COMPUTATIONAL VISION: Introduction

Master in Artificial Intelligence

Department of Mathematics and Computer Science

2019-2020

UNIVERSITAT DE BARCELONA

Outline:

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

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

What is Computer Vision?

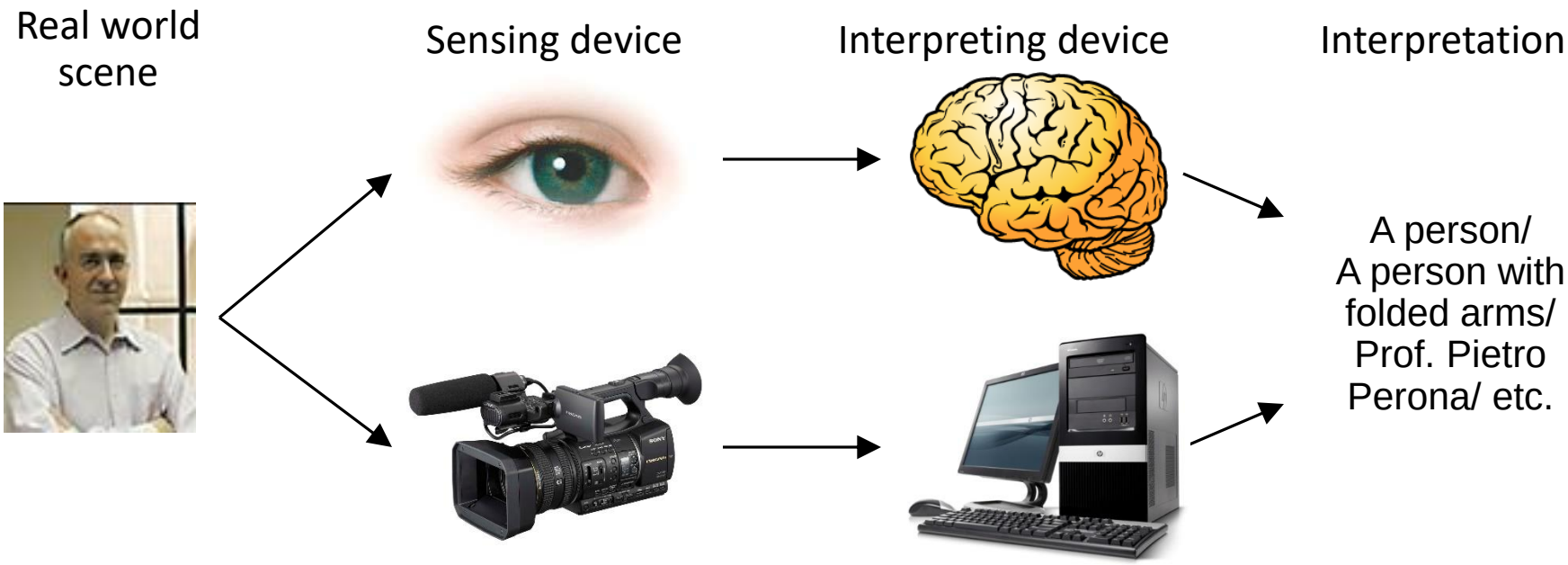
- *Vision* is about discovering from images what is present in the scene and where it is.
- In *Computer Vision* a camera (or several cameras) is linked to a computer.
- The computer interprets images of a real scene to:
  - perceive the story behind the image and
  - obtain information useful for tasks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

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

…

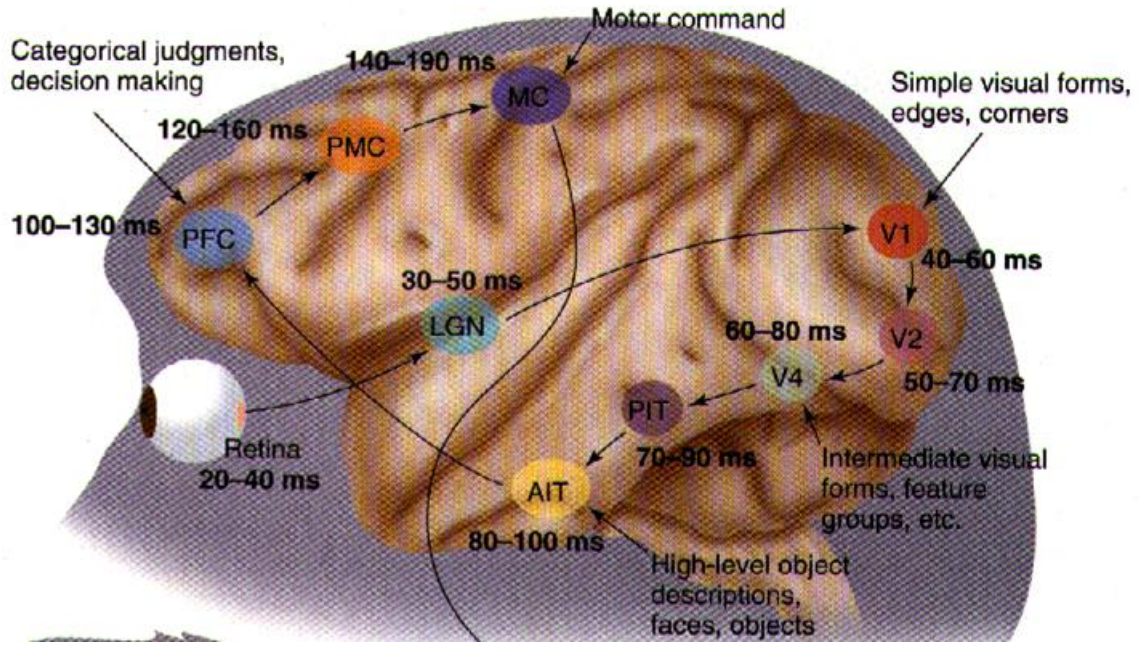
The problem

- Want to make a computer understand images
- We know it is possible – we do it effortlessly!



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.



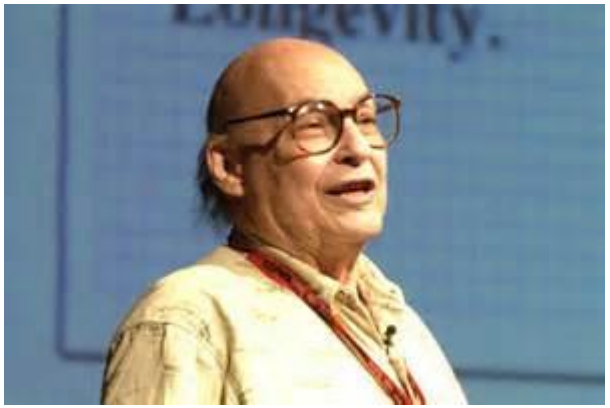
[Thorpe et. al.]

Outline:

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

The Vision Story Begins…

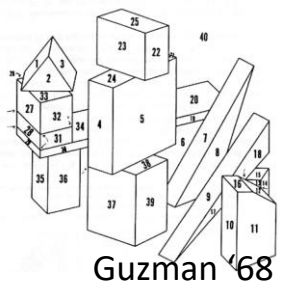
- In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”.



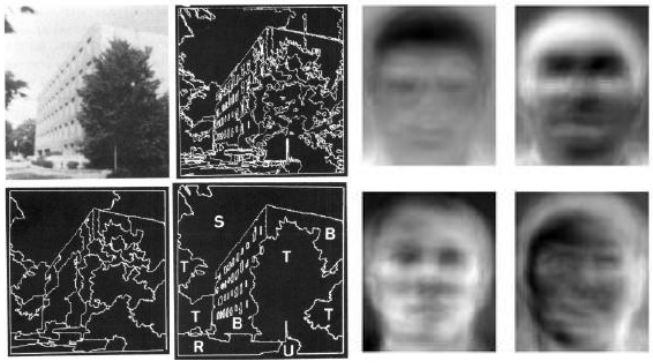
- We now know that the problem is slightly more difficult than that (Szeliski 2009, Computer Vision).

Ridiculously brief history of computer vision

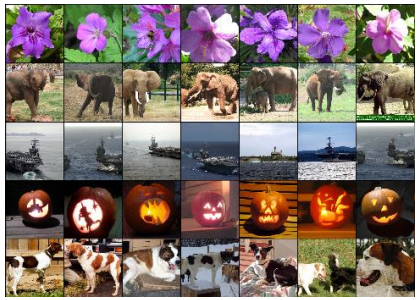
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds.
- 1970's: some progress on interpreting selected images (object contours and labelling parts for seg.)
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 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?



Guzman '68



Ohta Kanade '78 Turk and Pentland '91

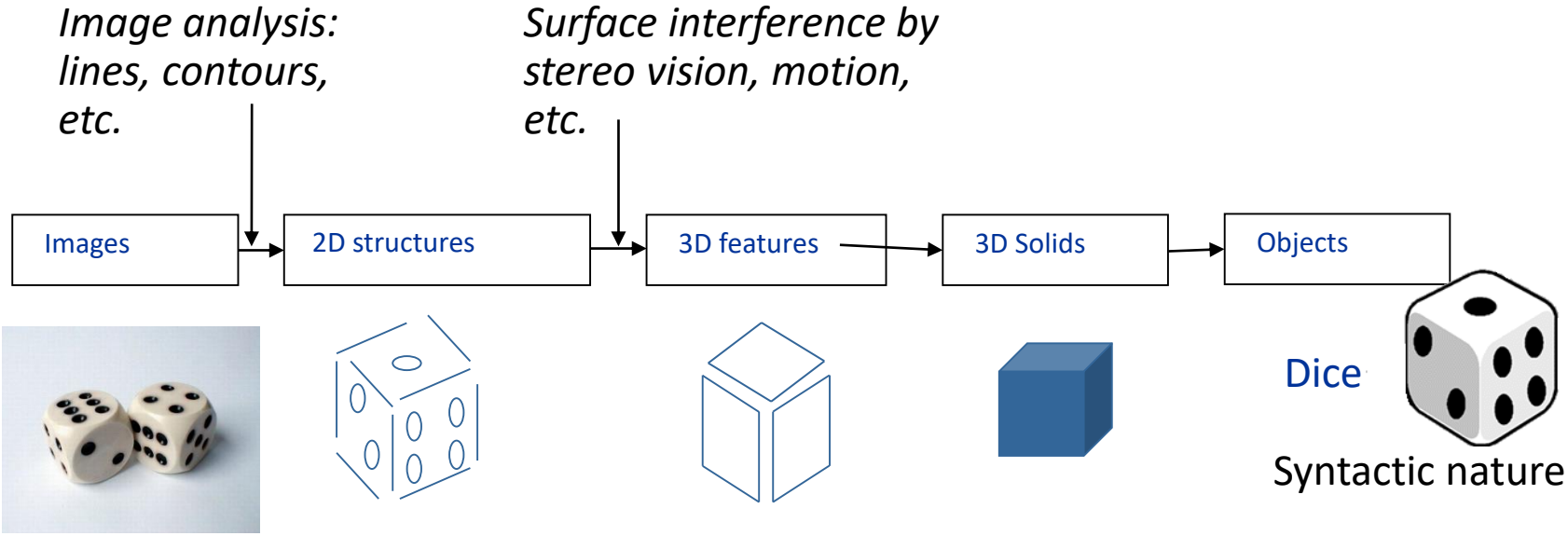


Imagenet: 15 mlln of images

[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.

Outline:

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

To bridge the gap between pixels and “meaning”

What we see

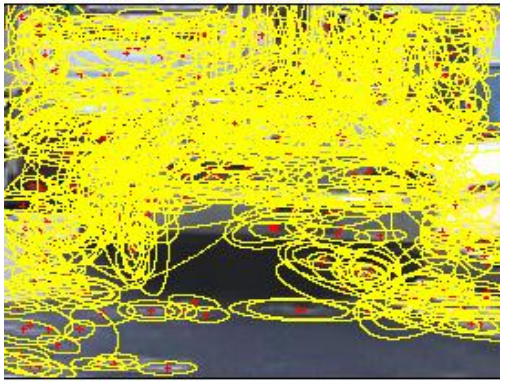
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Low-Level Feature Extraction

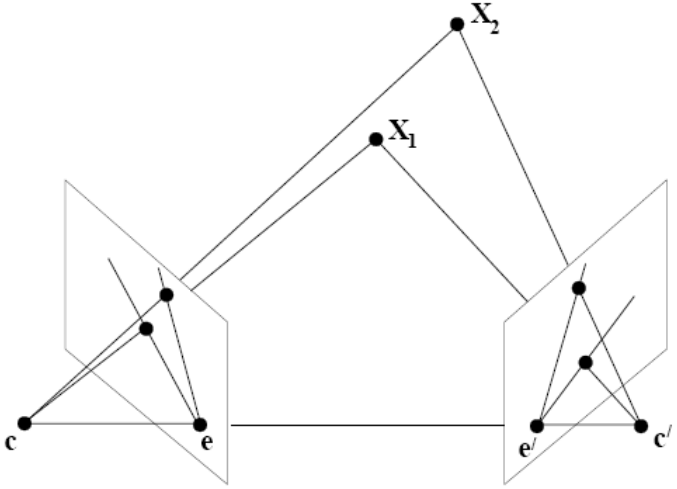
- Edges, corners

- Local regions



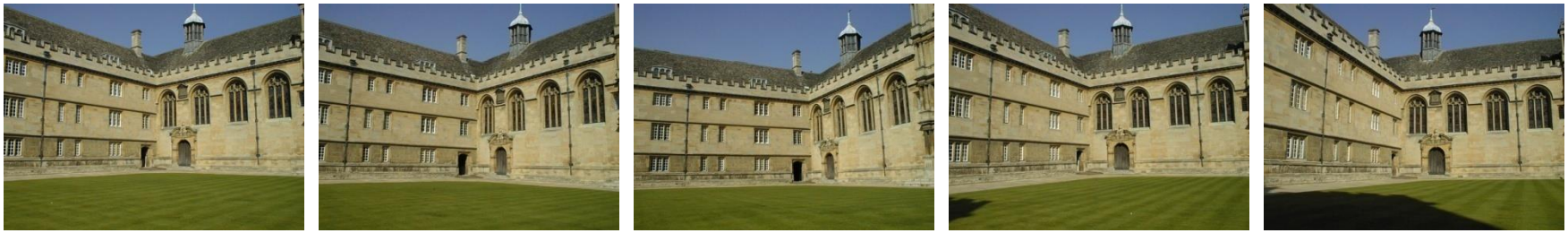
Stereo Vision

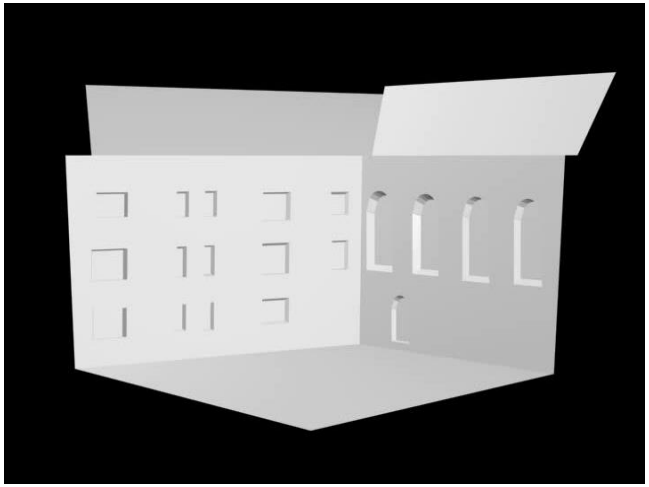
- One of the strategies to infer the 3D structure of the scene.
- By having two cameras, we can triangulate features in the left and right images to obtain **depth**.
- Need to match features between the two images:
  - *Correspondence Problem*



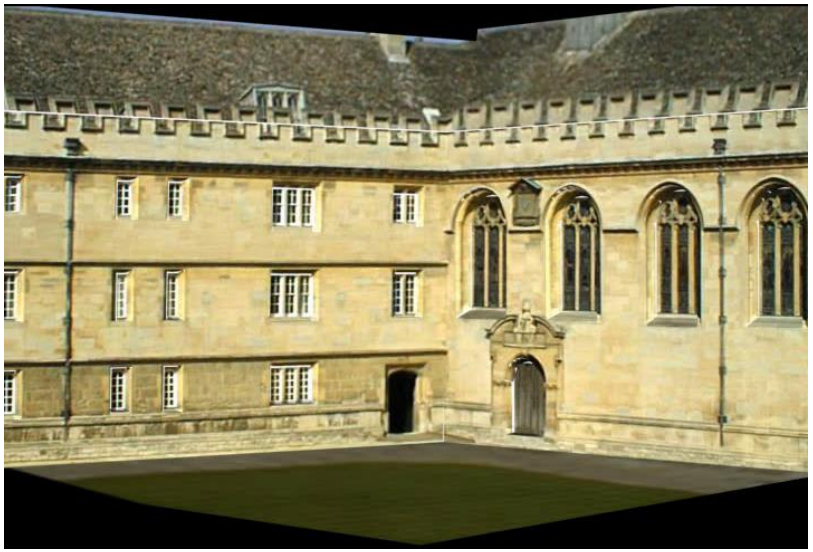
Scene reconstruction: 3D models of objects

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





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



Shape from Texture

- Texture provides a very strong cue for inferring surface orientation in a single image.
  - It is possible to infer the orientation of surfaces by analyzing how the texture statistics vary over the image.

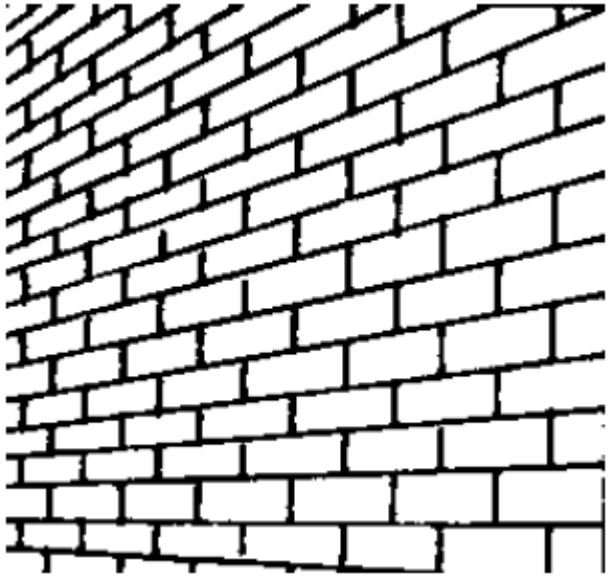


Image Segmentation

- It is the process of partitioning an image into multiple segments to simplify the representation and locate objects in the image, for instance.

Image

Object Segmentation

Medical image Segmentation

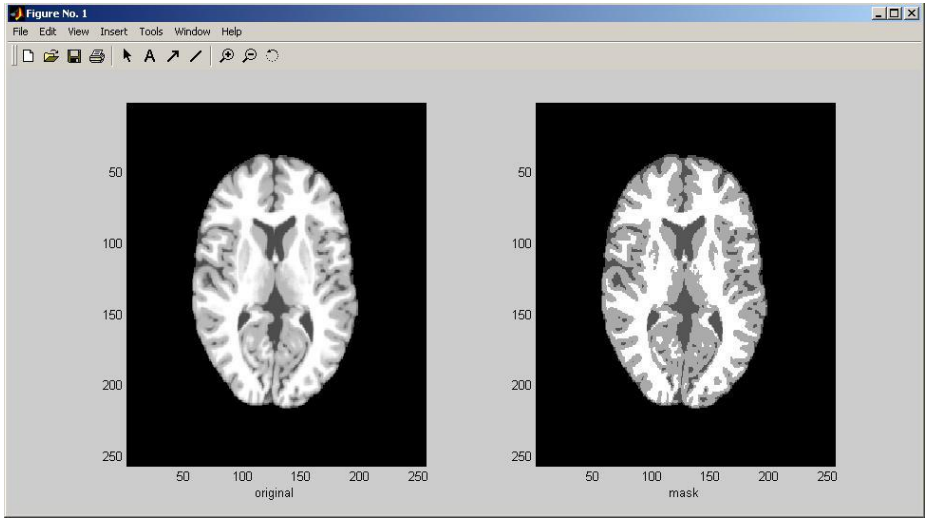
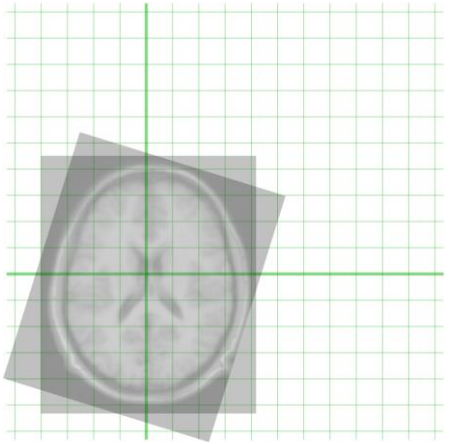


Image Registration

- It is the process of transforming different sets of images into one coordinate system.

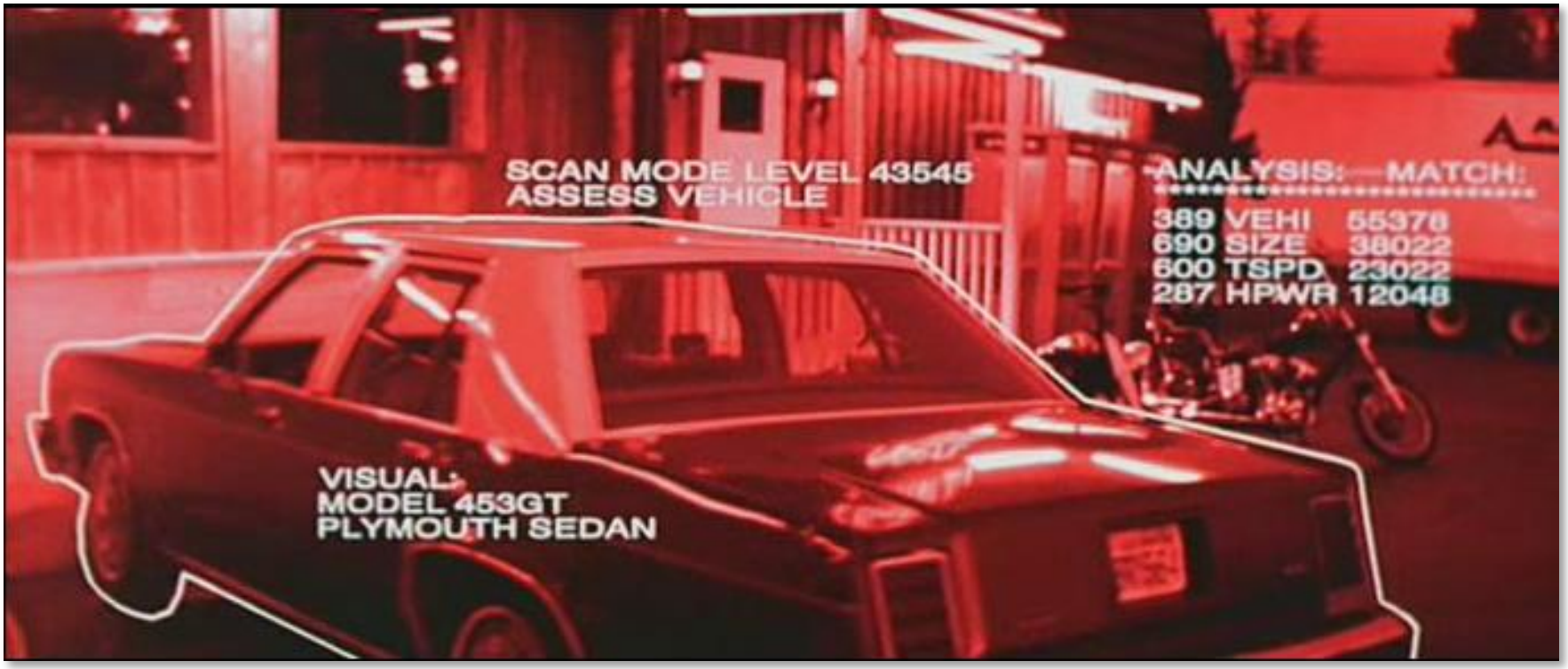


Motion estimation: human/camera

- Process of **determining motion vectors** that describe the transformation from one image to another; usually from adjacent frames in a video sequence.

Object and person recognition

- Process to find and identify objects in an image or video.



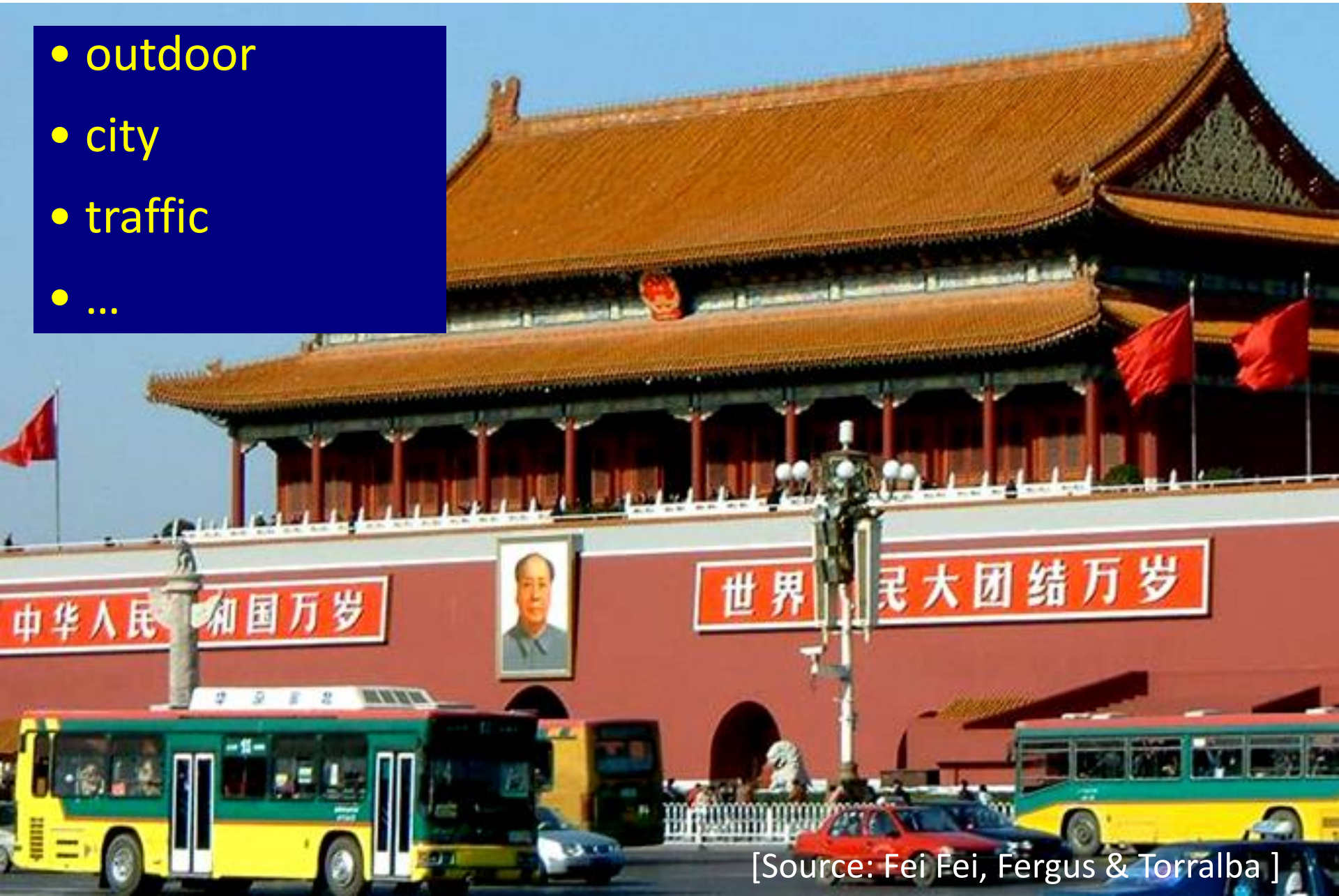
[Source: N. Snavely]

Object categorization

[Source: Fei Fei, Fergus & Torralba]

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 of Artificial Vision
4. Difficulties of the Artificial Vision
5. Applications

Why is it hard?

Variation of point of view

Illumination

Scale

Why is it hard?

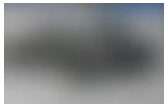
Intra-class variation

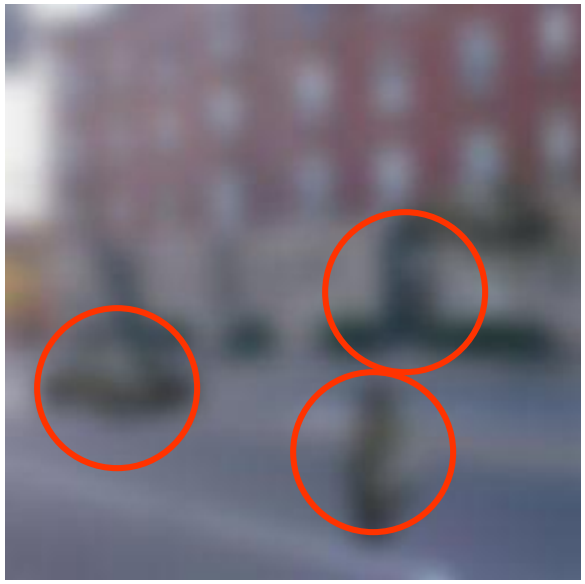
Movement(Font: S. Lazebnik)

Cluttered background

Occlusion

Why is it hard?





Local ambiguity

[Source: Fei Fei, Fergus & Torralba]

Bottom-up data analysis

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

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

[Image source: F. Durand]

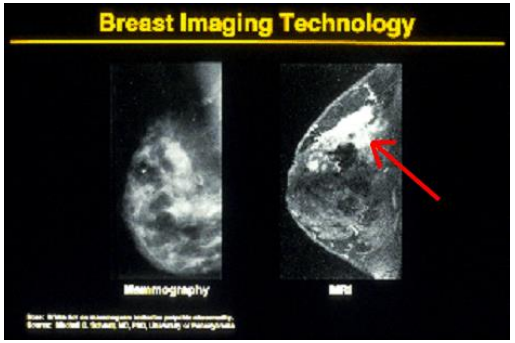
Outline:

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

Computer Vision Applications

Safety

- Smart cars with vision systems to prevent accidents.
- Autonomous driving

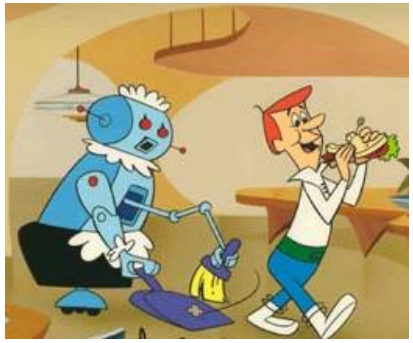


Health

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

Security

- Surveillance cameras to warn when a robbery happen



Comfort

- Vision system for robots who can clean or take care of (eldery) people.

Fun

- Vision-based interaction for games.
- Hands detection.
- Face detection.

Access

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

Computer Vision Today

- A very active and changing area.
- Deep Learning in CV.

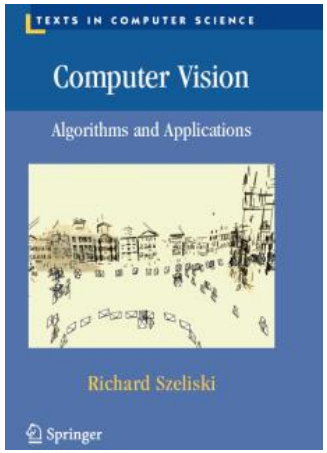
- Some articles:

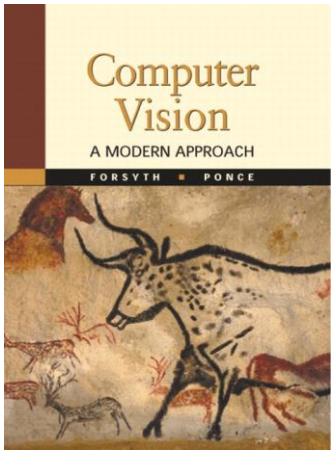
- https://www.forbes.com/sites/forbestechcouncil/2018/04/26/technology-that-sees-the-world-welcome-to-the-future-of-computer-vision/#7d0dbdcf33cb
- https://www.forbes.com/sites/cognitiveworld/2019/06/26/the-present-and-future-of-computer-vision/#59c98e52517d

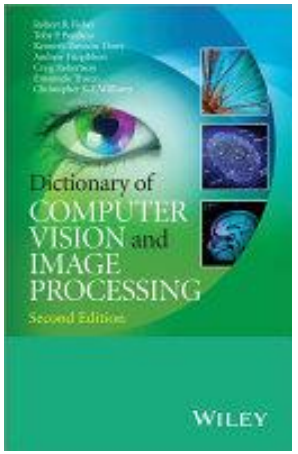
Material

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.







- 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/)

COMPUTATIONAL VISION: Introduction

Master in Artificial Intelligence

Department of Mathematics and Computer Science

2019-2020

UNIVERSITAT DE BARCELONA