COMPUTATIONAL VISION: Introduction

Master in Artificial Intelligence

Department of Mathematics and Computer Science

2023-2024

UNIVERSITAT DE BARCELONA

Outline:

1. What is Computer Vision?
2. A little bit of history
3. Main problems/tasks of Computer Vision
4. Difficulties of Computer Vision
5. Applications

- Szeliski, CV: A&A, Ch 1.0 (Introduction)

What is Computer Vision?

- *Vision* is about discovering from images what is present in the scene and where it is.
- In *Computer Vision* a camera (or several cameras) is linked to a computer.
- The computer interprets images of a real scene to:
  - perceive the story behind the image and
  - obtain information useful for tasks such as navigation, manipulation and recognition.

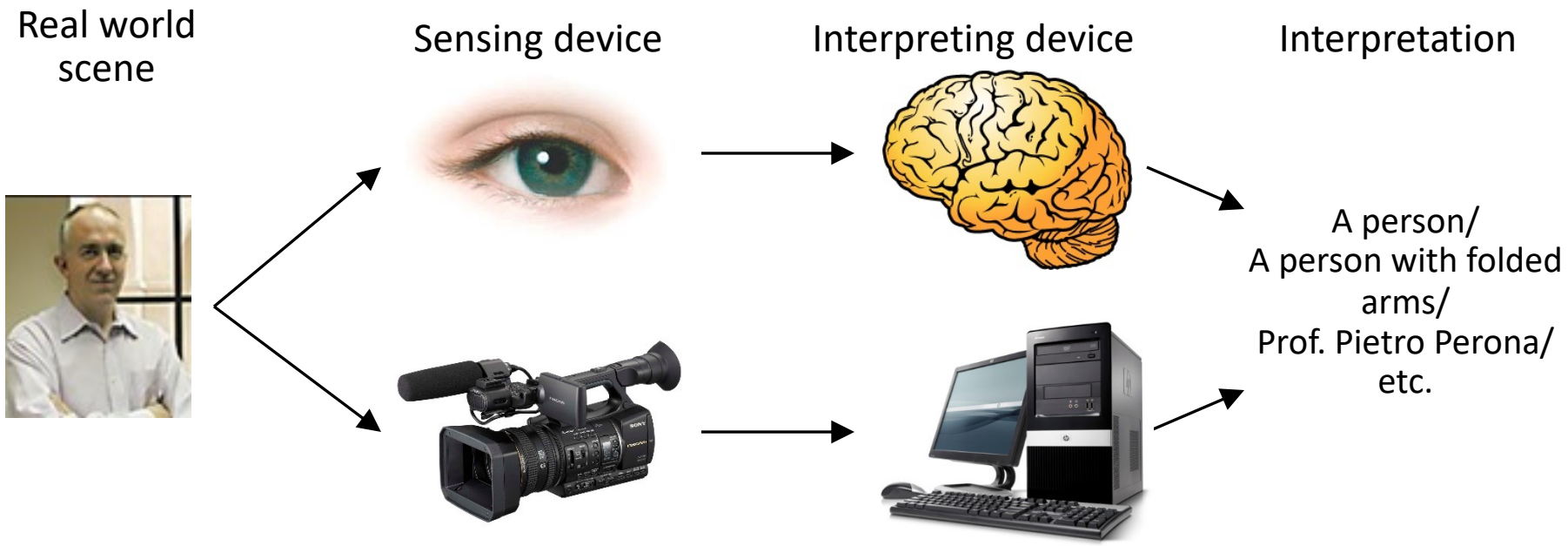
- What is happening?
- What kind of scene is there?
- Where are the cars?
- How far is the building?
- Can we track the cars?

…

The problem

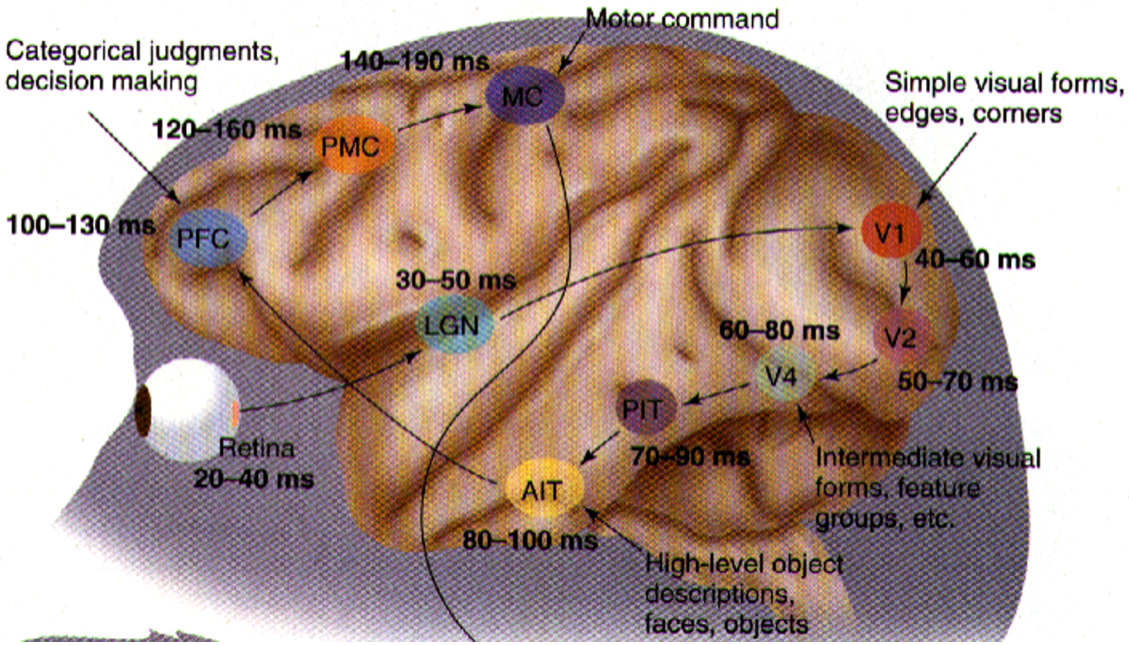
→ Want to make a computer understand images.

- We know it is possible – we do it without effort!



Human visual system

- Vision is the most powerful of our own senses.
- Around 1/3 of our brain is devoted to processing the signals from our eyes.
- The visual cortex has around $O(10^{11})$ neurons.



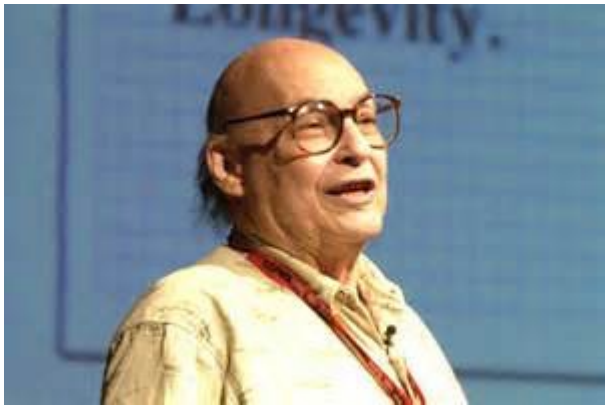
[Thorpe et. al.]

Outline:

1. What is Computer Vision?
2. A little bit of history
3. Main problems/tasks of Computer Vision
4. Difficulties of Computer Vision
5. Applications

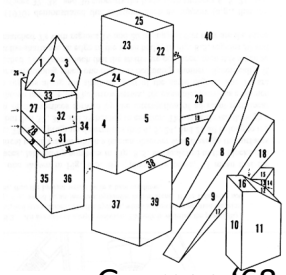
The Computer Vision Story Begins…

- In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”.



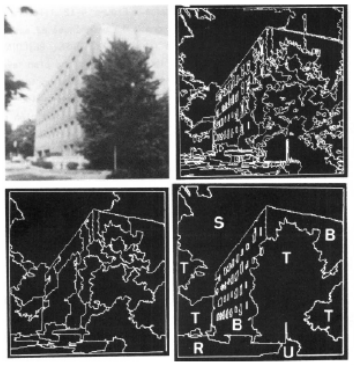
- We now know that the problem is slightly more difficult than that (Szeliski 2009, Computer Vision).

Ridiculously brief history of computer vision

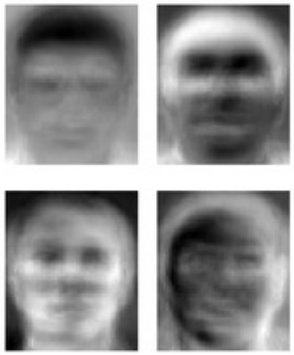


Guzman ‘68

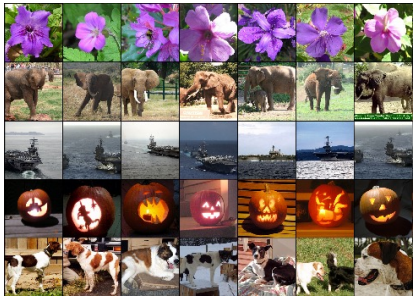
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960’s: interpretation of synthetic worlds as a composition of 3D objects (depth perception).
- 1970’s: some progress on interpreting selected images (object contours and labelling parts for seg.)
- 1980’s: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990’s: face recognition; statistical analysis in vogue
- 2000’s: broader recognition; large annotated datasets available; video processing starts
- 2010’s: ANN is back to stay, Deep learning.
- 2030’s: autonomous vehicles, robot uprising?



Ohta Kanade ‘78



Turk and Pentland ‘91

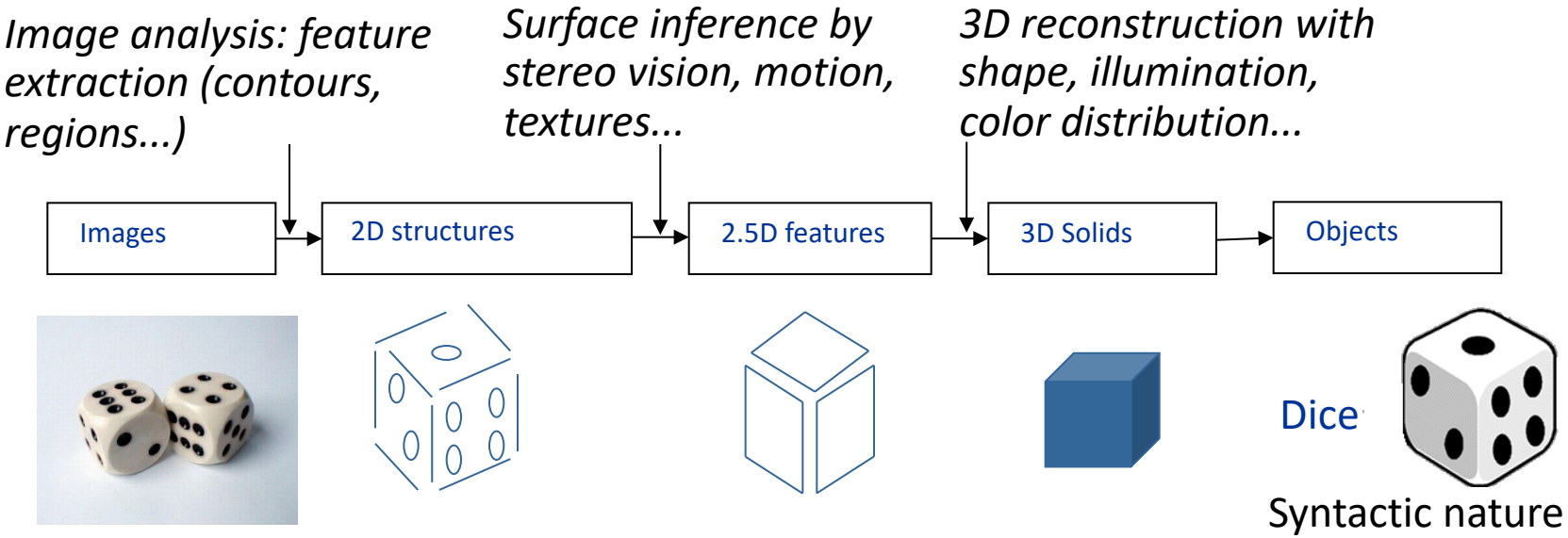


Imagenet: 15 mlln of images

[Source: Derek Hoiem]

Marr's Historical definition

- **Q:** How to answer: *What is this?*
- **R:** By a sequence of data transformations.



INVERSE PROBLEM: Recover some unknowns given insufficient information to fully specify the solution.

Vision...

But the environment is complex...

...not accessible, continuous, dynamic and non deterministic. And this is a BASIC component of the problem.

Outline:

1. What is Computer Vision?
2. A little bit of history
3. Main problems/tasks of Computer Vision
4. Difficulties of Computer Vision
5. Applications

Problems of Computer Vision

To bridge the gap between pixels and “meaning”

What we see

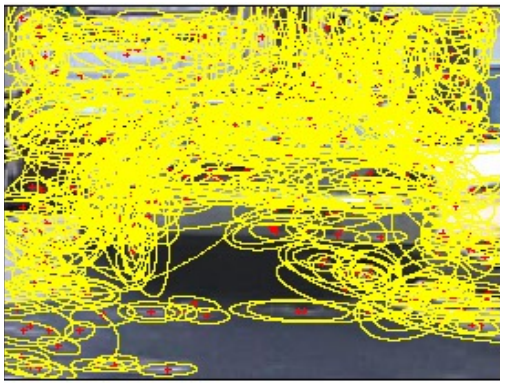
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Low-Level Feature Extraction

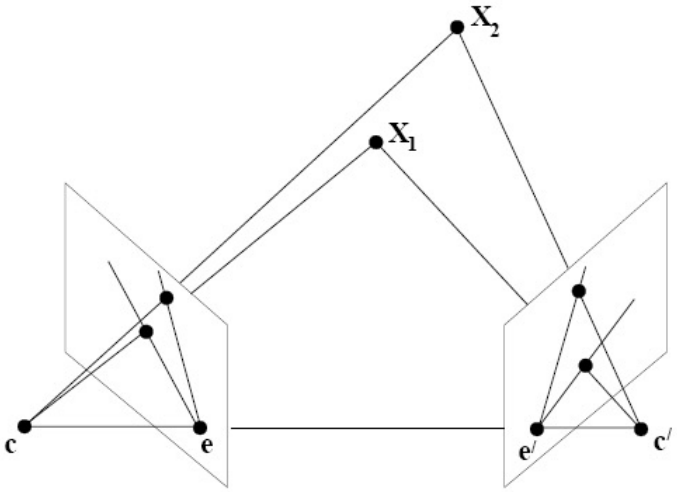
- Edges, corners

- Local regions



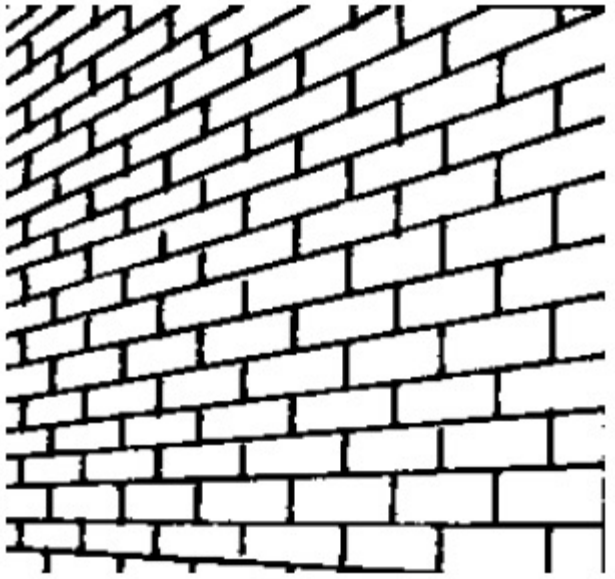
Stereo Vision

- Strategy to infer the 3D structure of the scene.
- By having two cameras, we can triangulate features in the left and right images to obtain **depth**.
- Need to match features/local points between the two images:
  - *Correspondence Problem*



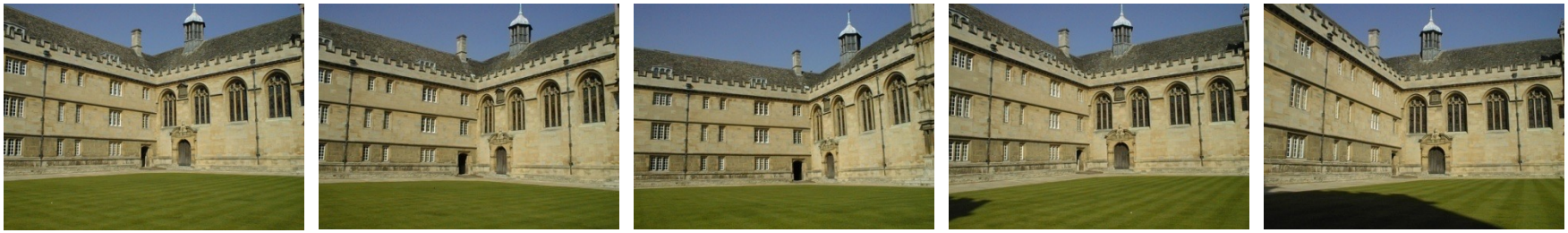
Shape from Texture

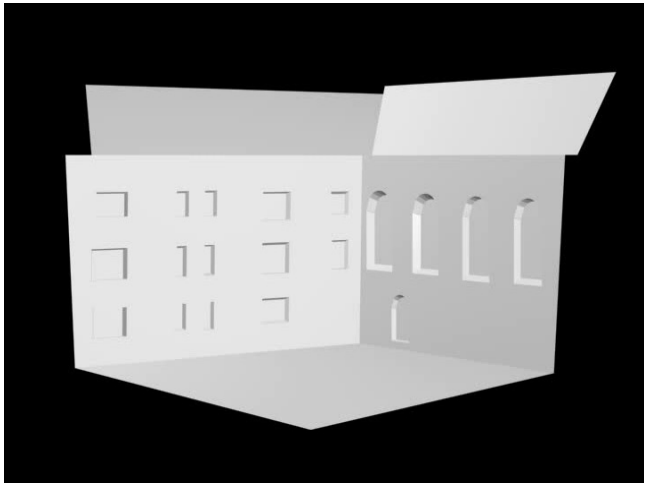
- Texture provides a very strong cue for inferring surface orientation in a single image.
  - By analyzing how the texture statistics vary over the image.



Scene Reconstruction

Given one or (typically) more images of a scene, or a video, scene reconstruction aims at computing a **3D model of the scene/objects**.





[Fitzgibbon et. al] [Zisserman et. al.]

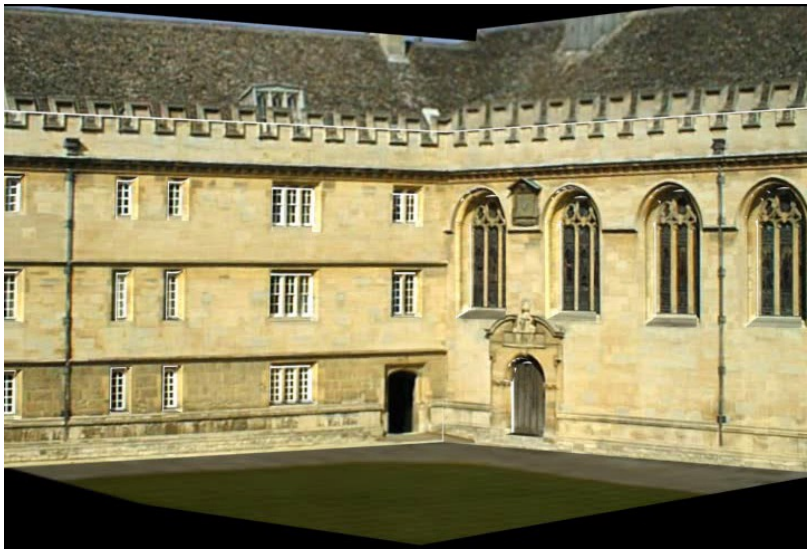
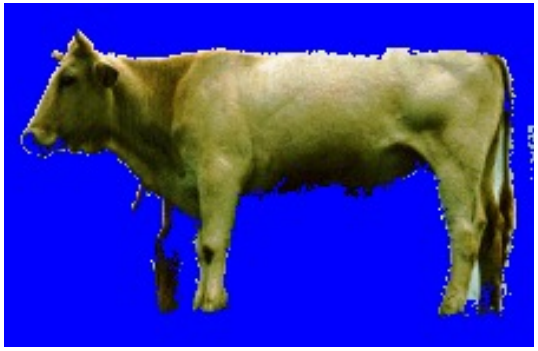


Image Segmentation

- Partition an image into multiple segments to simplify the representation and, for instance, locate objects in the image.

Image

Object Segmentation



Medical image Segmentation

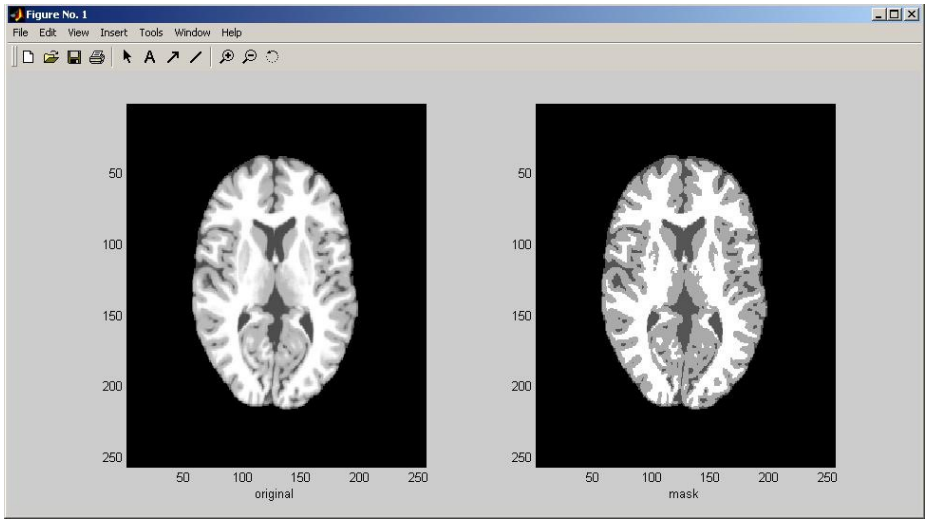
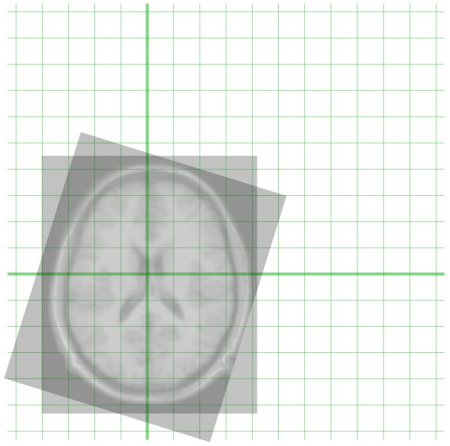


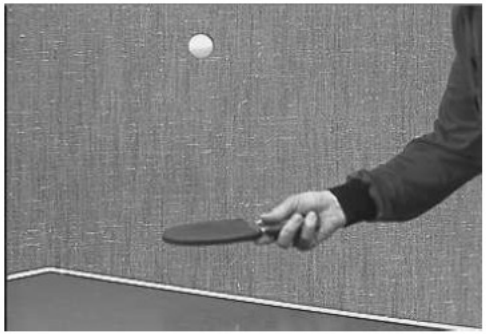
Image Registration

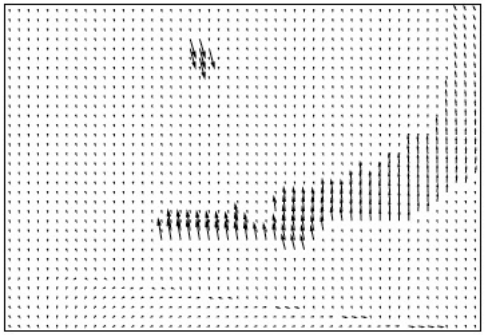
- Transform different sets of images into one coordinate system.

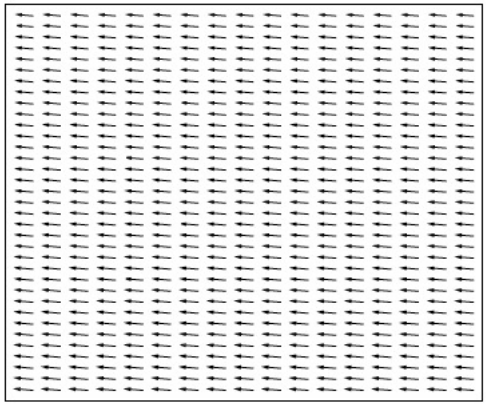


Motion Estimation: Human/Camera

- Determine motion vectors that describe the transformation from one image to another from adjacent frames in a video sequence.

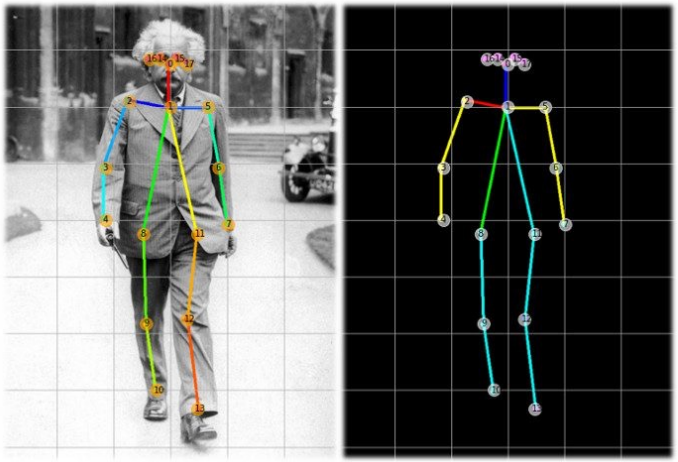


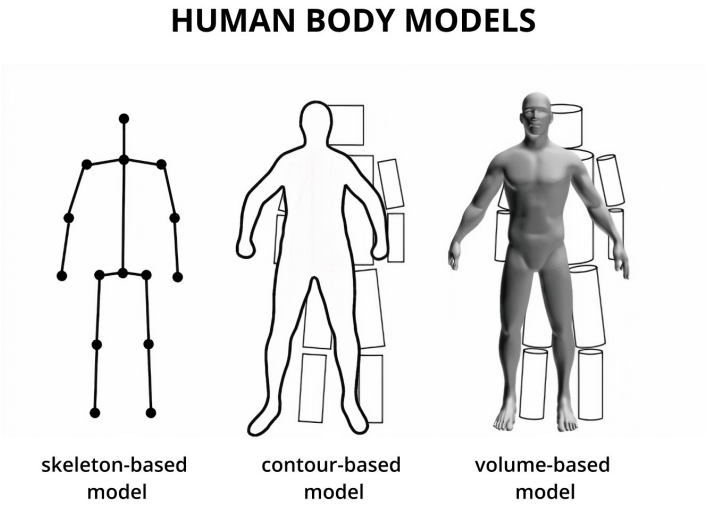


Human Pose Estimation

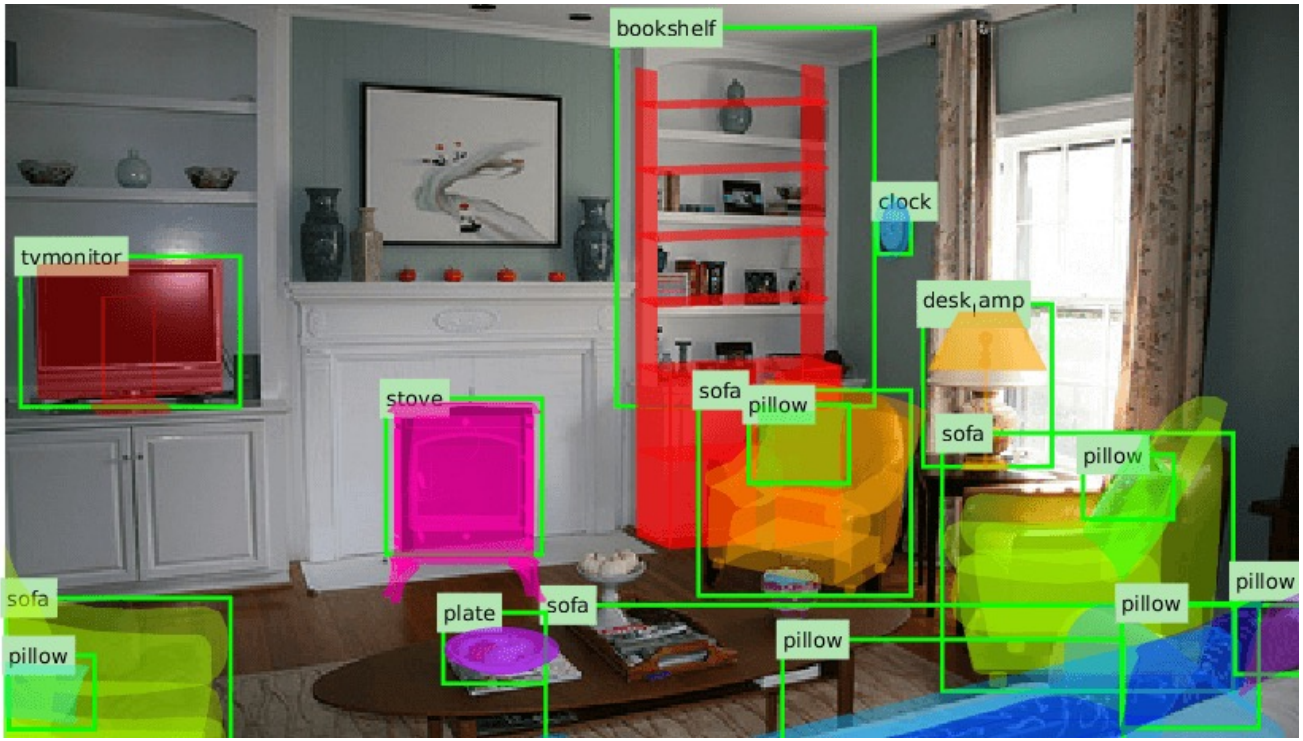
- Recognize and accurately track specific points on the human body which can be used to identify motion patterns, specific joint position, or the pose itself.





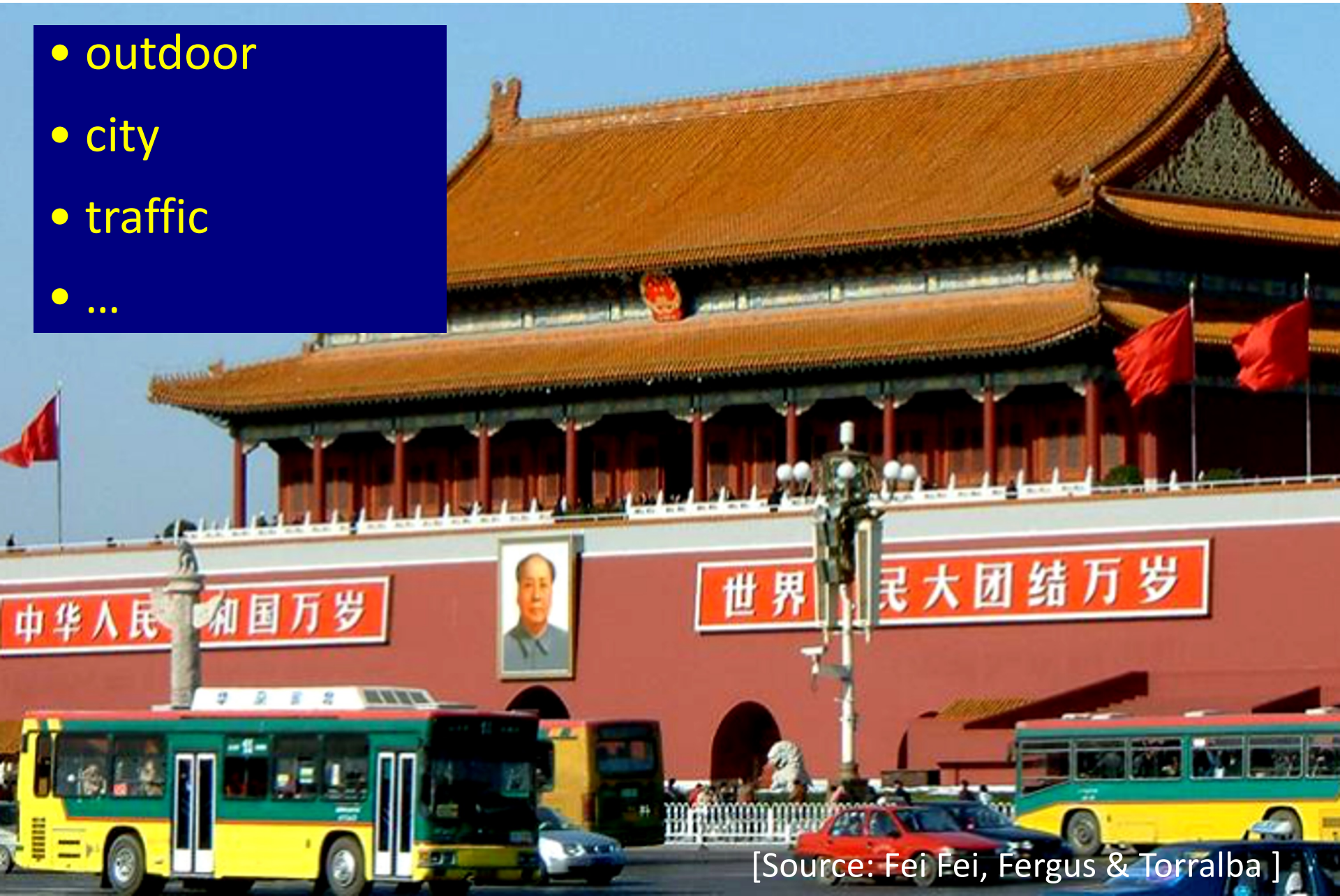
Object and Person Recognition

- Find and identify objects in an image or video.



Scene and context categorization

- outdoor
- city
- traffic
- …



[Source: Fei Fei, Fergus & Torralba]

Outline:

1. What is Computer Vision?
2. A little bit of history
3. Main problems/tasks of Computer Vision
4. Difficulties of Computer Vision
5. Applications

Why is it hard?

Variation of point of view

Illumination

Scale

Why is it hard?

Intra-class variation

Movement (Font: S. Lazebnik)

Cluttered background

Occlusion

Why is it hard?

- The ambiguity of the implicit perception
  - Many 3D scenes can give the same 2D scene

- We need information about the World to unambiguate the images.

[Image source: F. Durand]

Each image tells a story...

Can computers beat the humans?

Yes and no (mostly not!)

- Humans are better at "hard" things.
- Computers are better at "easy" things.

Outline:

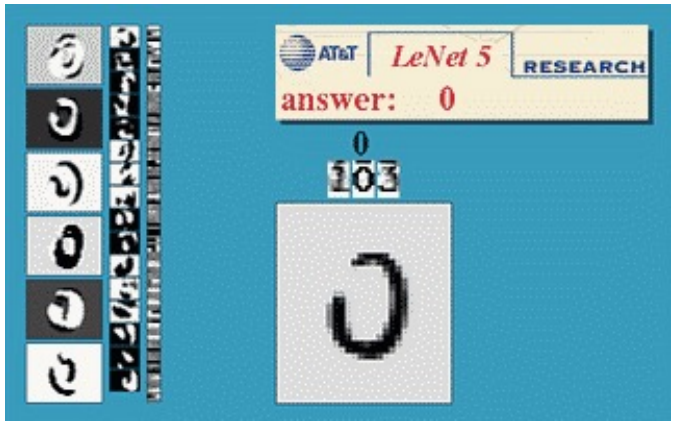
1. What is Computer Vision?
2. A little bit of history
3. Main problems/tasks of Computer Vision
4. Difficulties of Computer Vision
5. Applications

Why to study Computer Vision now?

- Each second million of images are captured and stored

- Numerous real applications: next slices…

Access

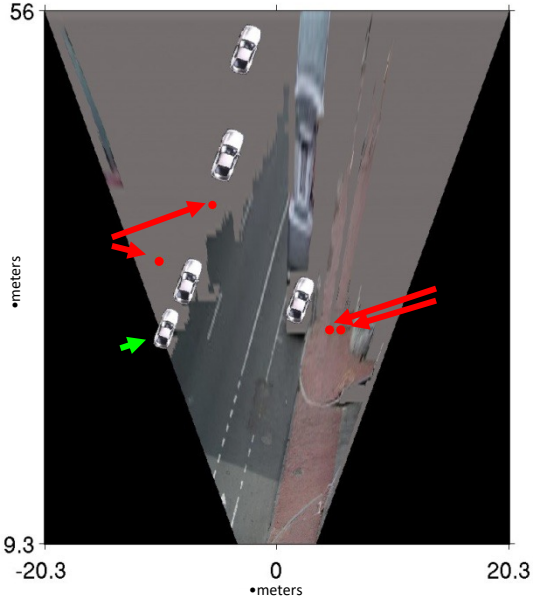


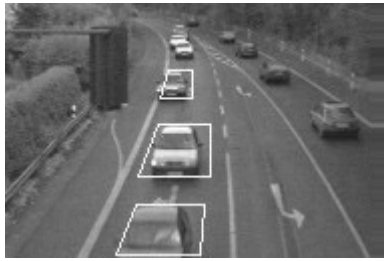
Digits recognition, AT&T labs

- Optical character recognition (OCR): Technology to convert scanned docs to text.

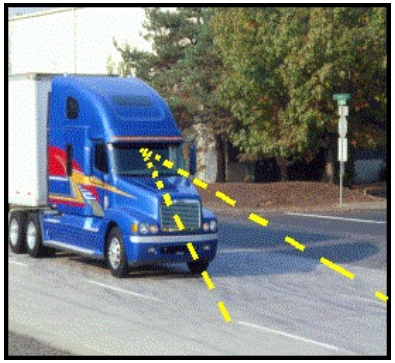
Safety and Autonomous Driving

- Pedestrian and car detection





- Lane detection



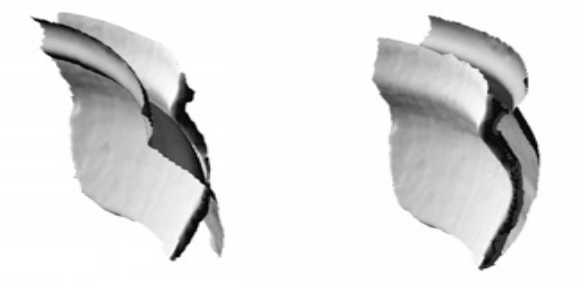
- Collision warning systems with adaptive cruise control
- Lane departure warning systems
- Rear object detection systems

Slide credit Fei-Fei, Fergus, Torralba

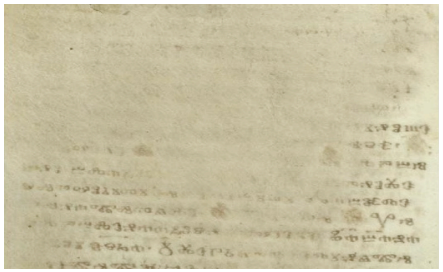
Cultural Inheritance

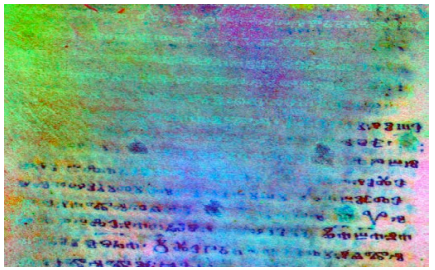
Arqueology

3D reconstruction of ceramics

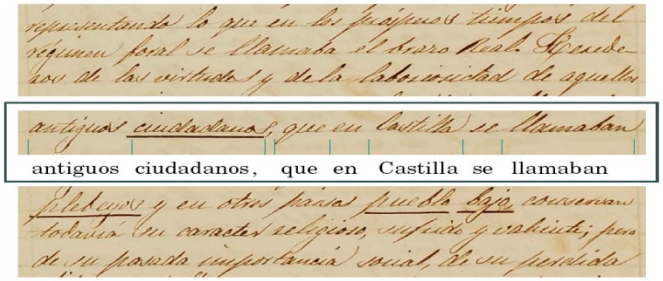


Old documents

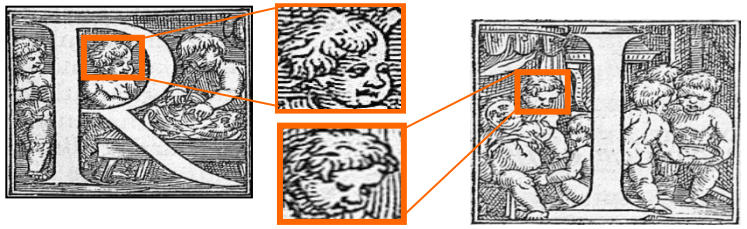




Manuscript recognition

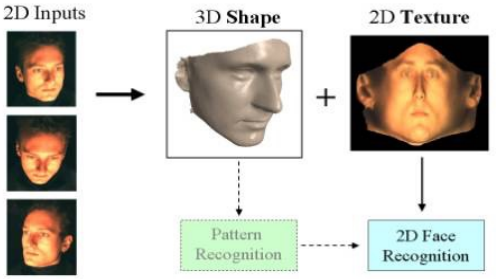


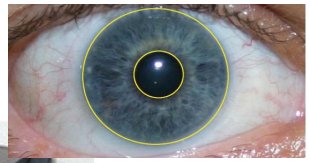
Restauration



Engravings analysis

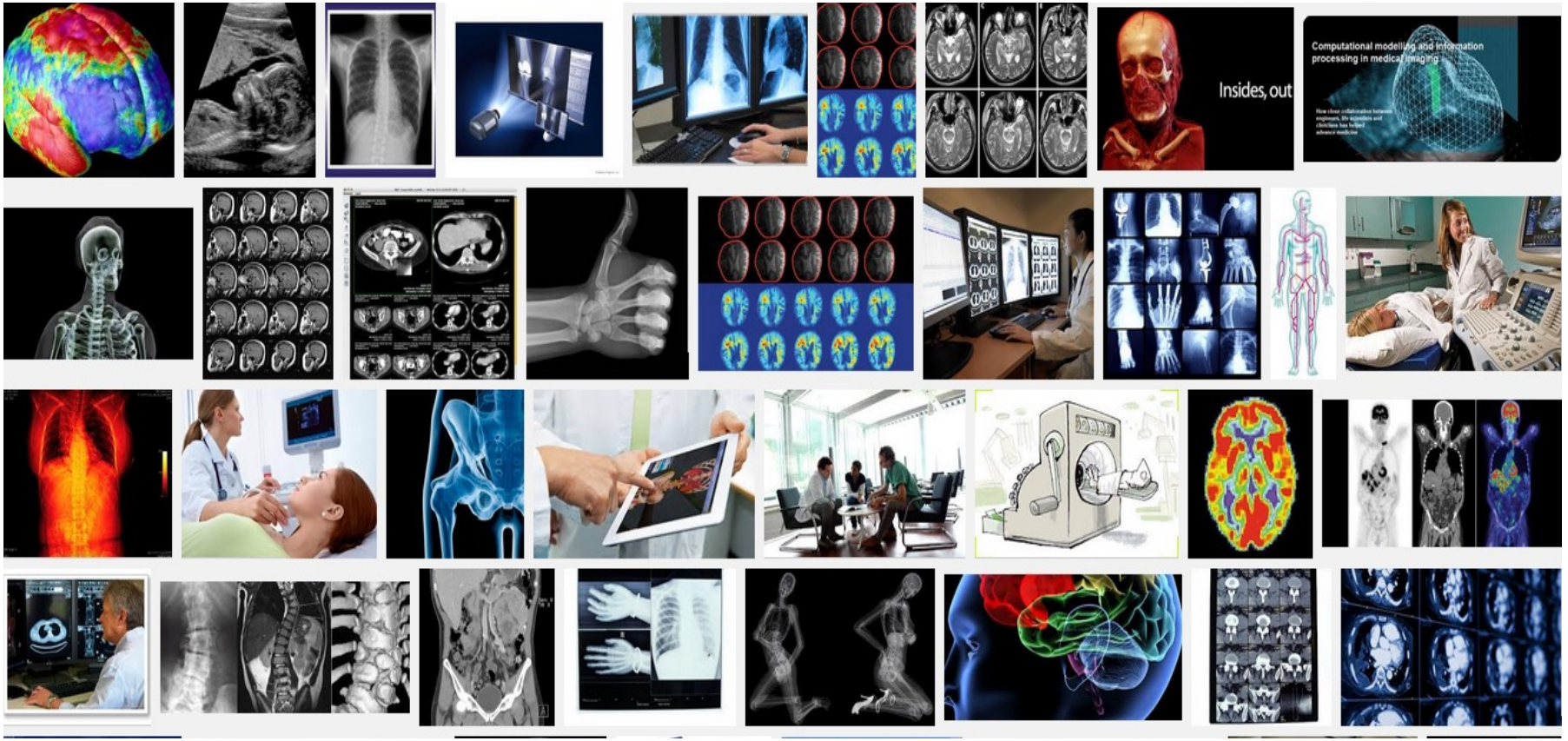
Security and Surveillance





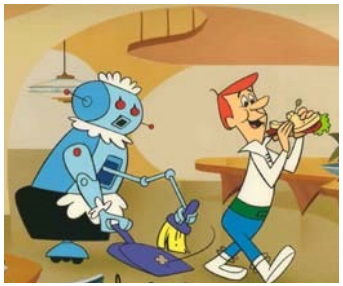
Fingerprint identification, facial recognition, iris and retinal scan, hand geometry, geometry of the ear, signature recognition, voice identification, identification of the DNS, the smell of human characteristics or recognition of typing motion (not a privilege of science fiction any more). Surveillance cameras to warn when a robbery happen.

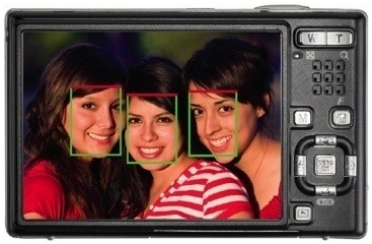
Medical Imaging and Health



- Automatic detection and recognition of different type of diseases or lessions in medical images.

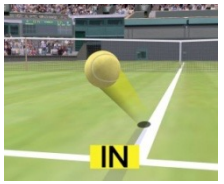
Comfort & Fun



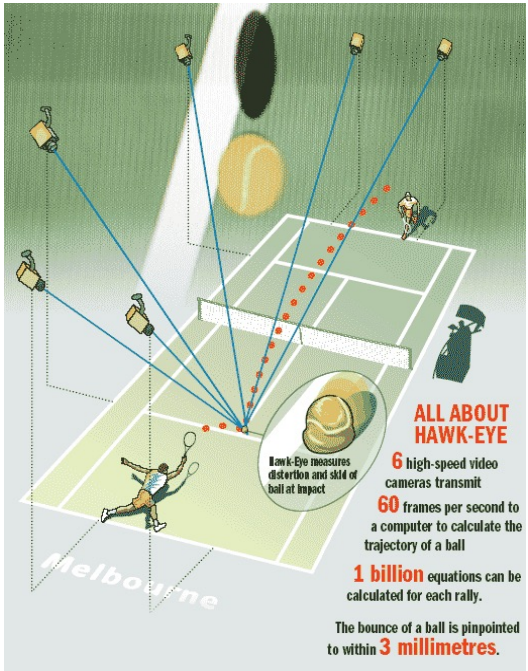


- Vision system for robots who can clean or take care of (eldery) people
- Vision-based interaction for games
- Hands detection
- Face detection

•Augmented TV



•Hawk Eye: tracking the ball in game sports.



Computer Vision Today

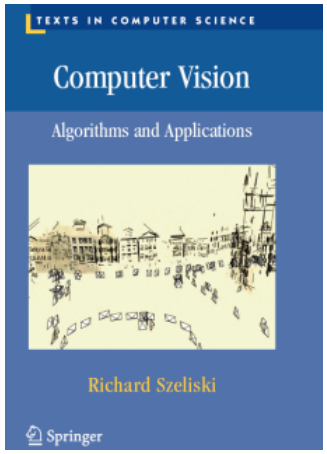
- A very active and changing area.

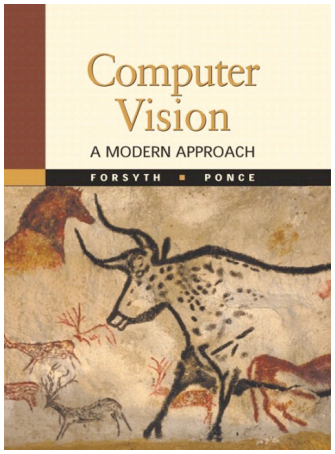
- Some articles:
- https://www.forbes.com/sites/bernardmarr/2022/03/04/the-5-biggest-computer-vision-trends-in-2022/
- https://medium.com/@nishantjethwa23122000/the-future-of-computer-vision-87623cddb865
- https://www.forbes.com/sites/forbestechcouncil/2020/06/16/from-computer-vision-to-deep-learning-how-ai-is-augmenting-manufacturing/#2163e9834829
- https://www.startus-insights.com/innovators-guide/computer-vision-startups-to-watch/

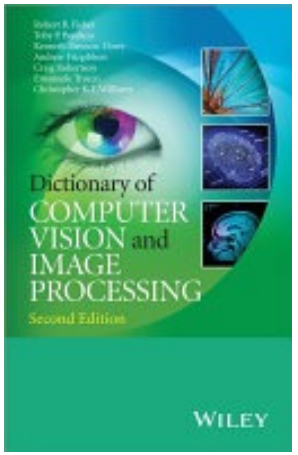
Material

Books:

- Rick Szeliski, Computer Vision: Algorithms and Applications,
- David Forsyth and Jean Ponce, Computer Vision: A Modern Approach,
- Dictionary of Computer Vision and Image Processing, by Fisher et al. Note: Full text is available in 'Online Resources' section.







http://szeliski.org/Book/

Material

More Books:

- Reinhard Klette, Concise Computer Vision An Introduction into Theory and Algorithms
- Ian Goodfellow and Yoshua Bengio and Aaron Courville, Deep Learning

Lot's of papers!!

Slice Sources

Thanks to other sources for slices:

- Derek Hoiem
- Thorpe et. al.
- Fei Fei, Fergus & Torralba
- F. Durand
- S. Lazebnik
- N. Snavely
- S. Narasimhan
- James Tompkin (http://cs.brown.edu/courses/cs143/)

Online courses and Lecture sites

Udacity courses:

- Introduction to Computer Vision

  → You can create an account in Udacity, start the free course and download all videos from the course to visualize the ones you are interested in.

- Computational Photography.

Lecture sites:

- CS231n: Deep Learning for Computer Vision (Stanford)
- Advances in Computer Vision (MIT)

Videos for the Introduction

- “How computer vision works”:

https://www.youtube.com/watch?v=OcycT1Jwsns

- “What Is Computer Vision & Why Does It Matter?”:

https://www.youtube.com/watch?v=OnTgbN3uXvw

- Lectures from Udacity course related with the Introduction to Computer Vision:
  - 4 - What is Computer Vision
  - Until 19 - Why is This Hard

  Total time: 15 minutes.

COMPUTATIONAL VISION: Introduction

Master in Artificial Intelligence

Department of Mathematics and Computer Science

UNIVERSITAT DE BARCELONA