



UNIVERSITÀ DI PARMA

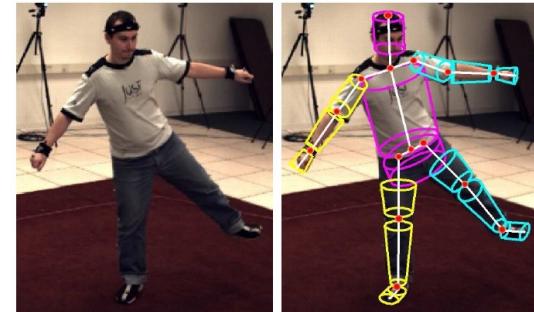
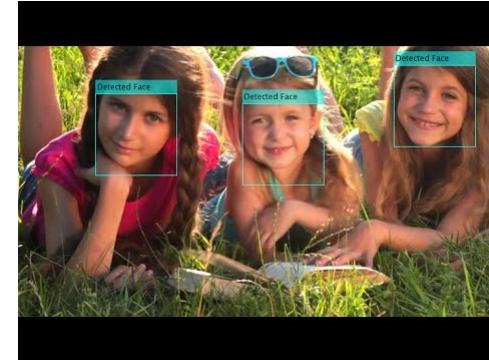
# Computer Vision

- Massimo Bertozzi
  - bertozzi@ce.unipr.it
  - +39 0521 905845
  - Tentatively: Tuesday Afternoon (>3:30 PM)

# Course Objectives



- Basic and Semi-advanced techniques for image processing & machine vision
- Computer vision is everywhere...





## 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?

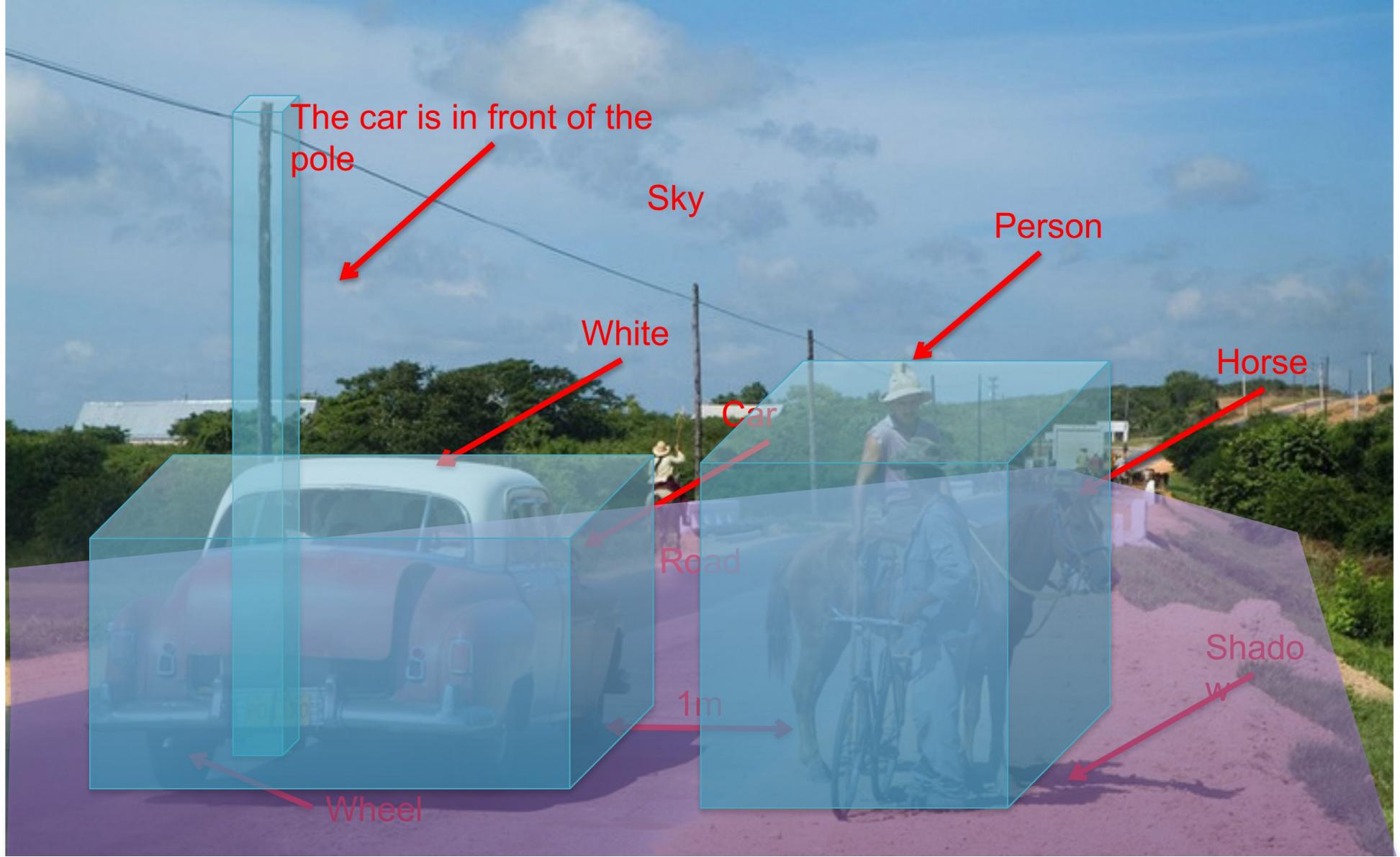
What's the use! Call and see what you are buying. One look at our display of automobile and motorcycle accessories will convince you of the fact.

THAT WE HAVE EVERYTHING FOR  
THE AUTO

# Piqua Auto Supply House

133 N. Main St.—Piqua, O.







# How hard is computer vision?

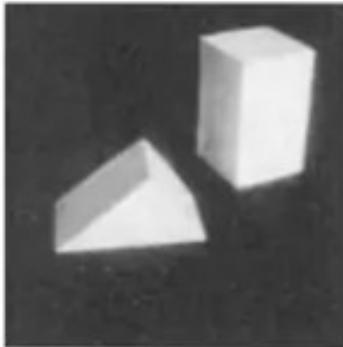
# Some history...



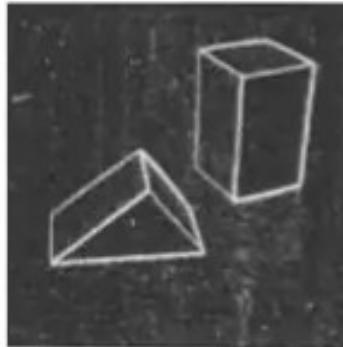
1960's: interpretation of synthetic worlds



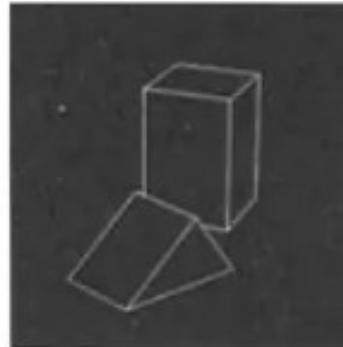
Larry Roberts  
"Father of Computer Vision"



Input image



2x2 gradient operator

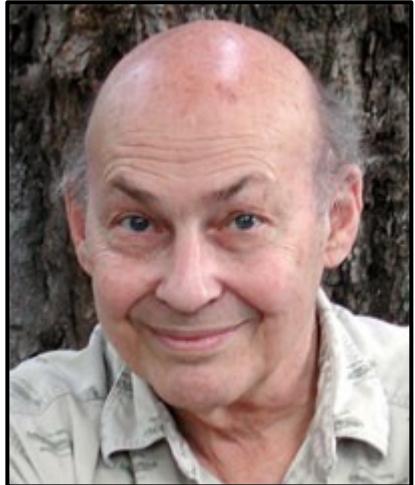


computed 3D model  
rendered from new viewpoint

Larry Roberts PhD Thesis, MIT, 1963,  
Machine Perception of Three-Dimensional Solids

Slide credit: Steve Seitz

# Some history...



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

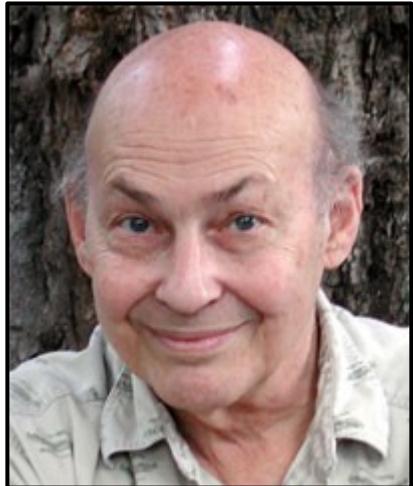
Seymour Papert

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".

# Some history...



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



Marvin Minsky, MIT  
Turing award, 1969



Gerald Sussman, MIT  
(the undergraduate)

# A little timeline...

- 1959: The very first digital image scanner was invented that converted images into number grids.
- 1963: Larry Roberts described the process of deriving 3D information of solid objects from 2D pictures.
- 1966: Marvin Minsky instructed a graduate student to attach a camera to a computer and describe what it saw.
- 1980: Kunihiko Fukushima created the neocognitron. It's considered the precursor of the modern convolutional neural network (CNN).
- 1998: VisLab tested a >2000 km trip on Italian Highways using an autonomous vehicle
- 2001: Paul Viola and Michael Jones, two researchers at MIT, created the first face detection framework that works in real time.
- 2009: Google started the self-driving car project.
- 2010: Google released Google Goggles, an image recognition app useful for searches based on pictures captured by mobile devices. The same year, Facebook started using facial recognition to tag people on photos effectively.
- 2011: Facial recognition technology was used to confirm the identity of Osama Bin Laden after he was killed.
- 2012: Google Brain created a neural network consisting of 16,000 computer processors that could recognize the pictures of cats with the help of a deep learning algorithm. The same year, AlexNet, a convolutional neural network, attained a top-5 error of 15.3% in the ImageNet 2012 Challenge.
- 2013: BRAiVE, VisLab's most advanced autonomous car, drove in downtown Parma in a fully autonomous mode
- 2014: Tesla introduced Autopilot in its Model S electric cars. The self-driving system not only worked offline but also parked with precision.
- 2015: Google launched TensorFlow, which is an open-source and free software library for machine learning. The same year, Google introduced FaceNet for facial recognition.
- 2017: Apple released the iPhone X with the face recognition feature.
- 2019: The UK HighCourt permitted the use of automated facial recognition technology to search for people in crowds.
- 2019: Italy allows testing autonomous vehicles on public roads!

Credits: (partly) Amal Joby

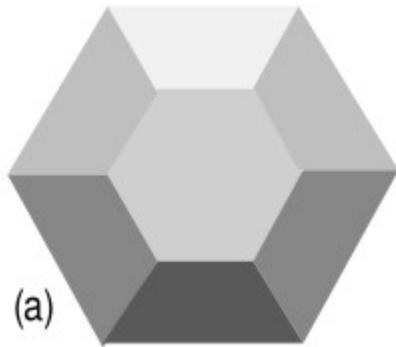


# Why vision is so hard?

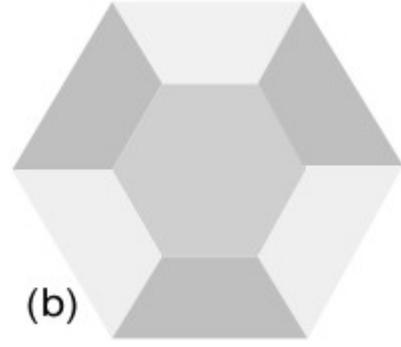
# Why vision is so hard?



Ill-posed problem

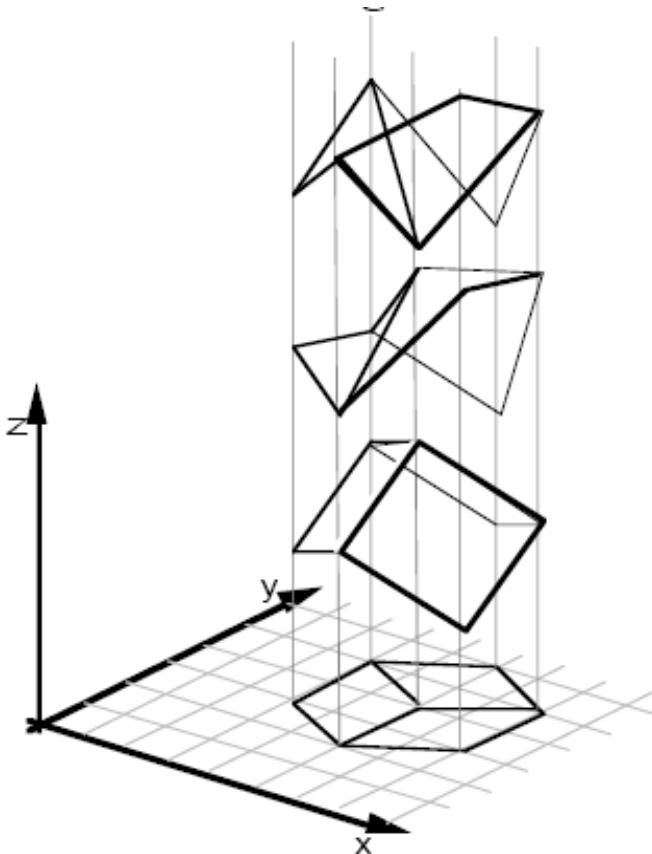


(a)



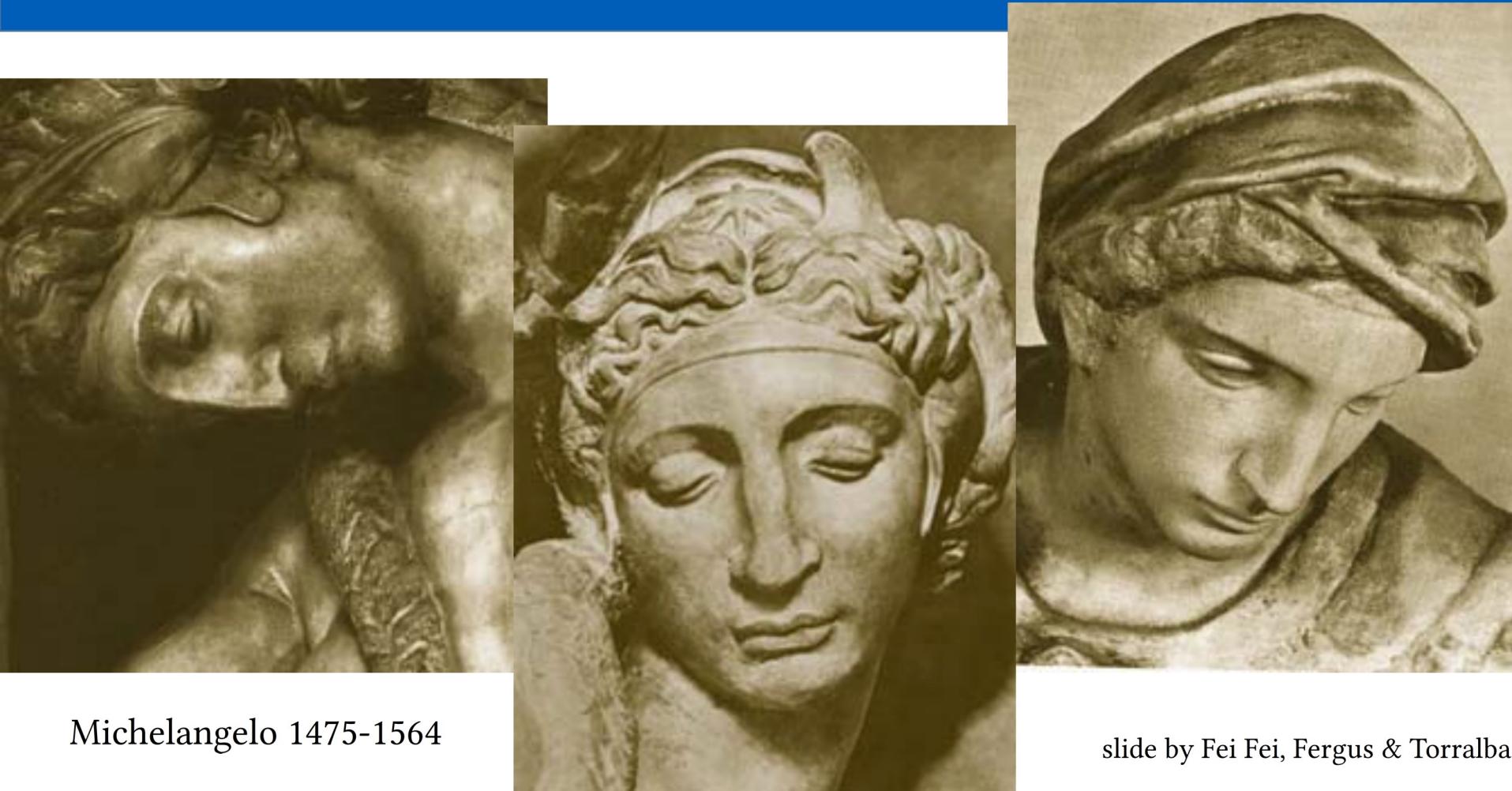
(b)

[Sinha and Adelson 1993]



# Challenges 1: view point variation

UNIVERSITÀ  
DI PARMA



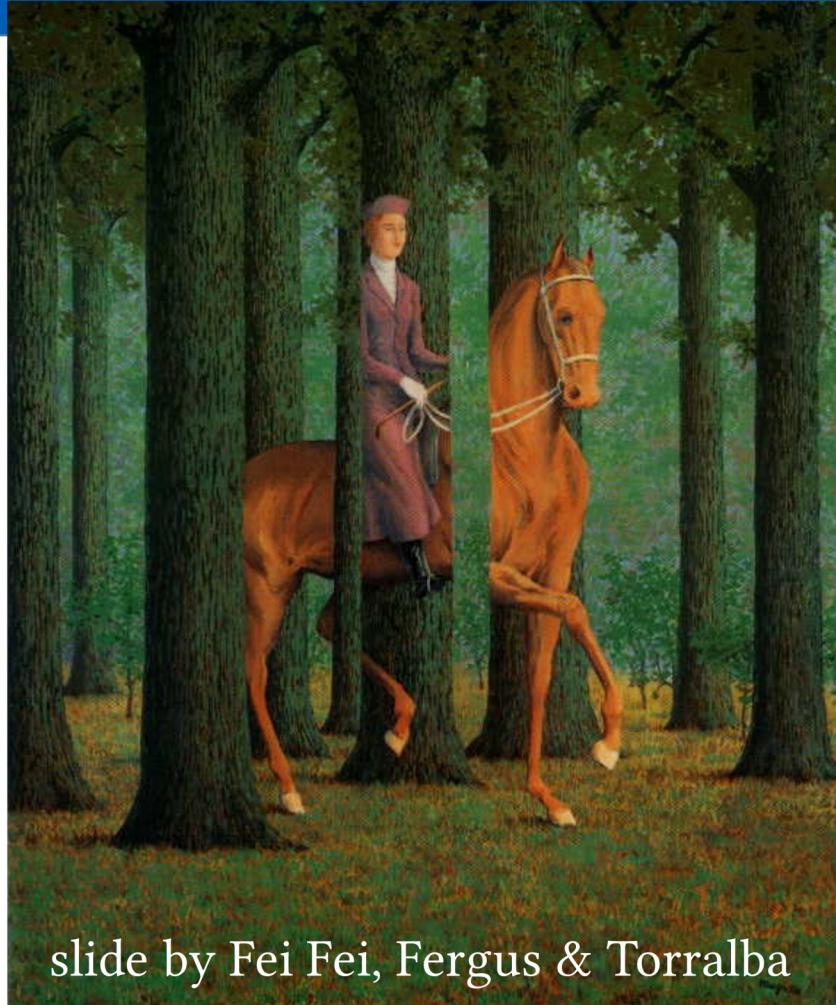
Michelangelo 1475-1564

slide by Fei Fei, Fergus & Torralba

# Challenges 2: illumination



# Challenges 3: occlusions



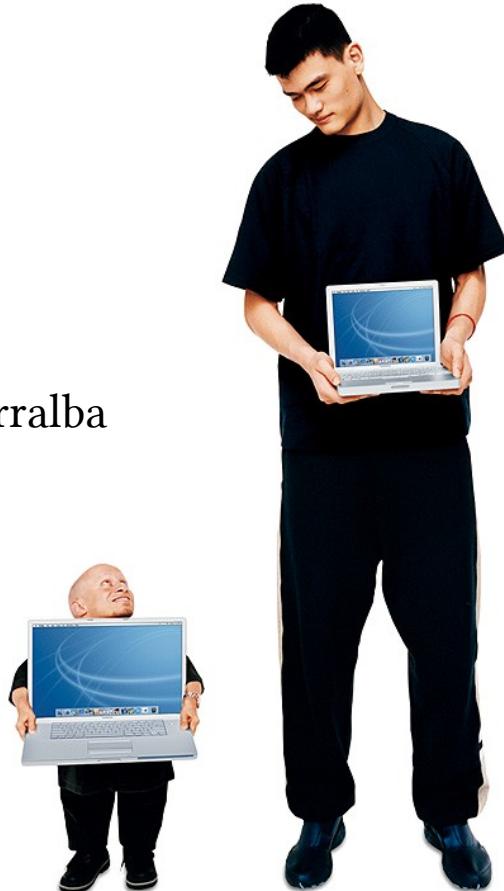
Magritte, 1957

slide by Fei Fei, Fergus & Torralba

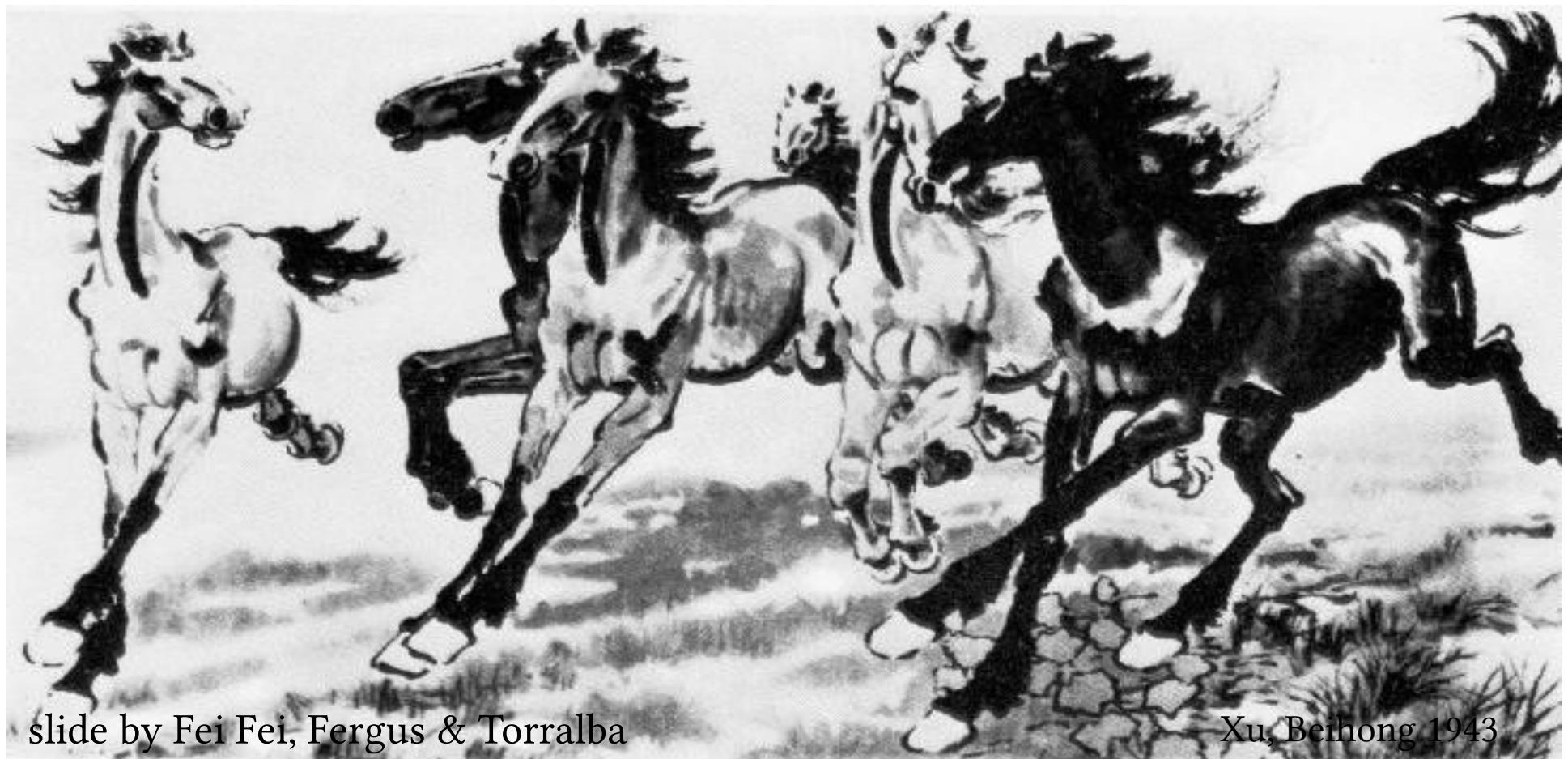
# Challenges 4: scale



slide by Fei Fei, Fergus & Torralba



# Challenges 5: deformation



slide by Fei Fei, Fergus & Torralba

Xu, Beihong 1943

# Challenges 6: background



# Challenges 7: intra class variation

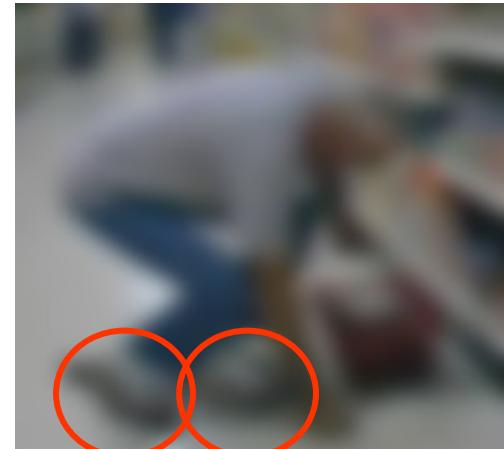
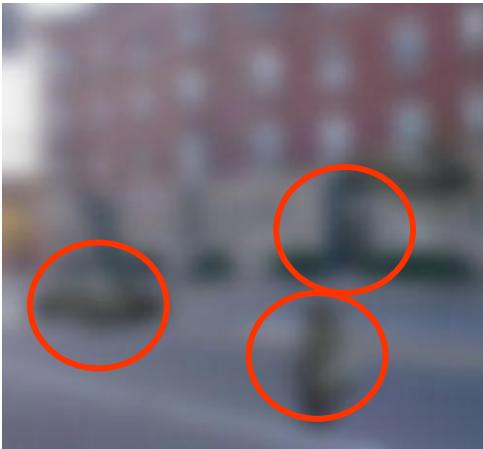
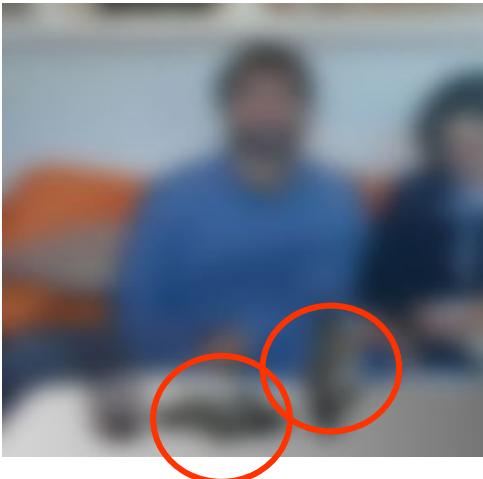


slide by Fei-Fei, Fergus & Torralba

# Challenges 8: local ambiguity



slide by Fei-Fei, Fergus & Torralba



# Challenges 9: the world behind the image



Slide Credit:  
Alyosha Efros



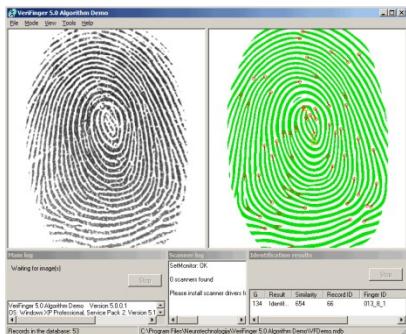
# What Works Today?

# What works today: OCR

Reading license plates, zip codes, checks

3 6 8 1 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 7 6 9 8 6 1

# What works today: biometrics



Fingerprint scanners on many new laptops, other devices

Source: S. Seitz



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

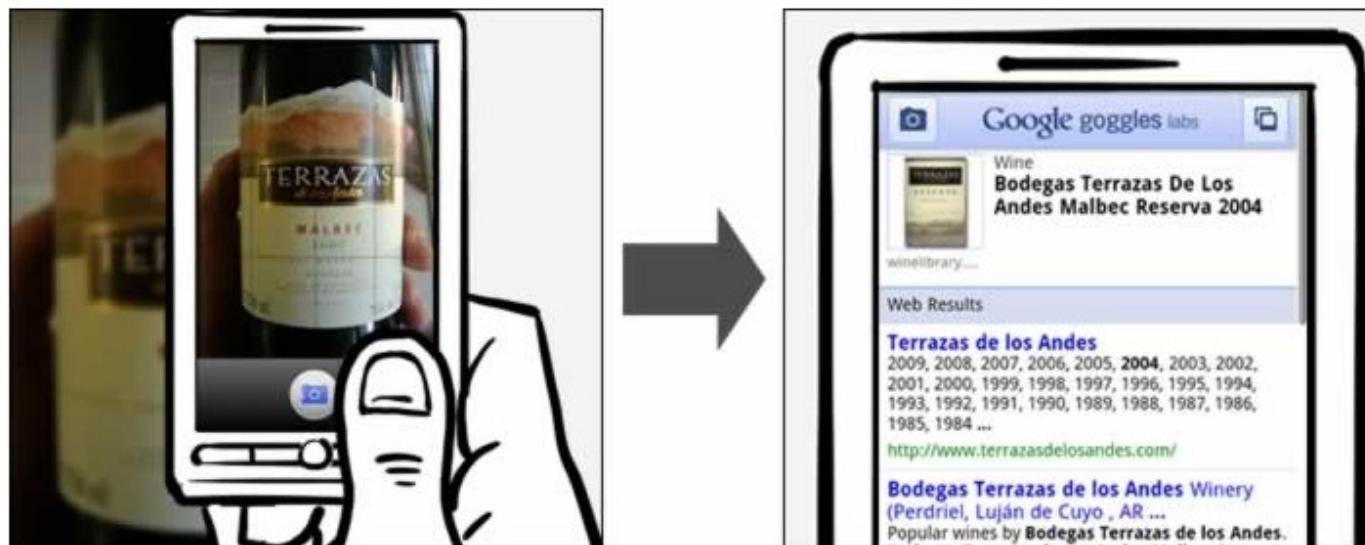
# What works today: mobile visual search

## Google Goggles in Action

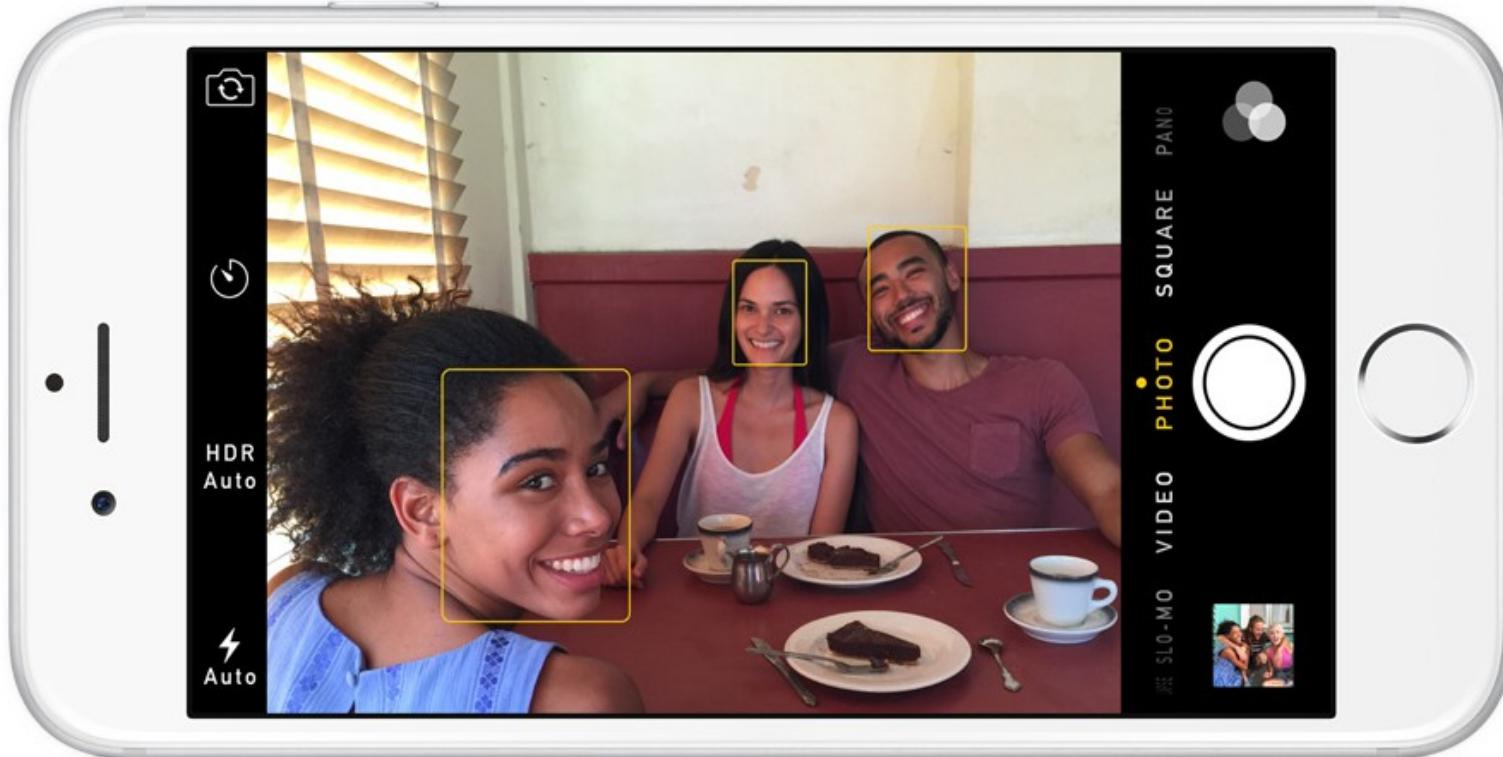
Click the icons below to see the different ways Google Goggles can be used.



Google goggles  
labs



# What works today: face detection



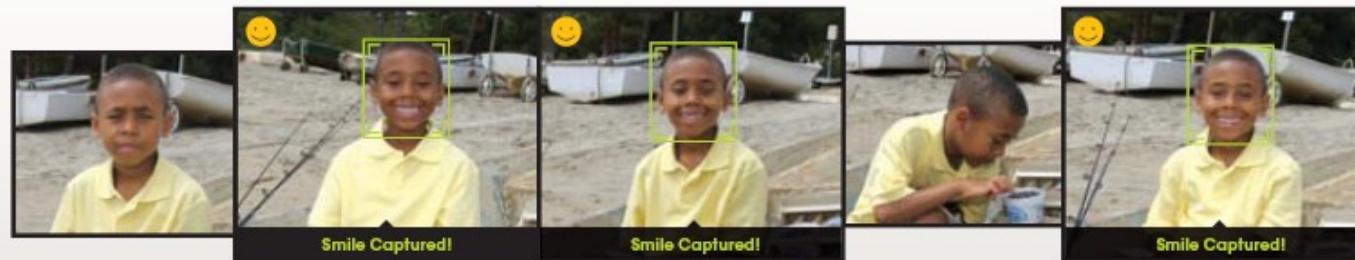
Many digital cameras now detect faces

# What works today: smile detection

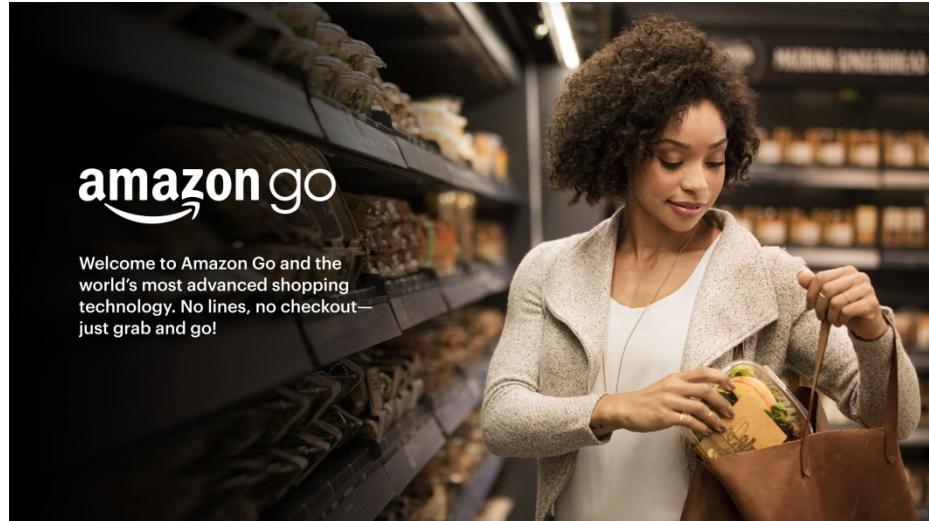


## The Smile Shutter flow

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.

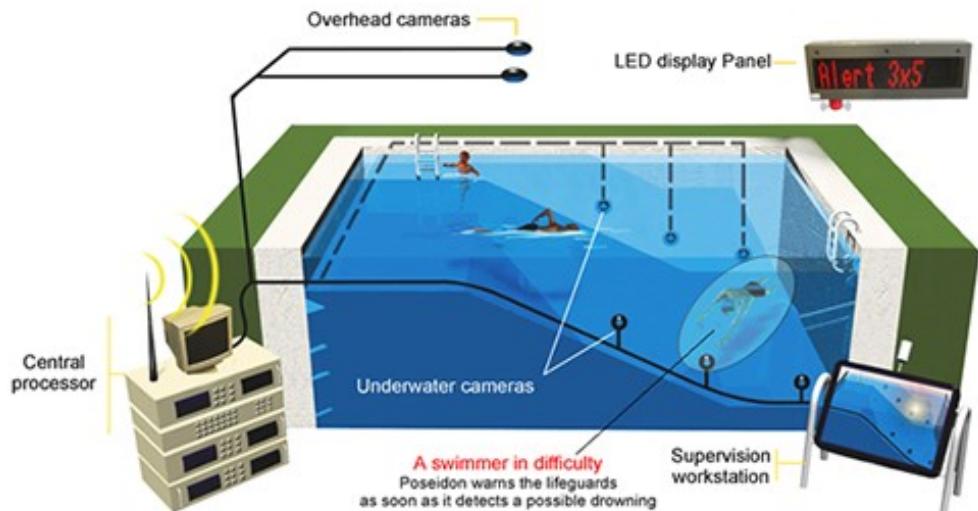


# What works today: supermarket object recognition



<https://youtu.be/NrmMk1Myrxc>

# What works today: safety



**BBC NEWS** [Watch One-Minute World News](#) 

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.



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

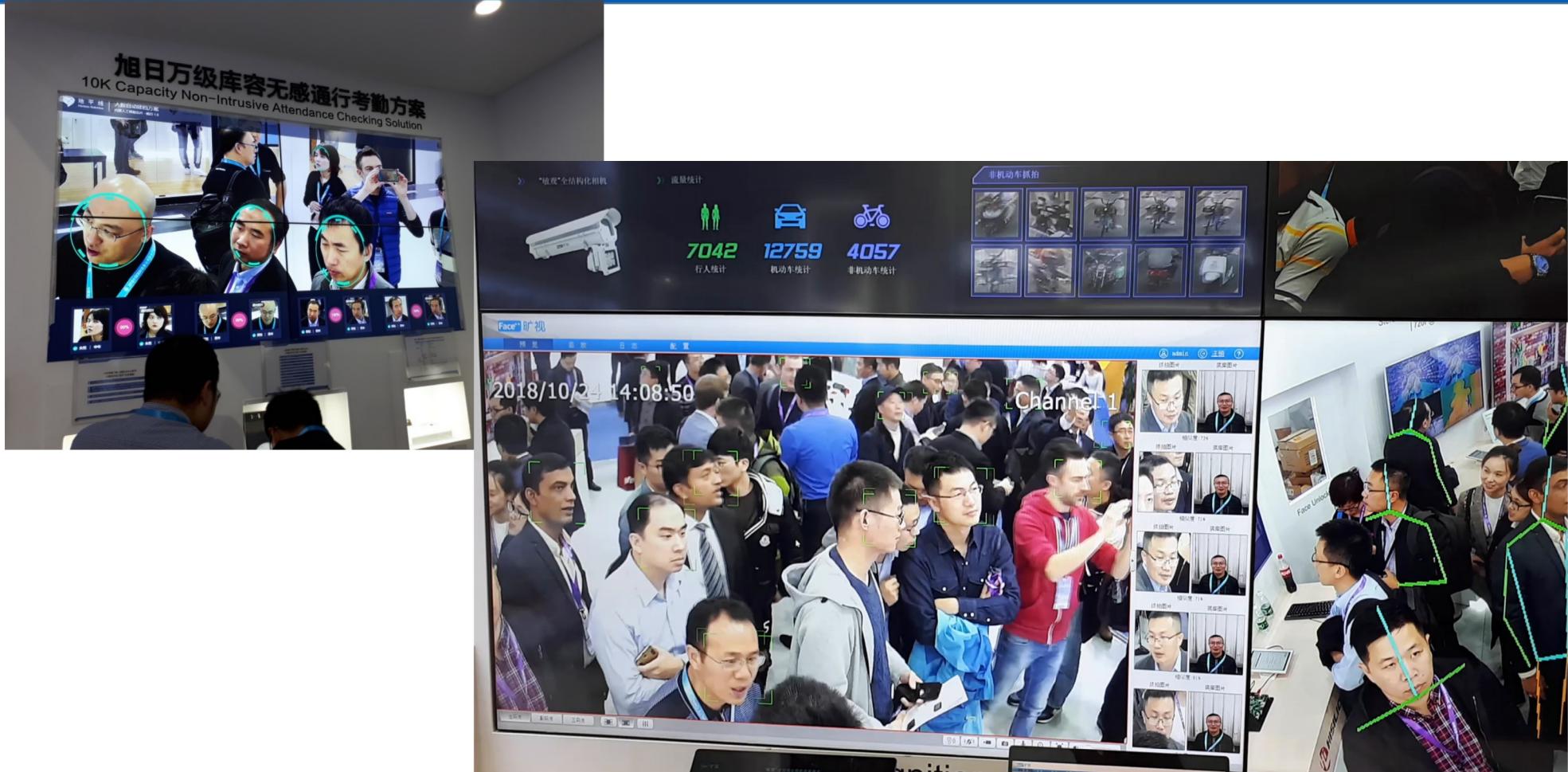
The £65,000 system, called Poseidon, detected her on the pool floor and sounded the alarm. A lifeguard pulled her out and she recovered in hospital.

[VIDEO Watch the rescue](#)

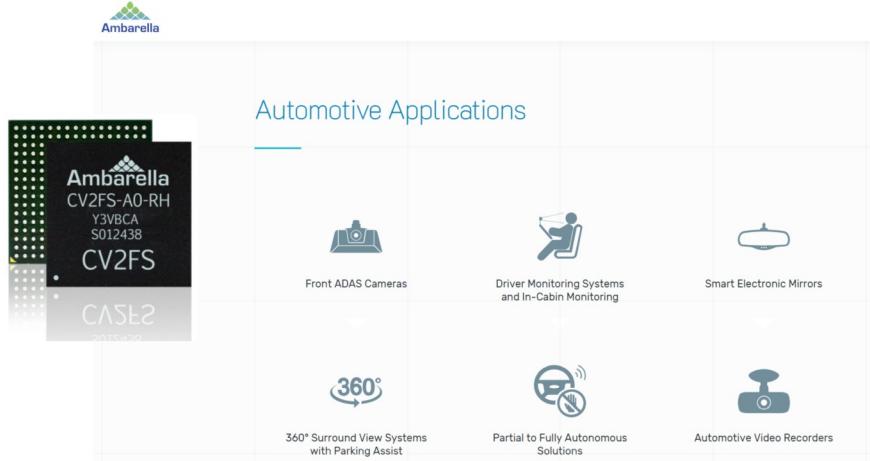
**News Front Page** 

[Africa](#) [Americas](#) [Asia-Pacific](#) [Europe](#) [Middle East](#) [South Asia](#) [UK](#) [England](#) [Northern Ireland](#) [Scotland](#) [Wales](#) [UK Politics](#) [Education](#) [Magazine](#) [Business](#) [Health](#) [Science &](#)

# What works today: security



# What works today: driving assistance/safety



Vision systems already in high-end car models

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



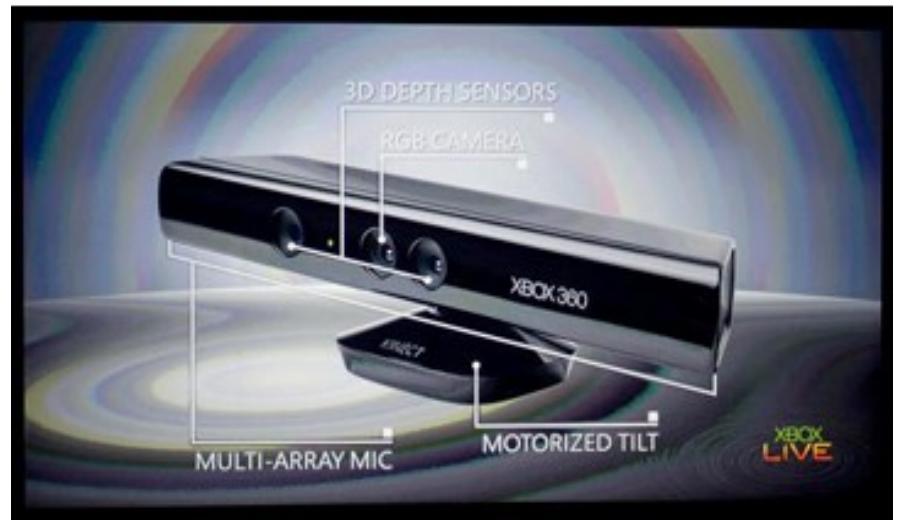
# What works today: GOOGLE, Tesla, UBER, Lyft, VisLab...



# What works today: vision interaction



Microsoft Kinect



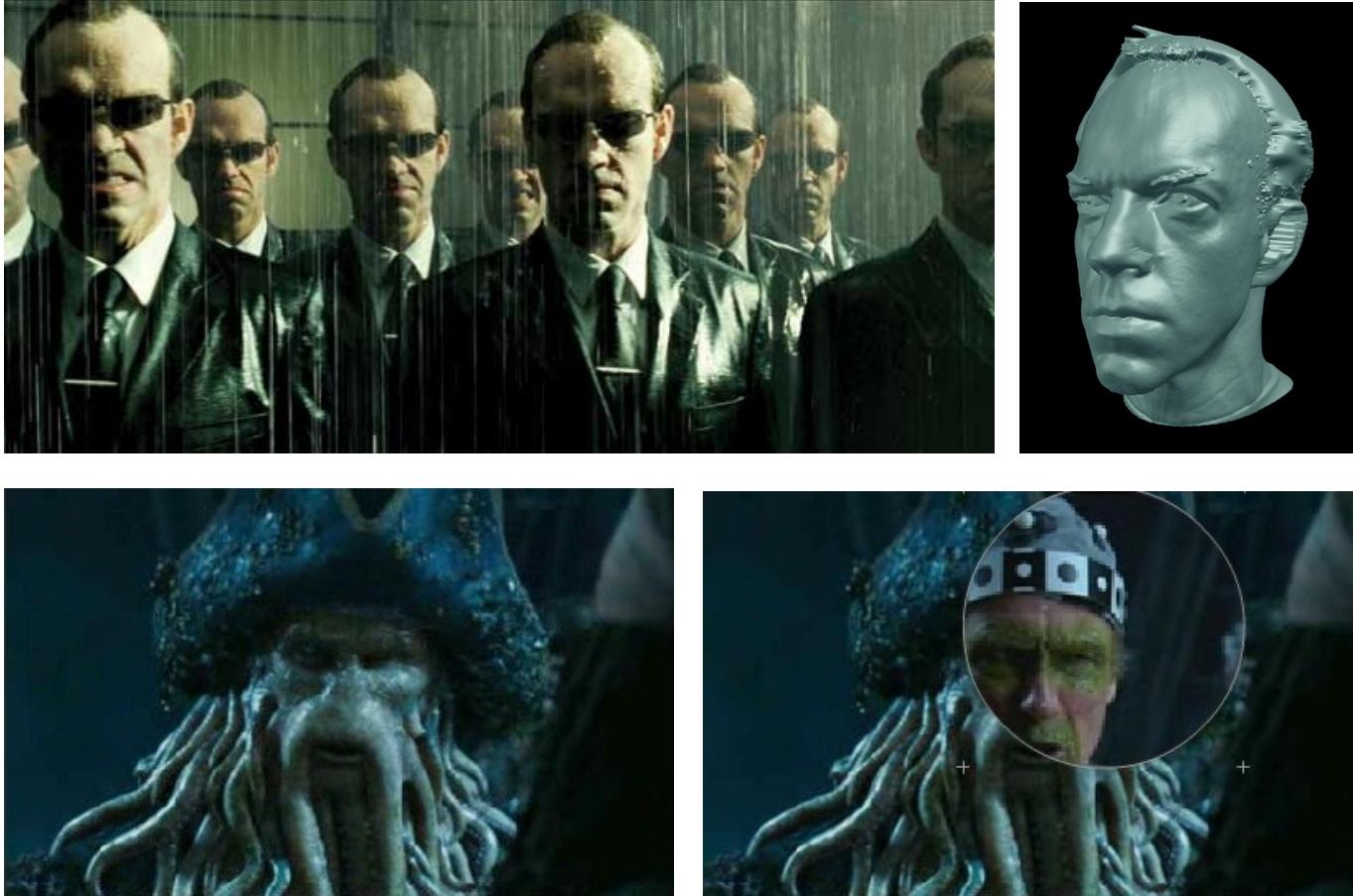
# What works today: augmented reality

UNIVERSITÀ  
DI PARMA



<http://nconnex.com/wp/>

# What works today: special effects (shape/motion capture)



Source: S. Seitz

# What works today: robotics

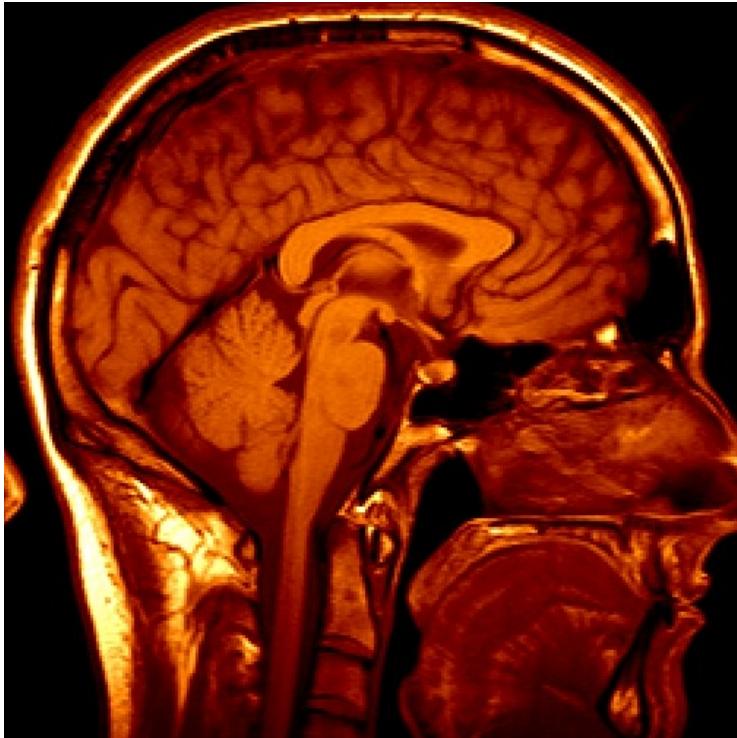


NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

## 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.

# What works today: medical imaging

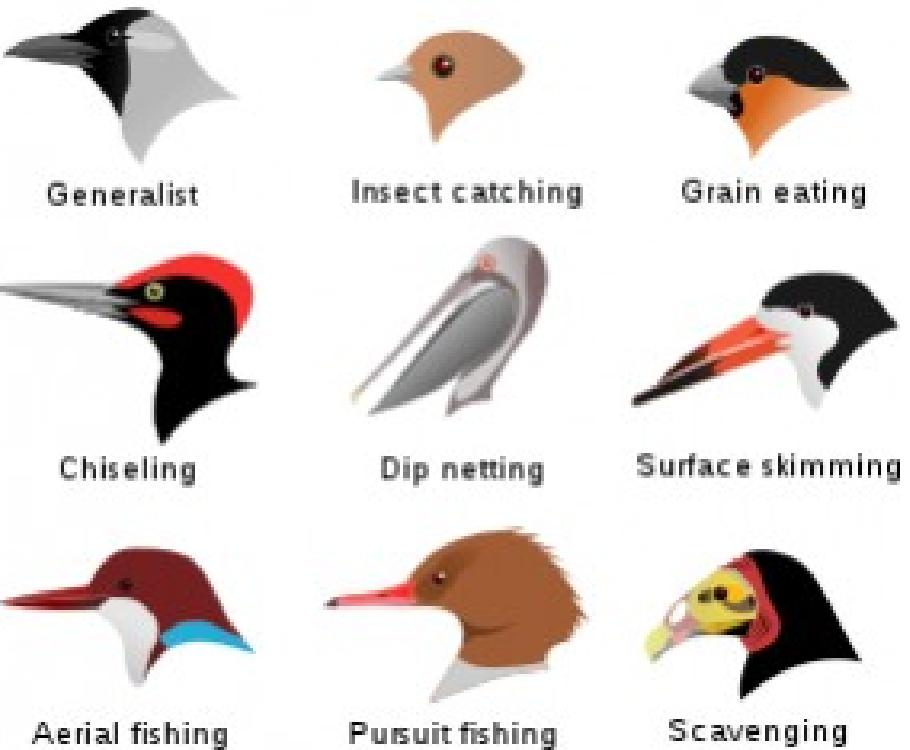


3D imaging  
MRI, CT



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

# What works today: biology

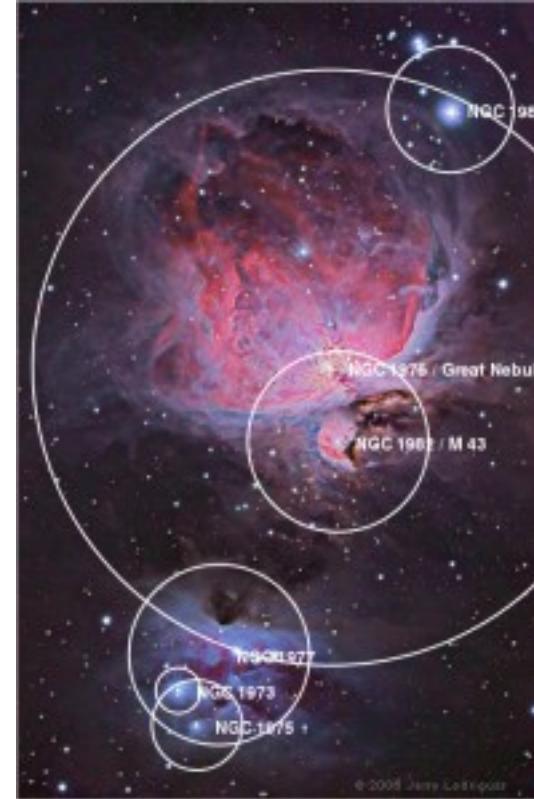
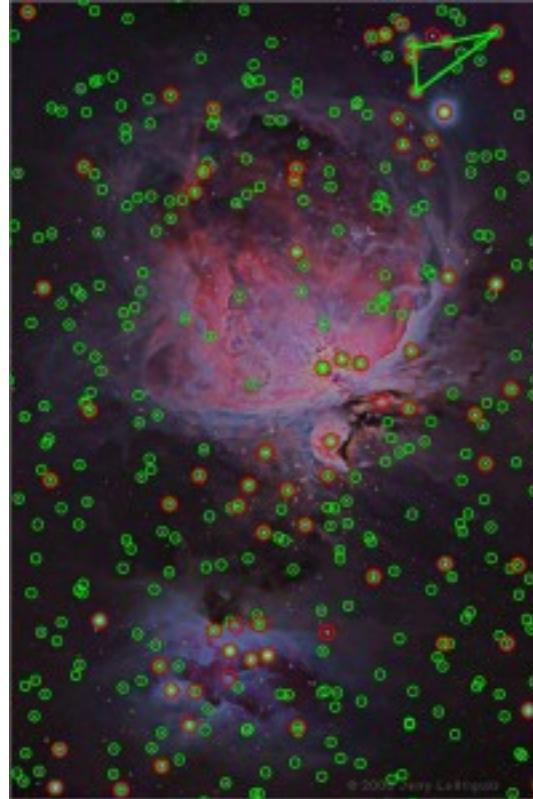


[Pl@ntNet](#)

[www.visipedia.org](http://www.visipedia.org)

# What works today: cosmology

UNIVERSITÀ  
DI PARMA



<http://astrometry.net/>

# What works today: healthcare



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



autism screening

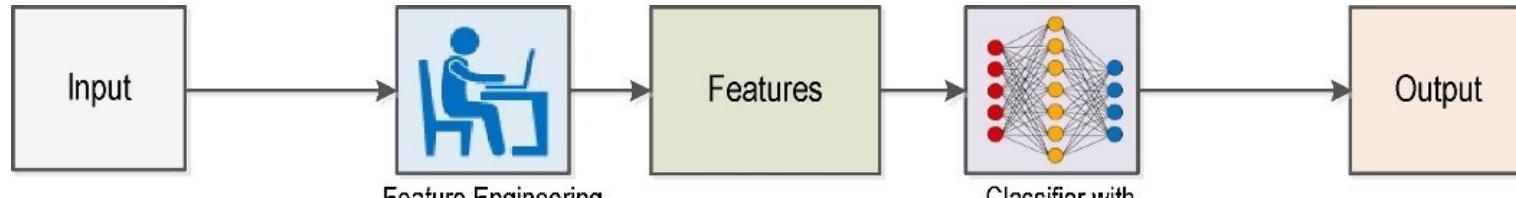
# Deep Learning vs *traditional* machine vision

- DL still not superseded traditional MV
- DL is not going to better solve all CV problems
- Sometime DL is a overkill
  - SIFT, coulor thersholding, not class specifric operations
  - Image stitching
  - 3D mesh reconstruction
  - Always when the probledm can be easily formalized
    - (!)

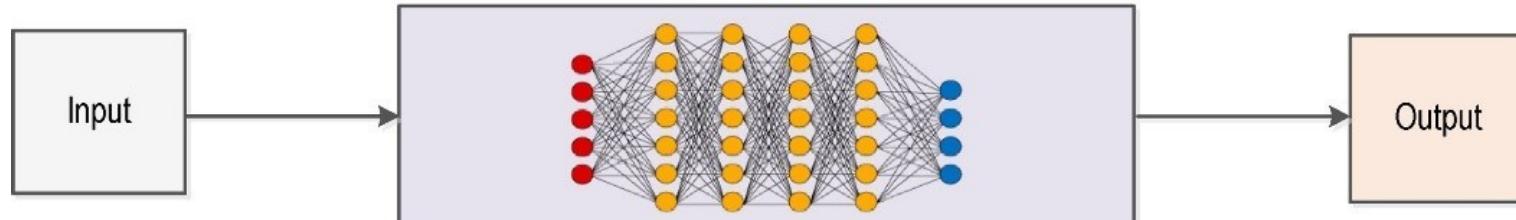
# Deep Learning vs *traditional* machine vision



- Often MV + Machine Learning classification



(a)



Feature Learning + Classifier  
(End-to-End Learning)  
(b)



UNIVERSITÀ DI PARMA

# Practical Information

# Requirements

- Expertise on what you should already study in:
  - Calcolatori Elettronici
  - Fondamenti di Informatica
  - Sistemi Operativi
- For projects/code we will use C++
  - OpenCV, Eigen
  - Ubuntu/Linux knowledge

- Images & their formats
- Image processing
  - Linear filtering
  - Non linear filtering
  - Binary images & Mathematical Morphology
  - Edges/Lines detection
- Machine vision
  - Pin hole model
  - Single view geometry
  - Stereo vision
  - Features extraction & matching
  - (Model fitting)
  - Segmentation

# Exams



- Theory
  - Written
  - 2 hours
  - Max 20 points, min 8 points
- Practical
  - Computer (YOUR laptop/notebook)
  - 2 hours
  - Max 10 points, min 8 points
- Assignements
  - Not mandatory
  - During laboratory
  - Max 2 points

- Slides, material & news: always look in elly
  - <https://elly2023.dia.unipr.it> → git
  - Please note that slides are for personal use only!
    - Many images/formulas/... are from copyrighted material...

- Machine vision and image processing books:
  - D.A. Forsyth and J. Ponce. Computer Vision: A Modern Approach (2° edition). Prentice Hall, 2011
  - R. Hartley and A. Zisserman. Multi View Geometry in Computer Vision. Cambridge University Press, 2003
  - Richard Szeliski. Computer Vision: Algorithms and Applications. [Link](#)
  - Shapiro and Stockman, Computer Vision, Prentice-Hall, 2001. Original chapters available at this URL [link](#)
- C++/OpenCV guides:
  - C++ Complete Reference
  - C++ STD Quick Reference
  - OpenCV Reference Manual
  - OpenCV Cheat Sheet [link](#)



UNIVERSITÀ DI PARMA

# Computer Vision