

Introduction

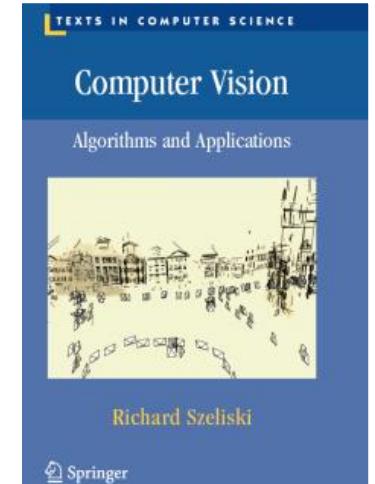
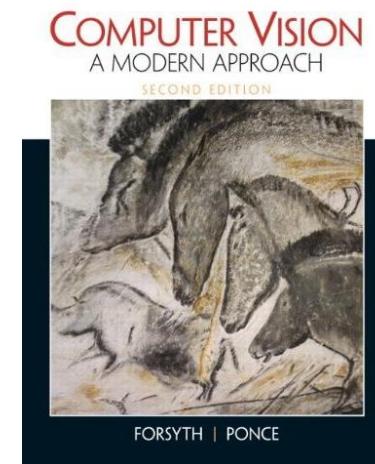
BIL719 – Computer Vision

Pinar Duygulu

Hacettepe University

Basic Info

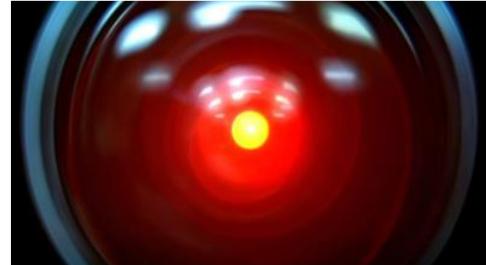
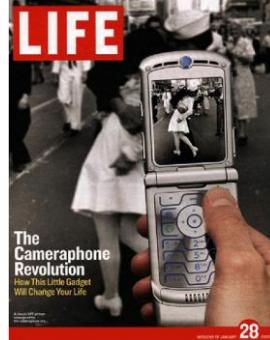
- Textbooks (suggested):
Forsyth & Ponce, *Computer Vision: A Modern Approach*
- Richard Szeliski, *Computer Vision: Algorithms and Applications* ([available online](#))



Course requirements

- Homeworks – 50%
- Quizzes – 20%
- Project – 30 %

Why study computer vision?



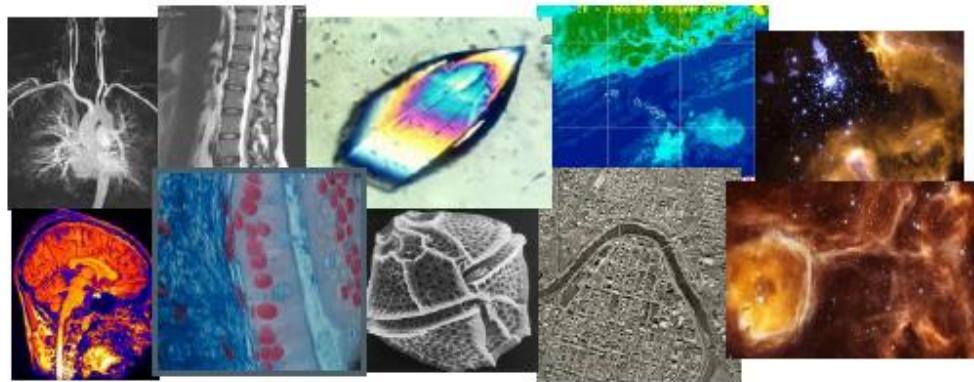
Why study computer vision?

An image is worth 1000 words

Images and movies are everywhere



Surveillance and security



Medical and scientific images

Massive amounts of visual data



flickr

YAHOO!
SCREEN

- For YouTube alone
 - More than 1 billion unique users
 - Hundreds of millions of hours are watched every day
 - 300 hours of video are uploaded every minute

What do you see in the picture?



What do you see in the picture?



Black background

Two objects

One teapot

One toy

There is a light coming from right

One object is shiny the other is not

Toy:

Consists of 5 layers, in different colors

There is a text : Fisher Price

The layers are in donut shape

Layers are plastic

Bottom is wood

Teapot:

Consists of body and handle

Body is metal

Handle is ceramic

Handle: Dark blue on white

Body : golden

Reflection of toy on the body

Challenge – What do you see in the picture?



Source: Octavia Camps, Penn State

Challenge – What do you see in the picture?



A hand holding a man

A hand holding a shiny sphere

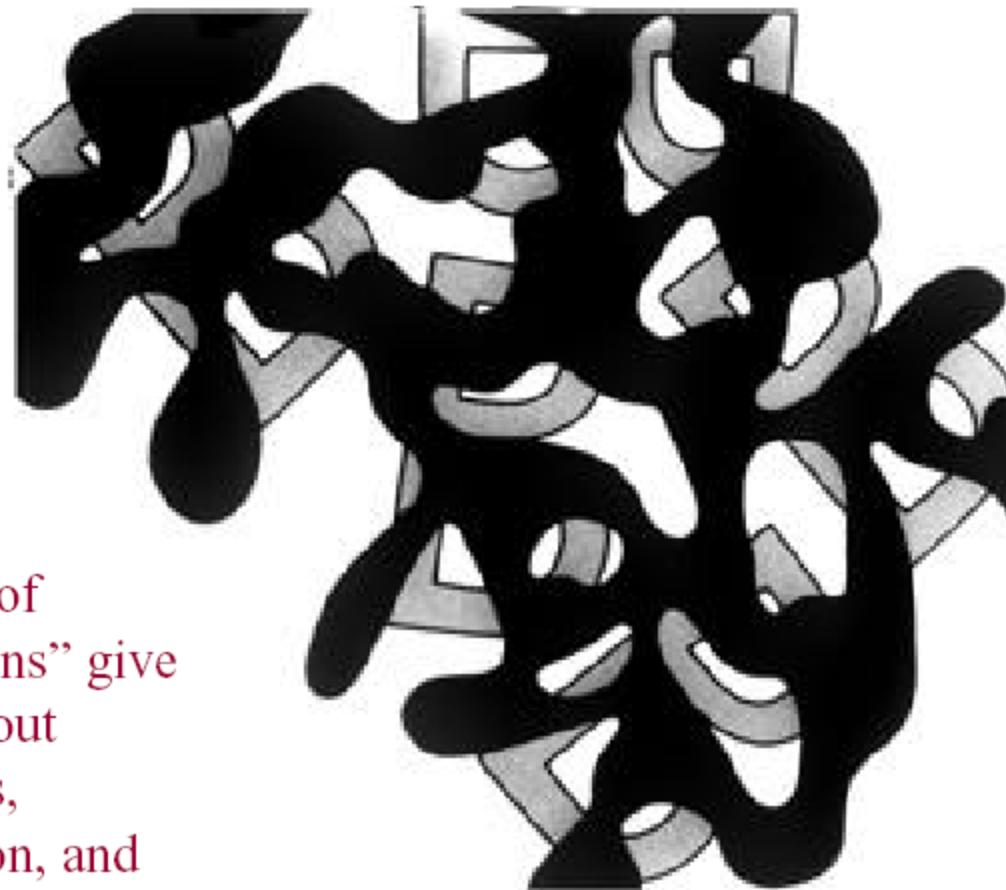
An Escher drawing

Perception and grouping



Subjective contours

Occlusion



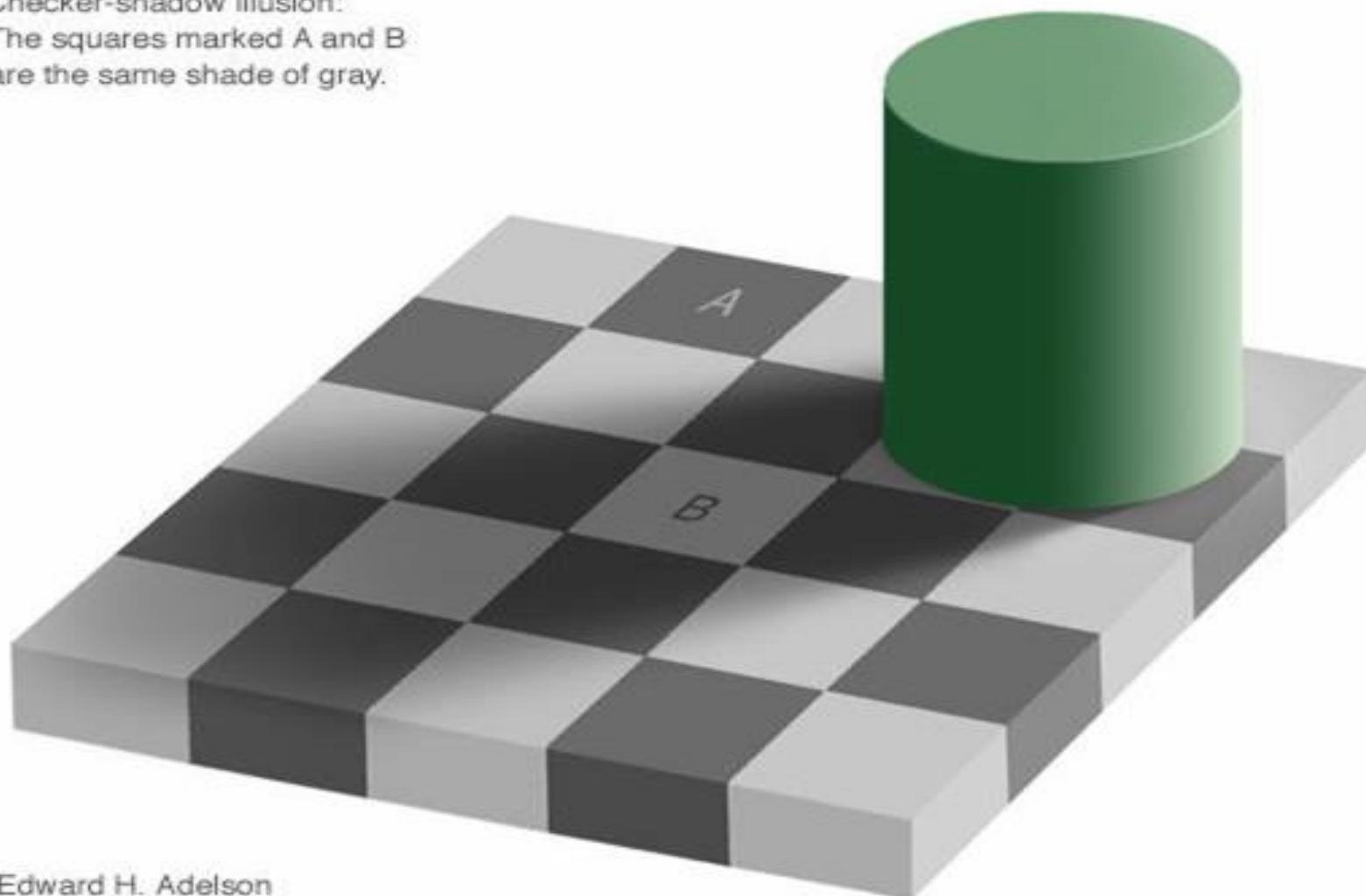
* types of
“junctions” give
cues about
surfaces,
occlusion, and
light.

I. Rock, The Logic of Perception, 1983.

Shading

Checker-shadow illusion:

The squares marked A and B
are the same shade of gray.



Edward H. Adelson

Shading



Parts and relations



How flexible are the spatial relations of the parts?

How good are our models?



Thompson, P. (1980). "Margaret Thatcher: a new illusion." *Perception* 9:483-484

How good are our models?



Thompson, P. (1980). "Margaret Thatcher: a new illusion." *Perception* 9:483-484

Is it only about matching?

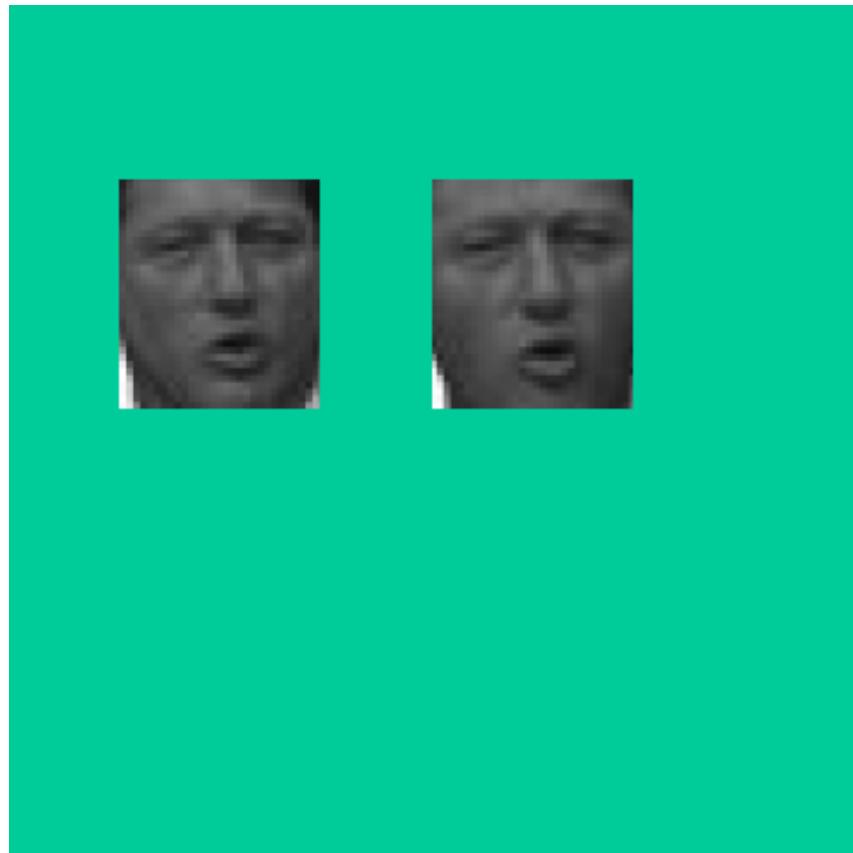
What are our
“models”?

How good
are they?



Ron Rensick

Is it only about matching?



Context



Antonio Torralba

Context



a person?

Antonio Torralba

Context

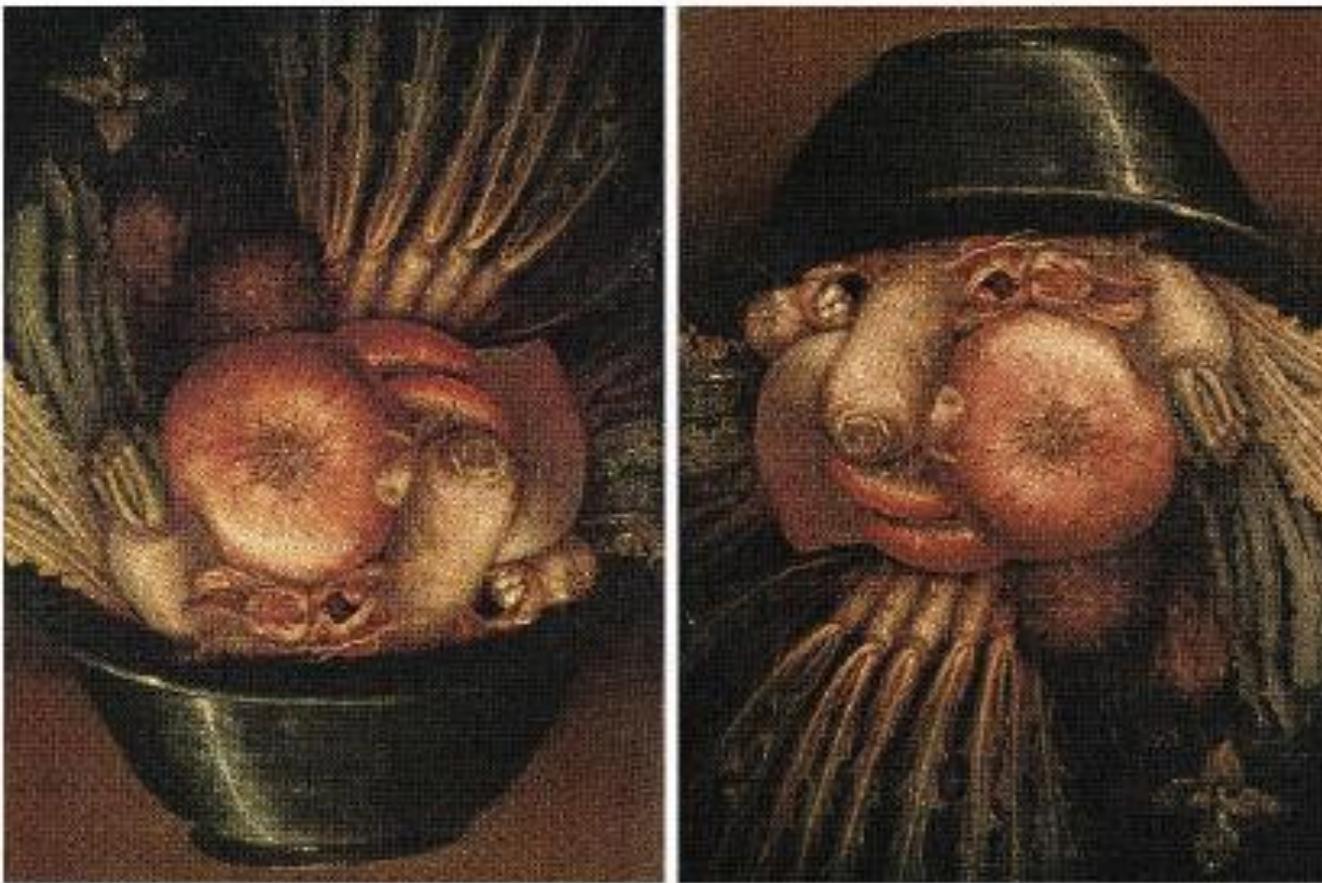


a person?

Antonio Torralba

the blob is identical to the one on the previous slide after a 90deg rotation

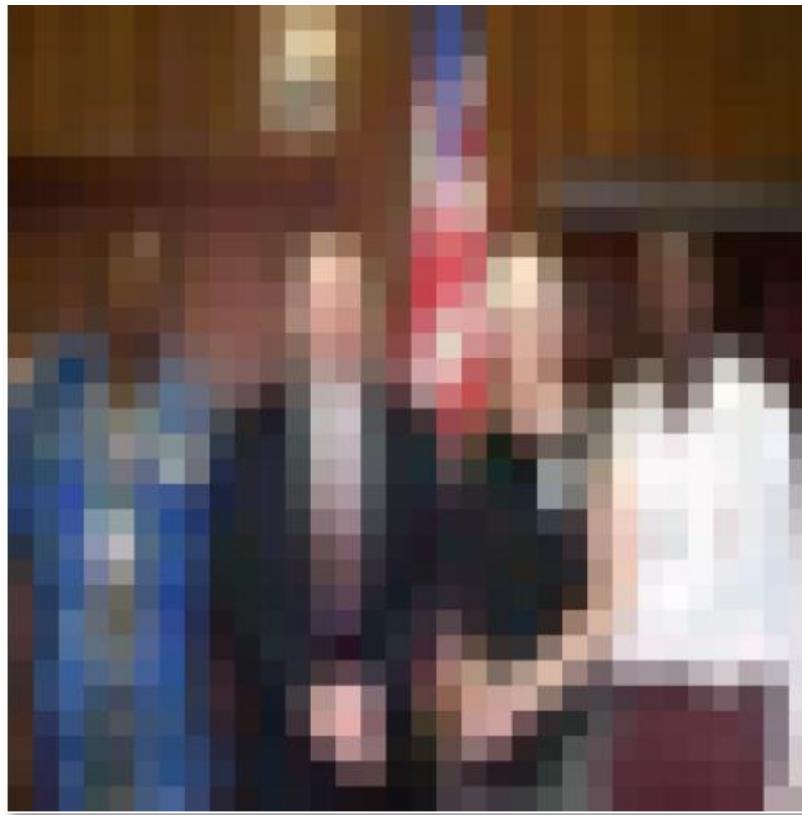
Prior Expectations



Giuseppe Arcimboldo

The goal of computer vision

- To extract “meaning” from pixels



Humans are remarkably good at this...

The goal of computer vision

- To extract “meaning” from pixels



La Gare Montparnasse, 1895

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

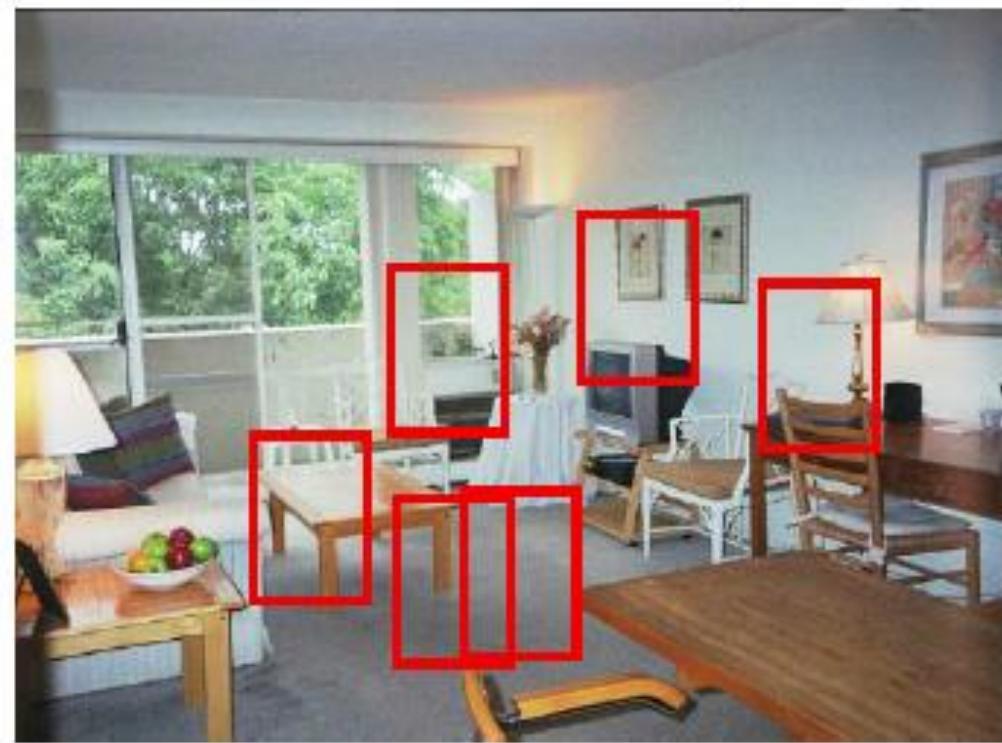
What a computer sees

Template matching

Find the chair in this image



Find the chair in this image

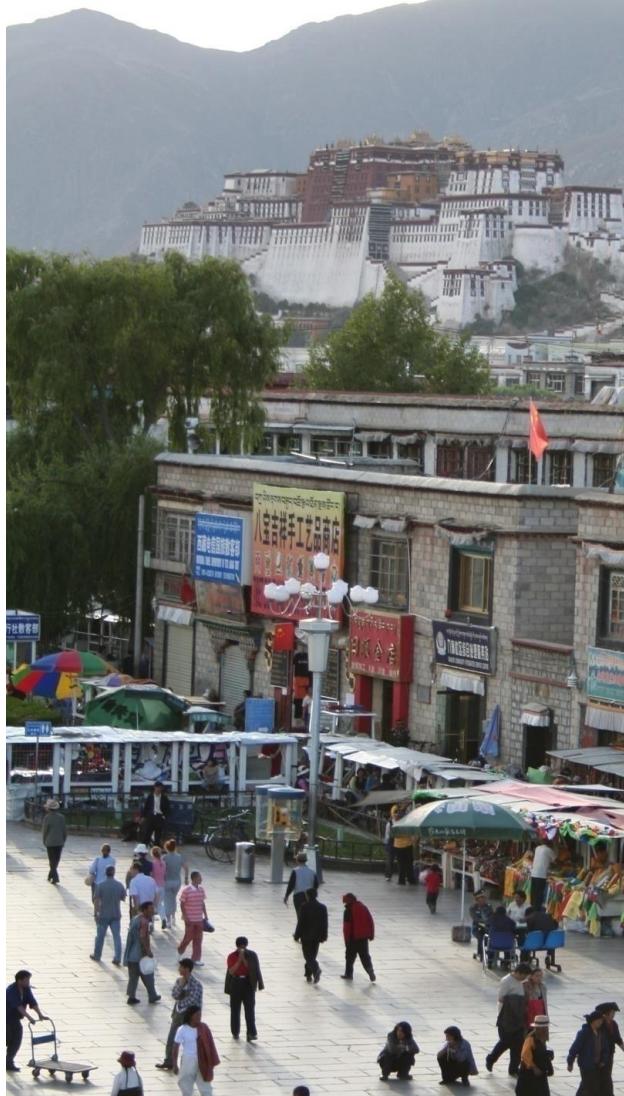


Scene categorization

- outdoor/indoor
- city/forest/factory/etc.



Image annotation/tagging

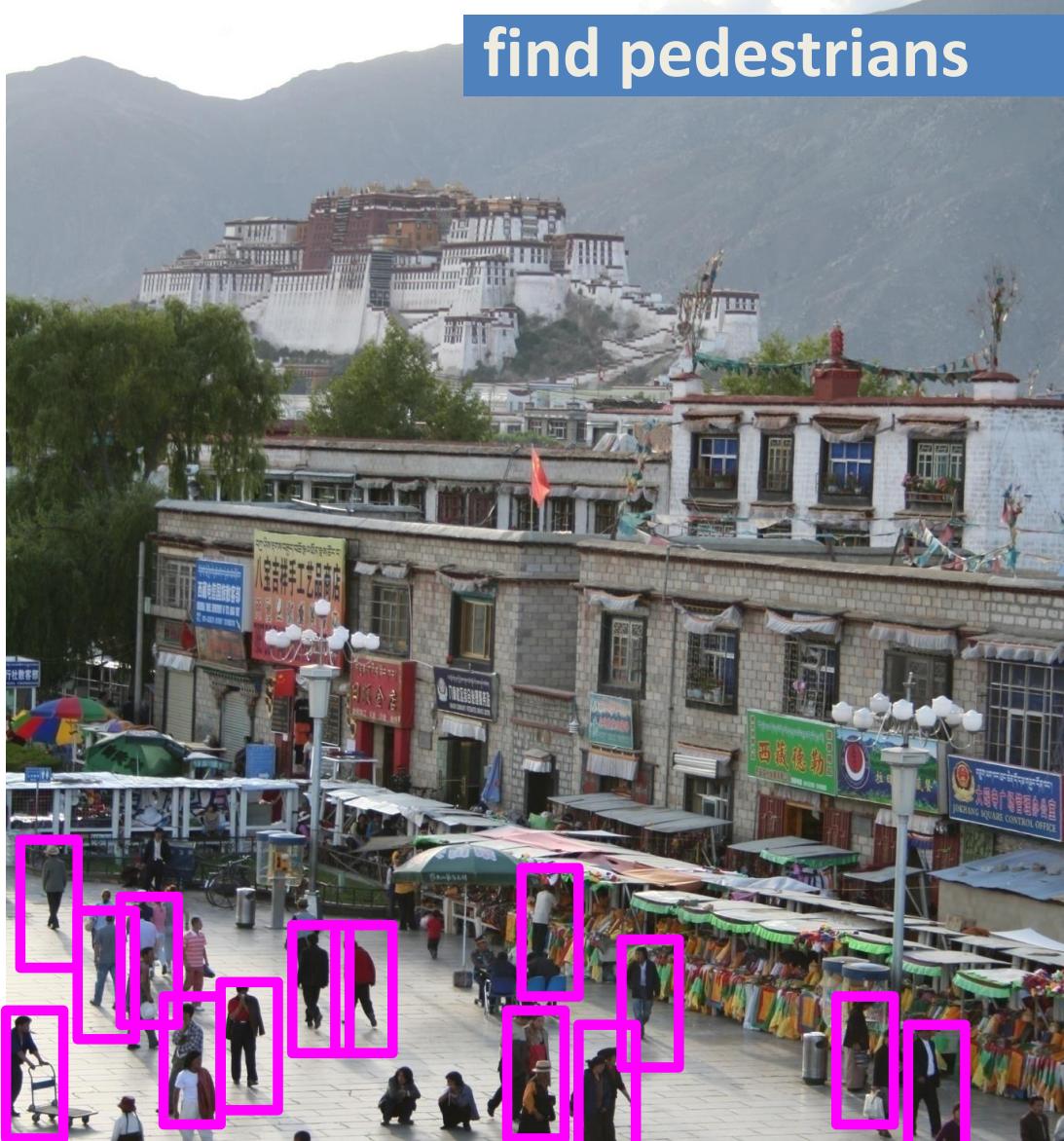


- street
- people
- building
- mountain
- ...

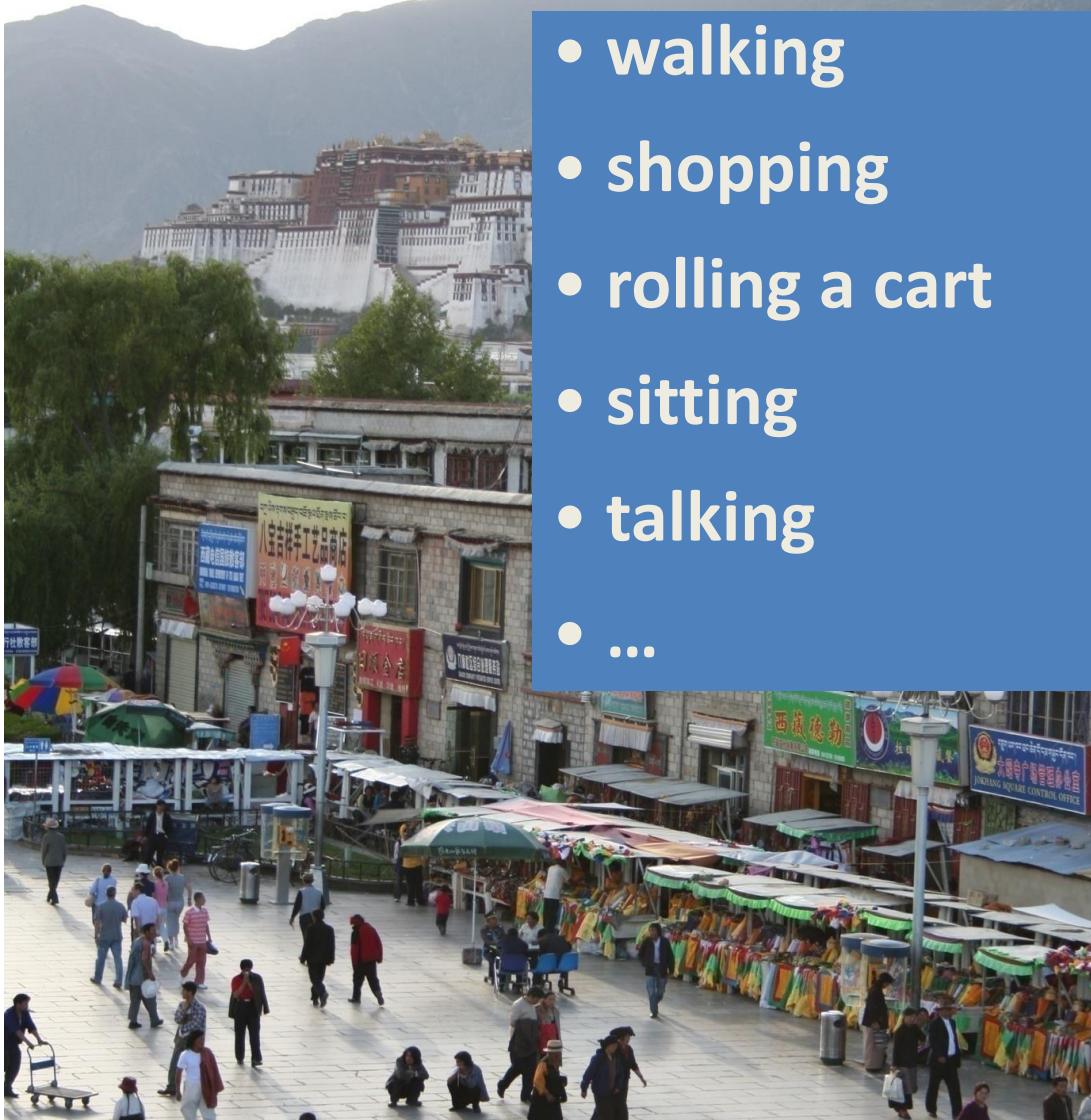


Object Detection

find pedestrians



Activity Recognition



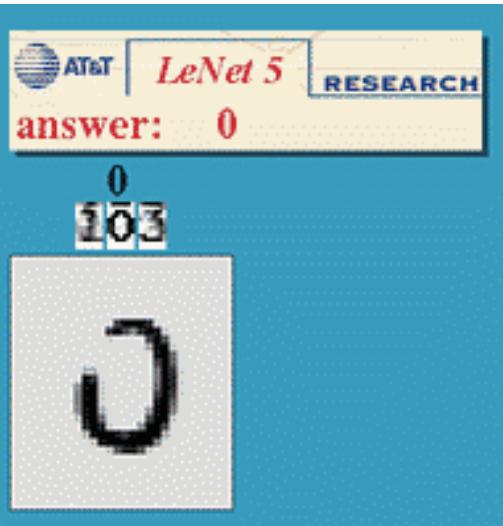
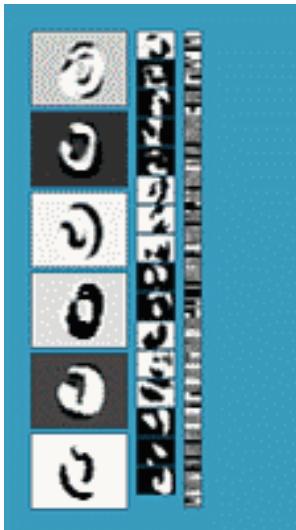
- **walking**
- **shopping**
- **rolling a cart**
- **sitting**
- **talking**
- ...

Object recognition



Applications

Optical character recognition (OCR)



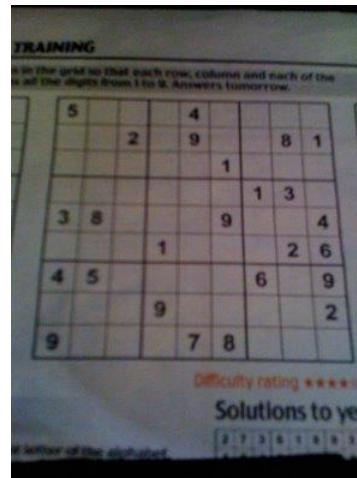
Digit recognition
yann.lecun.com



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



Automatic check processing

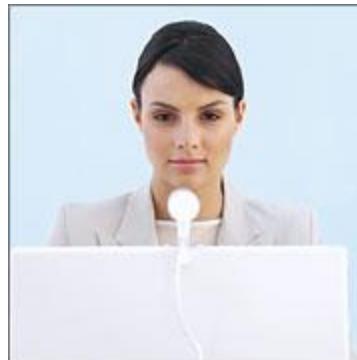


Sudoku grabber
<http://sudokugrab.blogspot.com/>

Biometrics



Fingerprint scanners on
many new laptops,
other devices



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

Face detection

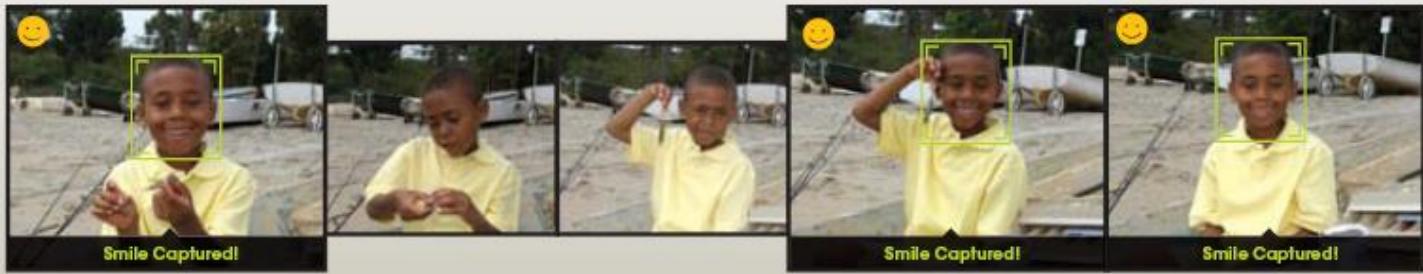
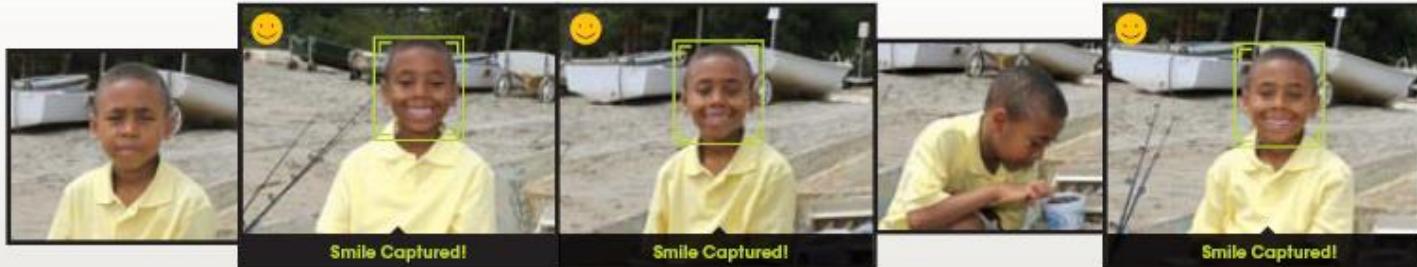


- Many consumer digital cameras now detect faces

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.



[Sony Cyber-shot® T70 Digital Still Camera](#)

Face recognition: Apple iPhoto software



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

Mobile visual search: Google Goggles

Google Goggles in Action

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



Automotive safety

▷▷ manufacturer products consumer products ◀◀

Our Vision. Your Safety.

rear looking camera forward looking camera side looking camera

EyeQ Vision on a Chip

Road, Vehicle, Pedestrian Protection and more

Vision Applications

AWS Advance Warning System

[read more](#) [read more](#) [read more](#)

News

- > [Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System](#)
- > [Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end](#)

[all news](#)

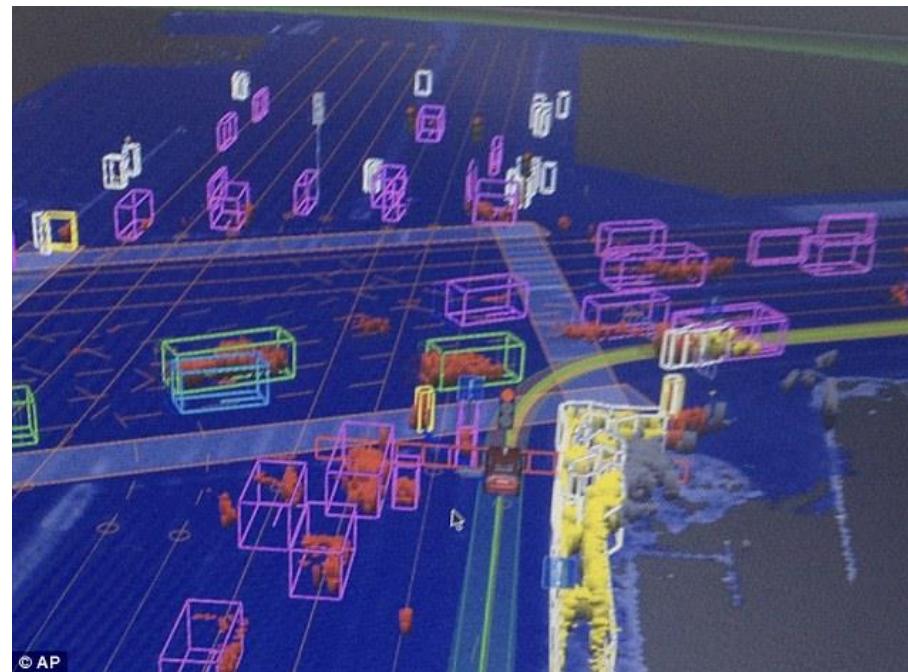
Events

- > [Mobileye at Equip Auto, Paris, France](#)
- > [Mobileye at SEMA, Las Vegas, NV](#)

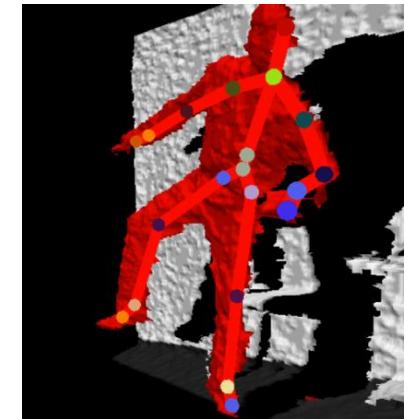
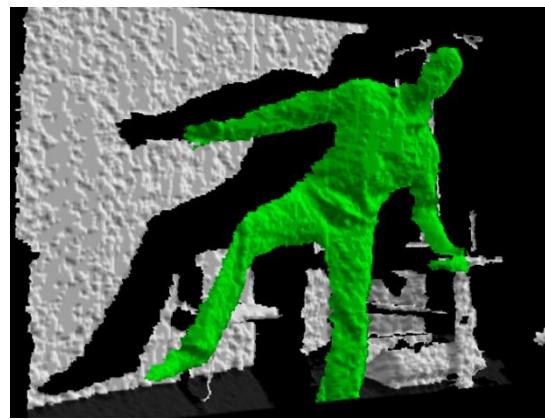
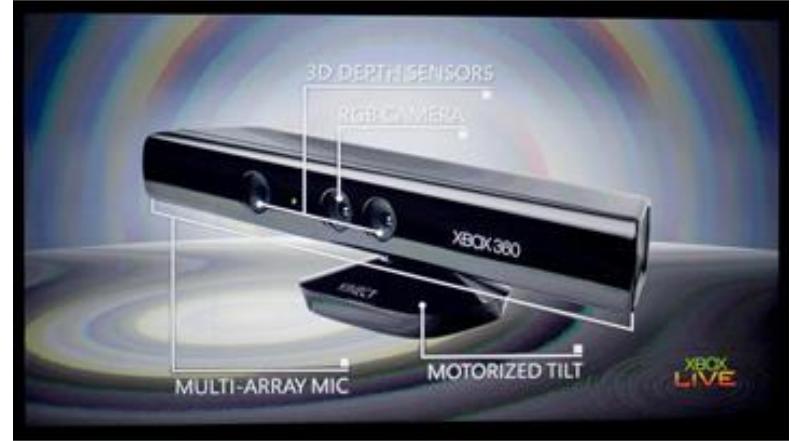
[read more](#)

- [Mobileye](#): Vision systems in high-end BMW, GM, Volvo models
 - Pedestrian collision warning
 - Forward collision warning
 - Lane departure warning
 - Headway monitoring and warning

Self-driving cars



Vision-based interaction: Xbox Kinect



3D Reconstruction: Kinect Fusion



[YouTube Video](#)

3D Reconstruction: Multi-View Stereo



[YouTube Video](#)



- Home
- Try it
- What is Photosynth?
- Collections
- Team blog
- Videos
- System requirements
- About us
- FAQ

*"What if your photo collection was an entry point into the world,
like a wormhole that you could jump through and explore..."*

Try it



Try the Tech Preview

The **Photosynth Technology Preview** is a taste of the newest - and, we hope, most exciting - way to **view photos** on a computer. Our software takes a large collection of photos of a place or an object, analyzes them for similarities, and then displays the photos in a reconstructed **three-dimensional space**, showing you how each one relates to the next.

<http://labs.live.com/photosynth/>

Based on Photo Tourism technology developed by
Noah Snavely, Steve Seitz, and Rick Szeliski

Google Maps Photo Tours



<http://maps.google.com/phototours>

Earth viewers (3D modeling)

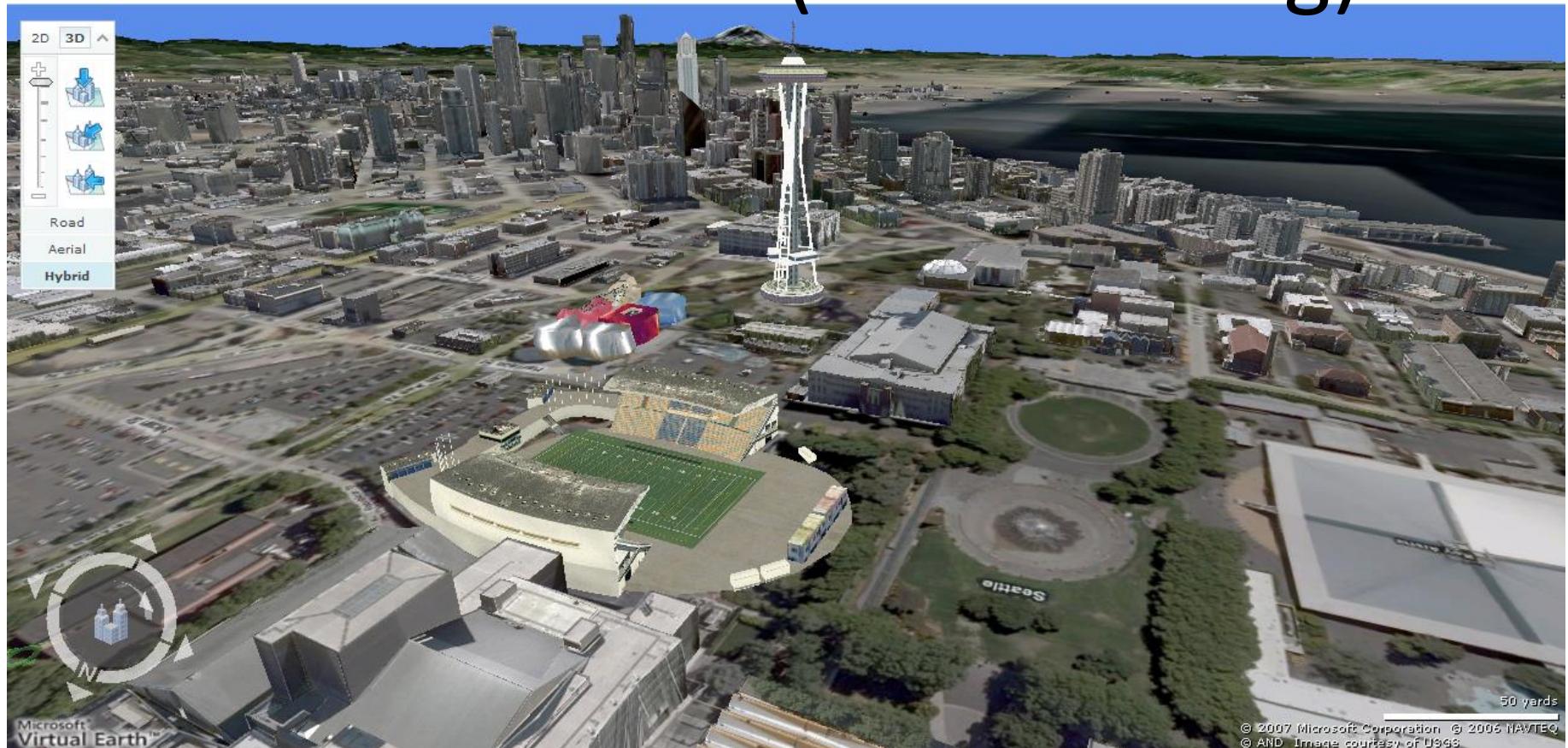


Image from Microsoft's [Virtual Earth](#)
(see also: [Google Earth](#))

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...”

Object recognition (in mobile phones)



- This is becoming real:
 - **Lincoln** Microsoft Research
 - [Point & Find, Nokia](#)

Special effects: shape and motion capture



Vision in space

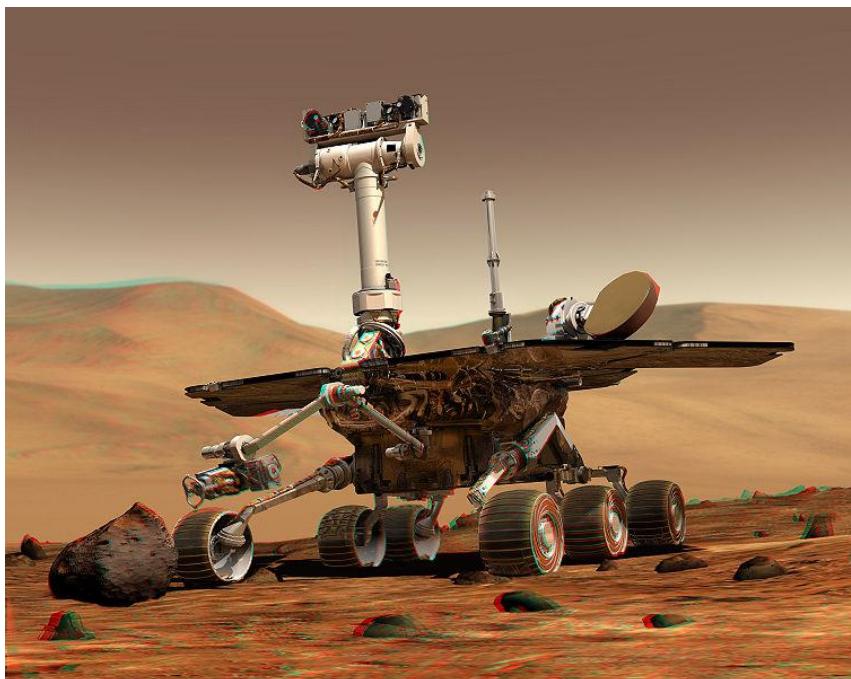


[NASA'S Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

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.

Robotics

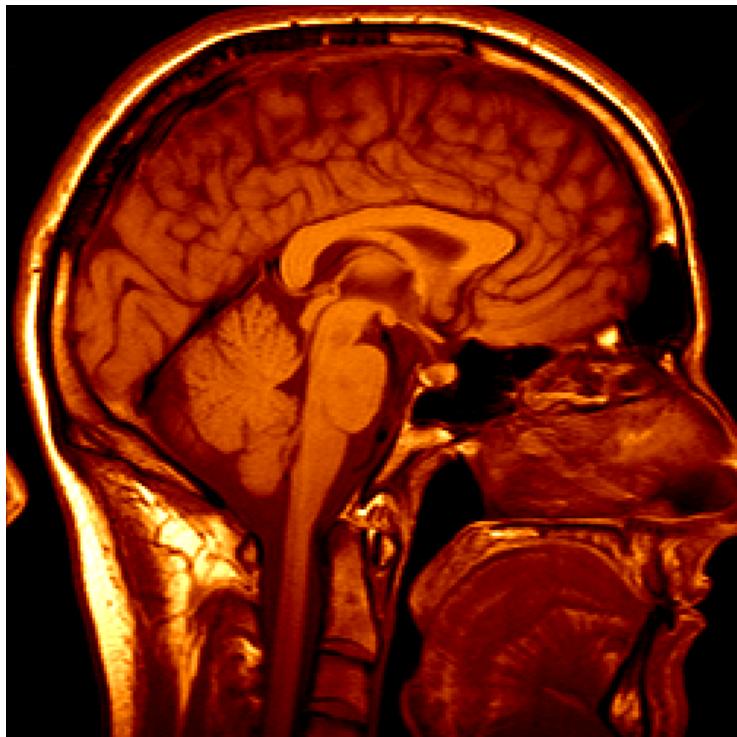


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



<http://www.robocup.org/>

Medical imaging



3D imaging
MRI, CT

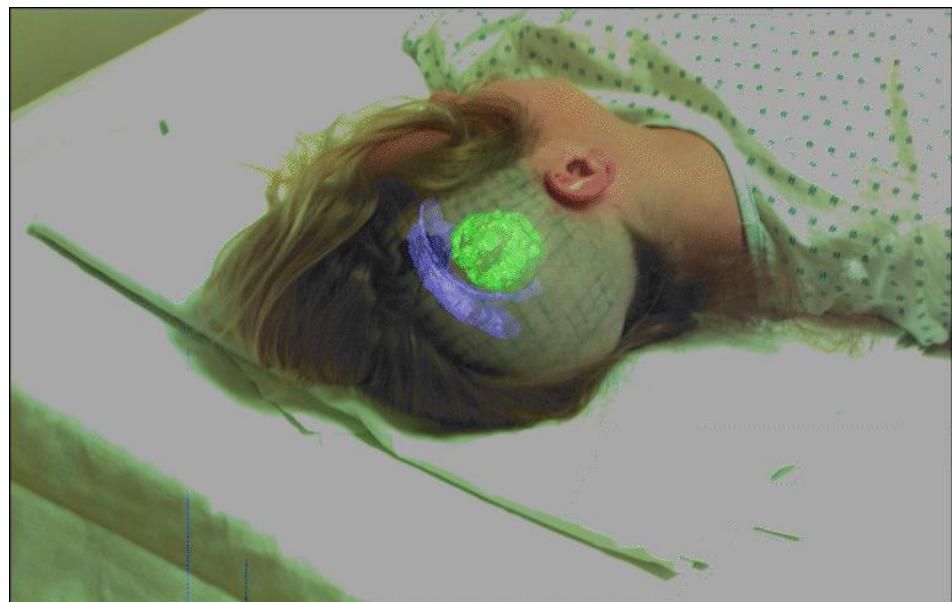
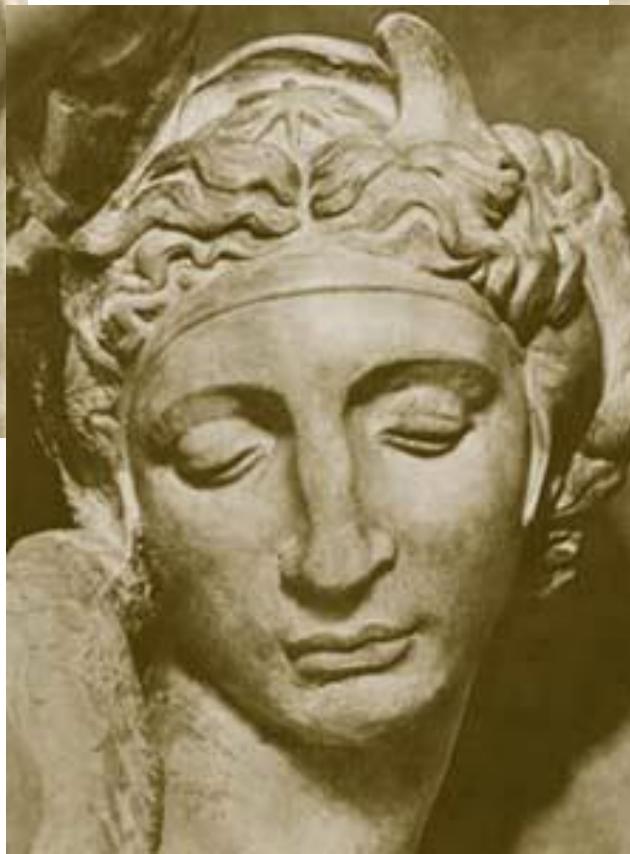


Image guided surgery
[Grimson et al., MIT](#)

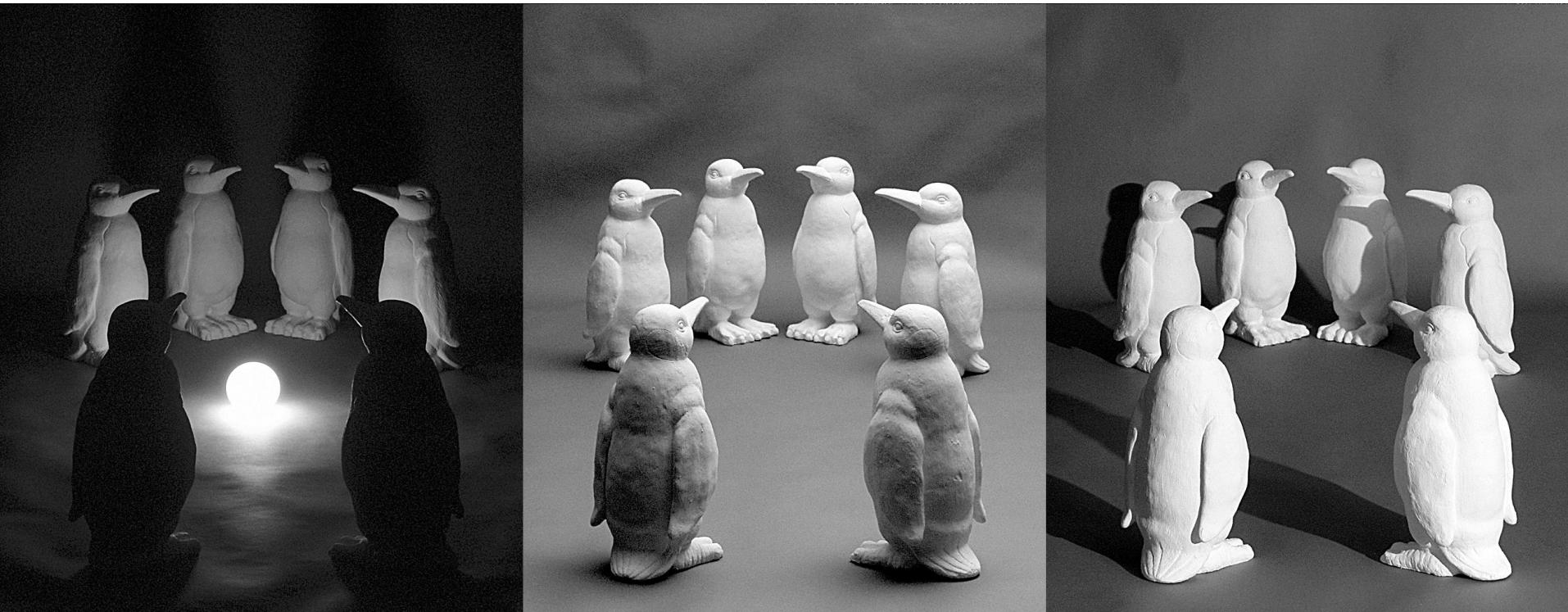
Why is computer vision difficult?

Challenges: viewpoint variation



Michelangelo 1475-1564

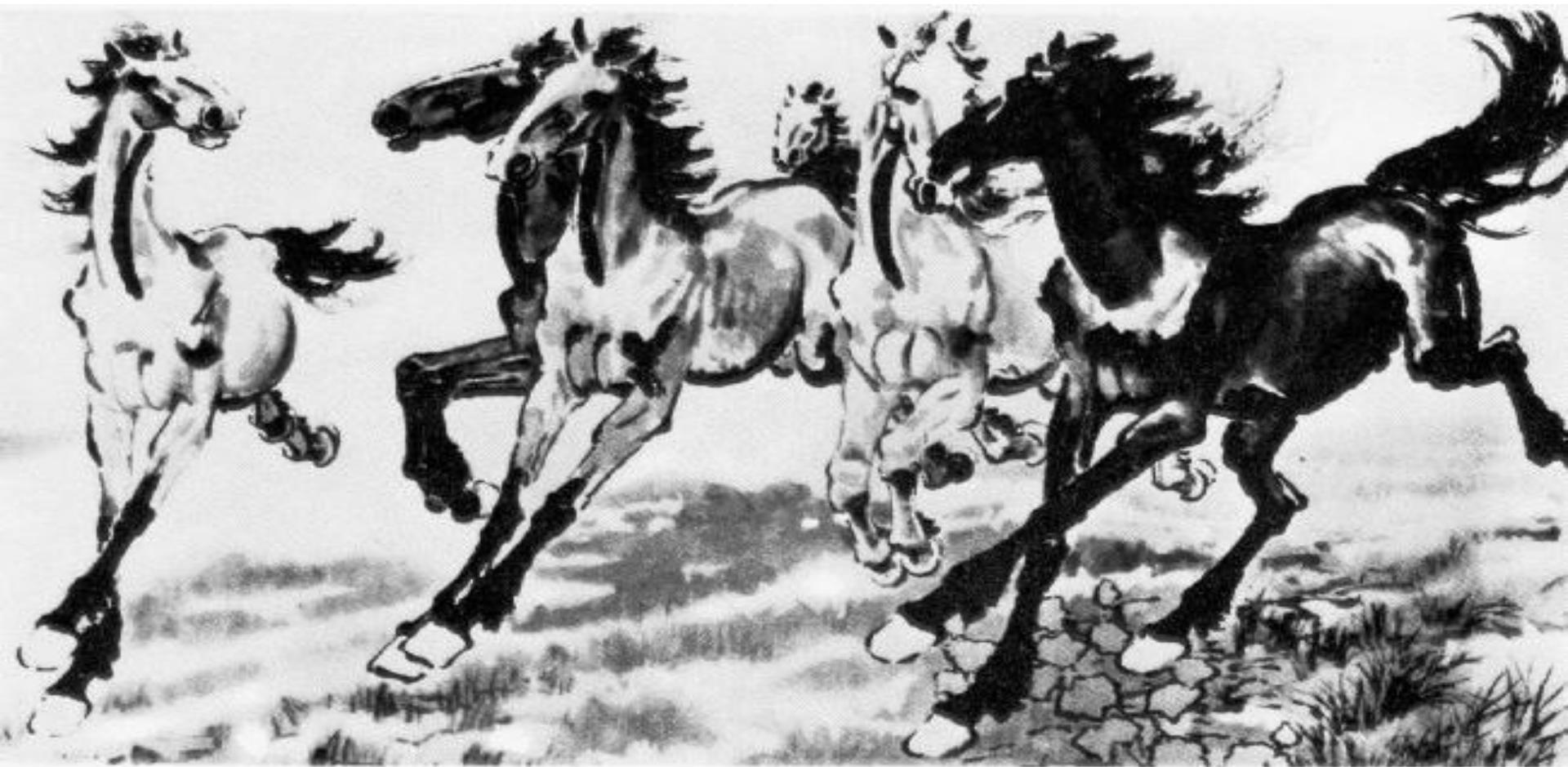
Challenges: illumination



Challenges: scale

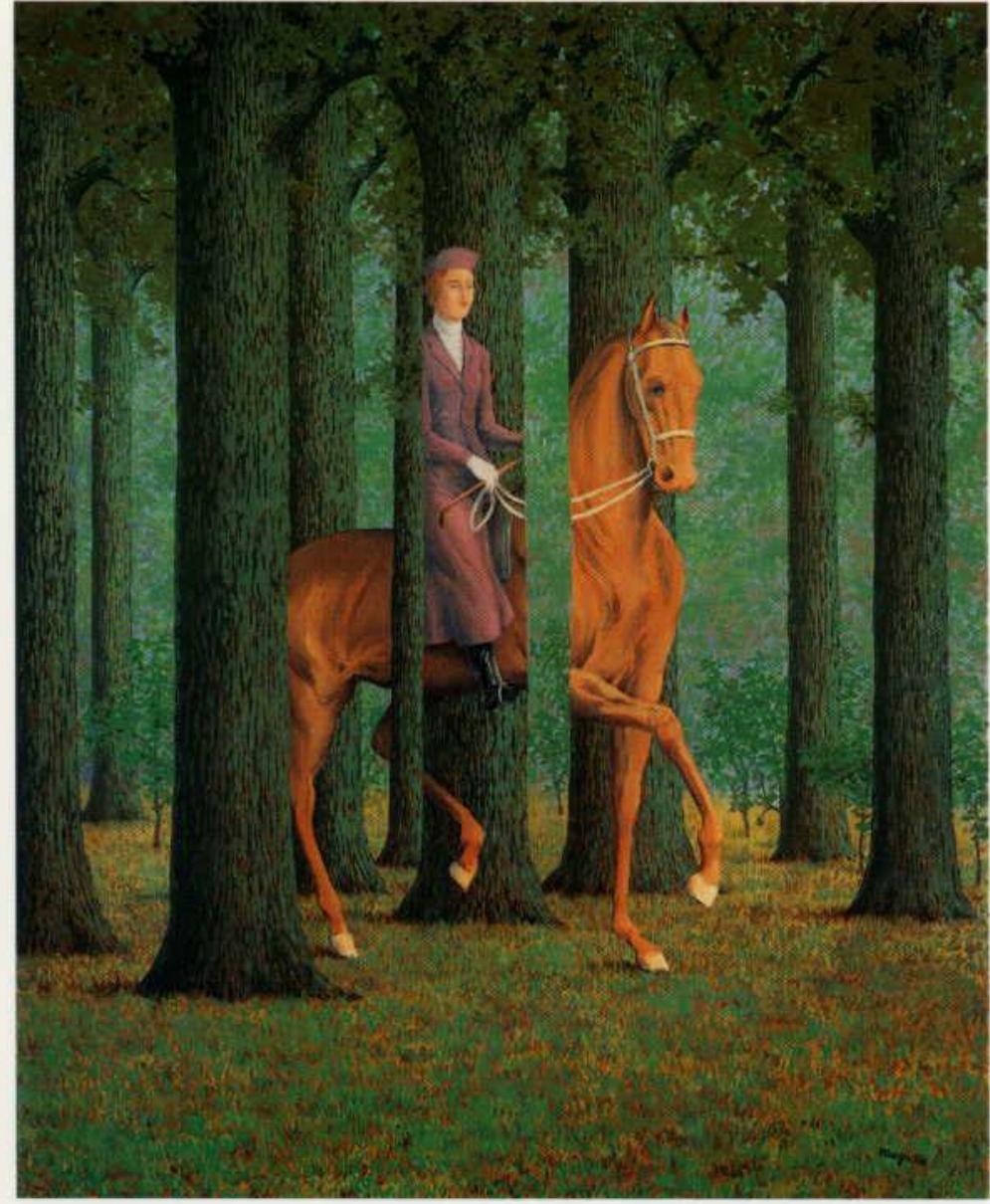


Challenges: deformation



Xu, Beihong 1943

Challenges: occlusion



Challenges: Background Clutter



Emperor shrimp and commensal crab on a sea cucumber in Fiji
Photograph by Tim Laman

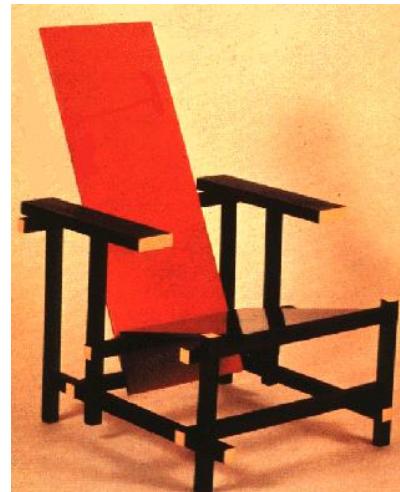
NATIONAL
GEOGRAPHIC

© 2007 National Geographic Society. All rights reserved.

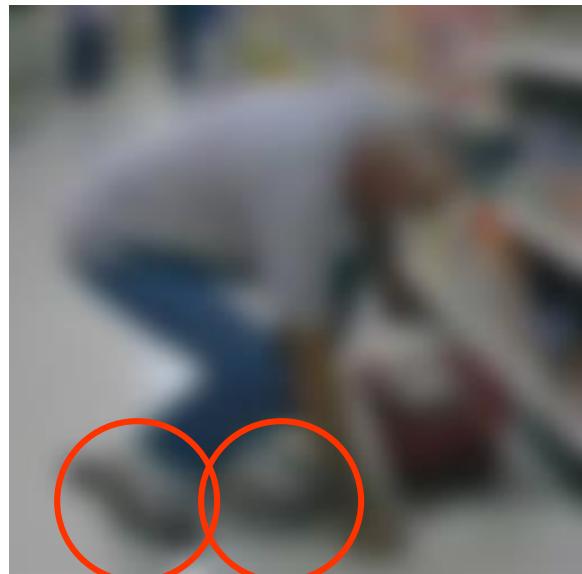
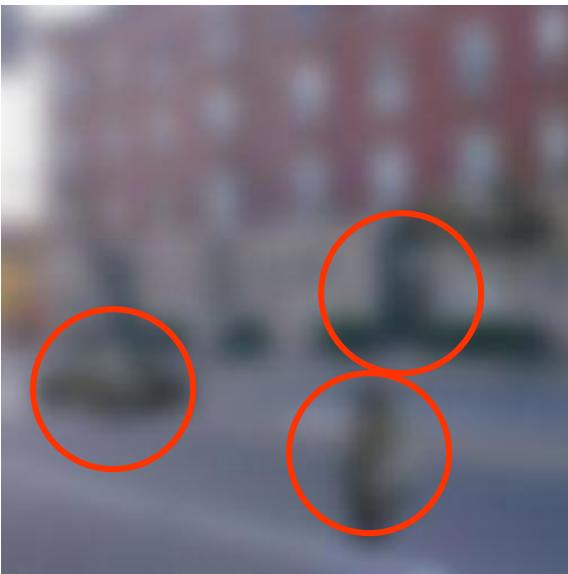
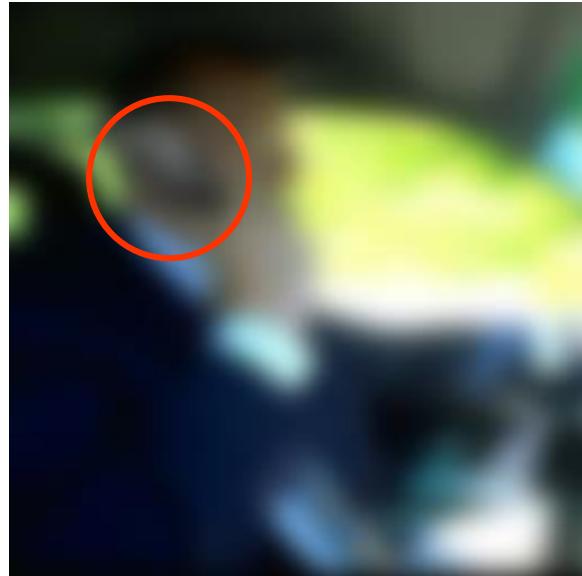
Challenges: Motion



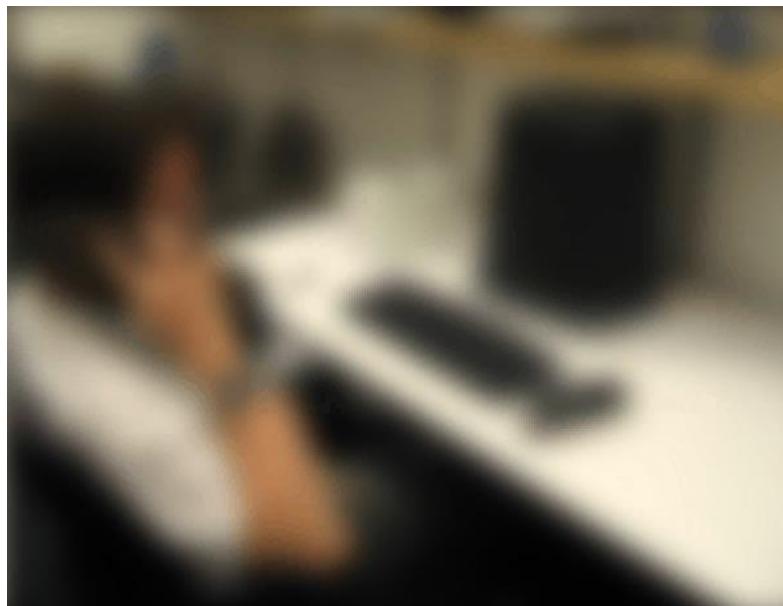
Challenges: object intra-class variation



Challenges: local ambiguity



Challenges: local ambiguity

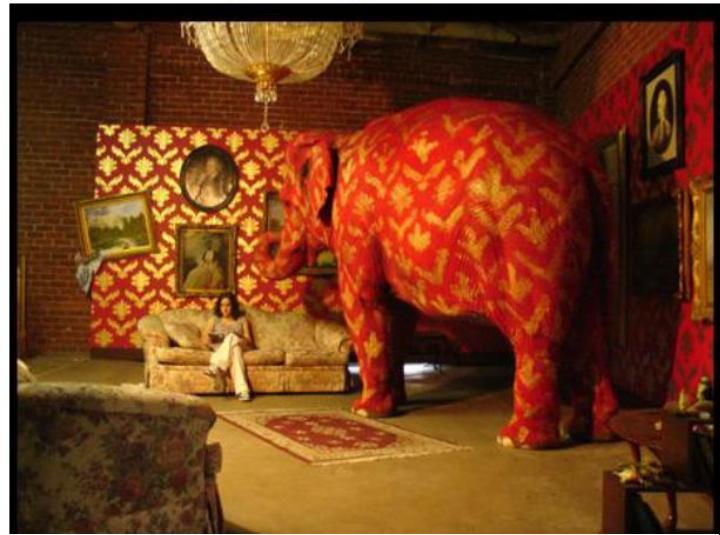


Challenges: local ambiguity

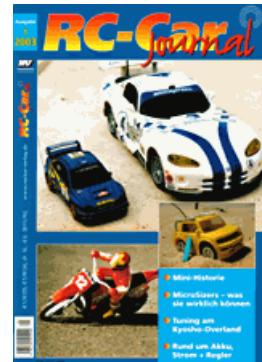


Source: Rob Fergus and Antoniou

Context



Look-Alikes by Joan Steiner



Banksy



Challenges: Inherent ambiguity

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



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!



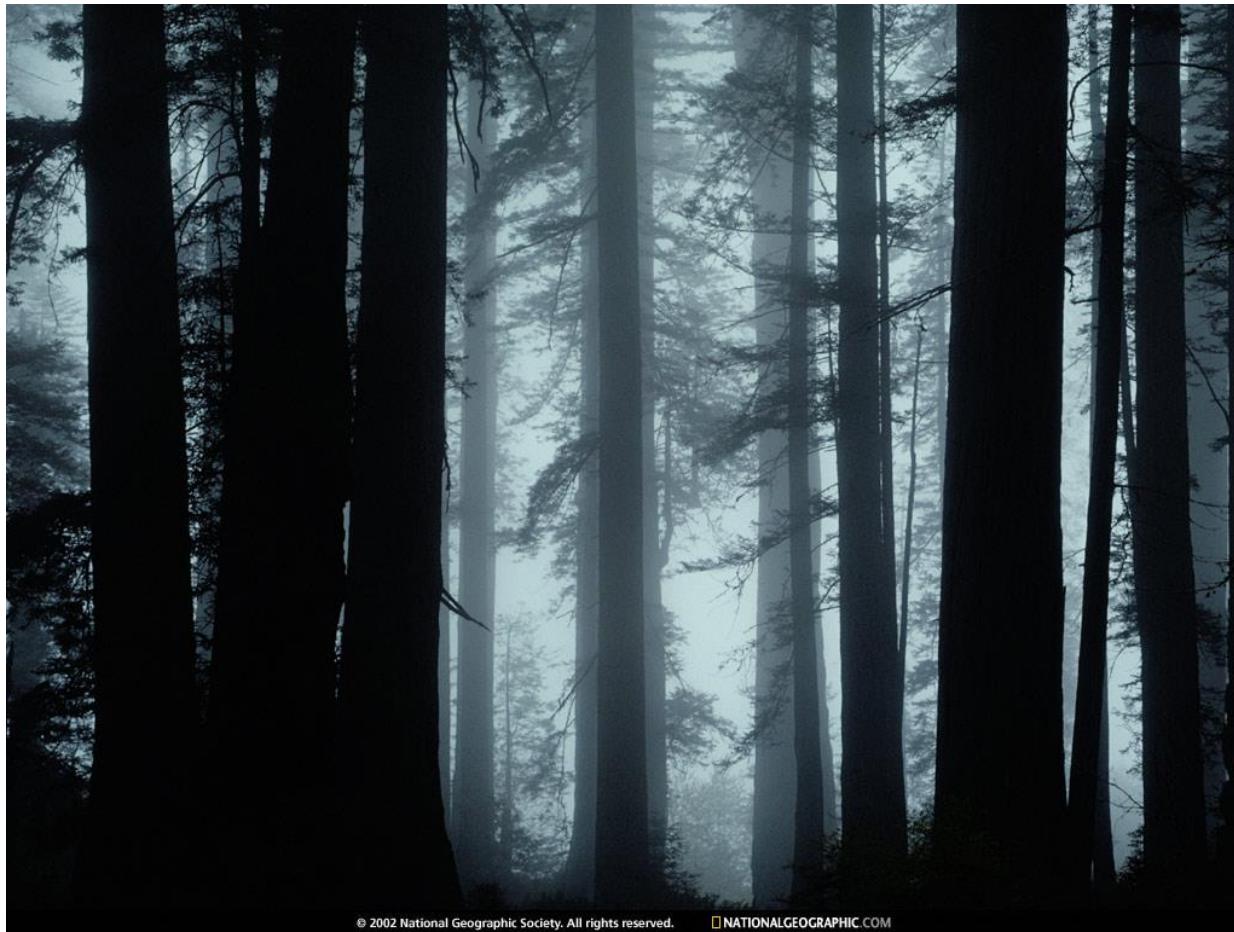
Depth cues: Linear perspective



NATIONAL GEOGRAPHIC.COM

© 2003 National Geographic Society. All rights reserved.

Depth cues: Aerial perspective



© 2002 National Geographic Society. All rights reserved. NATIONALGEOGRAPHIC.COM

Depth ordering cues: Occlusion



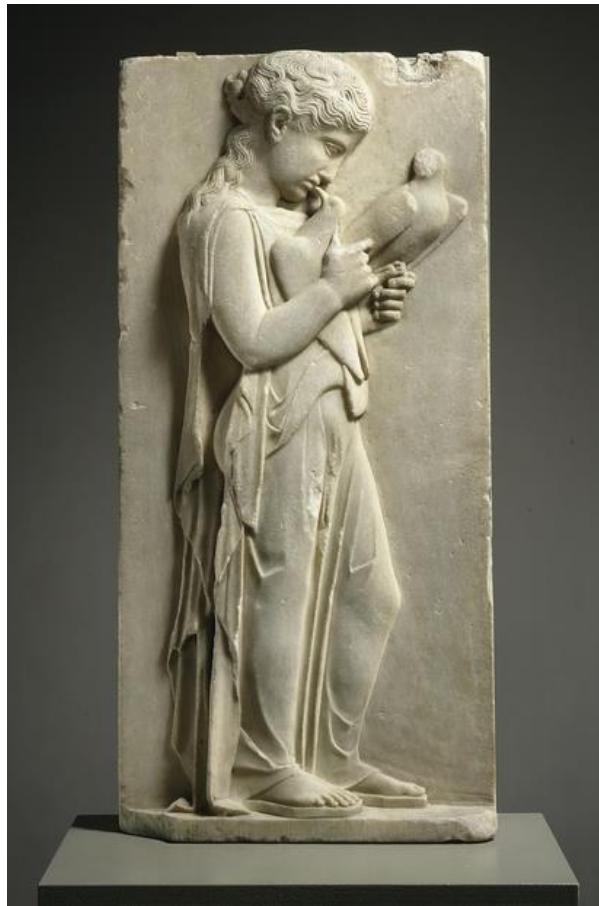
Shape cues: Texture gradient



© 2001 National Geographic Society. All rights reserved.

nationalgeographic.com

Shape and lighting cues: Shading



Position and lighting cues: Cast shadows



Grouping cues: Similarity (color, texture, proximity)



NATIONALGEOGRAPHIC.COM

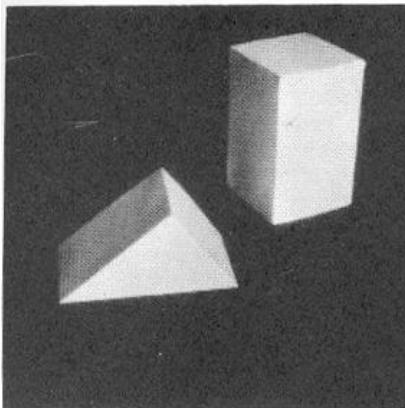
© 2003 National Geographic Society. All rights reserved.

Grouping cues: “Common fate”

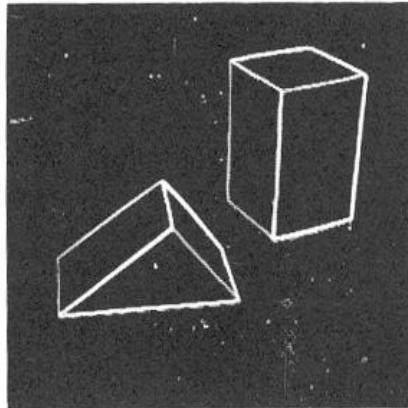


Origins of computer vision

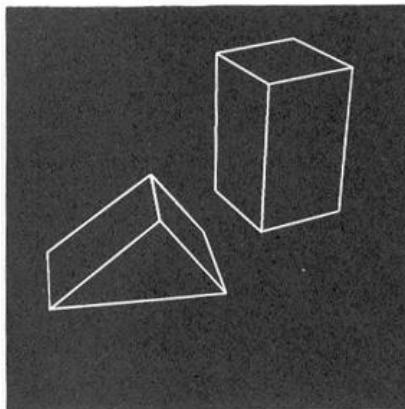
- 23 - 4445(a-d)



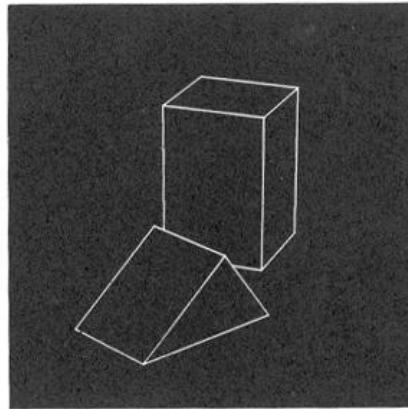
(a) Original picture.



(b) Differentiated picture.



(c) Line drawing.



(d) Rotated view.

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

Connections to other disciplines

