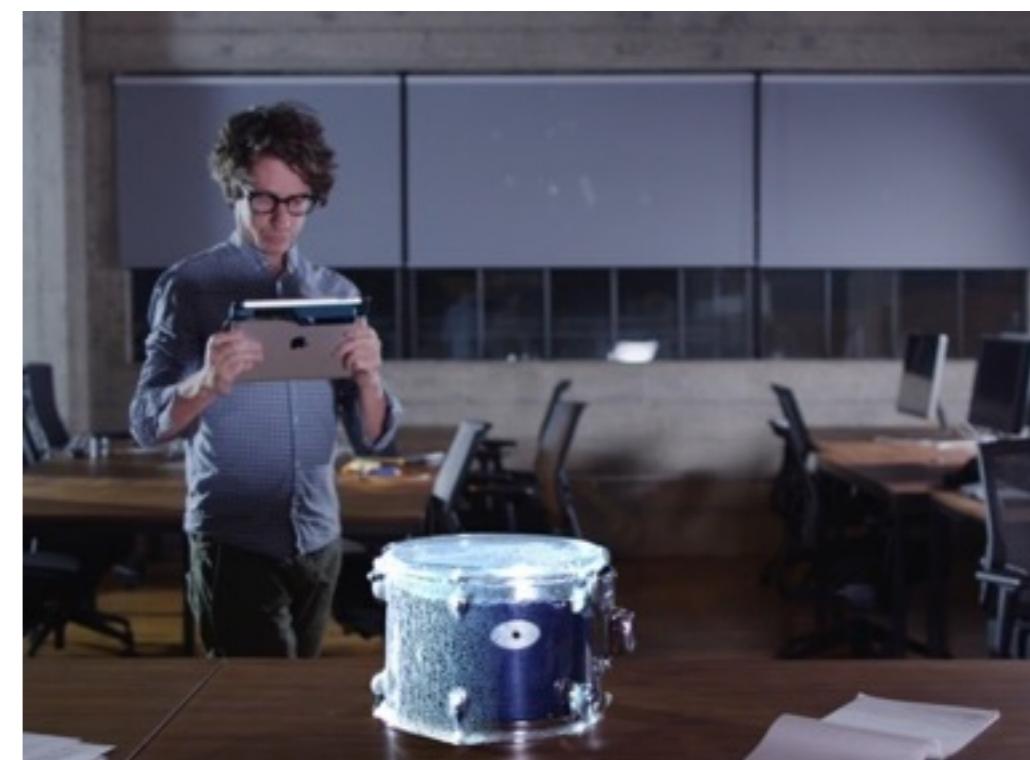
Occipital: 3D structure sensor for iPad

Structure Sensor: Capture the World in 3D

by Occipital

Home Updates 2 Backers 2,253 Comments 155

San Francisco, CA Hardware

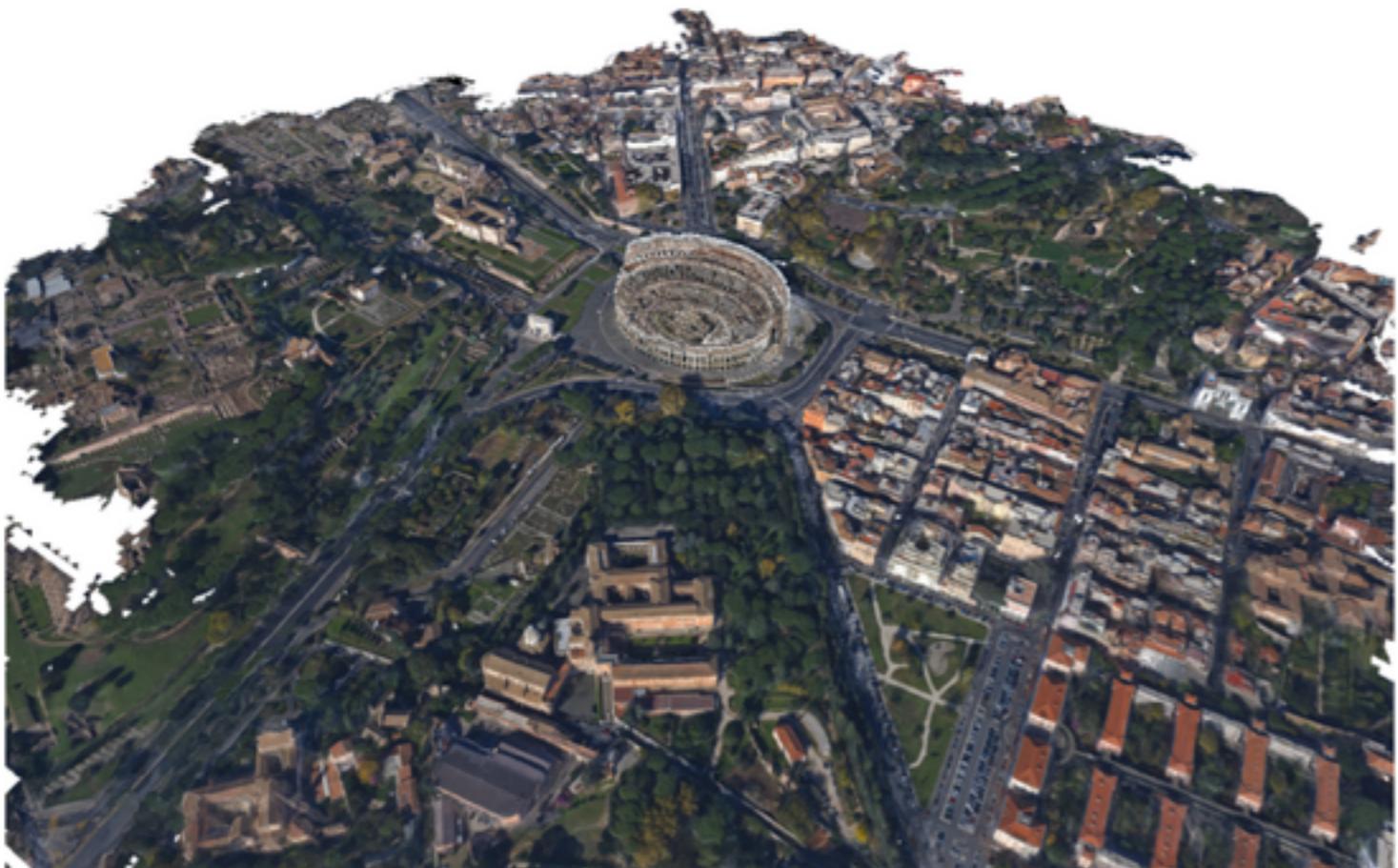


<http://www.kickstarter.com/projects/occipital/structure-sensor-capture-the-world-in-3d>

Source: S. Lazebnik

3D reconstruction from photo collections

Colosseum, Rome, Italy



San Marco Square, Venice, Italy

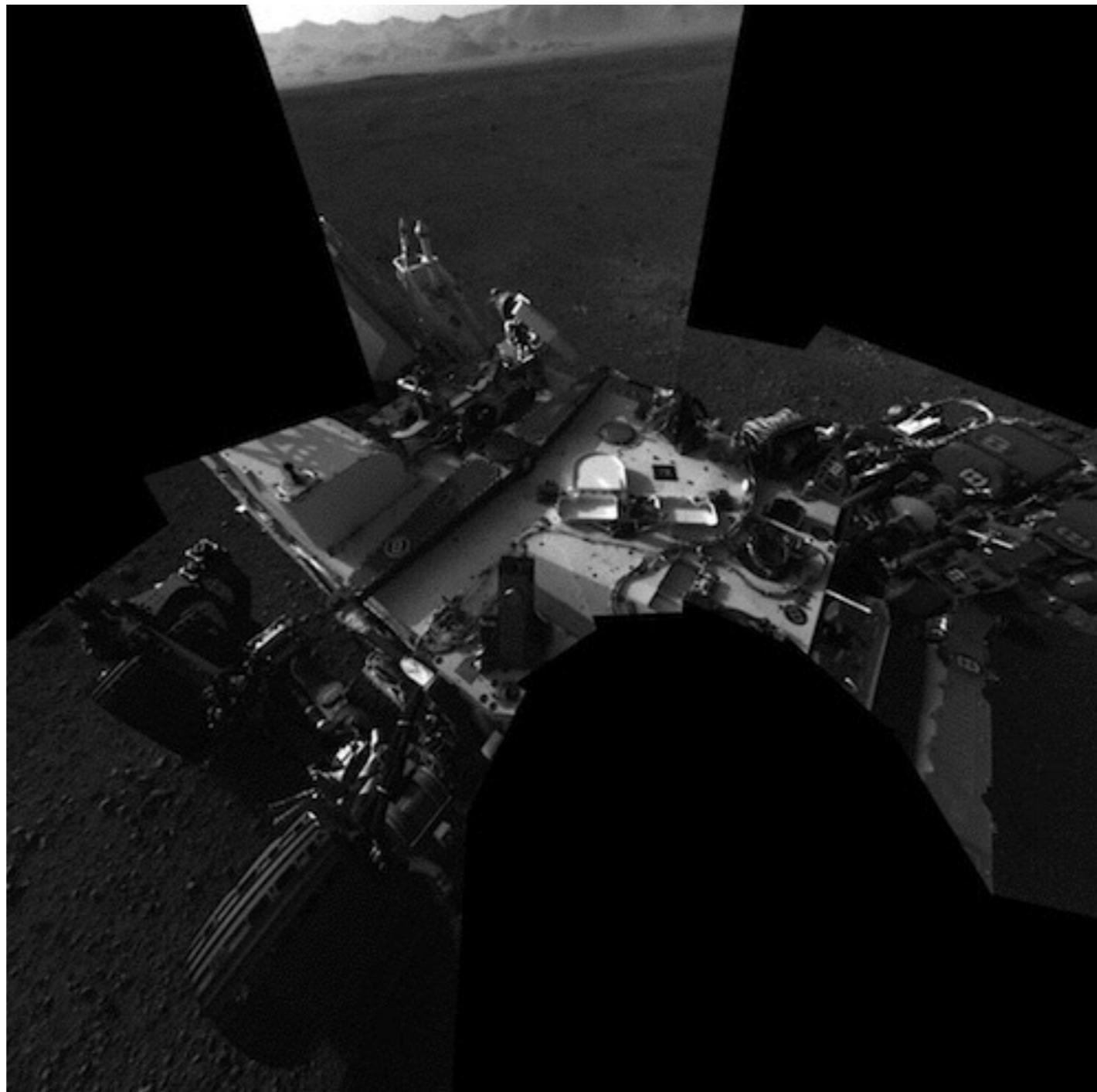


Q. Shan, R. Adams, B. Curless, Y. Furukawa, and S. Seitz, 3DV 2013

Special effects: shape and motion capture



Vision for robotics, space exploration

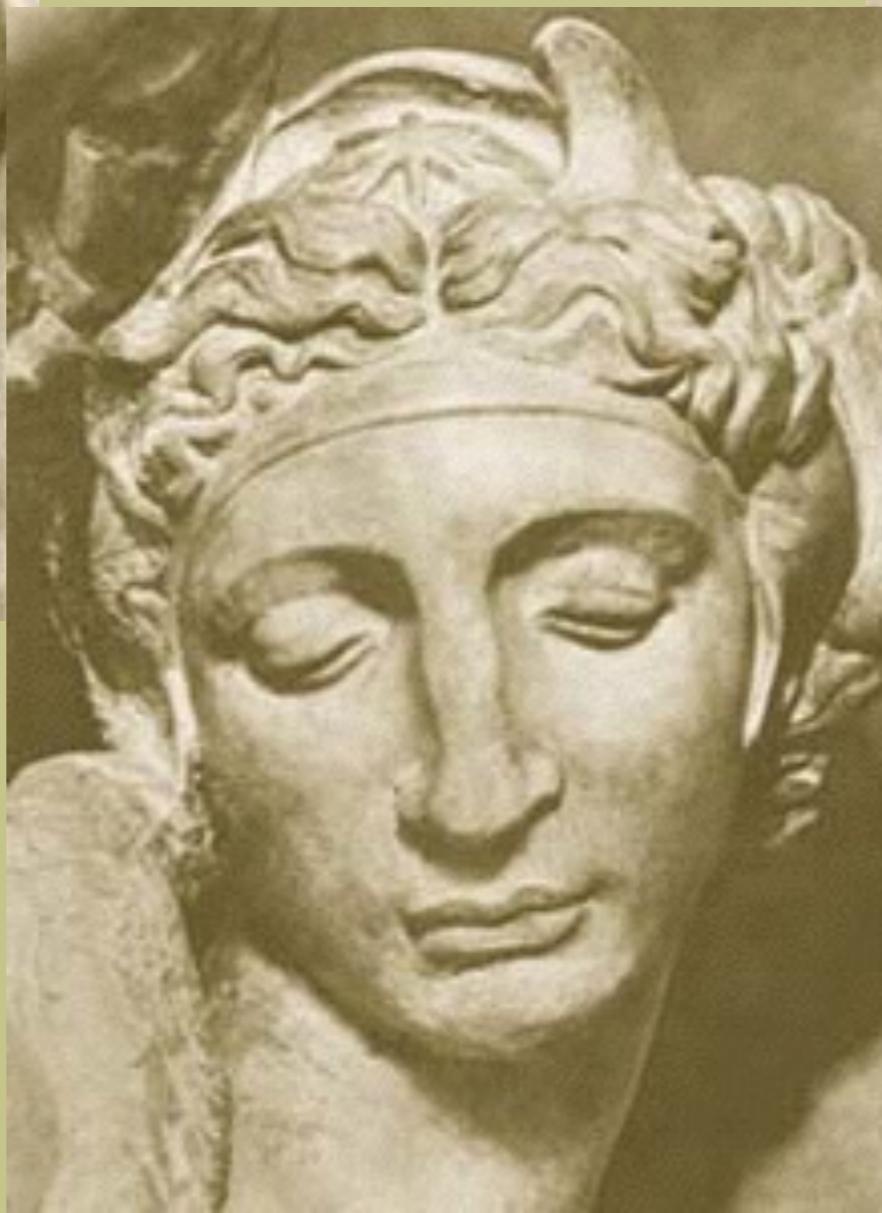


NASA's Curiosity Rover has a system consisting of 17 cameras

Source: S. Lazebnik

Why is computer vision difficult?

Challenges: viewpoint variation



Michelangelo 1475-1564

Source: S. Lazebnik
slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

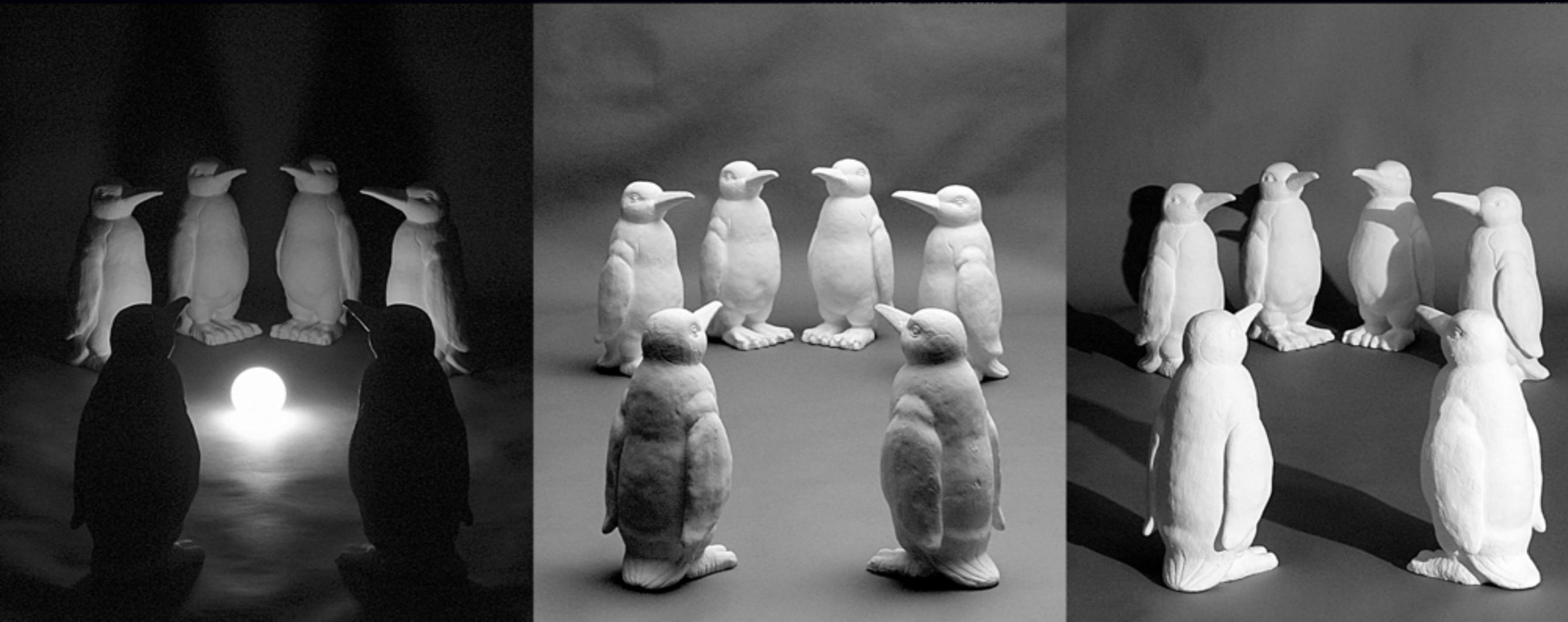


image credit: J. Koenderink

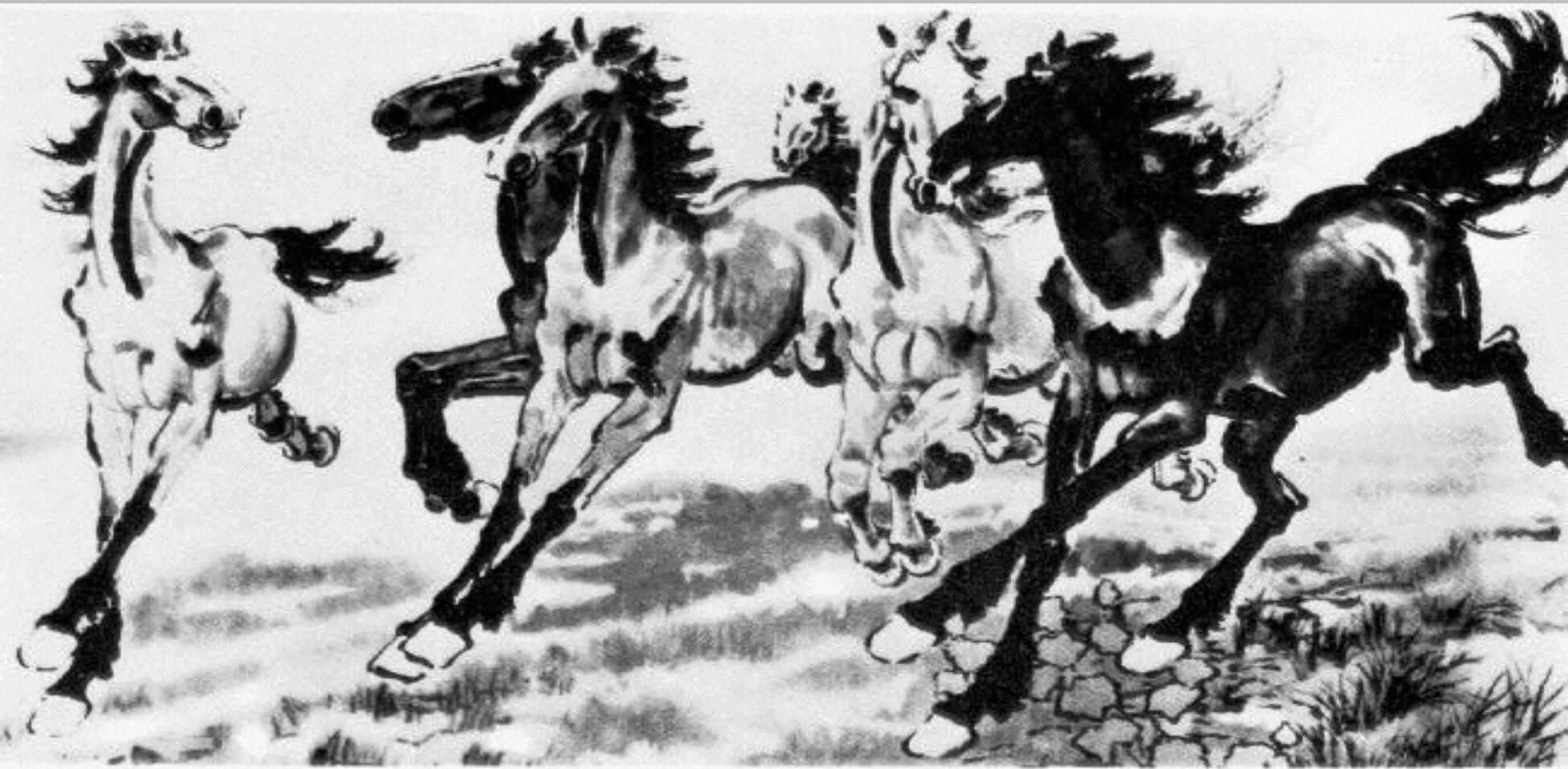
Challenges: scale



Source: S. Lazebnik

slide credit: Fei-Fei, Fergus & Torralba

Challenges: deformation



Xu, Beihong 1943

Source: S. Lazebnik
slide credit: Fei-Fei, Fergus & Torralba

Challenges: object intra-class variation



Source: S. Lazebnik
slide credit: Fei-Fei, Fergus & Torralba

Challenges: occlusion, clutter



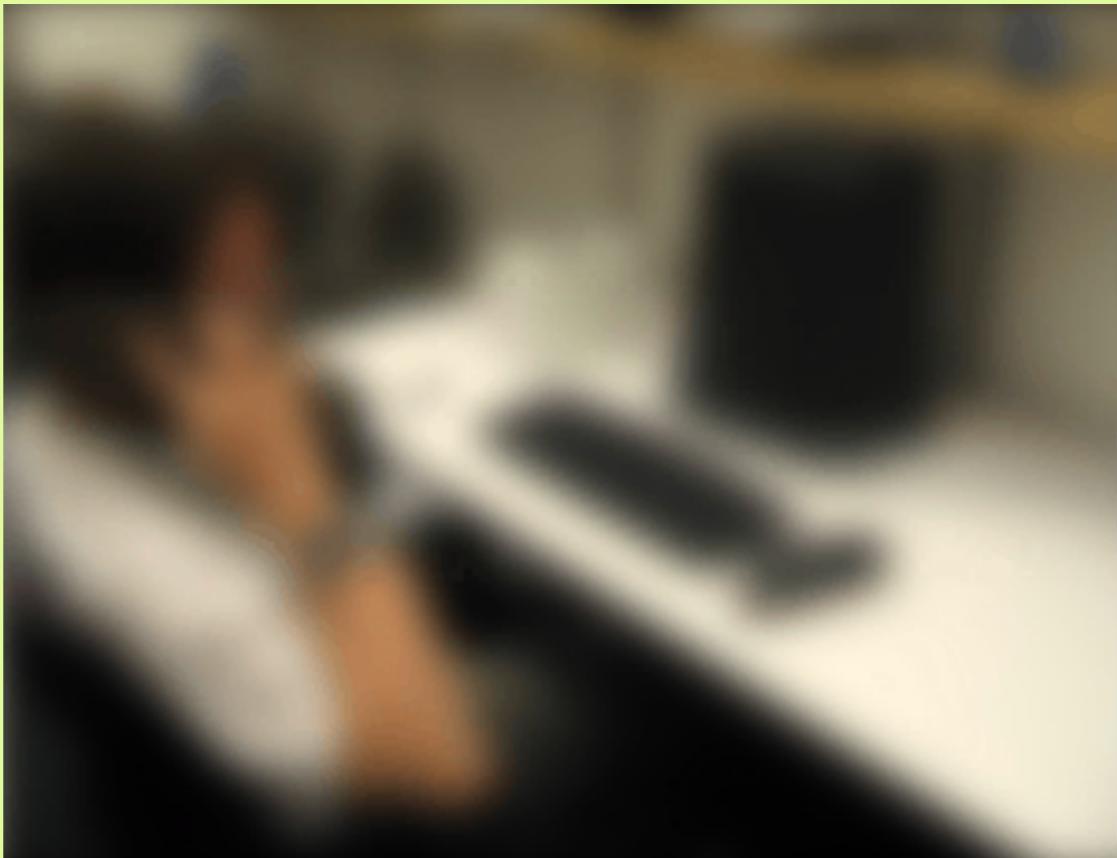
Image source: National Geographic
Source: S. Lazebnik

Challenges: Motion



Source: S. Lazebnik

Challenges: ambiguity



Source: Rob Fergus and Antonio Torralba

Challenges: ambiguity



Review: Intro to computer vision

- State-of-the-art applications
- Challenges of vision
 - Viewpoint and lighting variation
 - Intra-class variations: size, shape, deformation, etc.
 - Nuisances: motion, blur, noise, etc.
 - Intrinsic ambiguity

Challenges or opportunities?

- Images are confusing, but they also reveal the structure of the world through numerous cues



Depth cues: Linear perspective



Source: S. Lazebnik

Depth cues: Parallax



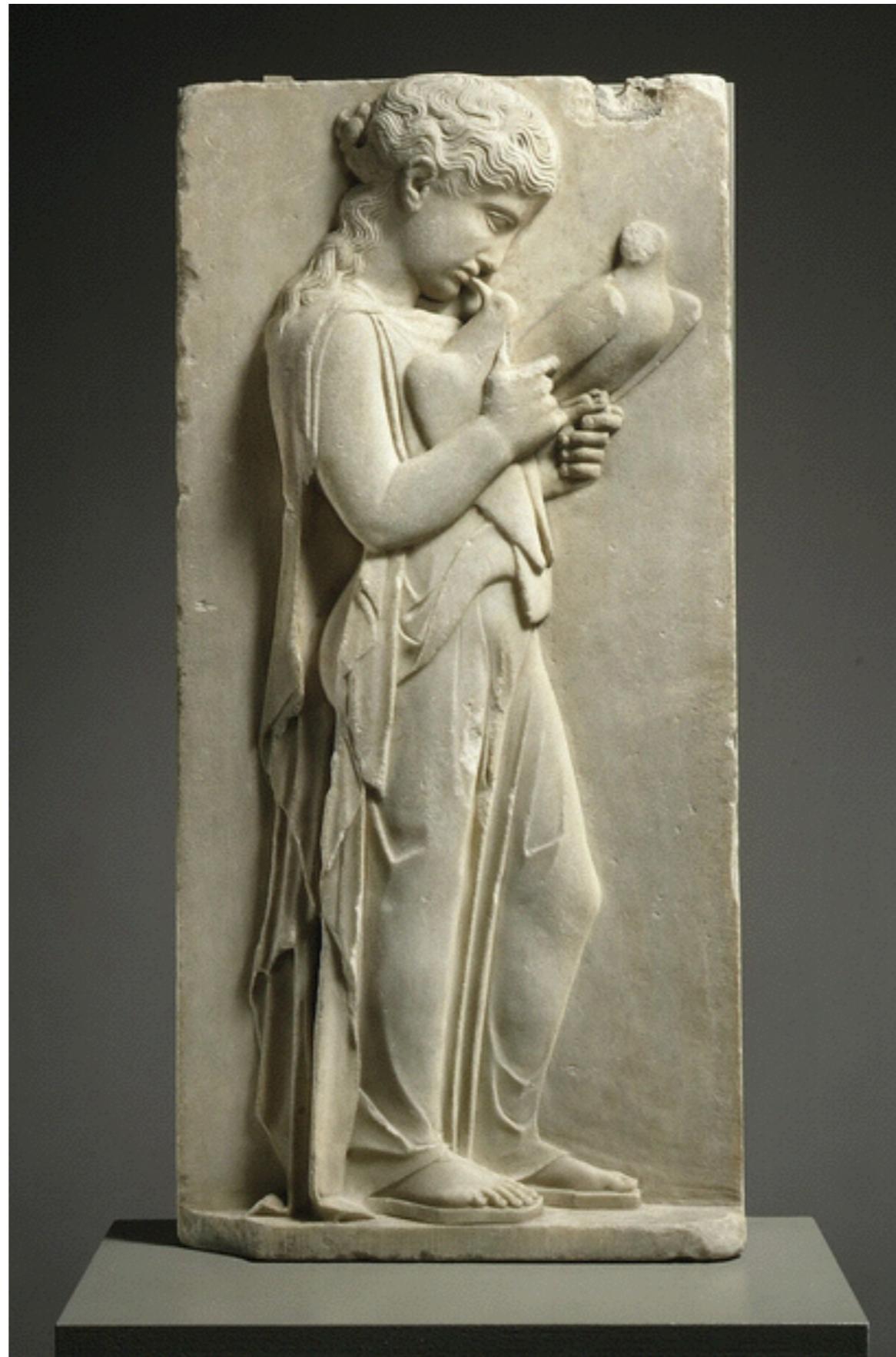
Source: S. Lazebnik

Shape cues: Texture gradient



Source: S. Lazebnik

Shape and lighting cues: Shading



Source: S. Lazebnik

Grouping cues: Similarity (color, texture, proximity)



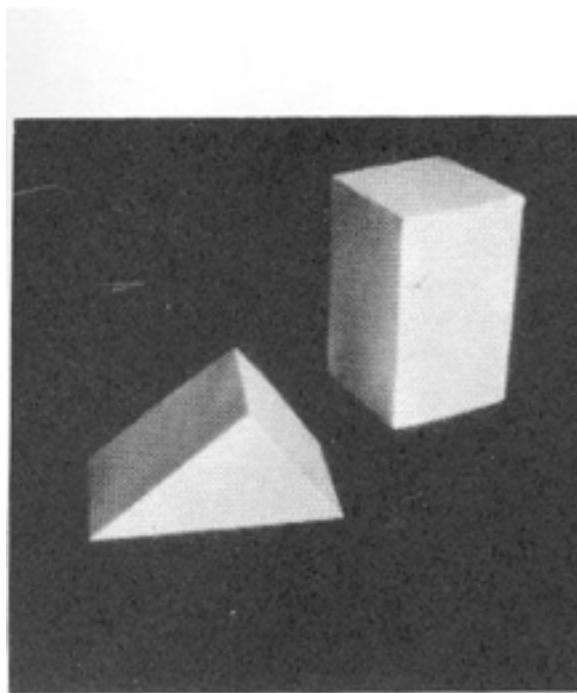
Source: S. Lazebnik

Grouping cues: “Common fate”

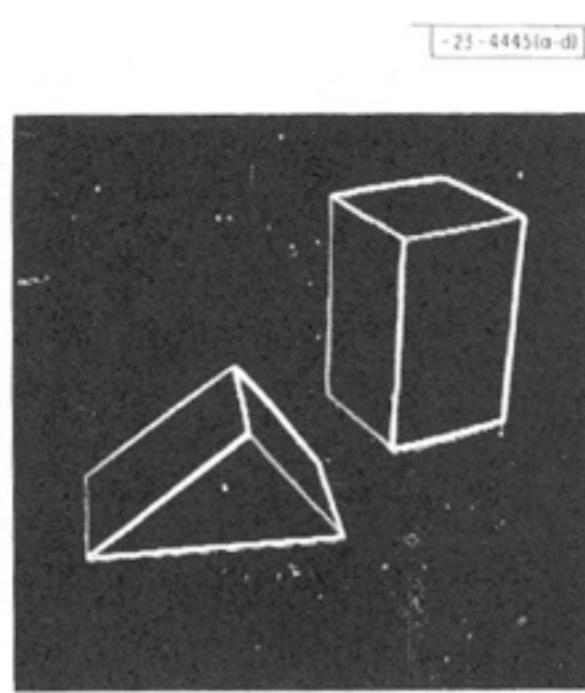


Source: S. Lazebnik

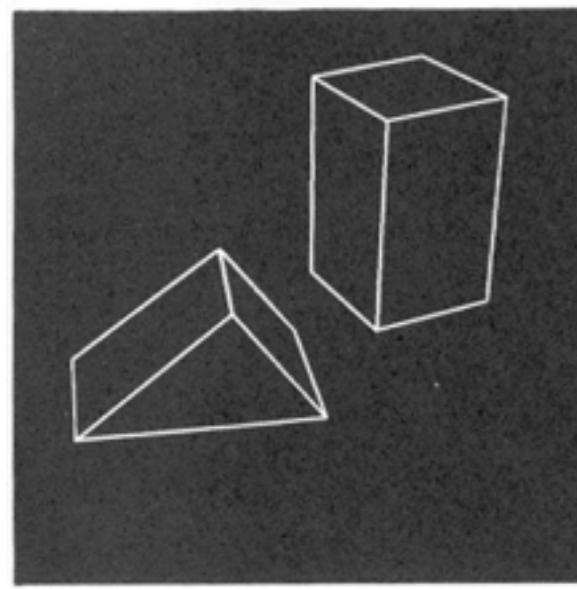
Origins of computer vision



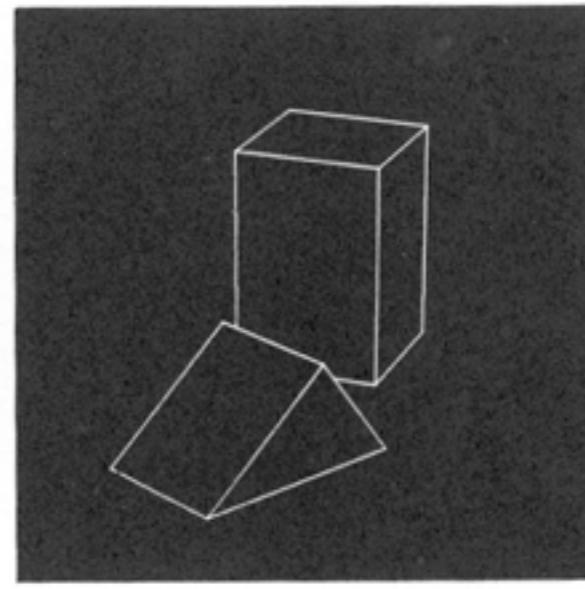
(a) Original picture.



(b) Differentiated picture.



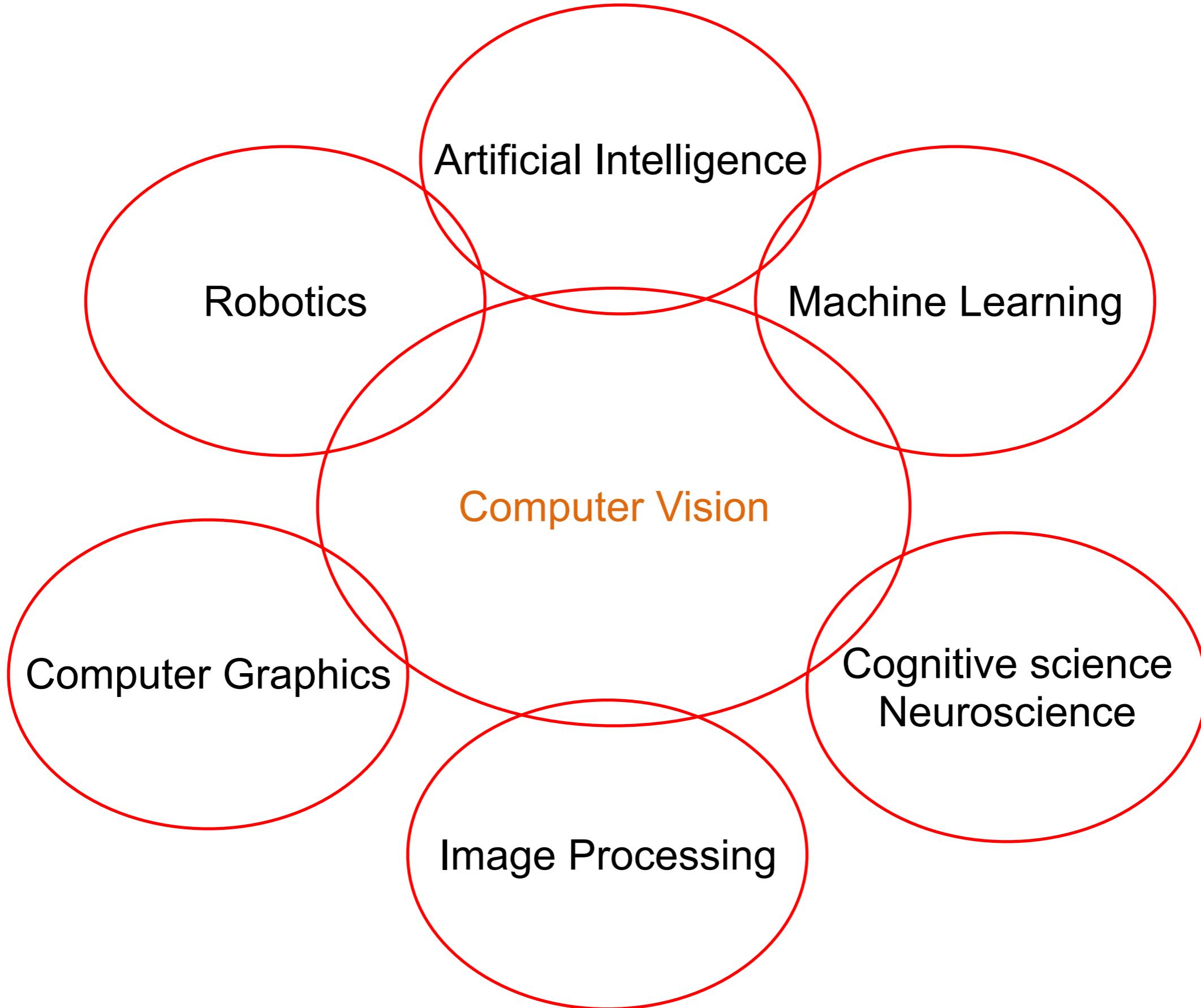
(c) Line drawing.



(d) Rotated view.

L. G. Roberts *Machine Perception
of Three Dimensional Solids*

Connections to other disciplines



The computer vision industry

- A list of companies here:

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