

# Computer Vision

CSE 576  
Ali Farhadi

Many slides from Steve Seitz, Larry Zitnick, Yang Wang

# Course Information

- Time:
  - Monday, Wednesday 1:30-2:50
- Location:
  - MGH 238
- Contact:
  - ali@cs.uw.edu , CSE 652
- TA:
  - Hessam Bagherinezhad (hessam@cs.uw.edu)
  - Yuguang Lee (ylee3@uw.edu)
- Website:
  - <http://courses.cs.washington.edu/cse576>

## **One Look Is Worth A Thousand Words--**

One look at our line of Republic, Firestone, Miller and United States tires can tell you more than a hundred personal letters or advertisements.

**WE WILL PROVE THEIR VALUE  
BEFORE YOU INVEST ONE DOLLAR  
IN THEM.**

Ever consider buying Supplies from a catalog?

What's the use! Call and see what you are buying. One look at our display of automobile and motorcycle accessories will convince you of the fact.

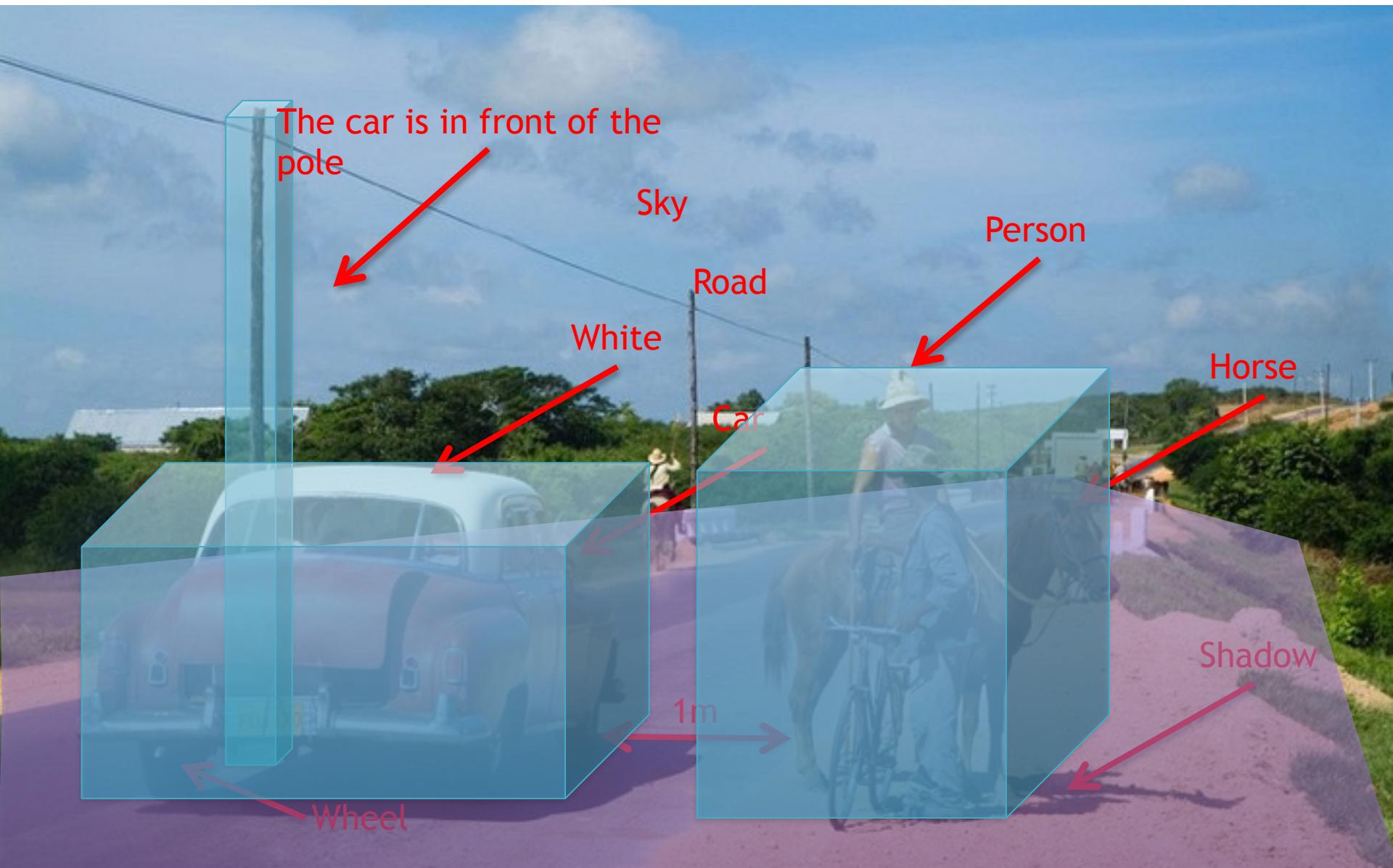
**THAT WE HAVE EVERYTHING FOR  
THE AUTO**

# **Piqua Auto Supply House**

133 N. Main St.—Piqua, O.

# What does it mean to see?

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



# Computer Vision

- **Low Level Vision**
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- **Mid Level Vision**
  - Reconstruction
  - Depth
  - Motion Estimation
- **High Level Vision**
  - Category detection
  - Activity recognition
  - Deep understandings



# Computer Vision

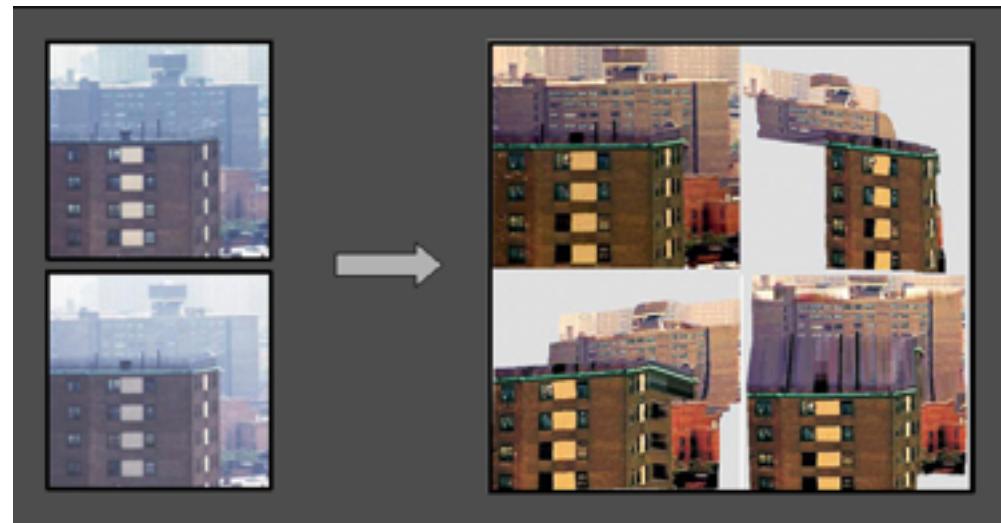
- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings



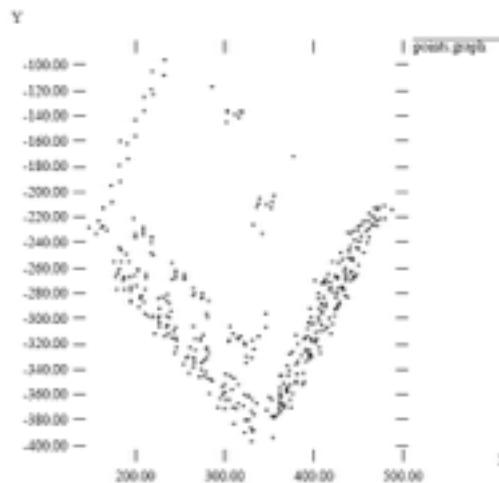
# Vision as Measurement Device



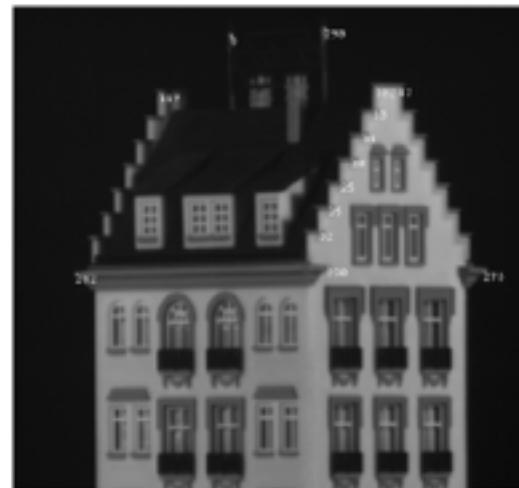
Real-time stereo on Mars



Physics-based Vision



Structure from Motion

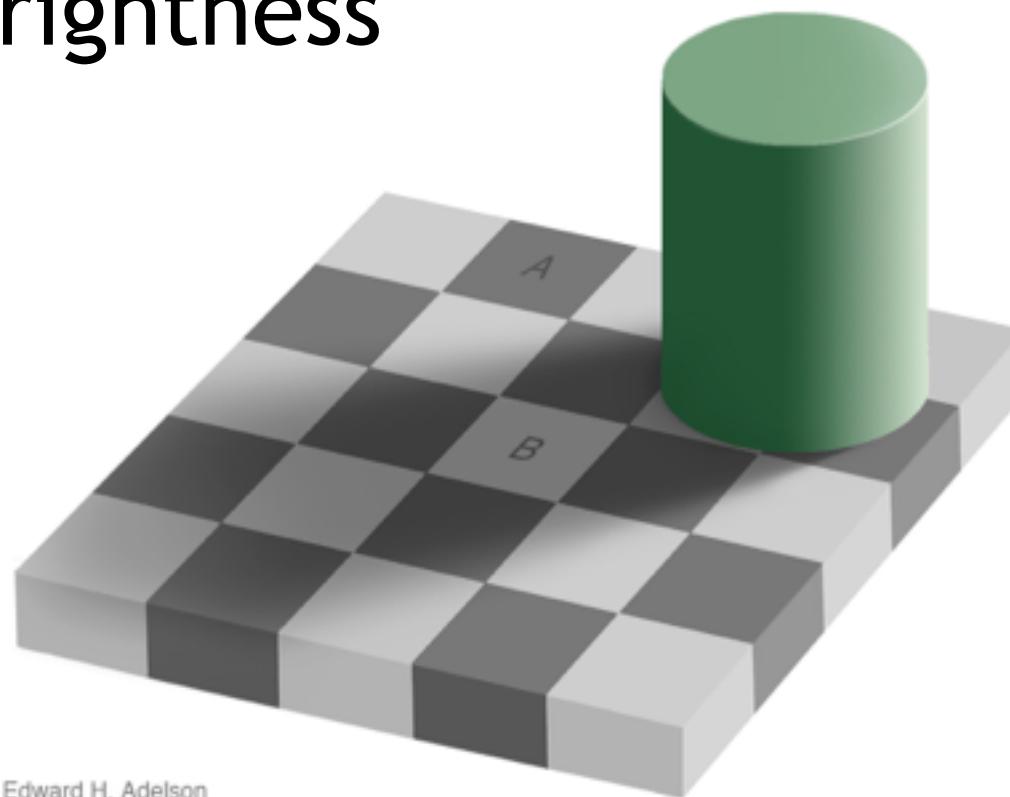


Virtualized Reality

Slide Credit: Alyosha

# Measurement

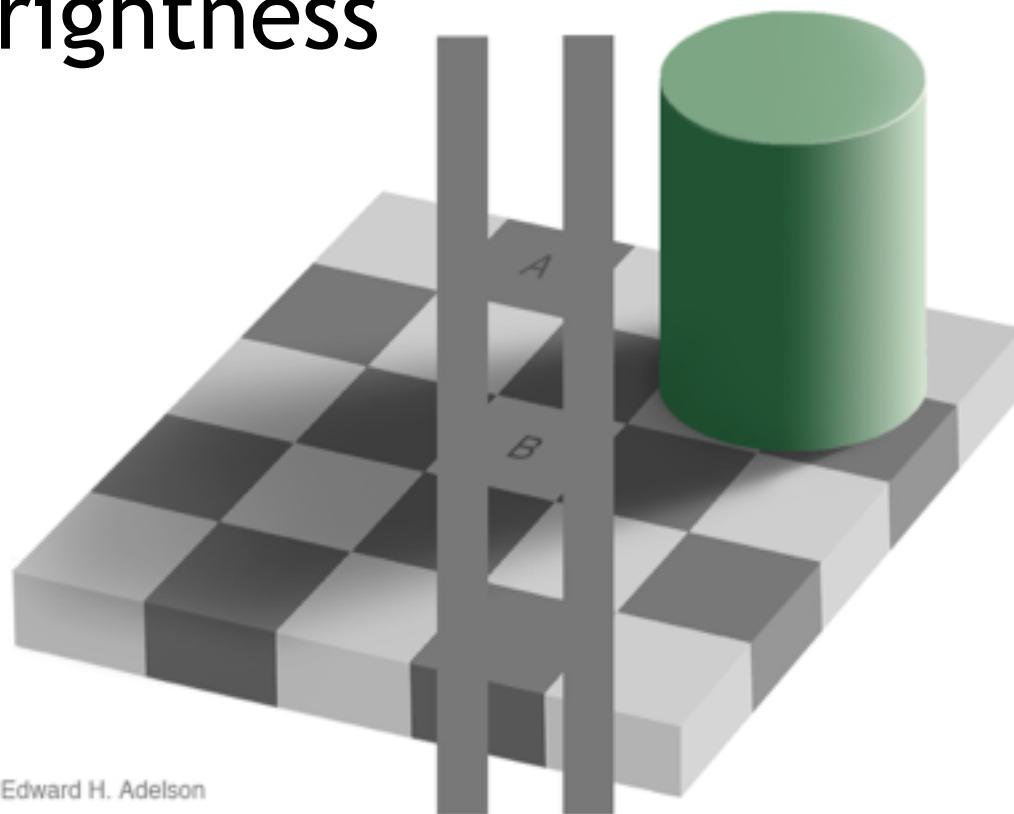
## Brightness



Edward H. Adelson

# Measurement

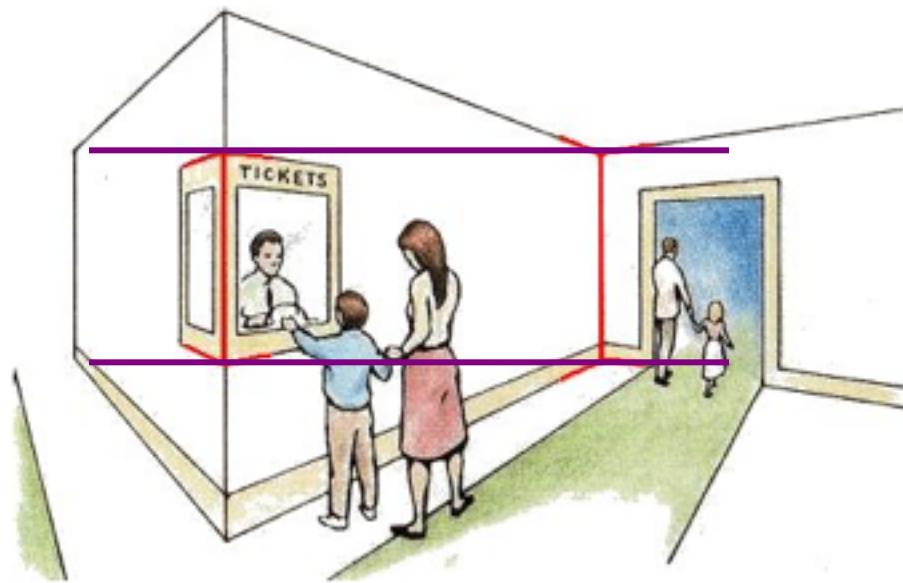
Brightness



Edward H. Adelson

# Measurement

## Length



Müller-Lyer Illusion

# Image Enhancement



*Image Inpainting*, M. Bertalmío et al.

<http://www.iua.upf.es/~mbertalmio//restoration.html>

# Image Enhancement



*Image Inpainting*, M. Bertalmío et al.

<http://www.iua.upf.es/~mbertalmio//restoration.html>

# Image Enhancement



*Image Inpainting*, M. Bertalmío et al.

<http://www.iua.upf.es/~mbertalmio//restoration.html>

# Seam Carving



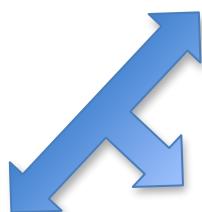
[Shai & Avidan, SIGGRAPH 2007]

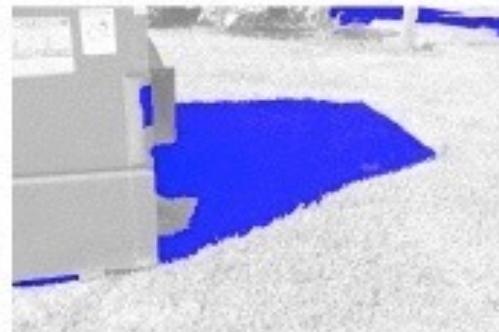


Traditional resizing



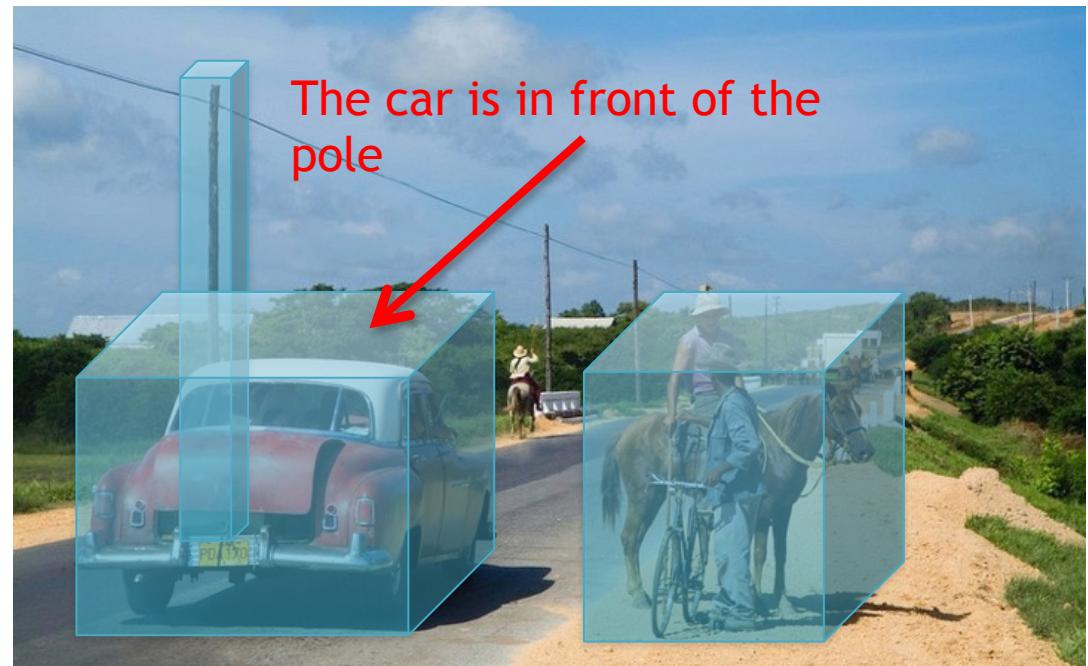
Content-aware resizing





# Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings





**Input Image (1 of 45)**



**Reconstruction**



**Reconstruction**



**Reconstruction**

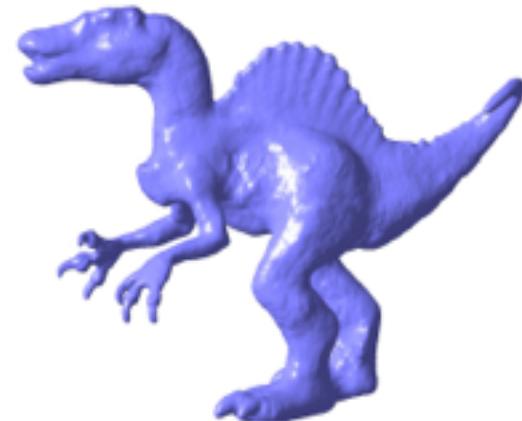
Source: S. Seitz



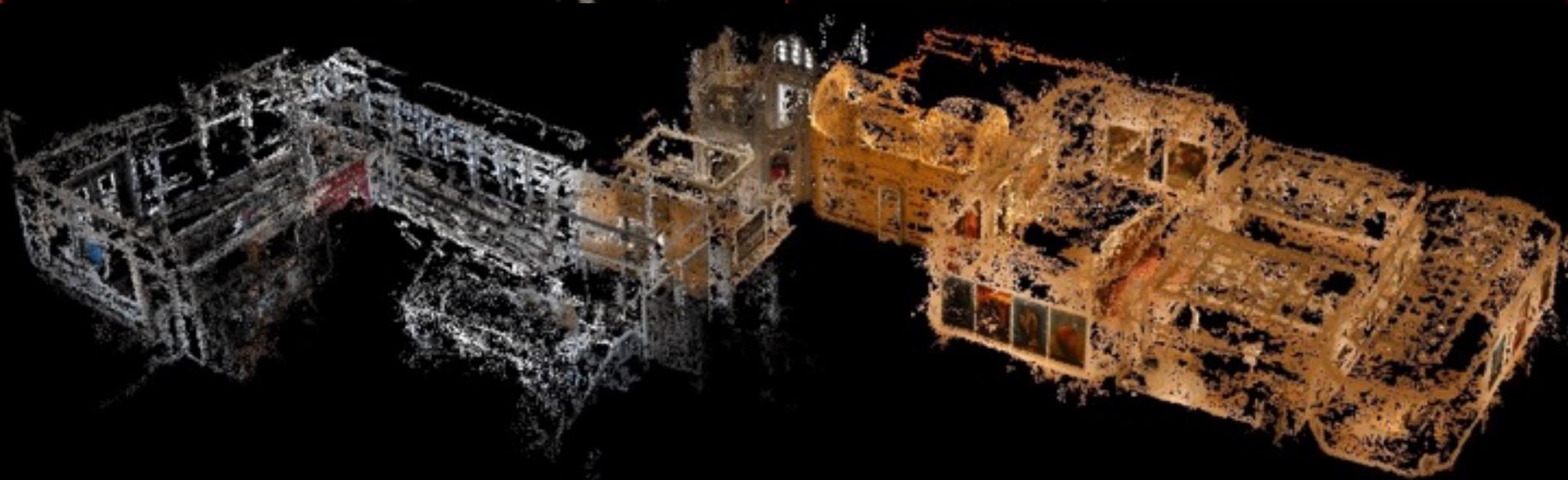
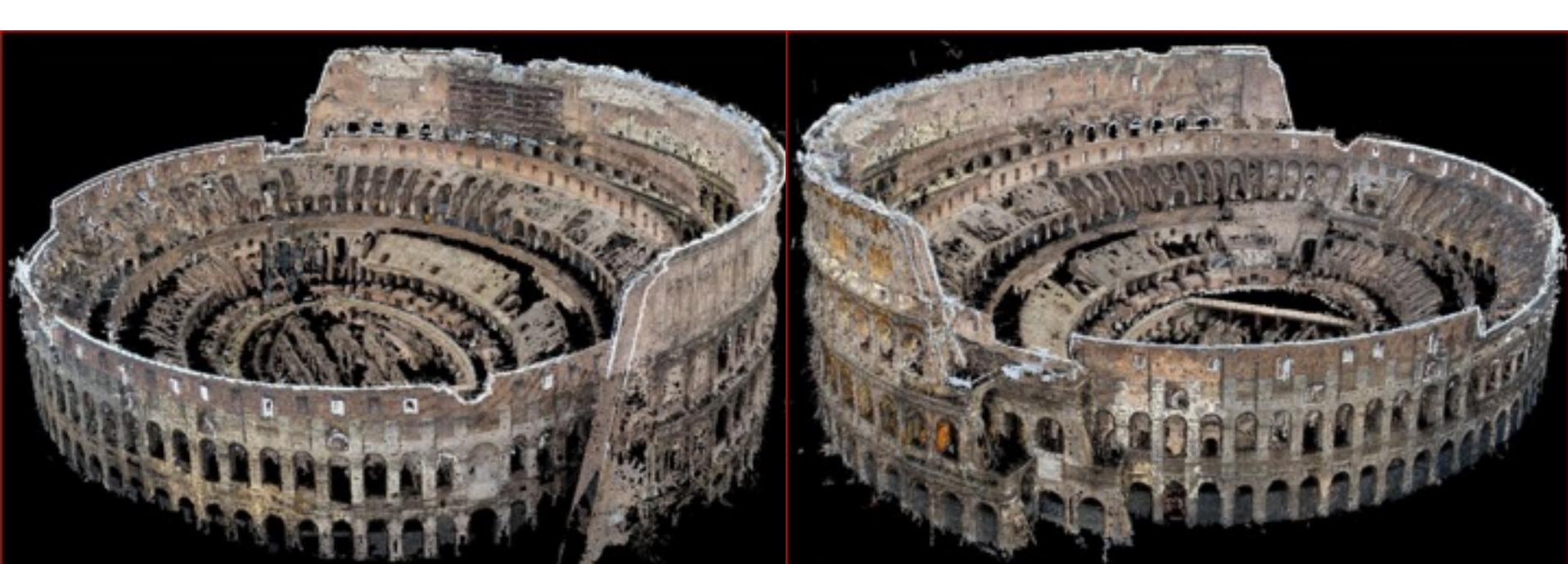
**Input Image  
(1 of 100)**



**Views of Reconstruction**

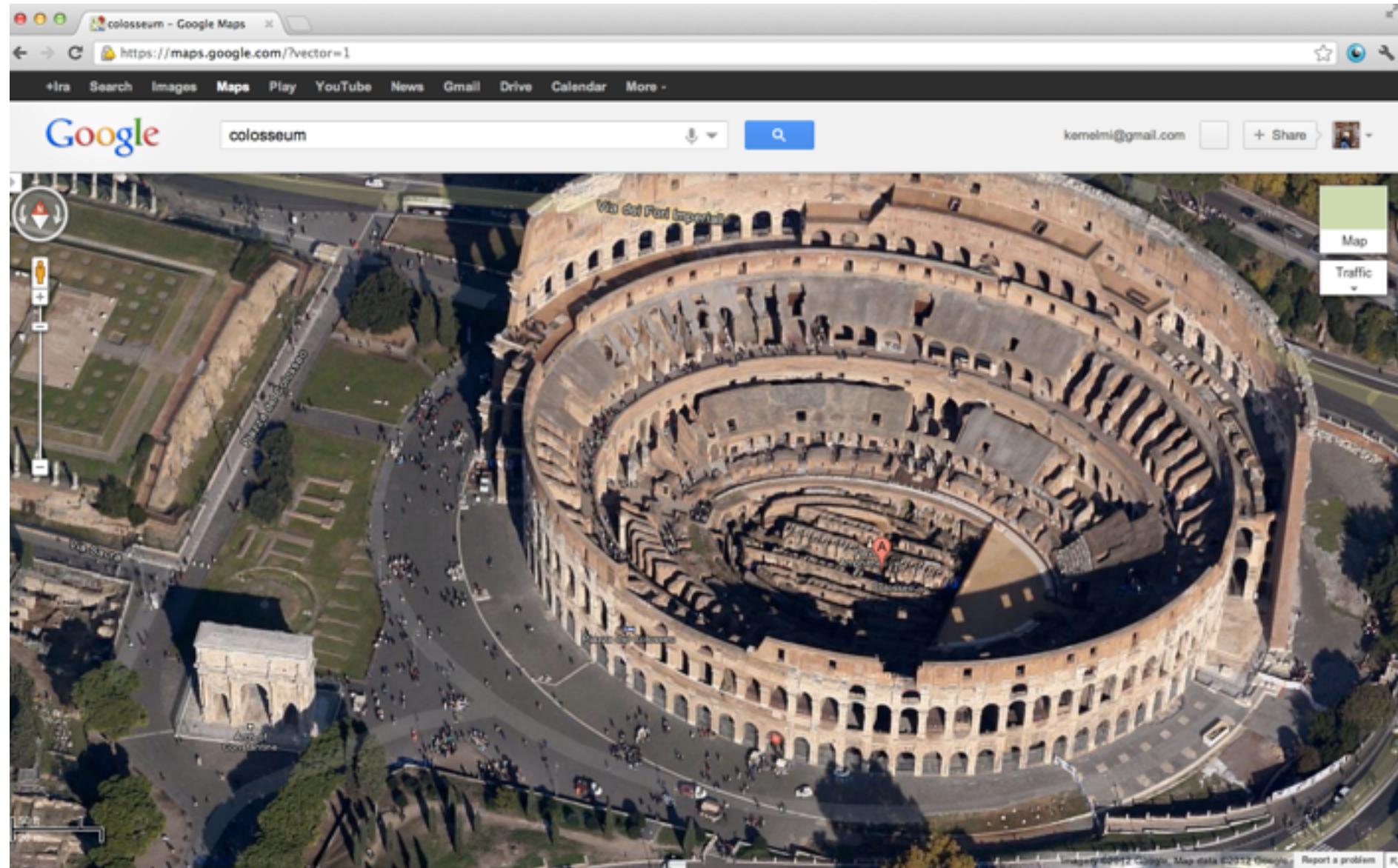


Yasutaka Furukawa and Jean Ponce, [Carved Visual Hulls for Image-Based Modeling](#), ECCV 2006.



# Google's 3D Maps

## Structure estimation from tourist photos

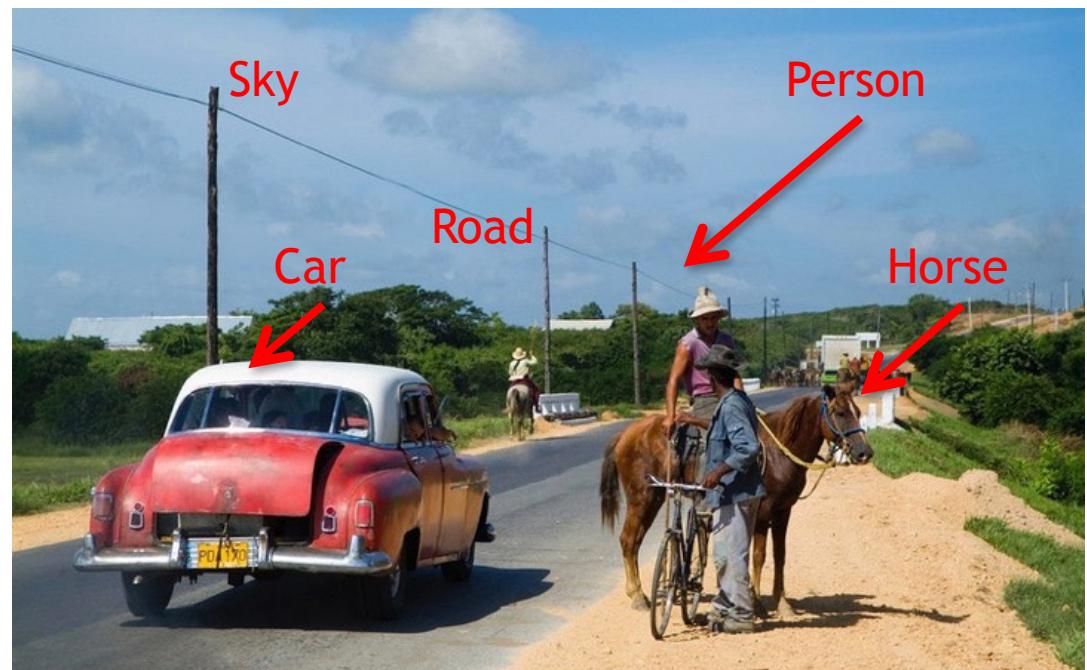


# Apple's 3D maps



# Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
- Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings
  - Pose estimation



# Visual Recognition?

- What does it mean to “see”?
  - “What” is “where”, Marr 1982
- Get computers to “see”

# Visual Recognition

Verification

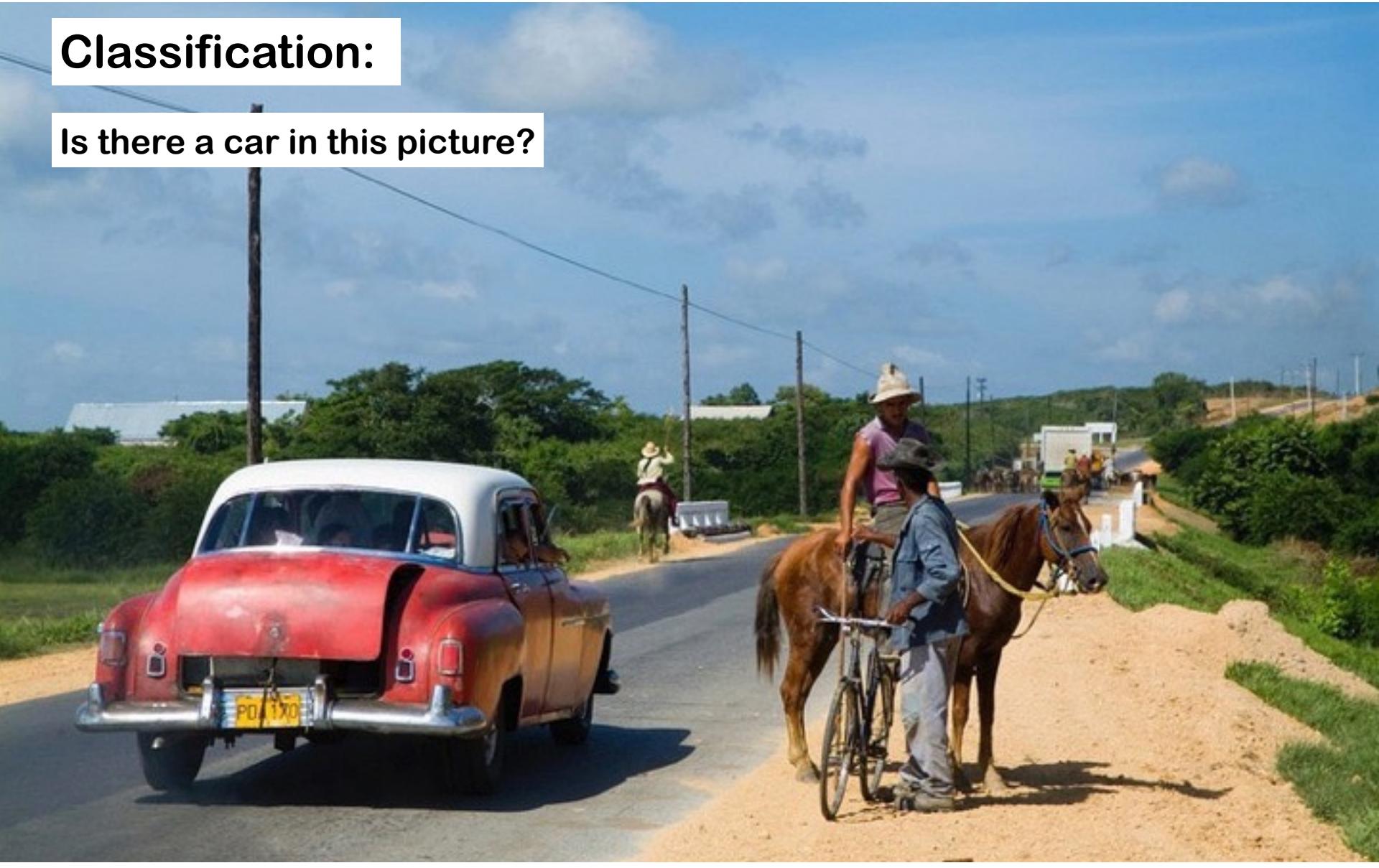
Is this a car?



# Visual Recognition

Classification:

Is there a car in this picture?



# Visual Recognition

Detection:

Where is the car in this picture?



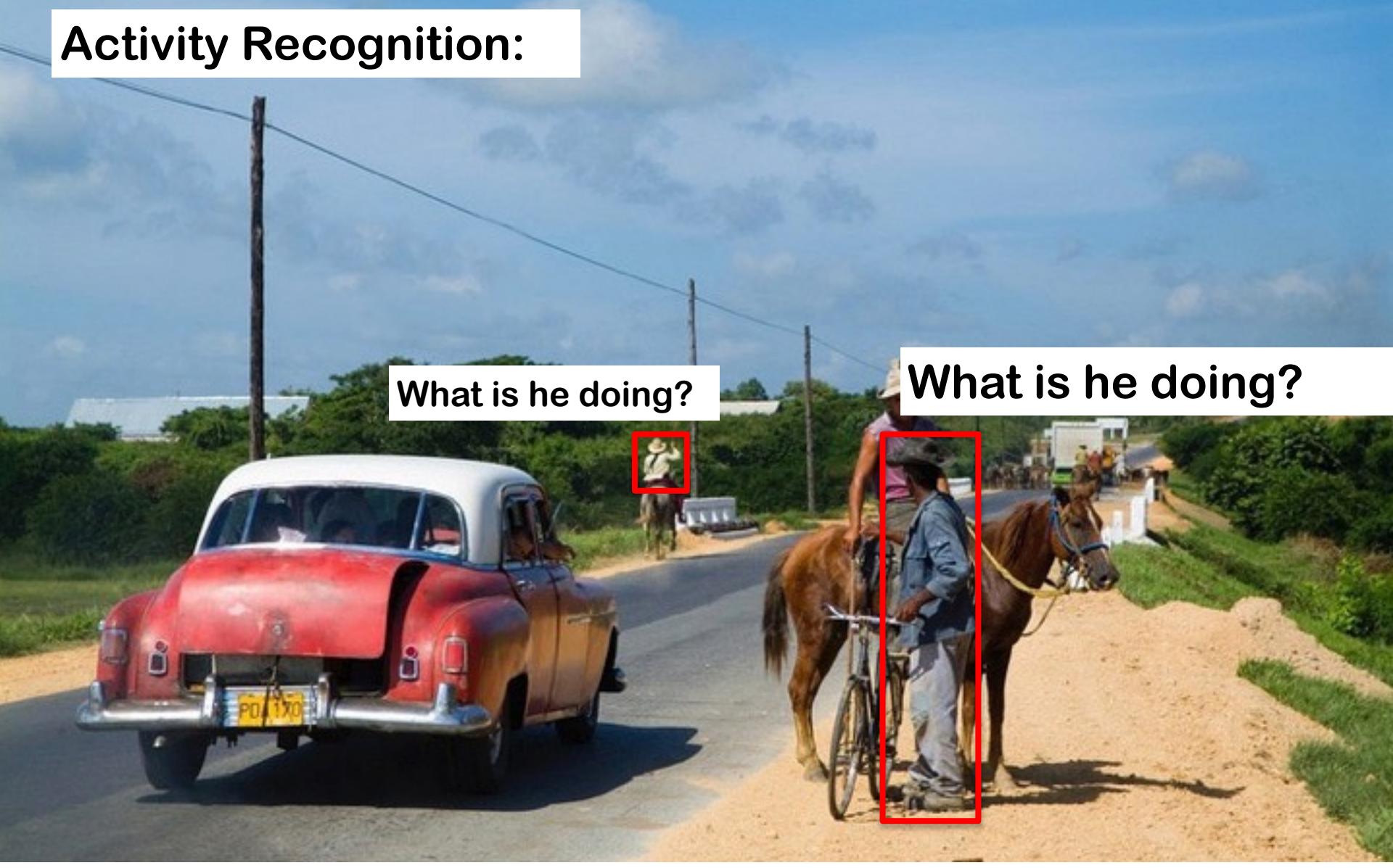
# Visual Recognition

Pose Estimation:



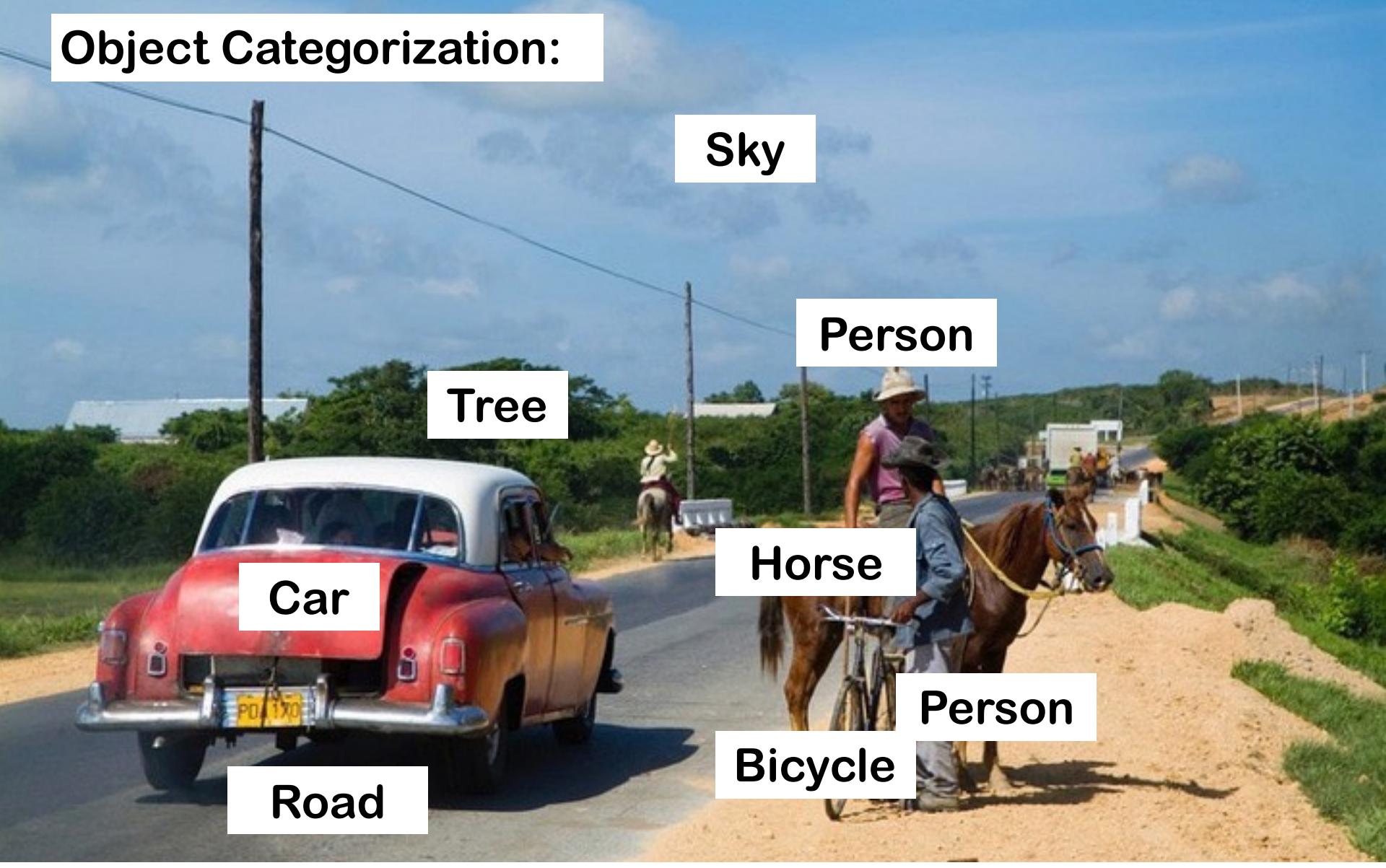
# Visual Recognition

Activity Recognition:



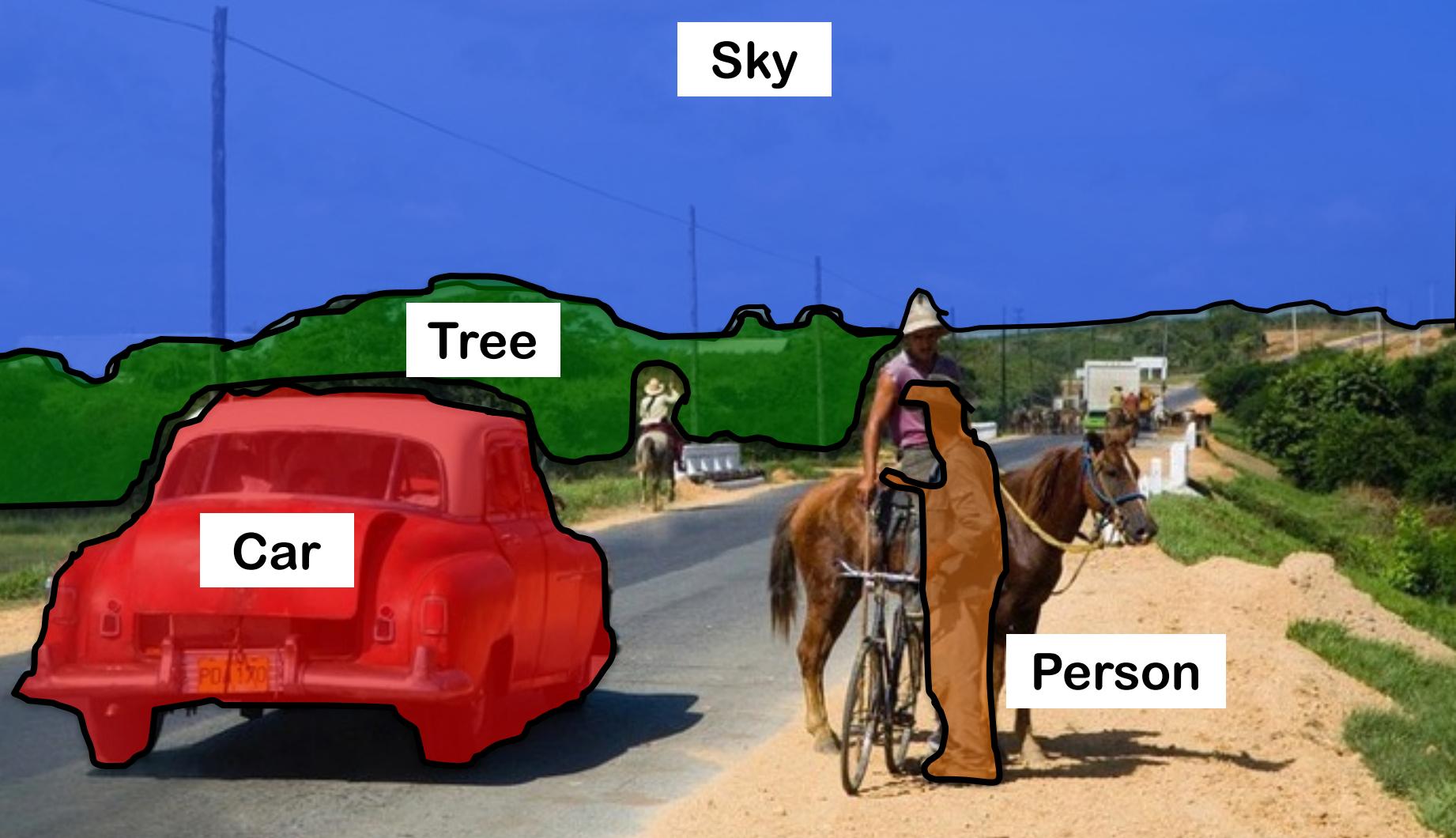
# Visual Recognition

Object Categorization:

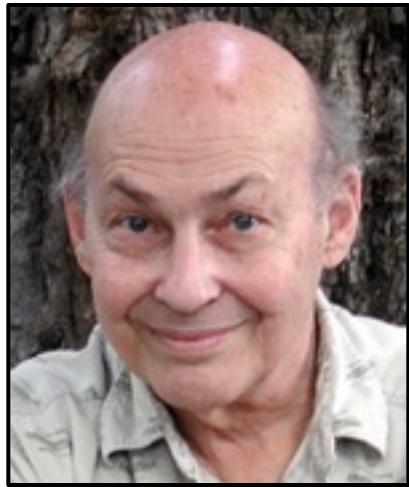


# Visual Recognition

Segmentation



# How hard is computer vision?



Marvin Minsky, MIT  
Turing award, 1969

“In 1966, Minsky hired a first-year undergraduate student and assigned him a problem to solve over the summer: connect a television camera to a computer and get the machine to describe what it sees.”

Crevier 1993, pg. 88

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
PROJECT MAC

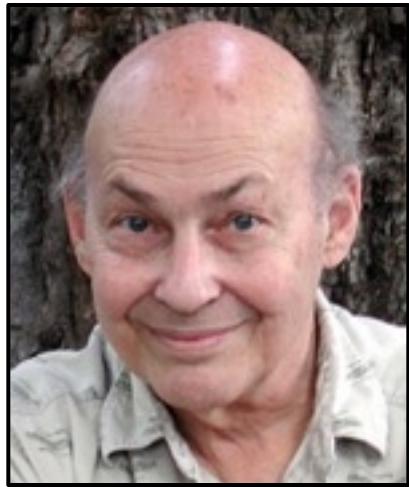
Artificial Intelligence Group  
Vision Memo. No. 100.

July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".



Marvin Minsky, MIT  
Turing award, 1969



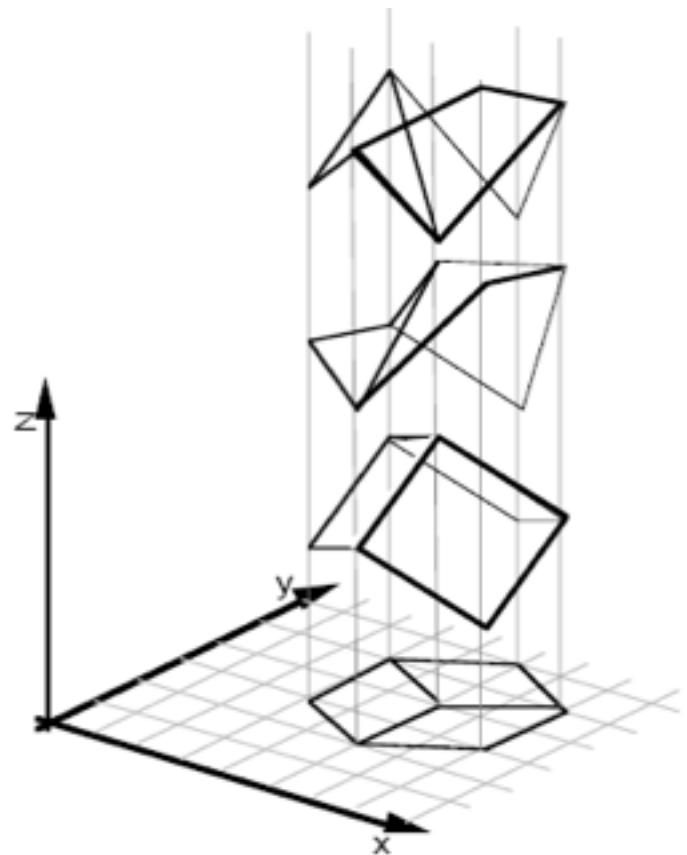
Gerald Sussman, MIT

“You’ll notice that Sussman never worked in vision again!” – Berthold Horn

# Why vision is so hard?

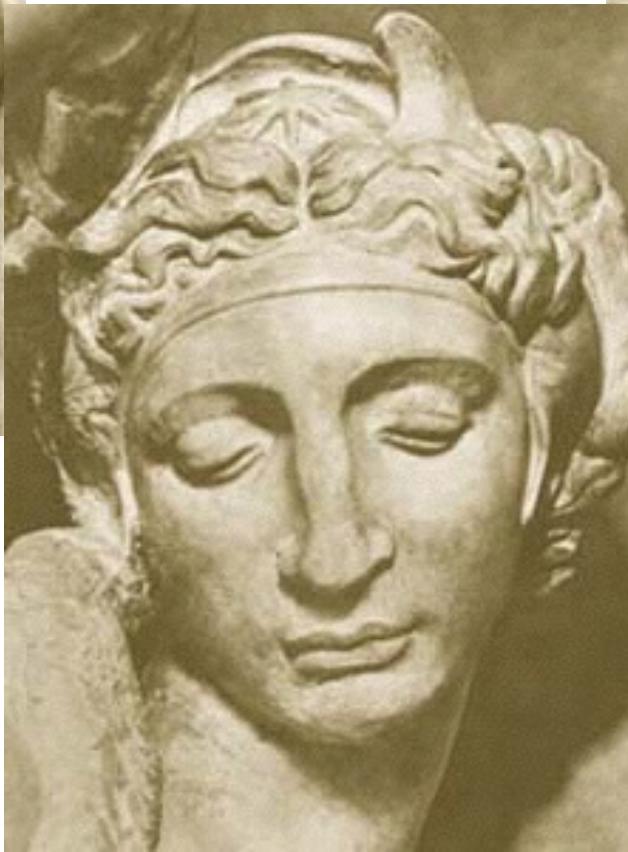
# Why is vision so hard?

- Ill-posed problem



[Sinha and Adelson 1993]

# Challenges 1: view point variation



Michelangelo 1475-1564

slide by Fei Fei, Fergus & Torralba

## Challenges 2: illumination



slide credit: S. Ullman

# Challenges 3: occlusion



Magritte, 1957

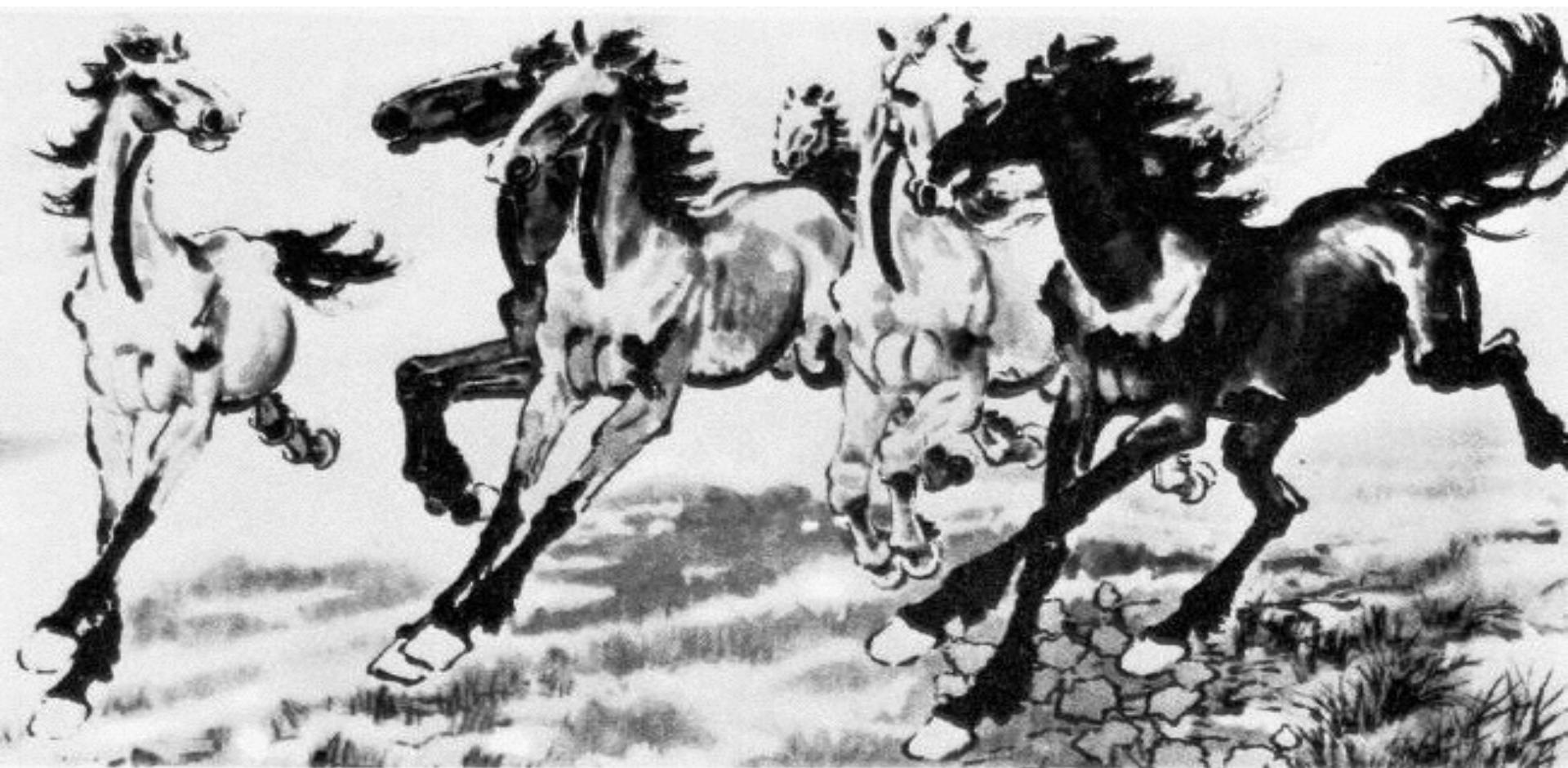
slide by Fei Fei, Fergus & Torralba

# Challenges 4: scale



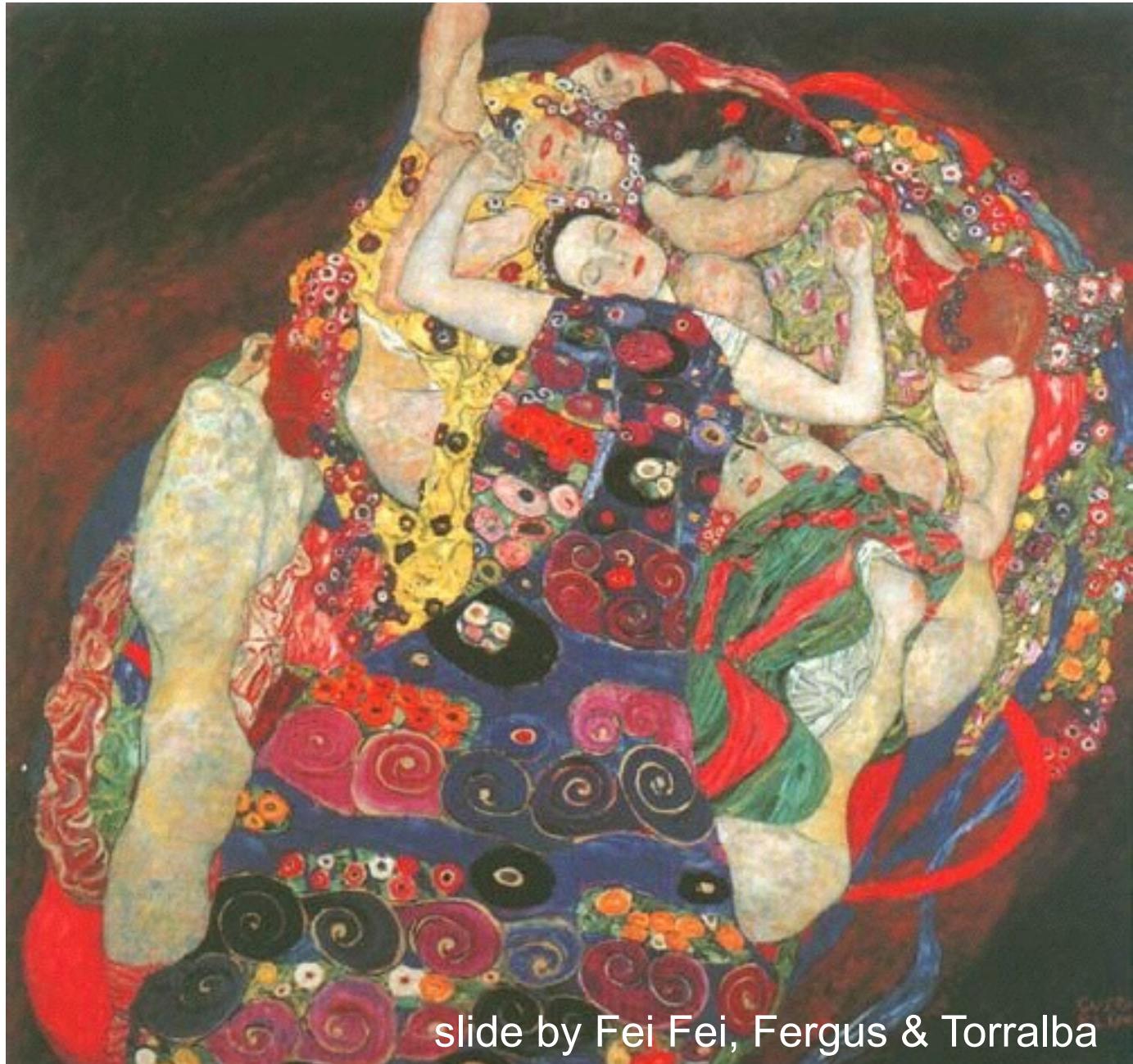
slide by Fei Fei, Fergus & Torralba

# Challenges 5: deformation



Xu, Beihong 1943

# Challenges 6: background clutter



Klimt, 1913

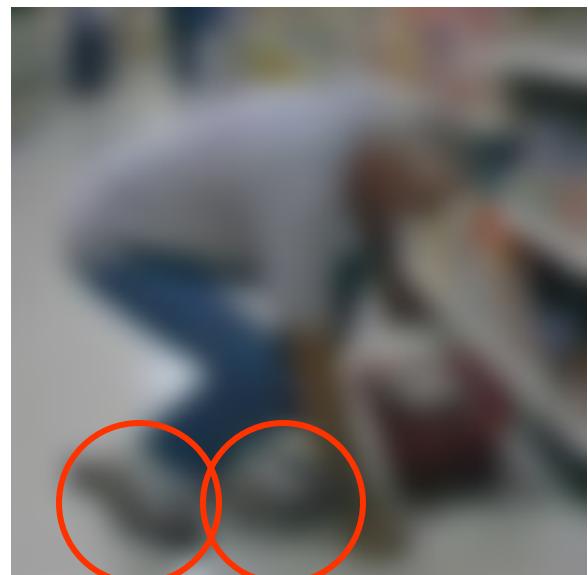
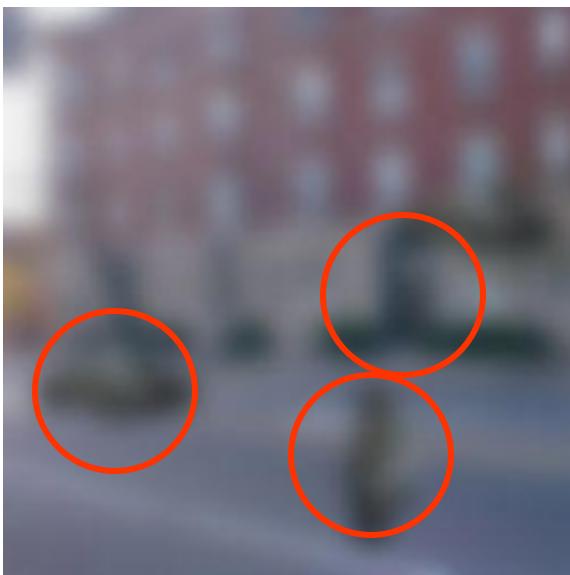
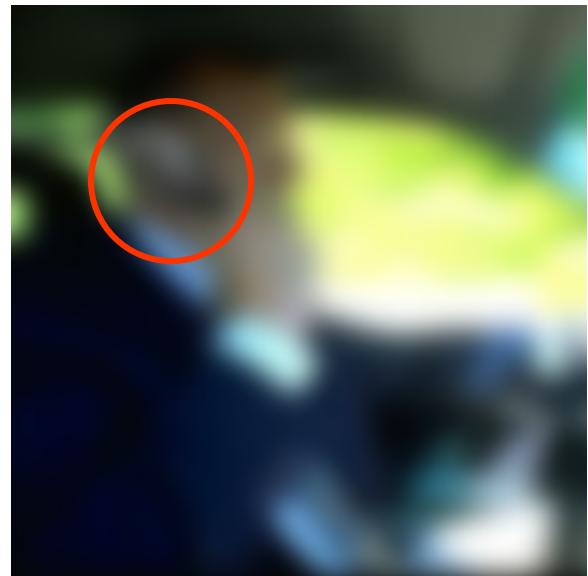
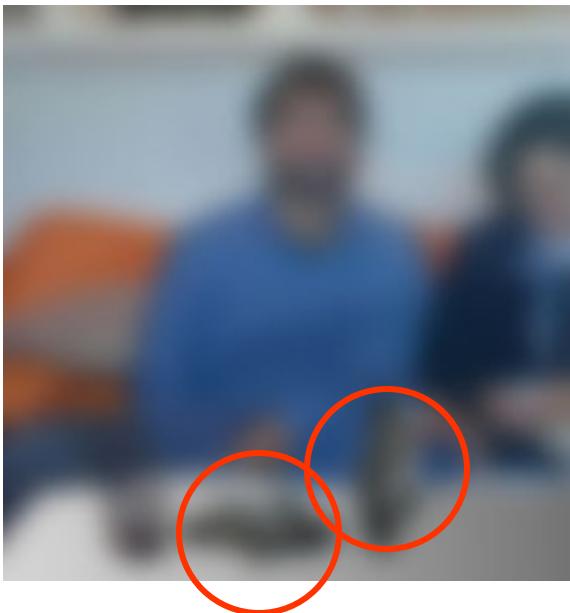
slide by Fei Fei, Fergus & Torralba

# Challenges 7: object intra-class variation



slide by Fei-Fei, Fergus & Torralba

# Challenges 8: local ambiguity



slide by Fei-Fei, Fergus & Torralba

# Challenges 9: the world behind the image



# What Works Today?

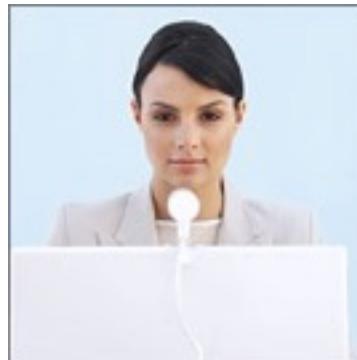
- Reading license plates, zip codes, checks

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

# 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

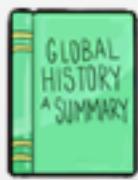
# Mobile visual search: Google Goggles

## Google Goggles in Action

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



Landmark



Book



Contact Info.



Artwork



Places



Wine



Logo

A photograph of a white smartphone held in a hand. The screen displays search results for "Terrazas de los Andes".

Google goggles labs

Wine  
Bodegas Terrazas De Los Andes Malbec Reserva 2004

winelibrary....

Web Results

**Terrazas de los Andes**  
2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002, 2001, 2000, 1999, 1998, 1997, 1996, 1995, 1994, 1993, 1992, 1991, 1990, 1989, 1988, 1987, 1986, 1985, 1984 ...  
<http://www.terrazasdelosandes.com/>

**Bodegas Terrazas de los Andes Winery**  
(Perdriel, Luján de Cuyo , AR ...)  
Popular wines by **Bodegas Terrazas de los Andes**.

# Face detection

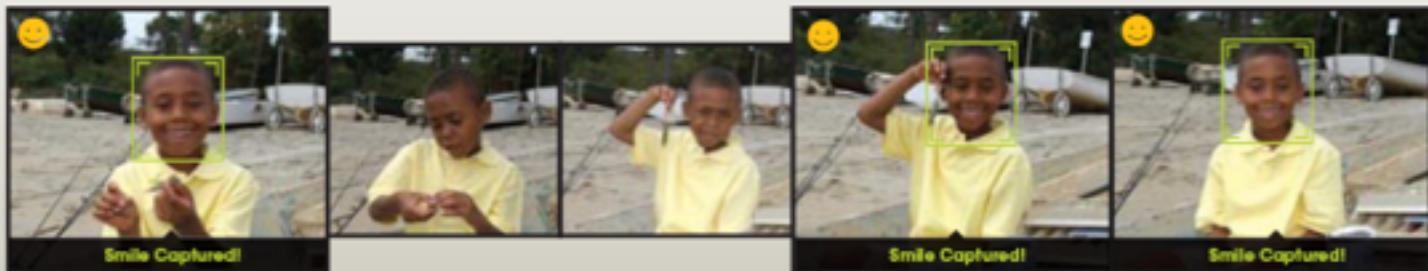
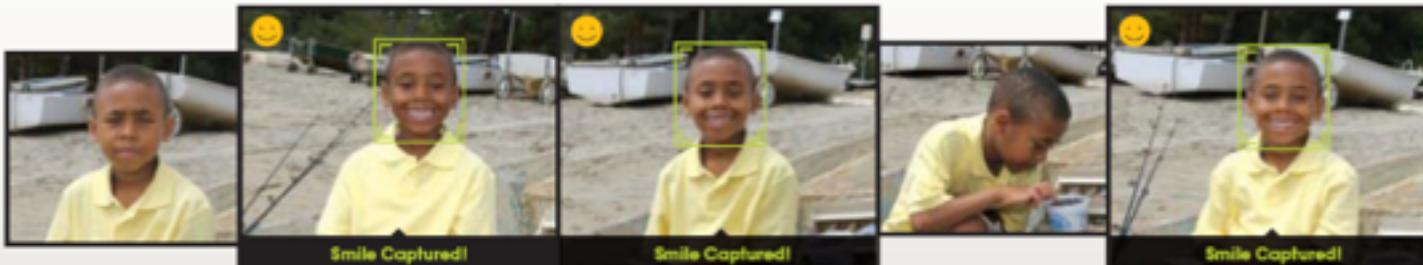


- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, ...

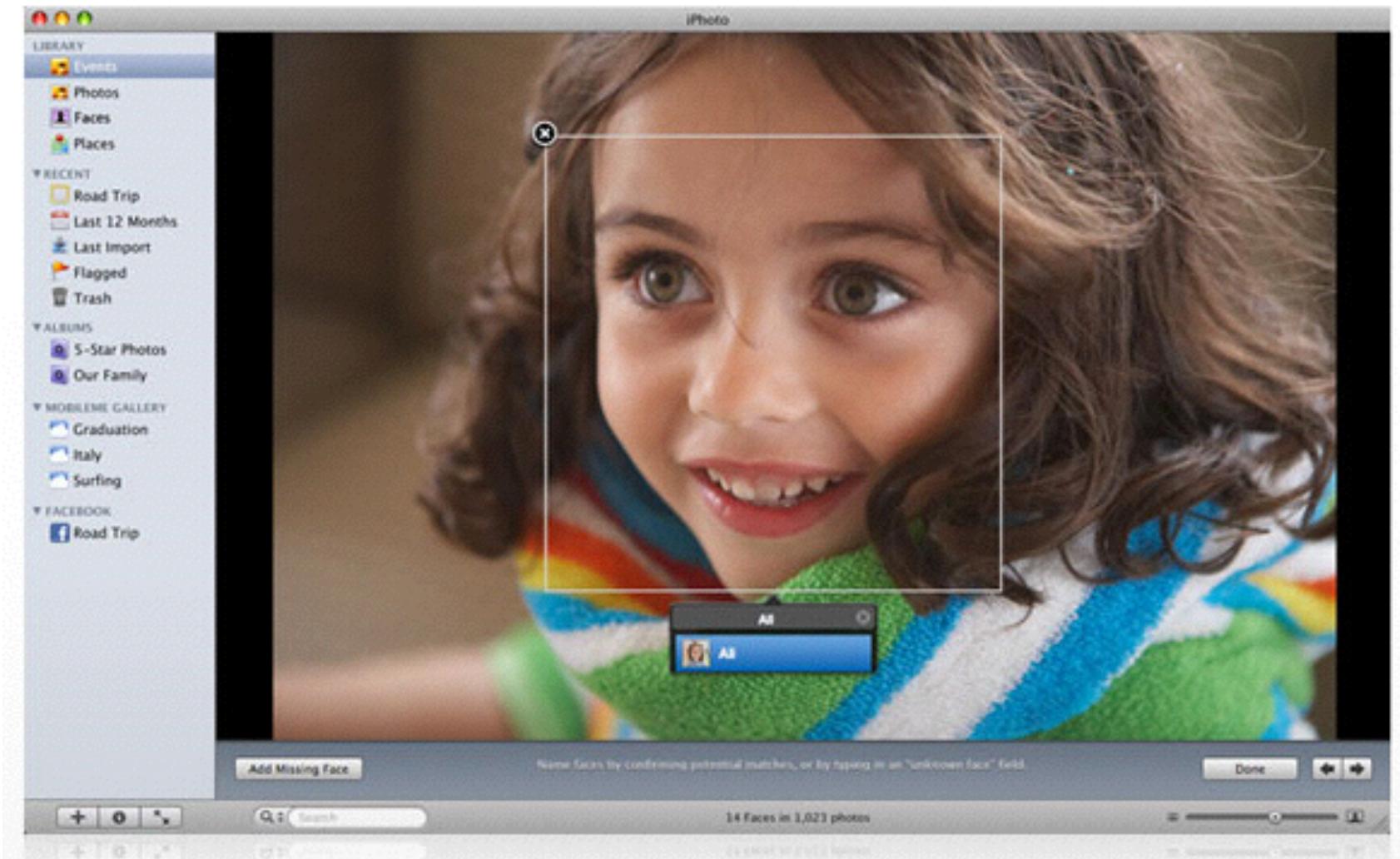
# Smile detection

## 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.



# Face recognition: Apple iPhoto, Facebook, Google, etc



# 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... “



[News Front Page](#)



[Africa](#)

[Americas](#)

[Asia-Pacific](#)

[Europe](#)

[Middle East](#)

[South Asia](#)

**UK**

[England](#)

[Northern Ireland](#)

[Scotland](#)

**Wales**

[UK Politics](#)

[Education](#)

[Magazine](#)

**Business**

[Health](#)

[Science & Environment](#)

Last Updated: Wednesday, 31 August 2005, 05:44 GMT 06:44 UK

[E-mail this to a friend](#)

[Printable version](#)

## Computer alert for drowning girl

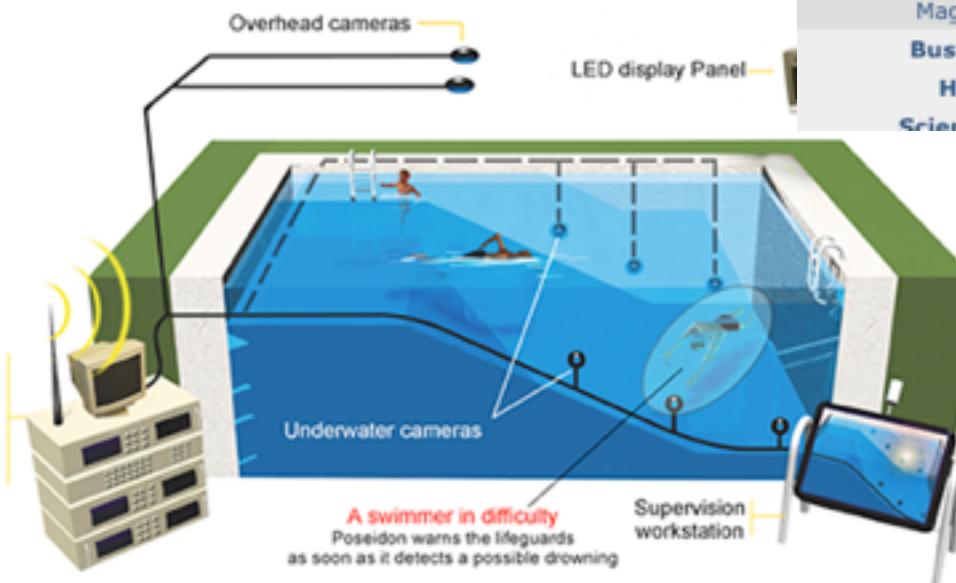
**A 10-year-old girl has been saved from drowning by a computer system designed to raise the alarm when swimmers get into difficulties.**



[▶ VIDEO Watch the rescue](#)

The girl, from Rochdale, was at the deep end of the pool in Bangor, north Wales, when she sank to the bottom.

The £65,000 system, called Poseidon, detected her on the pool floor and sounded the alarm. A lifeguard pulled her out and she recovered in hospital.



# Security

Local 

## Cameras help confirm Scott suicide ruling

Friday, December 04, 2009



TAGS: [local](#), [paul meincke](#)

 Comment Now  Email  Print  Report a typo      



**Paul Meincke**

More: [Bio](#), [News Team](#)

December 4, 2009 (CHICAGO) (WLS) -- Chicago police have closed the case in the death of Chicago School Board President Michael Scott.

Police Supt. Jody Weis says investigators used police cameras in the city to trace Scott's last steps in the hours before his body was found in November.

Scott's death has been ruled a suicide. The medical examiner's office concluded --not long after Scott's body was found -- that he had committed suicide. Police did not dispute the finding but wanted to pursue all the investigative leads they could. They say they have done that and have now reached the same conclusion.

### Share this Story



 Be the first of your friends to recommend this.



 Recommend this on Google

**News Headlines**  [Video](#) 

- 2 suspects arrested in volleyball star's murder 47 min ago
- BP Gas Recall: BP finds, fixes source of bad gas
- Teachers union, board resume negotiating
- Back to School
- 5 injured in South Side shooting 49 min ago
- Pastor: Stacy Peterson said she lied for Drew



# Automotive safety

The screenshot shows the Mobileye website interface. At the top, there are two tabs: "manufacturer products" (highlighted in blue) and "consumer products". Below the tabs, the slogan "Our Vision. Your Safety." is displayed. The main visual is a top-down view of a car with three camera systems highlighted: "rear looking camera" on the left, "forward looking camera" on the right, and "side looking camera" at the bottom. Below this, there are three sections: "EyeQ Vision on a Chip" featuring a close-up of a chip, "Vision Applications" showing a pedestrian crossing a street, and "AWS Advance Warning System" showing a dashboard display. To the right, there is a sidebar titled "News" with links to articles about Volvo's collision warning system and Mobileye's presence at Equip Auto and SEMA. There is also a section titled "Events" with a link to read more.

- > **EyeQ** Vision on a Chip
- > **Vision Applications** Road, Vehicle, Pedestrian Protection and more
- > **AWS** Advance Warning System

News

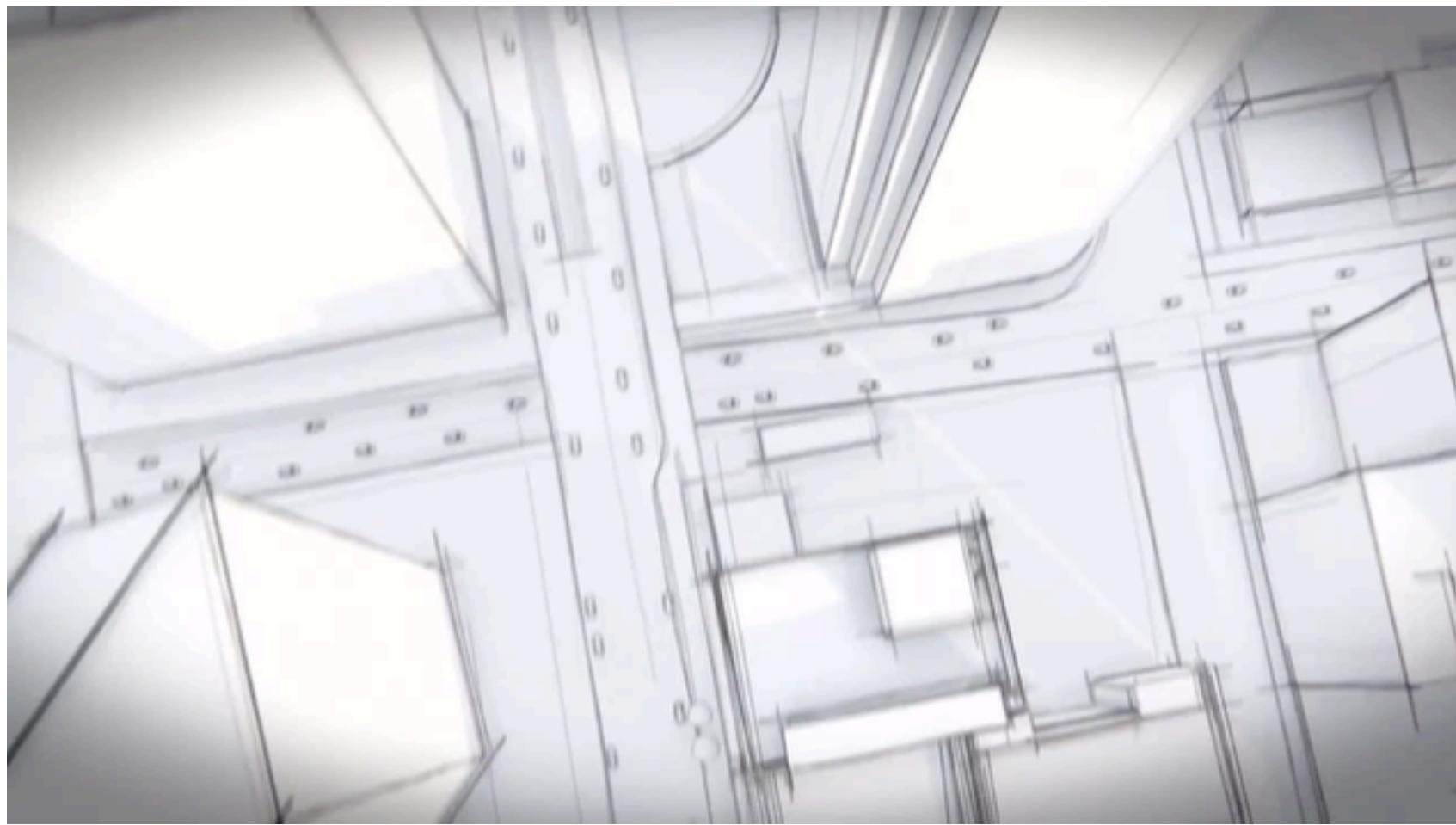
- > **Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System**
- > **Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end**

Events

- > **Mobileye at Equip Auto, Paris, France**
- > **Mobileye at SEMA, Las Vegas, NV**

- 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



# Google cars

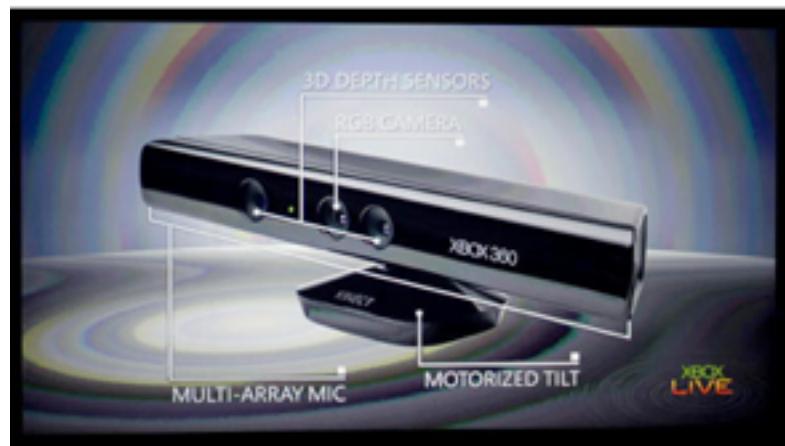


Oct 9, 2010. "[Google Cars Drive Themselves, in Traffic](#)". *The New York Times*. John Markoff

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)*

# Vision-based interaction: Xbox Kinect



# Kinect Fusion

SIGGRAPH Talks 2011

## KinectFusion:

**Real-Time Dynamic 3D Surface  
Reconstruction and Interaction**

**Shahram Izadi 1, Richard Newcombe 2, David Kim 1,3, Otmar Hilliges 1,  
David Molyneaux 1,4, Pushmeet Kohli 1, Jamie Shotton 1,  
Steve Hodges 1, Dustin Freeman 5, Andrew Davison 2, Andrew Fitzgibbon 1**

1 Microsoft Research Cambridge 2 Imperial College London  
3 Newcastle University 4 Lancaster University  
5 University of Toronto

# Augmented reality, consumer products



# Special effects: shape and motion capture



Source: S. Seitz

# Vision for robotics, space exploration

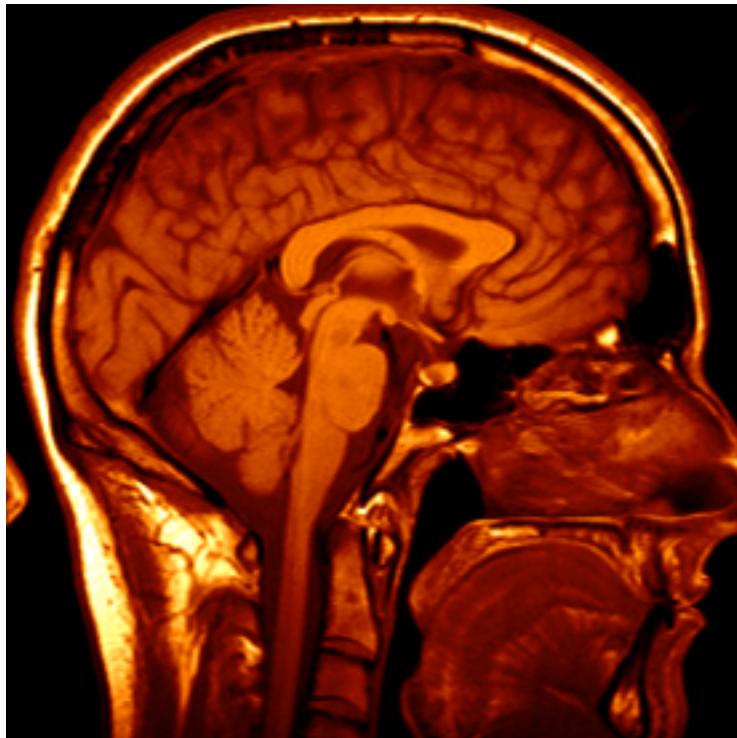


[NASA's Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

## 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.

# Medical imaging



3D imaging  
MRI, CT



Image guided surgery  
Grimson et al., MIT

## Computer vision in other scientific fields

# Computer vision research in biology

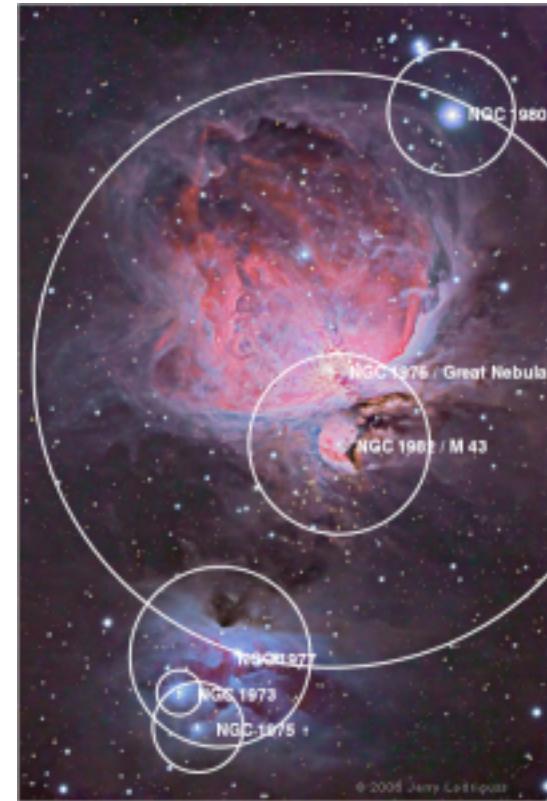
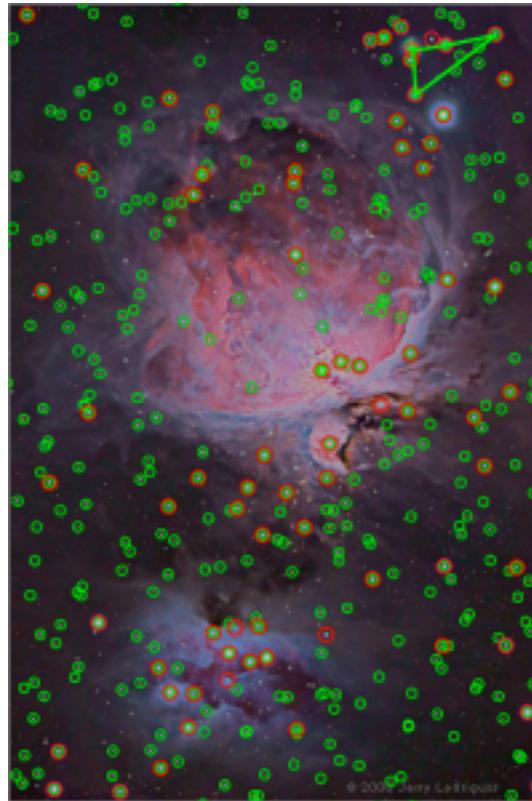
The screenshot shows the Leafsnap mobile application interface. At the top, there are buttons for 'First', 'Last', and 'Scientific'. Below that is a section labeled 'A' with a scroll bar. The first item listed is 'American Holly' (*Ilex opaca*), followed by 'American Linden' (*Tilia americana*), 'American Sycamore' (*Platanus occidentalis*), and 'Amur Corktree' (*Pheilodendron amurense*). At the bottom are navigation icons for 'Home', 'Browse', 'Collection', 'Options', and 'Snap It!'



<http://leafsnap.com/>

<http://www.vision.caltech.edu/visipedia/>

# Computer vision in cosmology



<http://astrometry.net/>

# Computer vision research in healthcare



assisted living, patient monitoring  
[Lan et al, PAMI 2012]



autism screening  
[http://www.gatech.edu/newsroom/  
release.html?nid=60509](http://www.gatech.edu/newsroom/release.html?nid=60509)

# Computer vision in the real-world

- Most examples are less than 5 years old
- Very active research area. Many new applications to come.
- A website of computer vision industries maintained by Prof. David Lowe (UBC):

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

# Tentative Syllabus

- **Image Processing (2 weeks)**
  - filtering, convolution
  - image pyramids
  - edge detection
  - feature detection (corners, lines)
  - hough transform
- **Image Transformation (2 weeks)**
  - image warping (parametric transformations, texture mapping)
  - image compositing (alpha blending, color mosaics)
  - segmentation and matting (snakes, scissors)
- **Motion Estimation (1 week)**
  - optical flow
  - image alignment
  - image mosaics
  - feature tracking

# Syllabus

## 3D Modeling (1 weeks)

- projective geometry
- camera modeling
- single view metrology
- camera calibration
- stereo

## • Computational Photography (1 week)

- Super resolution
- Alpha Matting
- Blur removal
- Poisson Blending

## • Visual Recognition (3 week)

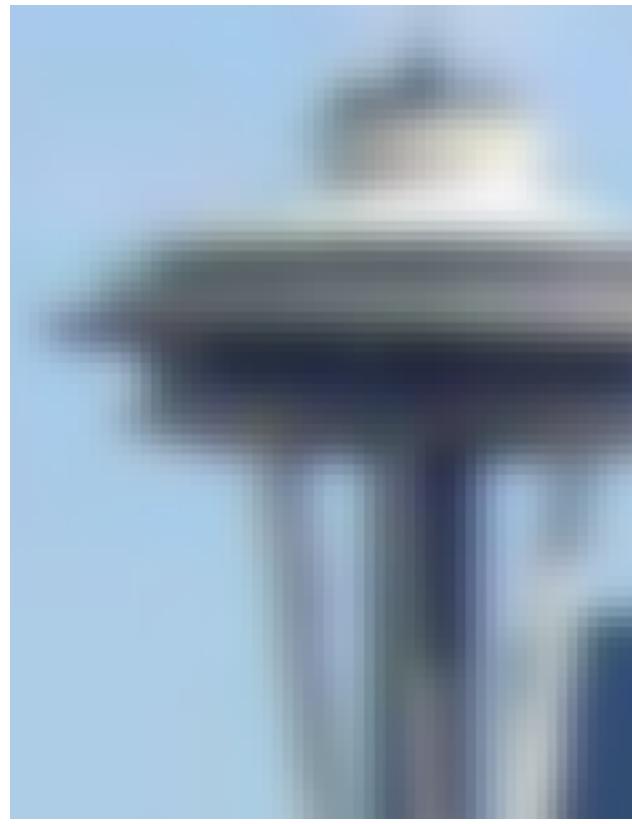
- Eigenfaces
- Category Recognition
- Object Detection
- Kinect

# Grading

- Four assignments (10 each+ extra points)
  - Mix of coding and written answers.
  - Using Qt (cross platform UI in c++) qt.nokia.com
  - Use of interactive UIs for exploring and gaining intuition
    - 1. Filters and edge detection
    - 2. Creating panoramas
    - 3. Computing depth from stereo
    - 4. Face detection
- **FINAL PROJECT (60 points + 20 extra points)**

# Assignment 1: Image Filtering

## 10 Points



# Assignment 2: Panorama Stitching

## 10 Points



# Assignment 3: Stereo Reconstruction

## 10 Points



# Assignment 4: Face Detection

## 10 Points



# Final Project

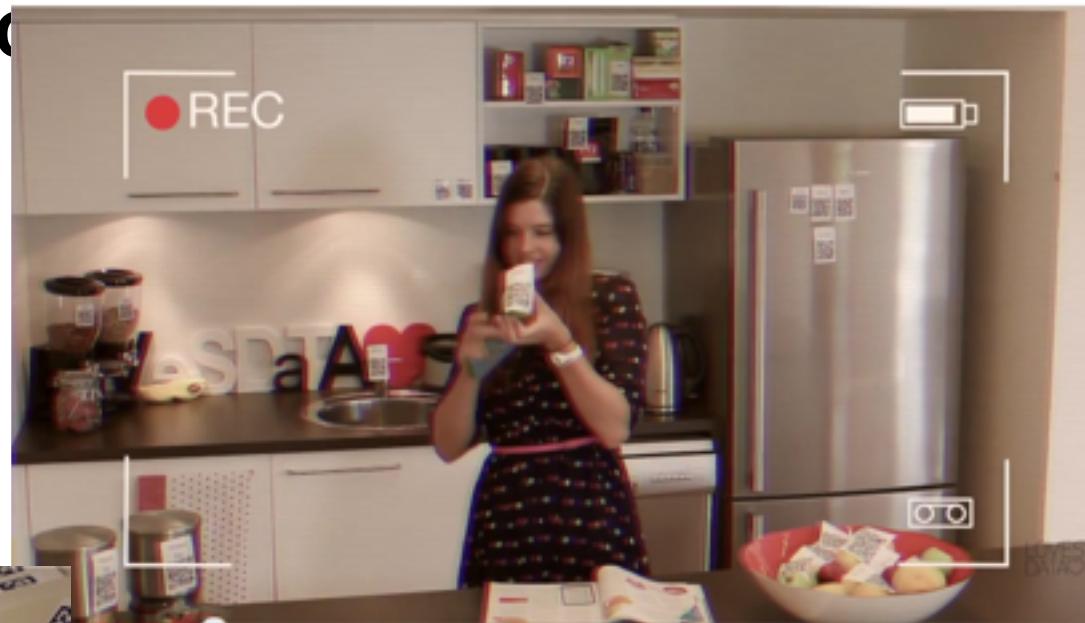
## 60 Points + 20 Extra points

- Big Project
  - Better if related to your own research
  - Demo is a **BIG** plus
- Proposal is due on 4/8
  - One Paragraph,
  - Crisp final outcome/deliverable
- Progress Reports are due on
  - 4/20, 5/4, 5/18, 6/1
  - What has changed since last report
- Final Presentation will be on 6/3,
  - Demo/Posters @ CSE atrium

# Sample Projects

From Taskar Center for Accessible Technology

Project: My Kingdom



# Sample Projects

From Taskar Center for Accessible Technology

**Project: Curb Alert**



# Sample Projects

From Taskar Center for Accessible Technology

## Project: Silent Movie



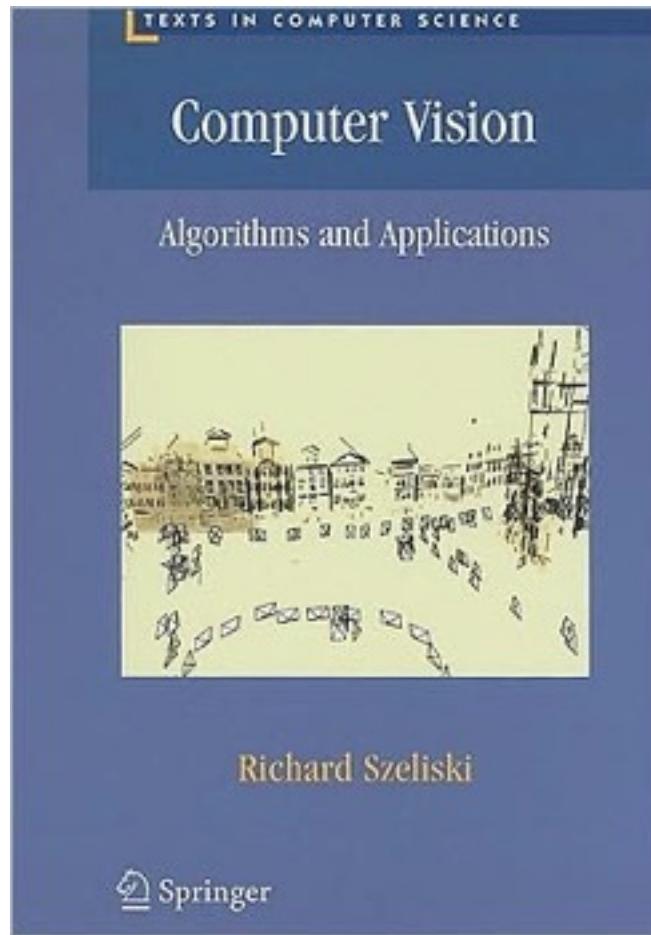
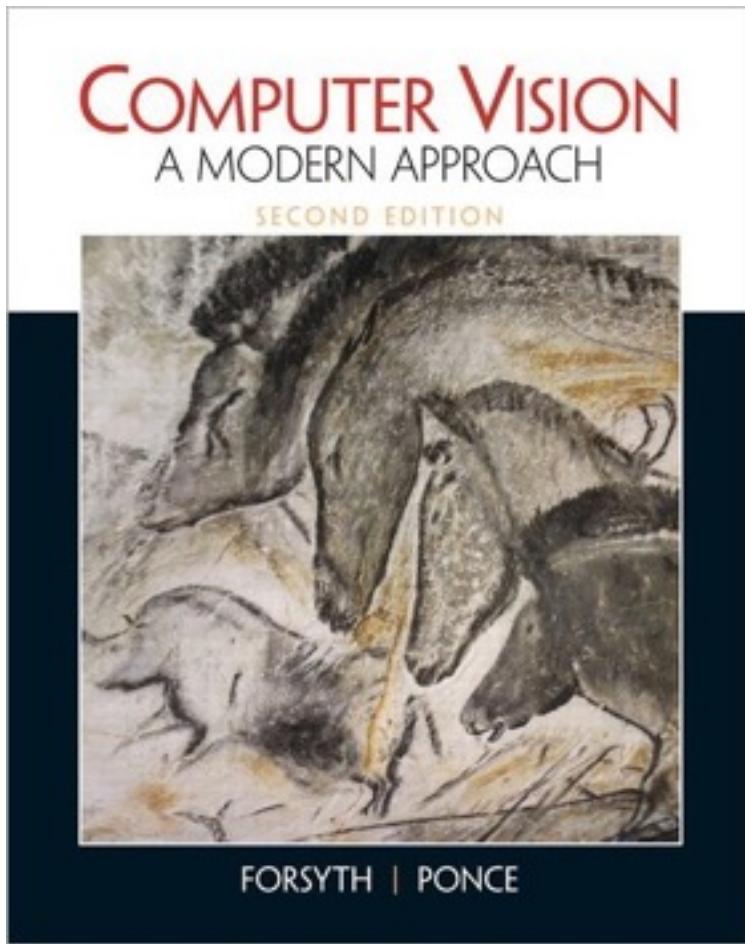
# Samples of Previous Projects

- Visual Calculator
- Seam Carving
- X-ray bone fracture detection
- Pipe leak detection
- Is it gonna be viral?
- Deep learning for object recognition
- ...

# Project Ideas

- Seam Carving
- Detecting Shadows
- Features
  - Learning Features
  - Features for regions
  - Comparison of features in the literature
- Action Recognition
  - Human pose
  - Objects and Interactions
  - Using Kinect
  - Detecting unattended luggages
  - Egocentric
- Grab cut
- Video Stabilization
- RGBD object Detection
- Object Detection in Videos
  - Video Google
- Matching Images and Videos in the wild
- Reading Street Signs
  - Wearable Cameras for visually impaired users
  - Auto Zooming
- Visual Odometer
  - Smart stop lights
- Language & Vision

# Books



# Calibration

- How many of you
  - have taken an undergrad vision course?
  - have taken an ML course?
  - have taken a Graphics course?
  - Remember your linear algebra course in your undergrad?
  - have any concerns about programming?

# Do these words remind you of something?

Interest Point	SIFT	
Laplacian	Eigenvalue	
SVD	SVM	
MRF	STEREO	
Random Forest	Graph cut	

# Preferences

- Low level vision?
- Mid level vision?
- High level vision?