

# CS795/895 Computer Vision

- Class time: Tues & Thurs 11:00am-12:15pm
- Location: ECSB 2120
- Instructor:
  - Vikas Ashok
  - Office hours: Tues & Thurs 4.30-5.30pm, ECSB 3107
  - Email: [vganjigu@odu.edu](mailto:vganjigu@odu.edu)
- TA: To be announced
- Slide credit: the content of most slides are by Prof. Dimitris Samaras and Prof. Minh Haoi (Stony Brook Univeristy)

# What is Computer Vision?

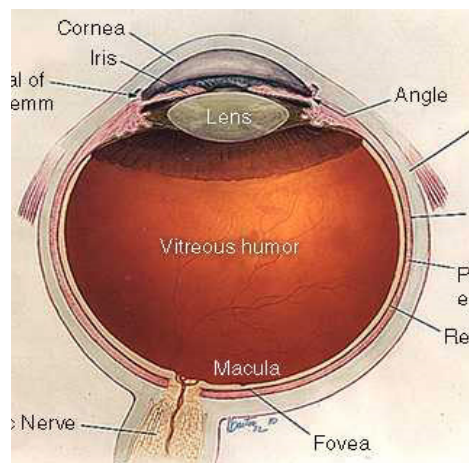
- Inverse Problem of Image Formation
- Compute properties of a world (either 2D or 3D from one or more digital images)
- Geometry
- Motion
- Recognition



# Why Vision? Light!



It is how we see other people, navigate our environment, communicate ideas, entertain, and **measure** the world around us.



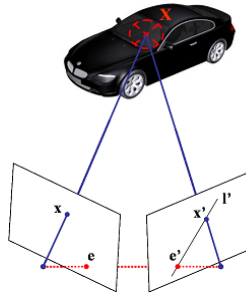
# Why is light good for measurement?



Microscopy



Surveillance



3D Analysis / Navigation



Remote Sensing

- Plentiful, sometimes free
- Interacts with many things, but not too many
- Goes generally straight over distance
- Very small  $\rightarrow$  high spatial resolution
- Fast, but not too fast  $\rightarrow$  time of flight sensors
- Easy to detect  $\rightarrow$  cameras work, are cheap
- Comes in flavors ( wavelengths )





# Why study Computer Vision?

- Images and videos are everywhere
- Fast-growing collection of useful applications
  - representations of the 3D world from pictures
  - automated surveillance (who's doing what)
  - movie post-processing
  - face finding
- Various deep and attractive scientific mysteries
  - how does object recognition work?
- Greater understanding of human vision

# Why study Computer Vision?

## Lots of Jobs!!!

facebook Adobe Microsoft amazon intel Google Athena eyes 智慧眼 MITSUBISHI ELECTRIC RESEARCH LABORATORIES, INC. NVIDIA A9 图森 Simple Prism

dji 大疆创新 TOYOTA RESEARCH INSTITUTE megvii Face++ BOSCH Invented for life DiDi

### Gold Sponsors

iRobot SKYDIO UBER ADVANCED TECHNOLOGIES CENTER Baidu 百度 DAQRI SENSETIME 商汤科技 Sighthound Linkface COGNEX IBM Research  
ebay here spare 5 360

### Silver Sponsors

HISCENE 英风台信息科技 MathWorks Disney Research SIEMENS SRI International DEEPLINT 格灵深瞳 Snapchat UISEE 驭势 QUALCOMM itseez  
Honeywell Tencent 腾讯 Panasonic STATS meitu M-LAB

### Bronze Sponsors

WOLFRAM COMPUTATION MEETS KNOWLEDGE Second Spectrum DATATANG POINT GREY MITEC OMRON rethink robotics Yandex Pinterest uSens Inc AiCure AVIGILON  
qb 清华大学 爱立信 Springer ELSEVIER DATA 61 magic leap Kitware IV IntelliVision OV ObjectVideo blippar CEVA Movidius

### Start-up Level Sponsors

nervana FIVEFOCAL DEXTRO MARKABLE SHOP ANY PHOTO umoo cv spotscale BUILDINGS DECODED SPORTLOGIQ wrnch  
MORPHEUS Neon BODYxLABS

# Computer Vision as a sensor

- Information about distant objects
- Passive Sensor
- High bandwidth
  - 1 picture = ? words
- Corresponds to the most complex human sensory function
  - Eat it? Run from it? Mate with it? +more...
- Computer Vision is not Animate Vision
  - Can be inspired though

# Properties of Vision

- One can “see the future”
  - Cricketers avoid being hit in the head
    - There’s a reflex - when the right eye sees something going left, and the left eye sees something going right, move your head fast.
  - Gannets pull their wings back at the last moment
    - Gannets are diving birds; they must steer with their wings, but wings break unless pulled back at the moment of contact.
    - Area of target over rate of change of area gives time to contact.





# Properties of Vision

- 3D representations are easily constructed
  - There are many different cues
  - Useful
    - to humans (avoid bumping into things; planning a grasp; etc.)
    - in computer vision (build models for movies).
  - Cues include
    - multiple views (motion, stereopsis)
    - texture
    - shading

# Properties of Vision

- People draw distinctions between what is seen
  - “Object/scene recognition”
    - “is this a fish or a bicycle?”
    - “is this George Washington?”
    - “is this poisonous or not?”
    - “is this slippery or not?”
    - “will this support my weight?”
  - Great mystery
    - How to build programs that can draw useful distinctions based on image properties.

# Why Computational Visual Recognition?

# Why Computational Visual Recognition?





# Why Computational Visual Recognition?



seal

Search Images

# Why Computational Visual Recognition?



Google  
images

seal

Search Images

Billions of images, many with no text / meta data

# Why Computational Visual Recognition?



Billions of images, many with no text / meta data

Google  
images

seal

Search Images

Google

seal

Search

SafeSearch: [Moderate](#) ▼



**Seal**  
400 x 400 - 48k - jpg  
puckthmedia.wordpress.com  
[Find similar images](#)



**Seal's really**  
300 x 300 - 54k - jpg  
arjanwrites.com  
[Find similar images](#)



**Seal and Heidi: sex**  
325 x 325 - 32k - jpg  
lightscamerscaption...  
[Find similar images](#)



**of Seal on**  
388 x 357 - 11k - jpg  
east-coast-bias.com  
[Find similar images](#)



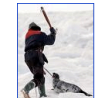
**harp-seal-baby**  
470 x 324 - 29k - jpg  
journeymen47.wordpress.com  
[Find similar images](#)



**the seal**  
400 x 391 - 24k - jpg  
bulletproof.wordpress.com  
[Find similar images](#)



**Graphic image of**  
500 x 400 - 38k - jpg  
giavasan.diludovico.it  
[Find similar images](#)



**The Canadian Seal**  
348 x 450 - 29k - jpg  
advocacy.britannica.com  
[Find similar images](#)



**kill baby seals**  
500 x 333 - 67k - jpg  
liberationbc.org  
[Find similar images](#)



**clubbing with Seal**  
324 x 400 - 29k - jpg  
justwilliam1969.  
[Find similar images](#)



**SEAL OF APPROVAL**  
300 x 400 - 51k - jpg  
people.com  
[Find similar images](#)



**Harp Seal**  
500 x 375 - 20k - jpg  
flanature.org  
[Find similar images](#)



**Weddell Seal**  
763 x 510 - 65k - jpg  
codantarctica.com  
[Find similar images](#)

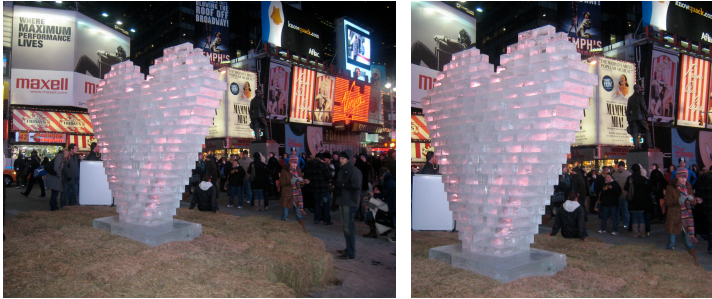


**Hawaiian Monk Seal**  
350 x 354 - 33k - jpg  
mobi.org  
[Find similar images](#)



**Antarctic seals and**  
436 x 316 - 43k - jpg  
gdargaud.net  
[Find similar images](#)

# Range of recognition tasks



- Duplicate detection
- Edge detection
- Same (rigid) object
- Face detection
- Face Identification
- General category recognition



# Range of recognition tasks



- Duplicate detection
- Edge detection
- Same (rigid) object
- Face detection
- Face Identification
- General category recognition

# Range of recognition tasks



- Duplicate detection
- Edge detection
- Same (rigid) object
- Face detection
- Face Identification
- General category recognition

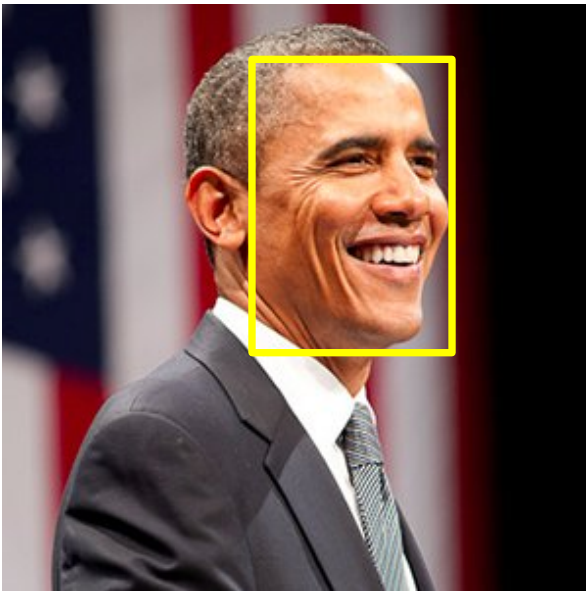
# Range of recognition tasks

- Duplicate detection
- Edge detection
- Same (rigid) object
- Face detection
- Face Identification
- General category recognition



# Range of recognition tasks

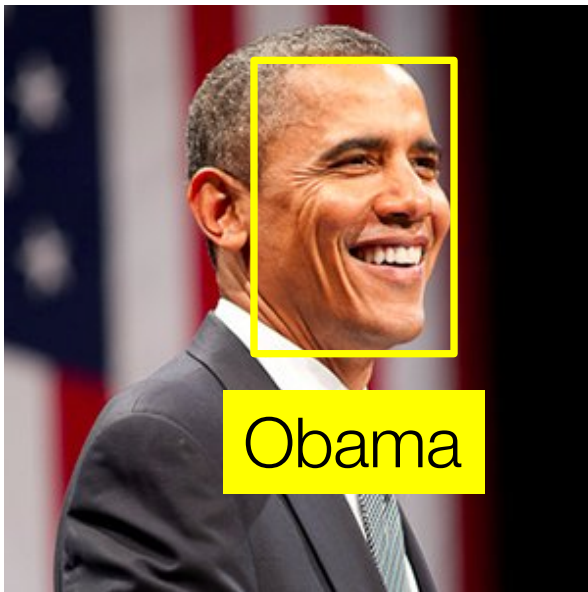
- Duplicate detection
- Edge detection
- Same (rigid) object
- Face detection
- Face Identification
- General category recognition





# Range of recognition tasks

- Duplicate detection
- Edge detection
- Same (rigid) object
- Face detection
- Face Identification
- General category recognition



# Range of recognition tasks

- Duplicate detection
- Edge detection
- Same (rigid) object
- Face detection
- Face Identification
- General category recognition

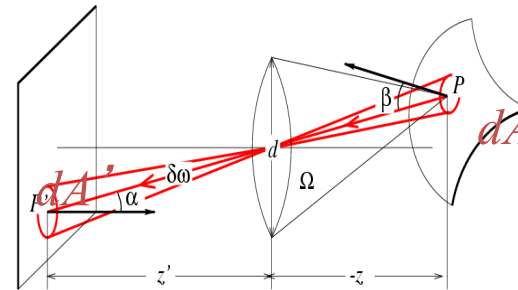


# Image Science

Image processing  
Image to Image



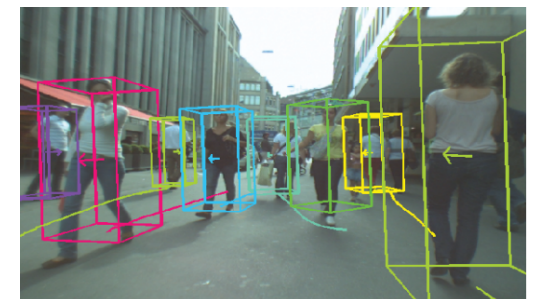
Imaging  
Physics to Image



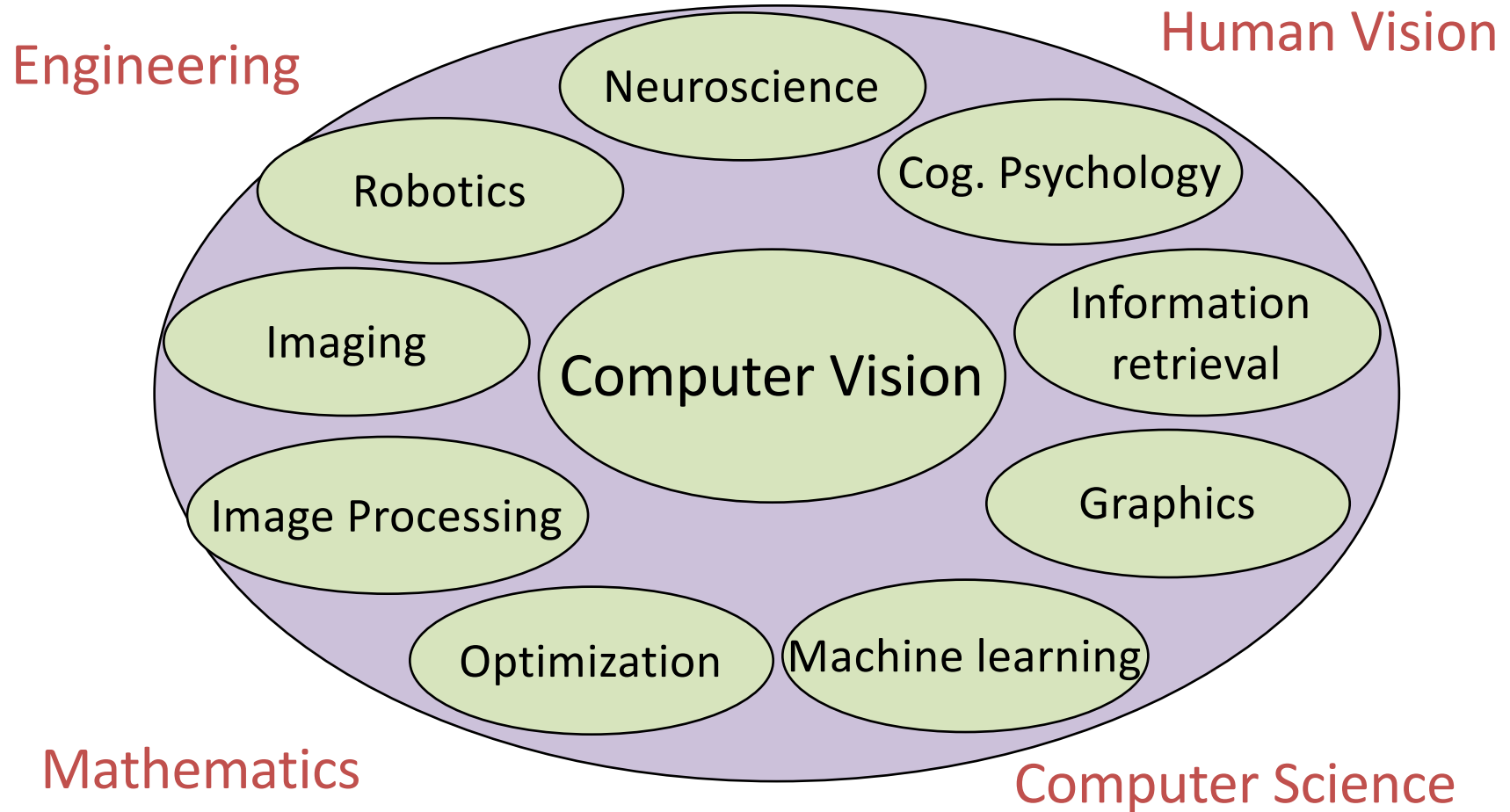
Graphics  
Symbols to Image



*Computer Vision*  
*Image to Symbols*



# Relationship with other fields



# Computer Vision vs. Computer Graphics

- Graphics
  - Produce “plausible” images
  - You choose the models, conditions, imaging parameters, etc.
- Computer Vision
  - Given real images with noise, sampling artifacts ...
  - Estimate physically quantities
  - Ill-posed – what is the minimum world knowledge we need?

# Computer Vision vs. Image Processing

- Image Processing
  - Mostly concerned with *image-to-image* transformations
  - Filtering
  - Enhancement
  - Compression
- Computer Vision
  - Concerned with how images *reflect the 3D world*
  - Filtering *for feature extraction*
  - Enhancement *for recognition/detection*
  - Compression *that preserves geometric information in images*

# Where is Computer Vision useful?

Medical Image Analysis

Robotics

Automobile industry

Optical Character Recognition

Visual aids for the blind

Industrial inspection

Surveillance

Military

Film industry

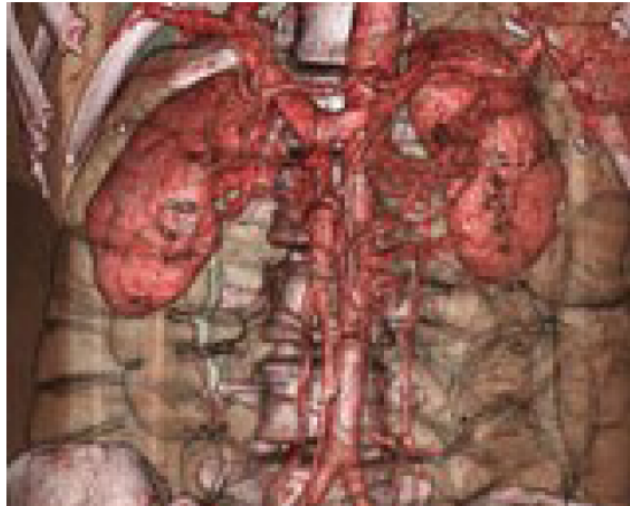
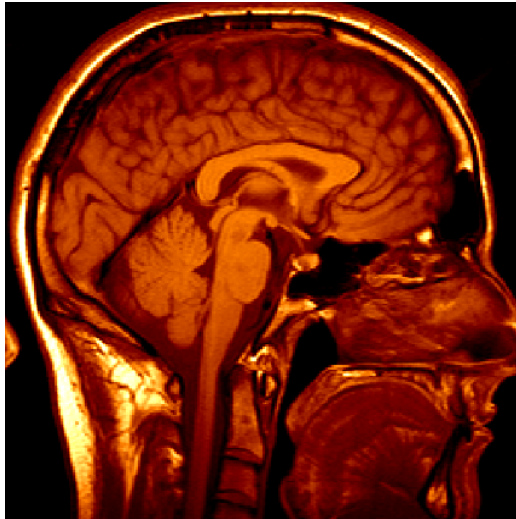
Entertainment

Human-Computer Interaction

Image Search

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

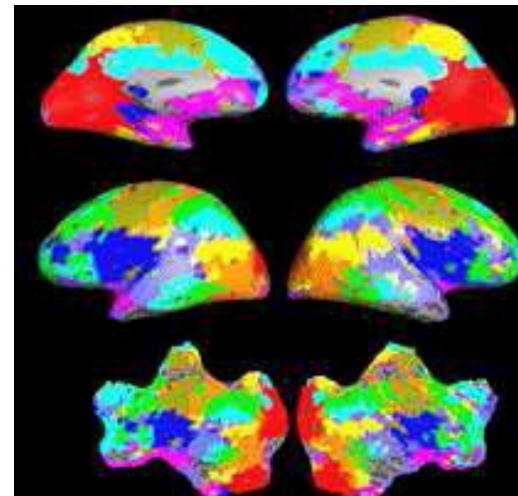
# Medical Image Analysis



Analysis of images acquired with Computerized Tomography, Magnetic Resonance Imaging, Ultrasound...



Image Guided Surgery



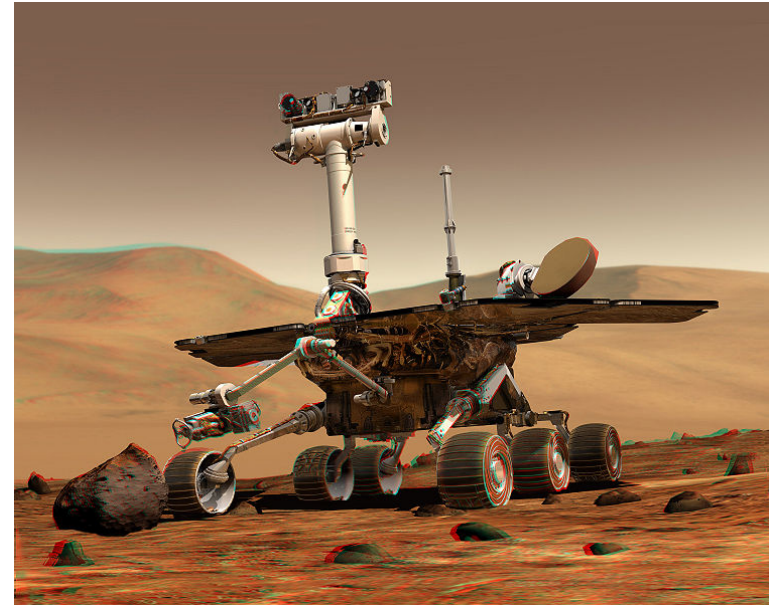
FMRI data analysis



# Robotics



DARPA Grand Challenge



NASA's Mars Rover



Toy industry

# Biometrics



Iris Recognition

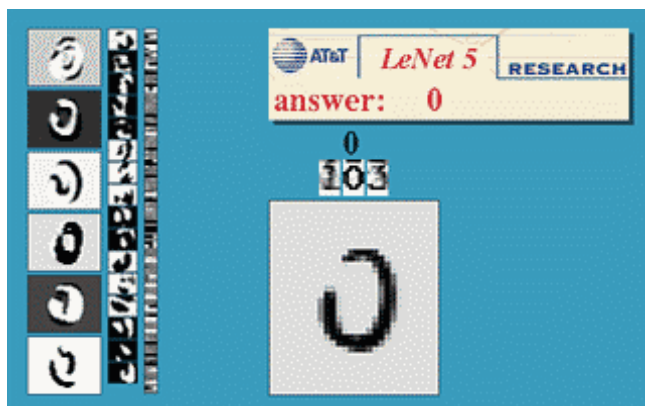


Fingerprint Recognition



Face Recognition

# Text Detection and Recognition



Optical Character Recognition (OCR)



Licence Plates



'Blindsight'



# Automobile Industry



Pedestrian and Car Detection



Monitoring Driver Alertness



Lane Detection

manufacturer products consumer products

## Our Vision. Your Safety.

rear looking camera forward looking camera side looking camera

► **EyeQ** Vision on a Chip

► **Vision Applications**  
Road, Vehicle, Pedestrian Protection and more

► **AWS** Advance Warning System

► **Mobileye at Equip Auto, Paris, France**

► **Mobileye at SEMA, Las Vegas, NV**

► 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

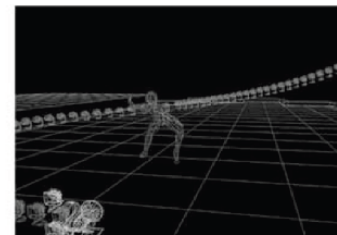
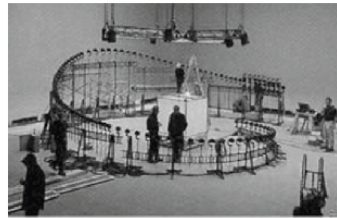
Events

► **Mobileye at Equip Auto, Paris, France**

► **Mobileye at SEMA, Las Vegas, NV**

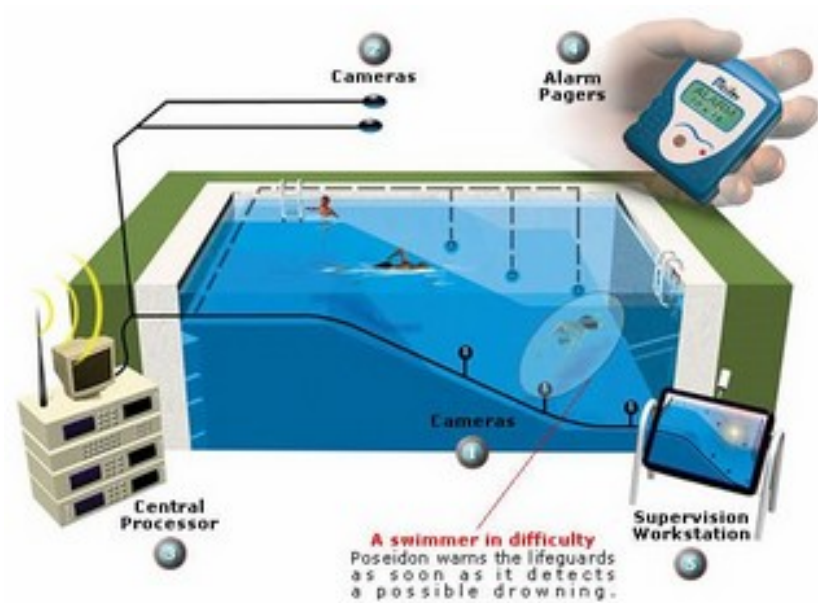
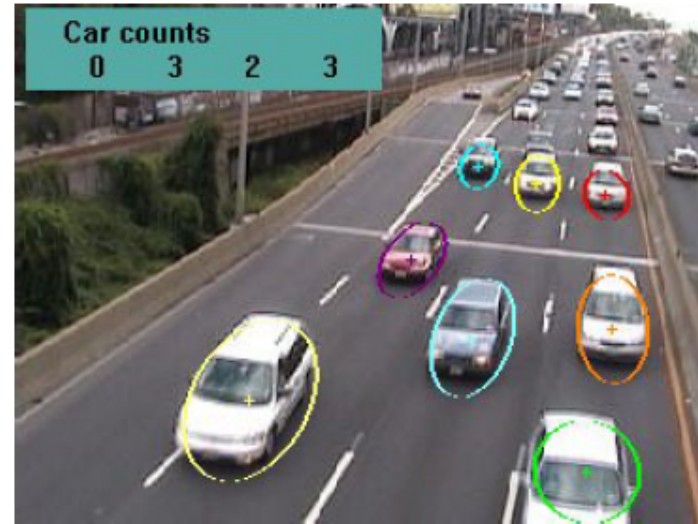
► read more

# Visual Effects





# Security & Surveillance



# Object recognition





# Added value to commercial products

- Digital Cameras



- Mobile phones



Point & find



kooaba

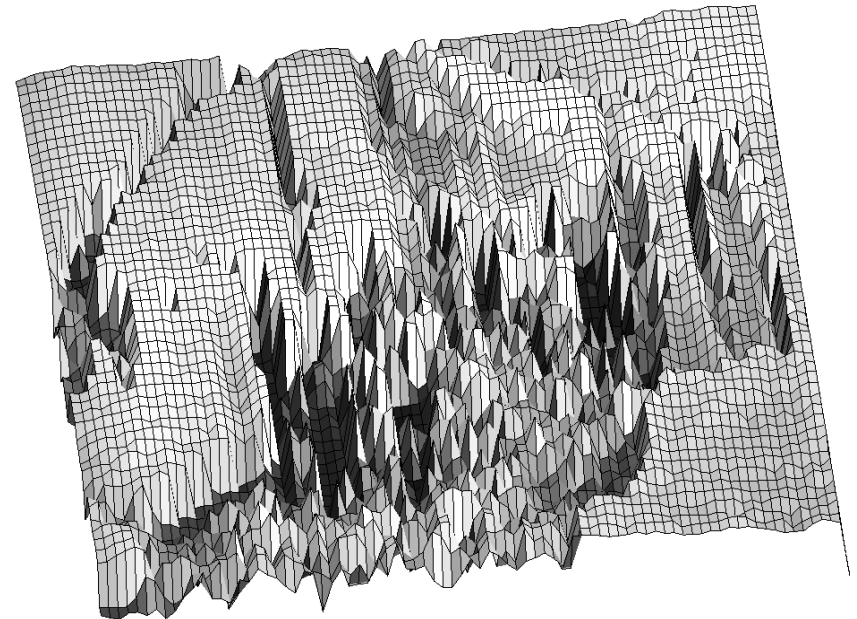
# Motion capture



Microsoft's Xbox Kinect

# Why can't computers see (yet)?

- Imagine describing 'red' or 'ugly' to a blind man
- Input to a computer: 2D/3D function



# Current state of the art

- You just saw examples of current systems.
  - Many of these are less than 5 years old
- This is a very active research area, and rapidly changing
  - Many new apps in the next 5 years
- To learn more about vision applications and companies
  - [David Lowe](http://www.cs.ubc.ca/spider/lowe/vision.html) maintains an excellent overview of vision companies
    - <http://www.cs.ubc.ca/spider/lowe/vision.html>



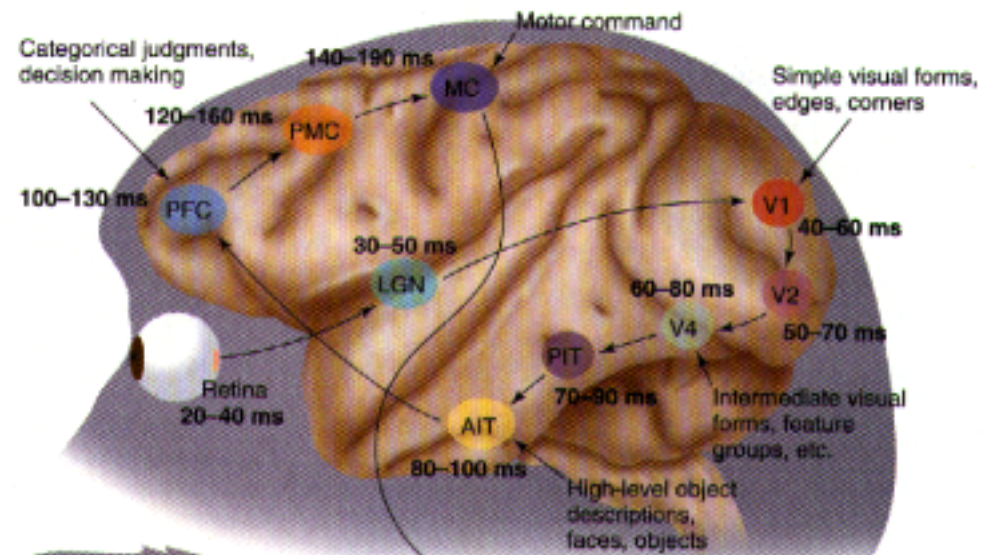
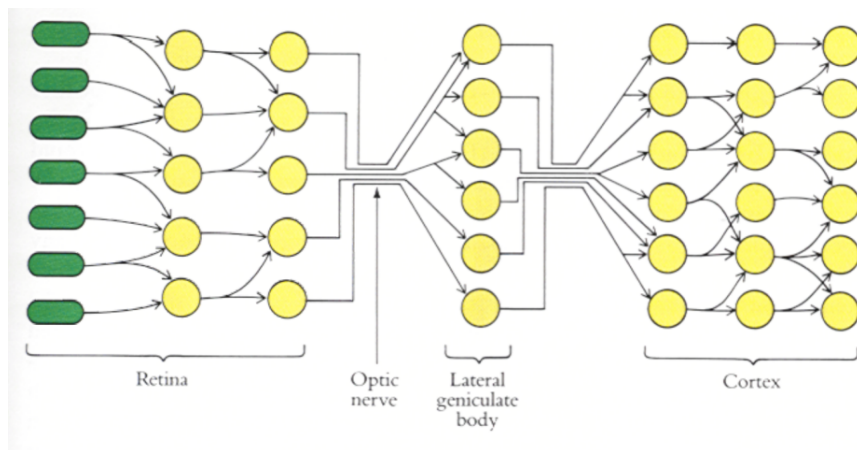
# How do we solve vision?

We perform the vision task with amazing speed and accuracy

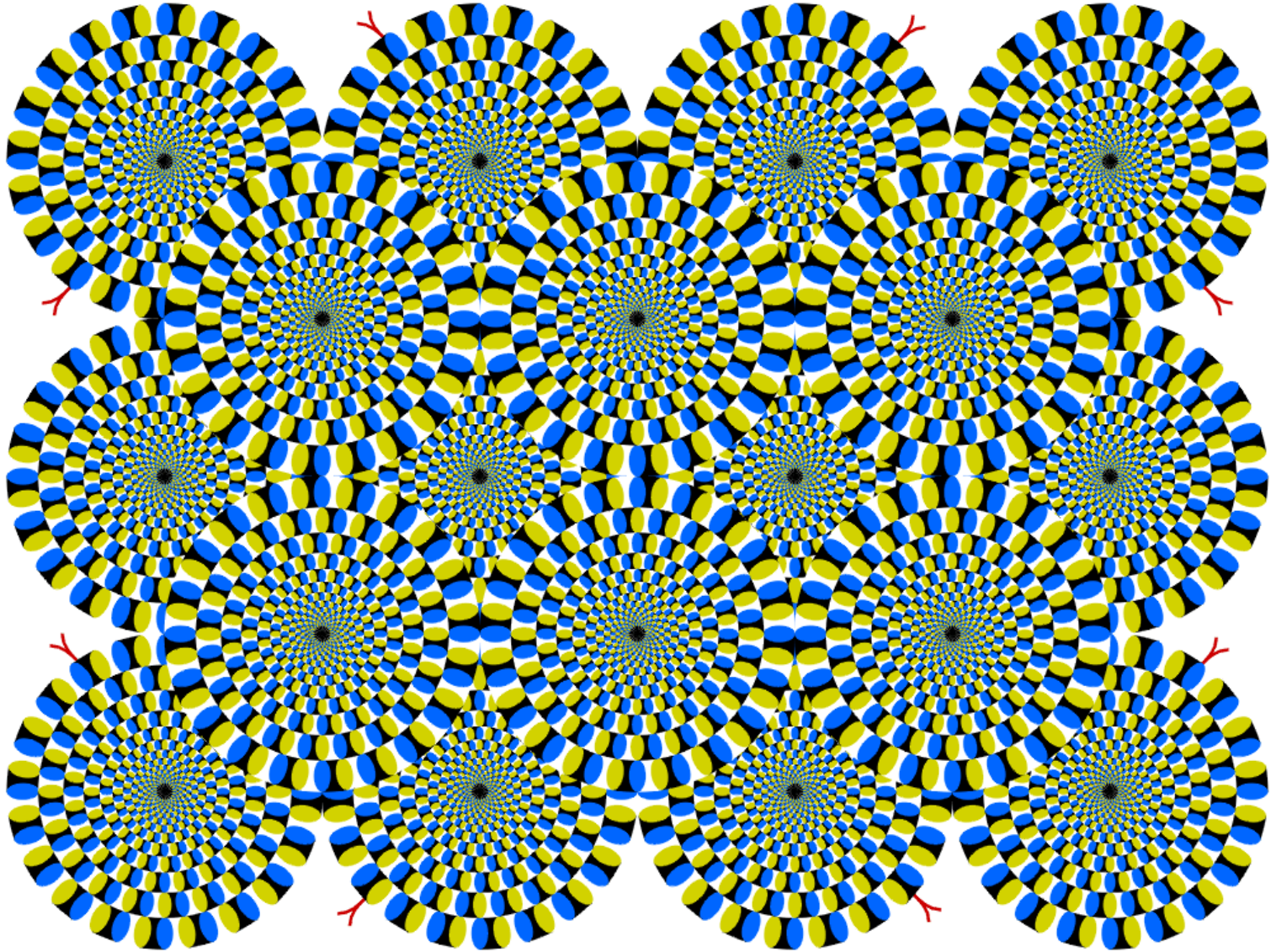
But not effortlessly:

Almost 50% of your brain is doing vision

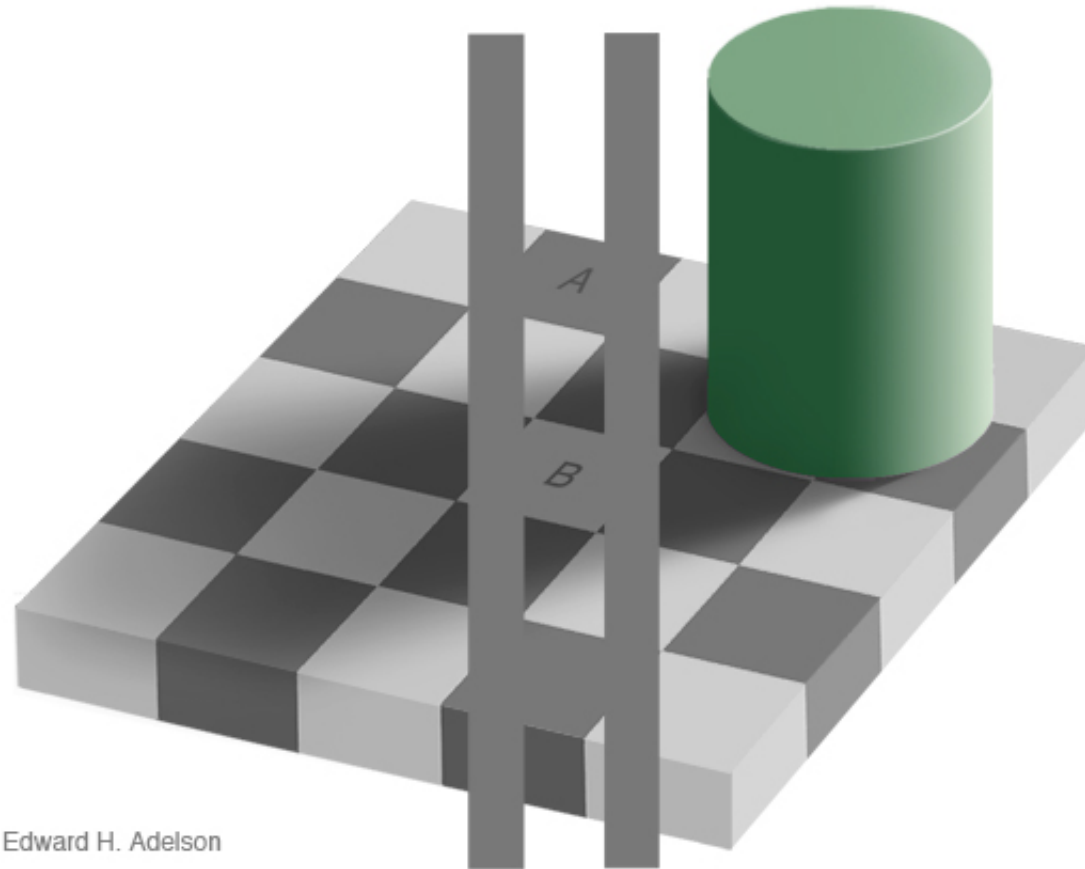
Substantially more than what is involved in doing math!



Is seeing trivial?



# Is seeing trivial?

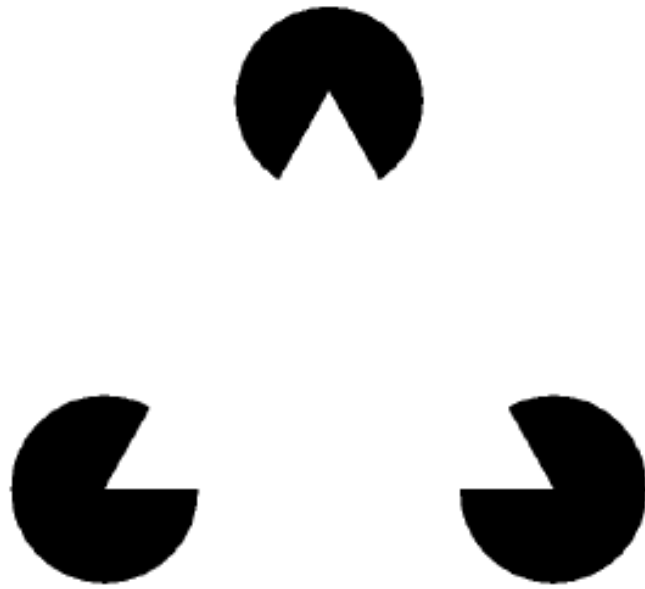


Edward H. Adelson

[http://web.mit.edu/persci/people/adelson/checkershadow\\_illusion.html](http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html)



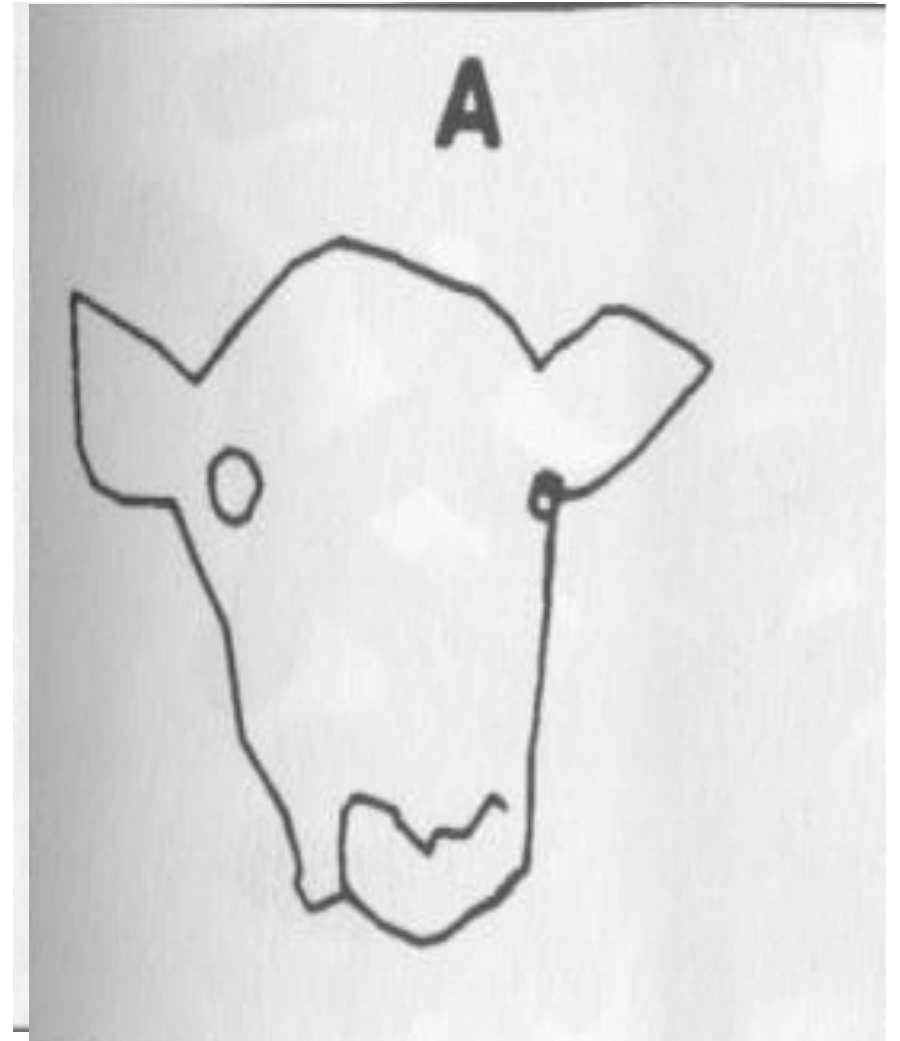
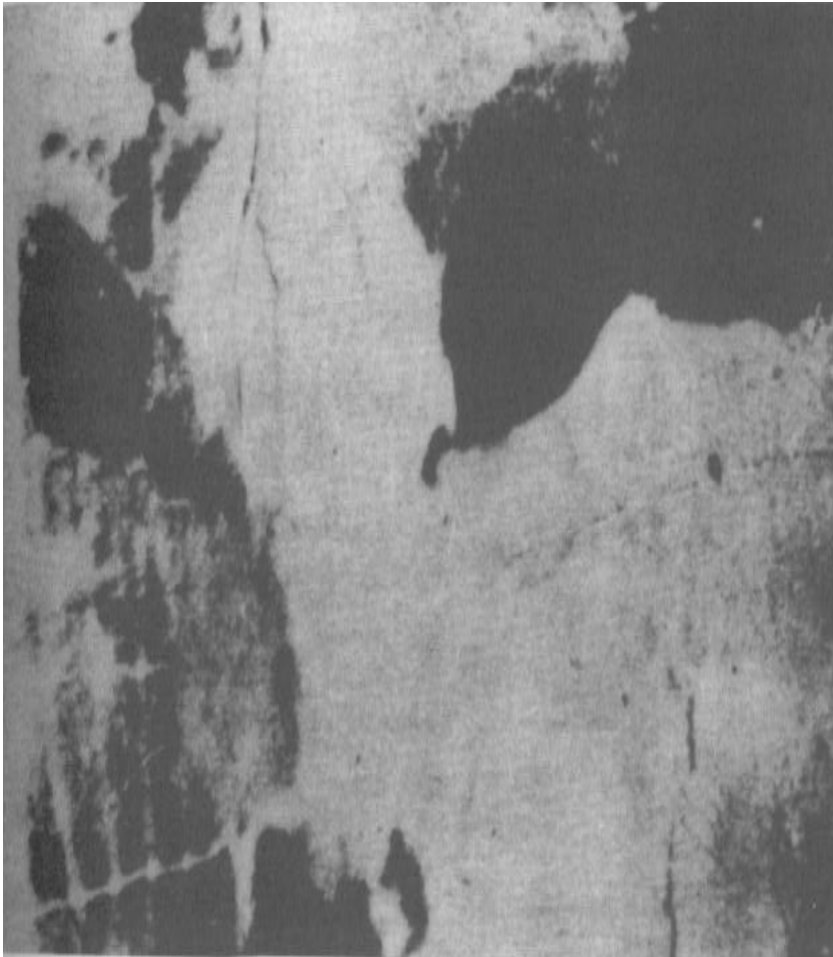
Is seeing trivial?



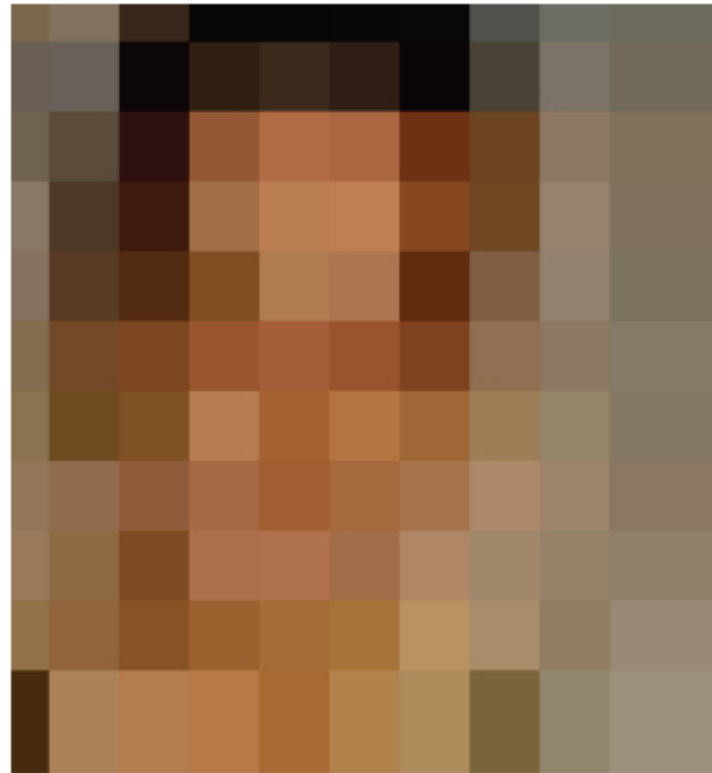
Is seeing trivial?



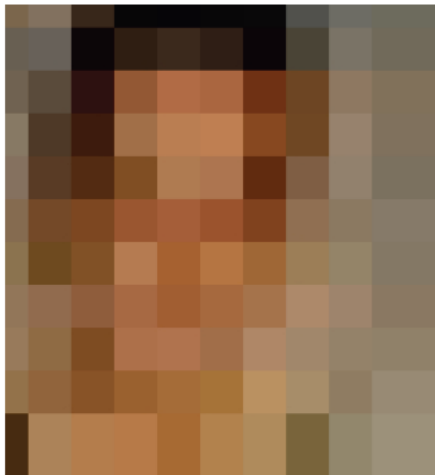
# Is seeing trivial?



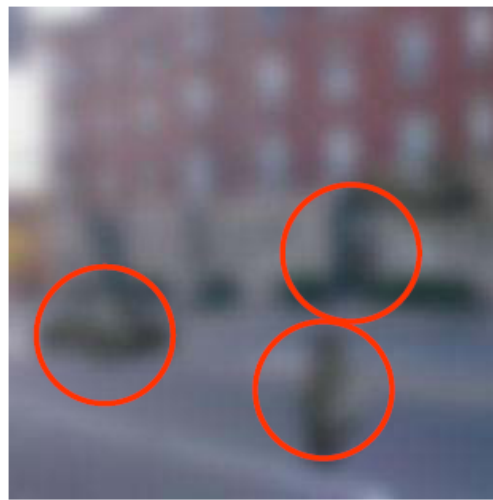
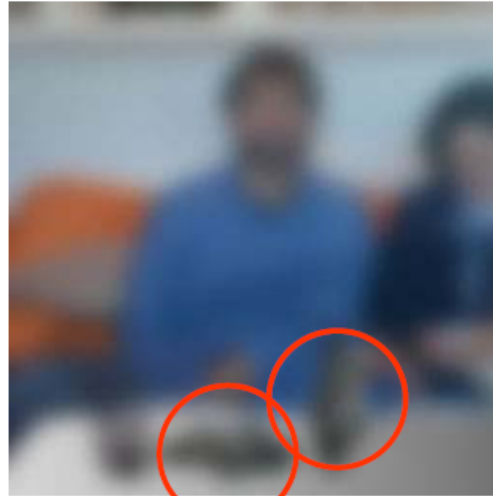
# Face or non-face?



# Face or non-face?



# Context

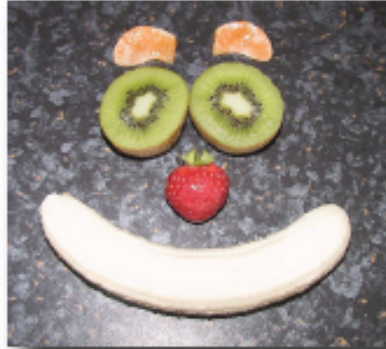


# Low-level & high-level vision: Chicken & egg





# Challenges: What is an object?



# Challenges: Many Nuisance Parameters



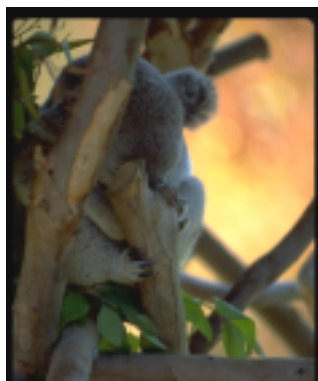
Illumination



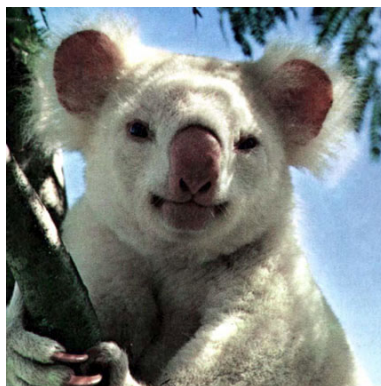
Object pose



Clutter



Occlusions



Intra-class  
appearance



Viewpoint

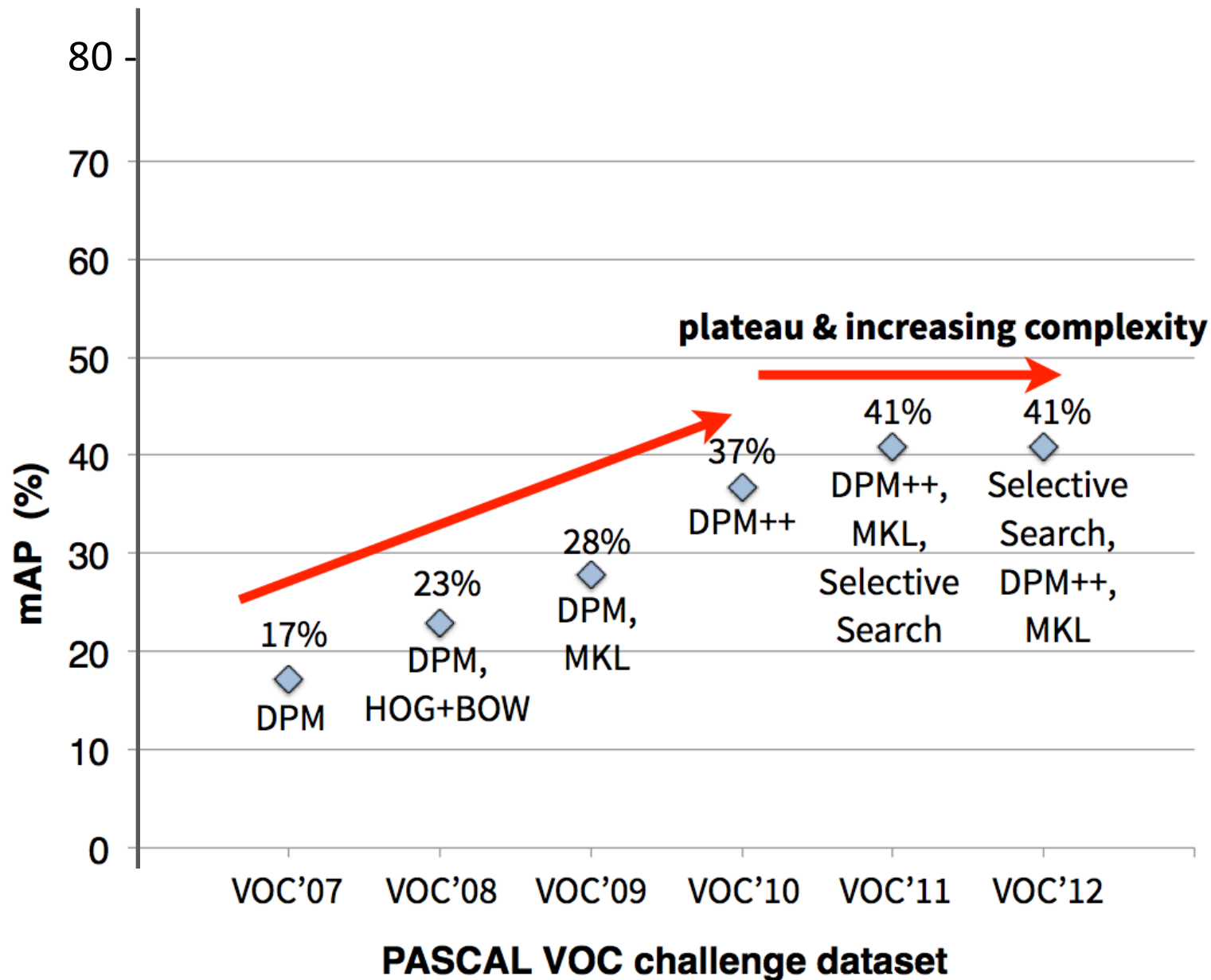


# Challenges: Intra-Category Variation

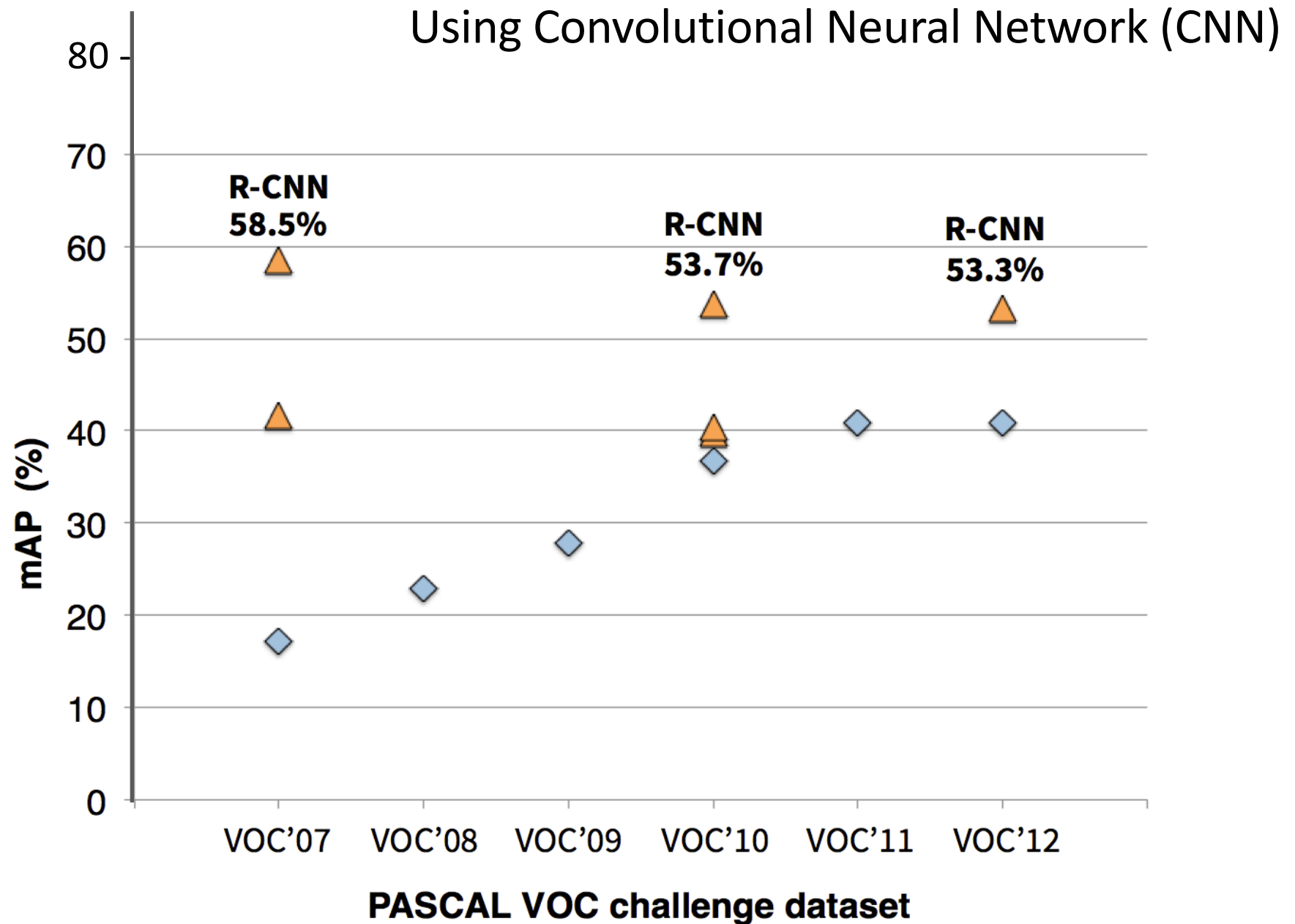




# Object detection 2007-2012

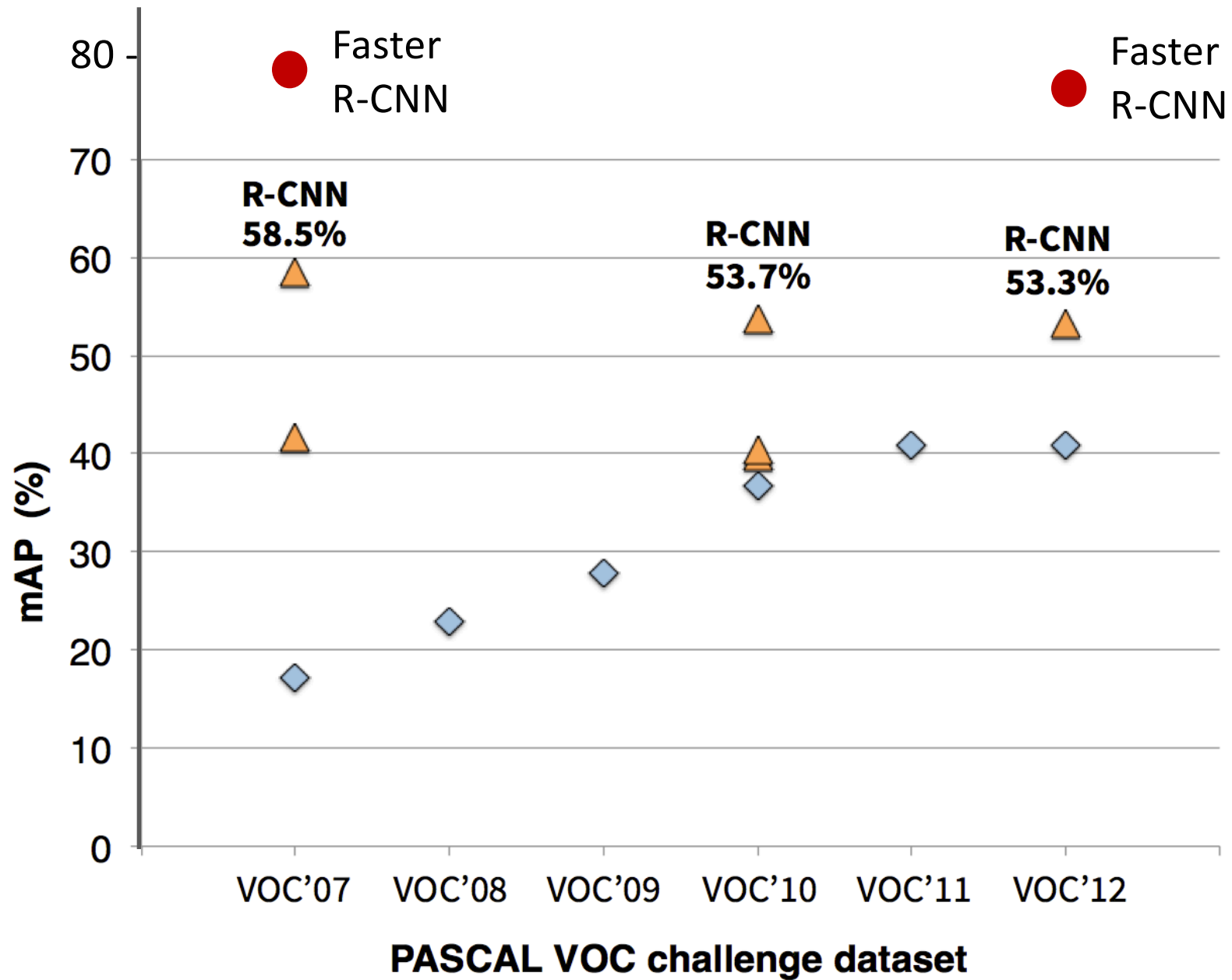


# Object detection 2013

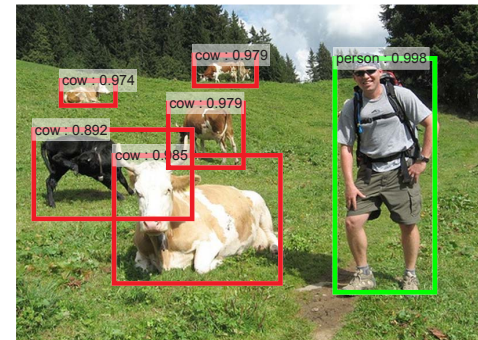
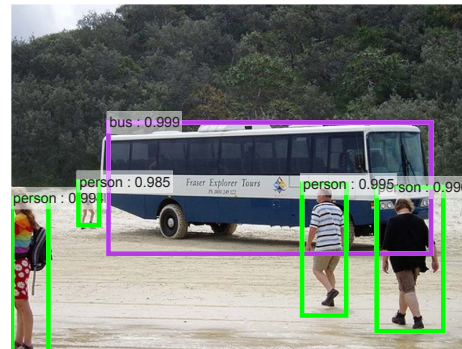
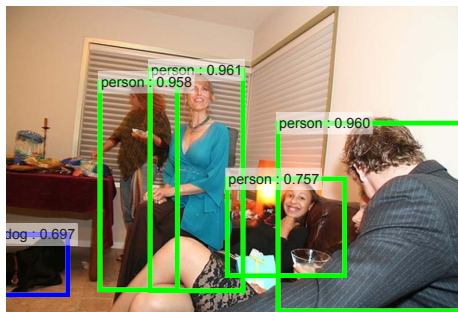
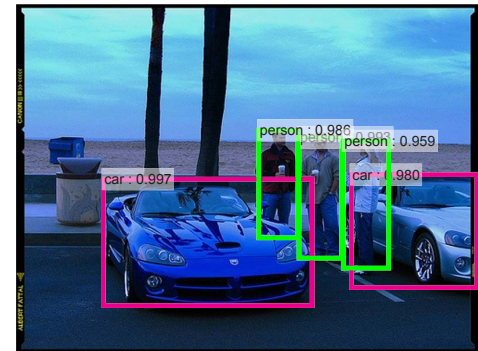
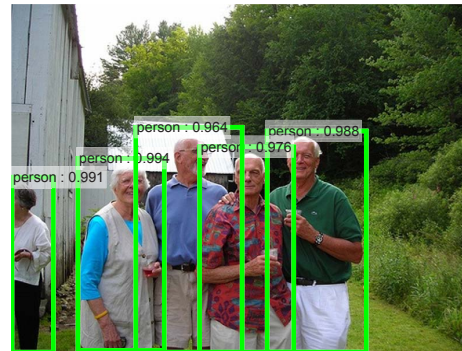
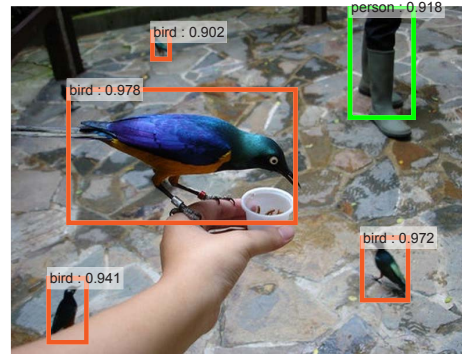
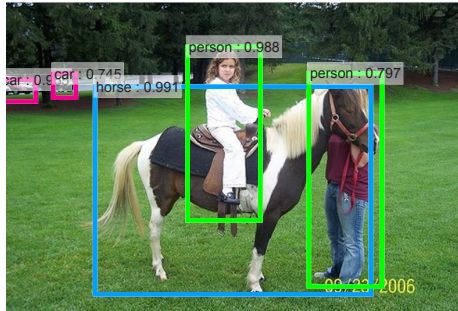




# Object detection 2016



# Object detection



# This Course Covers

- A wide range of Computer Vision techniques – from basic to state-of-the-art
- You will learn about:
  - Camera and Image formation, filters and features, image recognition, object detection, 3D reconstruction, motion and tracking, activity recognition, segmentation and grouping, physics-based vision
- Covers algorithms, theory and applications
- It's going to be fun and hard work 😊

# Grading

- Homework (4): 40 pts
- Midterm: 10
- Final Exam: 20
- Project: 30
- Grading philosophy:
  - Give A to everyone who deserves it
  - Use Human Vision to figure out the rest

# Late Policy for Homework

- Suppose:
  - A homework is worth  $N$  points
  - You submit  $K$  days late
- The maximum score you can get is:  $N^*(1 - K/4)$
- Allowance:
  - A total of 3 late days (for all homeworks) without penalty
  - Use them wisely



# Prerequisites

- Ability to do maths
  - Linear Algebra
  - Some knowledge about probability & optimization
- Ability to program
  - Matlab
  - Or ability to learn Matlab
- Must have access to your own GPU
  - If you want to do a project on Deep Learning

# Academic Misconduct

- Don't Cheat. It's not fun and you won't learn
- You need to do homework by yourself
  - Discussion is acceptable but you must do your own work
  - You **MUST** tell us who you discuss with. Write their names on your submission file.
- Existing solutions:
  - You **cannot** search for solutions online
  - You **cannot** look at previous years' solution
- Algorithm implementation:
  - Many algorithms exist online
  - Do not download/use them unless we say you can
  - When in doubt, ask

# Finally

- Computer Vision is one of the hottest topics in academia and industry today
- This class should give you the basic foundation for applying CV
- Enjoy!