Introduction to Computer Vision

Ivor Simpson

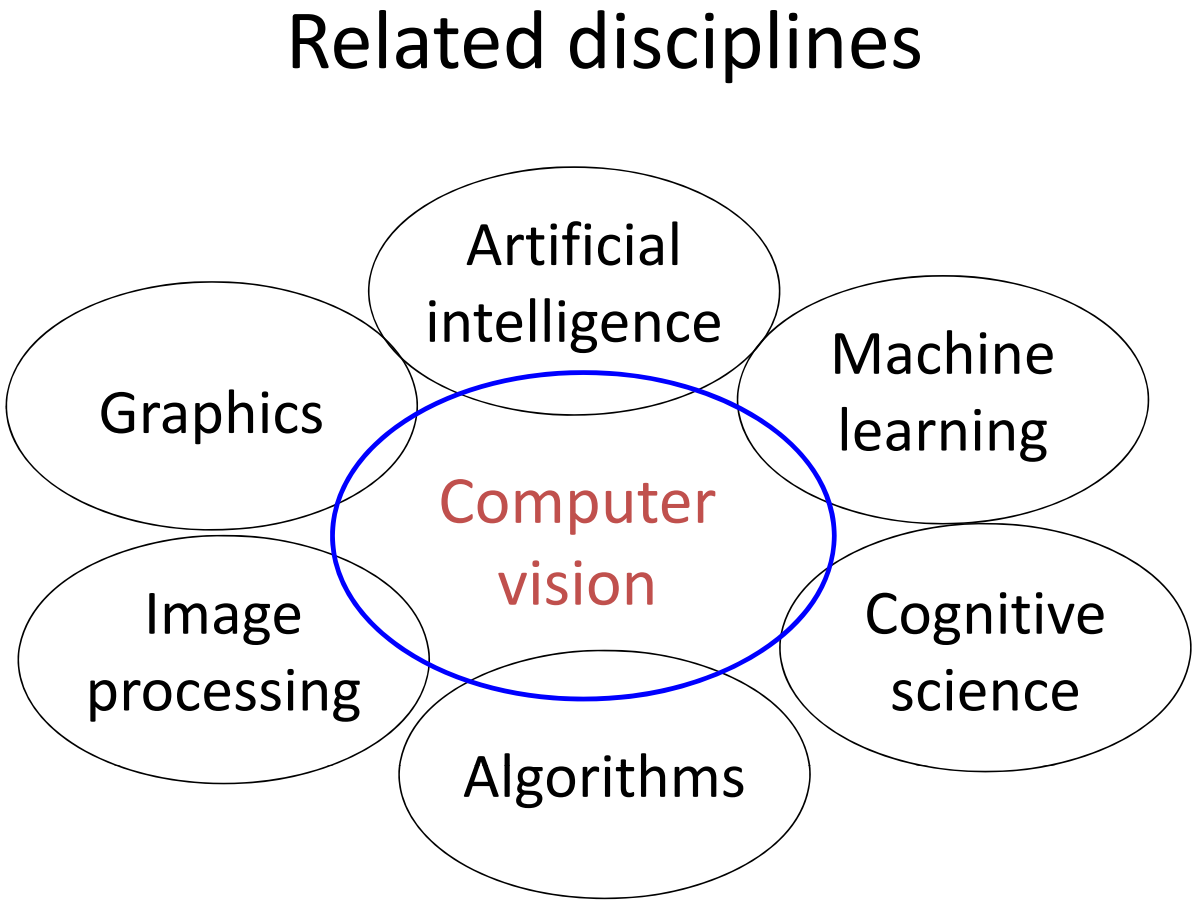
24th January 2023

What is Computer Vision?

Using computers to understand the contents of image and video data acquired from the real world.

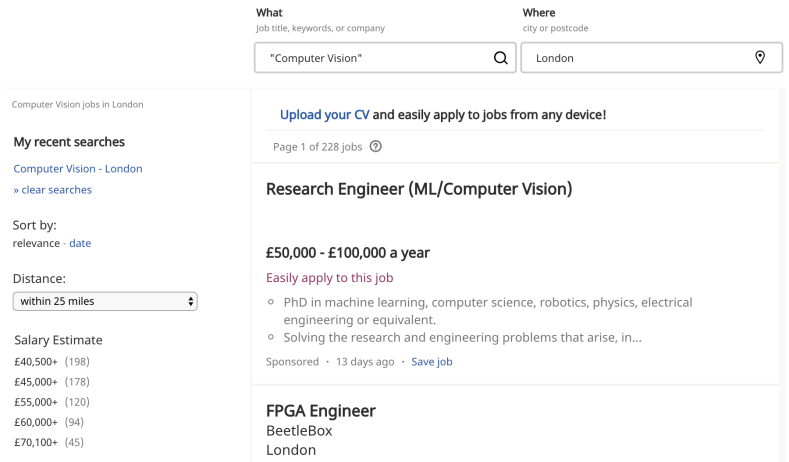
- Computing properties of the 3D world from visual data (measurement)
- Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (perception and understanding)

Where does CV fit in with CS?



Why should Computer Vision be interesting to you?

- Curiosity - computer vision is tasked with understanding a fundamental human sense.
  You also interact with a number of vision systems in day-to-day life.
- Combination of programming, applied mathematics and machine learning should fit well within your skillset at graduation.
- Job prospects are very good in this field at the moment!



University of Sussex

How do you interact with Computer Vision in your life?

- Image and video data are everywhere in modern society.
- Let's do a quick survey of how everyone here interacts with computer vision in the real world? - https://www.pollev.com/ivorsimpson490
- Think about what you do on your phones? Social Media? Anything else?

Outline for Today

- About the module: logistics, evaluation, and syllabus overview
- What can we use Computer Vision for?
- What are the challenges (aka. research opportunities!) in Computer Vision?

Module Learning Outcomes

- Learn to write computer programs that extract useful information from image data.
- Propose designs for simple computer vision systems.
- Determine the applicability of a variety of computer vision techniques to practical problems.
- Describe and recognise the effects of a variety of image processing operations.

Logistics

- Class webpage:
  - All teaching materials (lecture notes, lab exercises, assessment details, additional reading, etc.) will be made available at `https://canvas.sussex.ac.uk/courses/23320`.
  - **Module Discussion Forum** at Canvas: Please use this forum to ask any questions, or raise any matters you'd like to discuss in connection with the course.
- Lectures:
  - Weekly lectures will cover key concepts and techniques.
  - Tuesday **9-10**, **Wednesday 12-13** in person
  - This may be supplemented by pre-recorded segments.
- **Office hours**:
  - Tues 10-11, Weds 9-10 (f2f or Zoom).
    Please send me an email to book a slot.

Labs

- Support the lecture content.
  - They will cover the **previous weeks material**.
  - Take much longer than the 1 hour slot.
  - Prepare in advance.
  - Attempt the extensions if you can.
- Labs will take place in person!
- Solutions will be made available for the labs on Fridays and walkthroughs will be posted on Canvas.

Evaluation

Two multiple choice exams (20% total):

- 1 hour long. End of week 4. 10% weighting
- 1 hour long. End of week 78 10% weighting.

Report and assignment (80%):

- A programming project and a report detailing the implementation, critical evaluation, and creative solution of a computer vision application.
- Due for submission on the 11th May. The assignment will be released in week 9.

Source Materials

Some of the good options:

- **Simon J. D. Prince**. Computer Vision: Models, Learning, and Inference. Cambridge University Press. 2012. The online version of the book is available at Simon's homepage: http://www.computervisionmodels.com/. This book has more machine learning view points.
- Richard Szeliski. Computer Vision: Algorithms and Applications. Springer. 2010. The online version of the book is available at Richard's homepage: https://szeliski.org/Book/1stEdition.htm.

Syllabus Overview (Part 1)

- Computer vision in Python using Colaboratory
- Basic image (point/pixel) operations: histogram equalization
- Linear filters and convolution operation
- Edge detection
- Binary image analysis: morphological operation, blobs and regions
- Feature detection and descriptors.
- Camera models and stereo geometry.
- Video processing: Tracking

Syllabus Overview (Part 2)

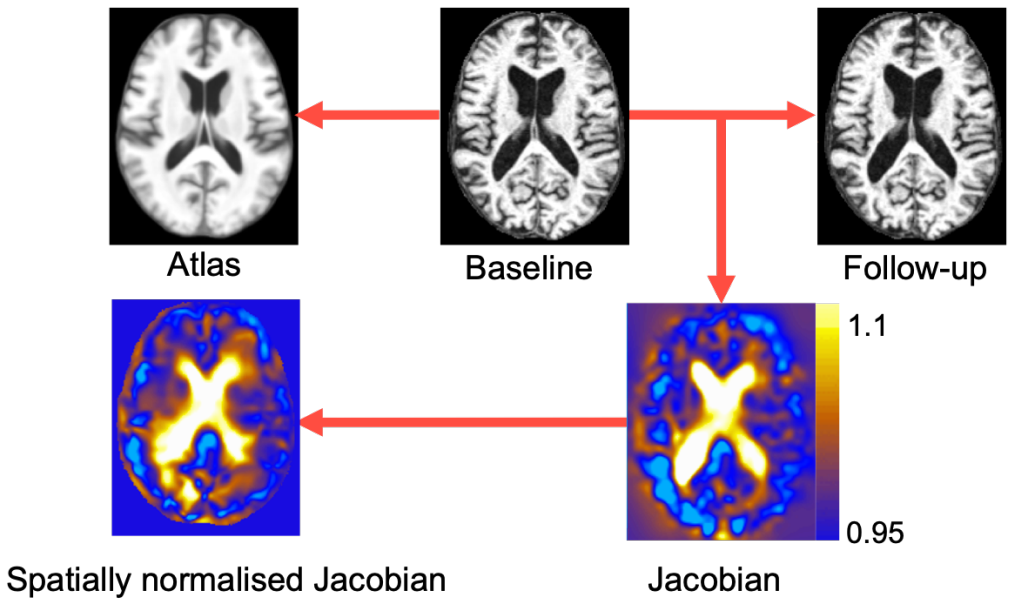
- Image classification
- Zero shot learning
- Object detection
- Face alignment
- Segmentation
- Generative models
- Uncertainty in Computer Vision

Why am I Teaching You Computer Vision?

- I've been working as an academic and industrial researcher in computer vision since I finished my undergraduate.
- I've also built and deployed Computer Vision systems in the real world! I worked for 5 years at Anthropics Technology, a small tech company making intelligent image editing / virtual try-on software.
- I've interviewed many undergraduates, like you, wanting to work in computer vision.
- I've published many paper at top conferences in medical image analysis and computer vision.

Examples of my research topics: Modelling changes in human brain morphology

- Problem: How can we quantify effectiveness of dementia treatments and make accurate diagnoses?
- Solution: Measure and compare changes in brain shape over time in healthy controls and disease patients.



University of Sussex

Examples of my research topics: Generative models

- Problem: Editing high resolution photos.
- Solution: Build a system that learns shape changes.

(a) Input image (previously unseen) (b) User requested edit: "beak larger than head" (c) User requested edit: "beak smaller than head"

Dorta et al., " The GAN that warped", CVPR 2020

University of Sussex

Why Do People Study Computer Vision?

So many reasons!:

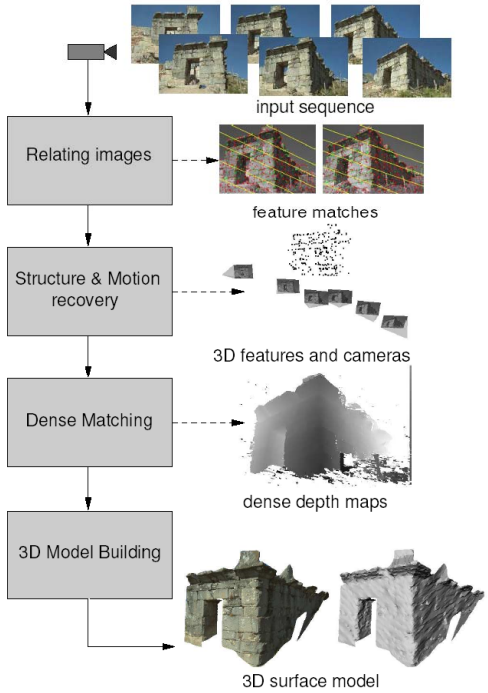
- To relieve, or assist, humans with big data tasks.
- To enable augmented reality applications.
- To allow robots and autonomous vehicles to perceive the world.
- To organise and explore large datasets.
- To build "digital twins" of biology and the world.
- etc.

Vision for measurement

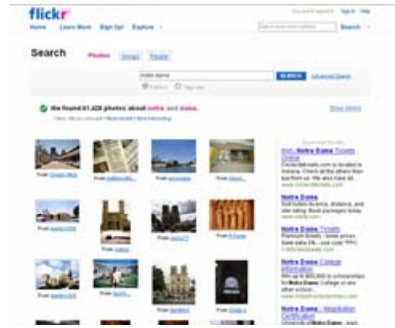
Real-time stereo

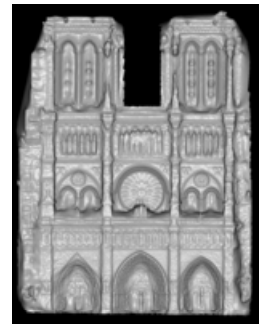
Pollefeys et al.

Structure from motion



Multi-view stereo for community photo collections



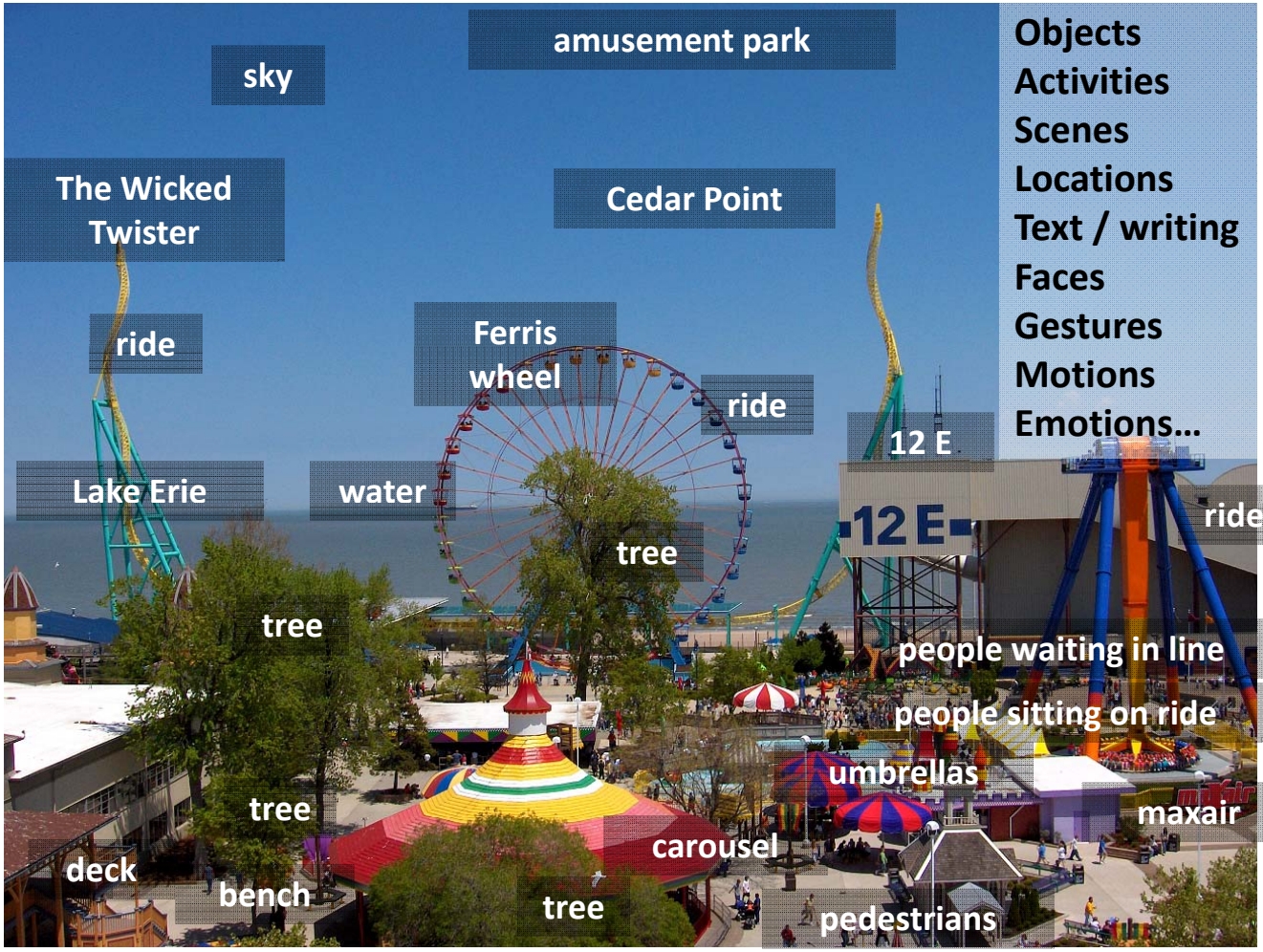


Goesele et al.

Slide credit: L. Lazebnik

University of Sussex

Vision for perception, interpretation

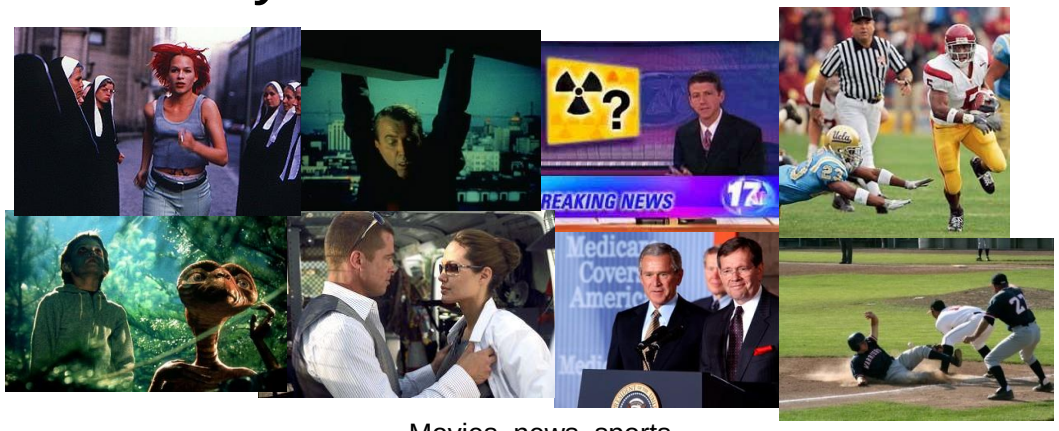


University of Sussex

Why vision?

- Images and video are everywhere!

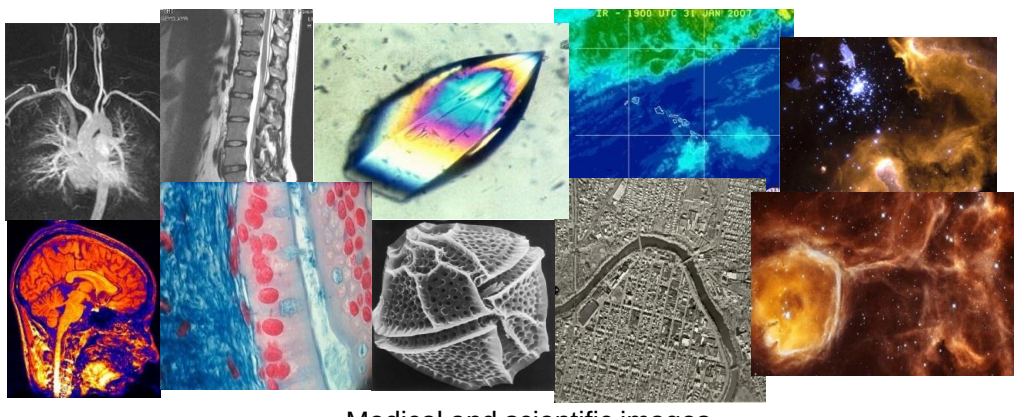
Personal photo albums



Movies, news, sports

Google Image Search Picasa™ flickr webshots picsearch™ YouTube Broadcast Yourself™

Surveillance and security

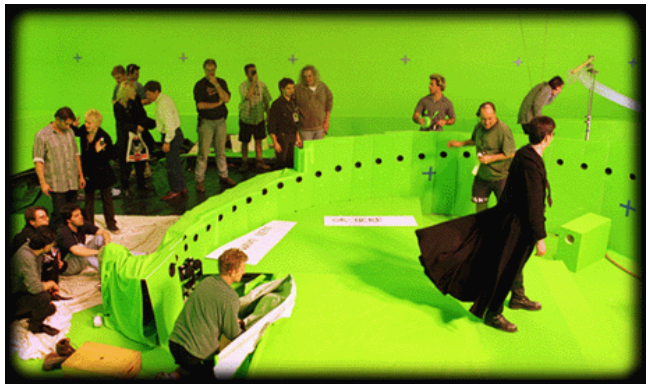


Medical and scientific images

Lana Lazebnik

University of Sussex

Special visual effects

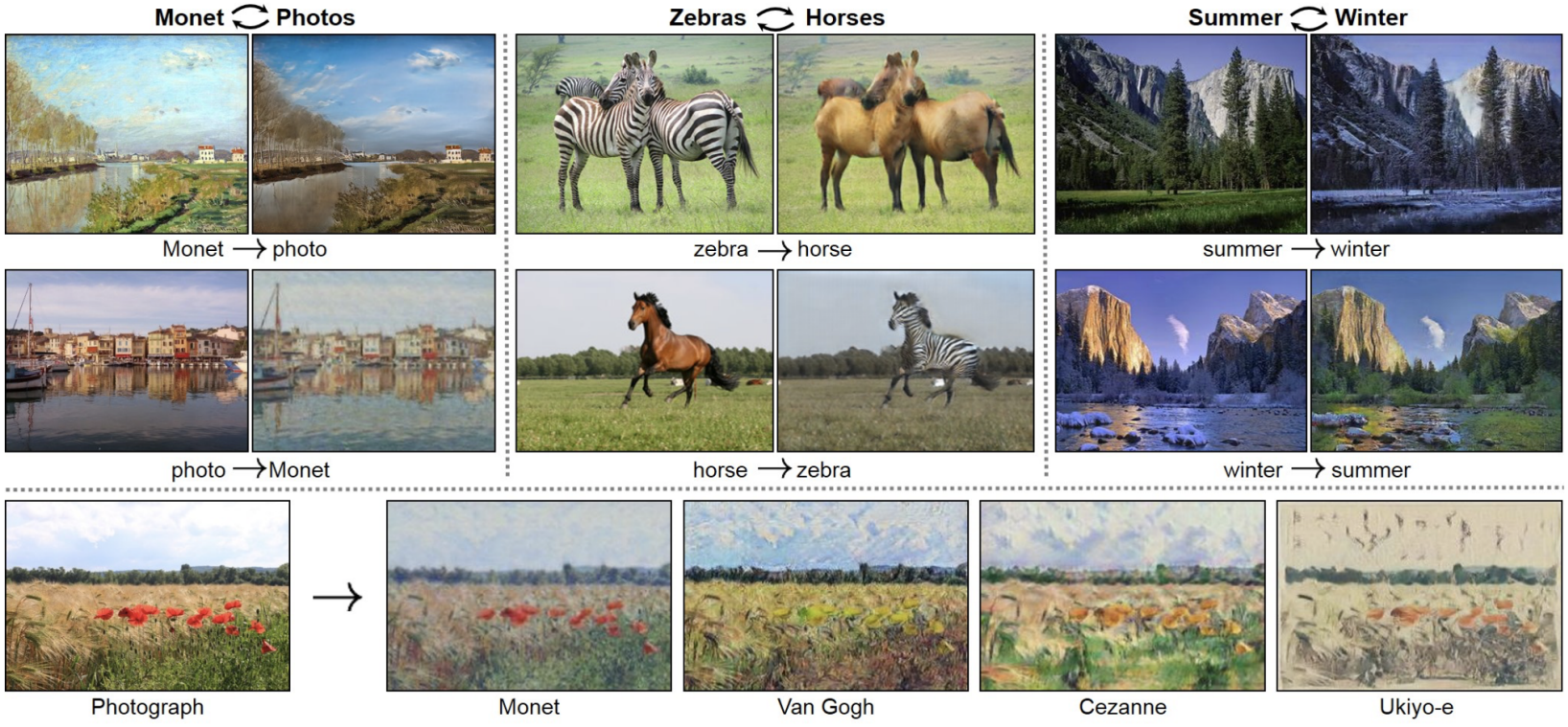


The Matrix

What Dreams May Come

Mocap for *Pirates of the Carribean*, Industrial Light and Magic
Source: S. Seitz

Special Effects with Cycle-Consistent Adversarial Networks



Link: https://junyanz.github.io/CycleGAN/

Source: Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, ICCV, 2017.

University of Sussex

Interactive systems

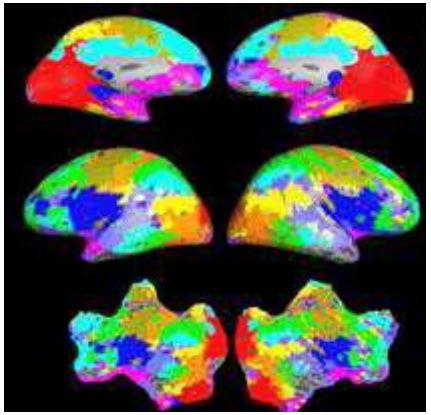
Shotton et al.

Yong Jae Lee

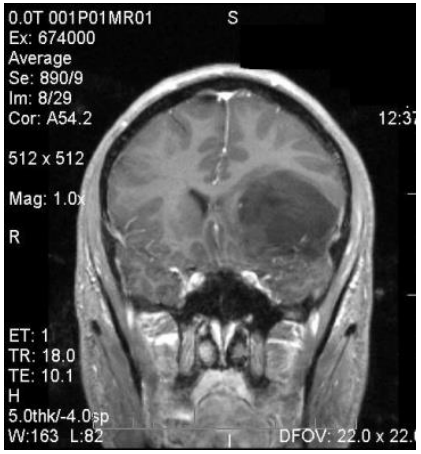
University of Sussex

Vision for medical & neuroimages

Image guided surgery
MIT AI Vision Group



fMRI data
Golland et al.

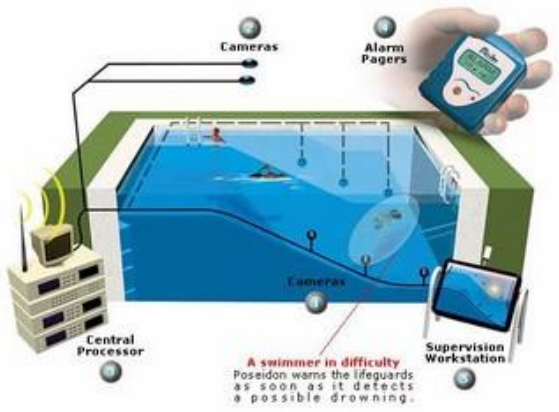


Kristen Grauman

University of Sussex

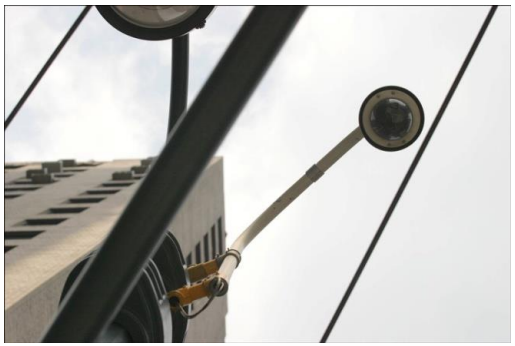
Safety & security

Navigation,
driver safety



Monitoring pool
(Poseidon)

Pedestrian detection
MERL, Viola et al.



Surveillance

Obstacles?

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

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".

Kristen Grauman

Ethical Questions

- In recent years issues relating to bias and unexplainable errors have become more important.
  - In a large part because these systems can be usefully deployed in certain contexts.
- It's important to consider how and why a method might not work as intended, and then work to mitigate this.

Why is Computer Vision Difficult?

- What major challenges can you think of?
  - It's an ill-posed problem, we receive a 2D projection of a 3D world.
  - We cannot "invert" the image formation process to create a 3D world from a single 2D image.
  - This (and sensor accuracy) leads to a loss of information.

Data variability

- Imagine you are tasked with building a detection system for koalas?

- What type of variability might you get between your images?
- Answers in the Poll!
  https://www.pollev.com/ivorsimpson490

Challenges: many nuisance parameters

Illumination

Object pose

Clutter

Occlusions

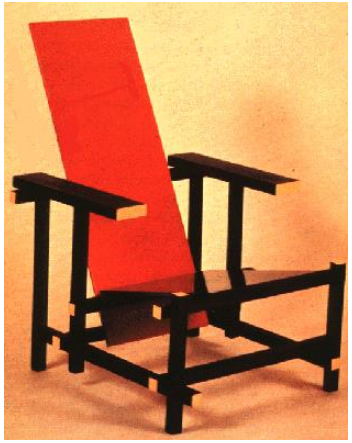
Intra-class appearance

Viewpoint

Kristen Grauman

Challenges: intra-class variation



slide credit: Fei-Fei, Fergus & Torralba

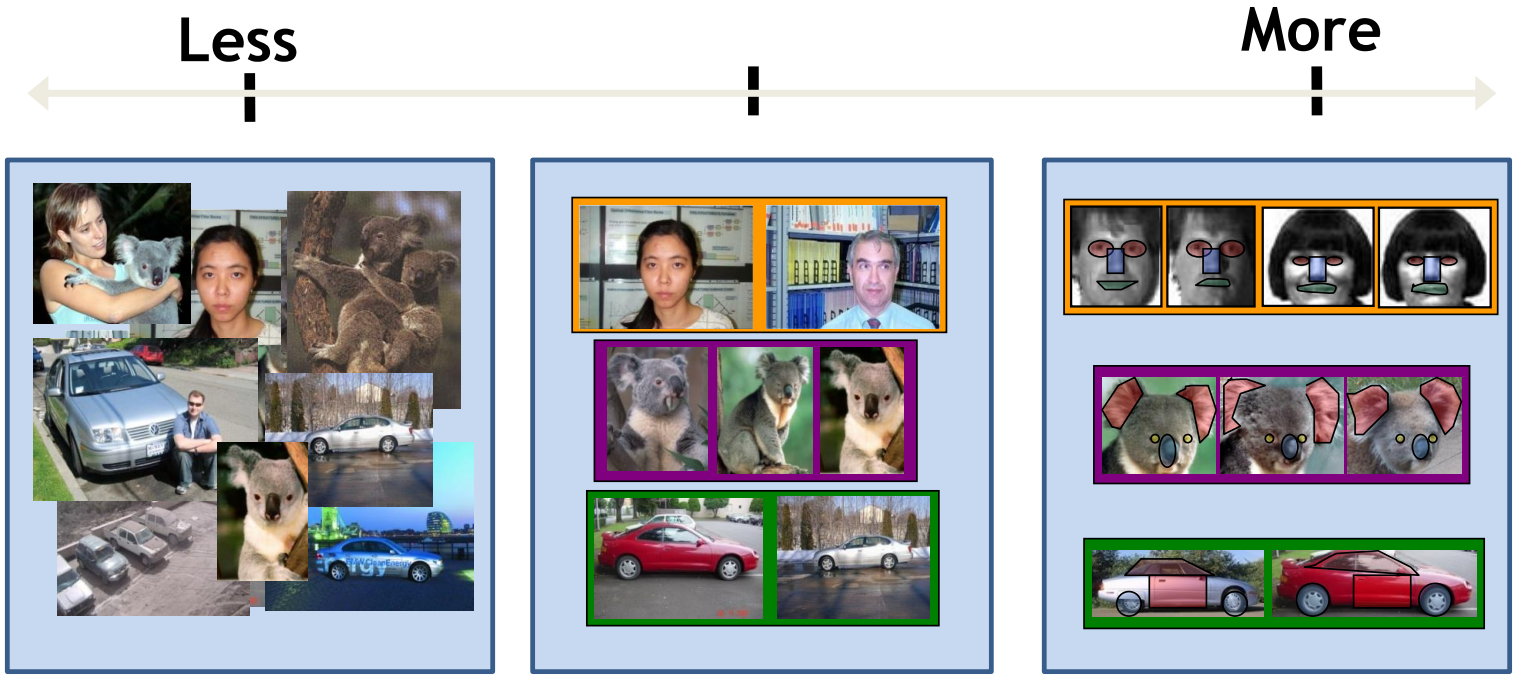
Challenges: importance of context

slide credit: Fei-Fei, Fergus & Torralba

Challenges: Complexity

- Millions of pixels in each image.
  - with complicated correlations between them.
- 3000-30,000 human recognisable object categories.
- More than 10 billion images indexed by google image search.
- 1.52 billion smartphones sold in 2019.
- It is estimated that half of the cerebral cortex in primates is devoted to processing visual information [Felleman and van Essen 1991]

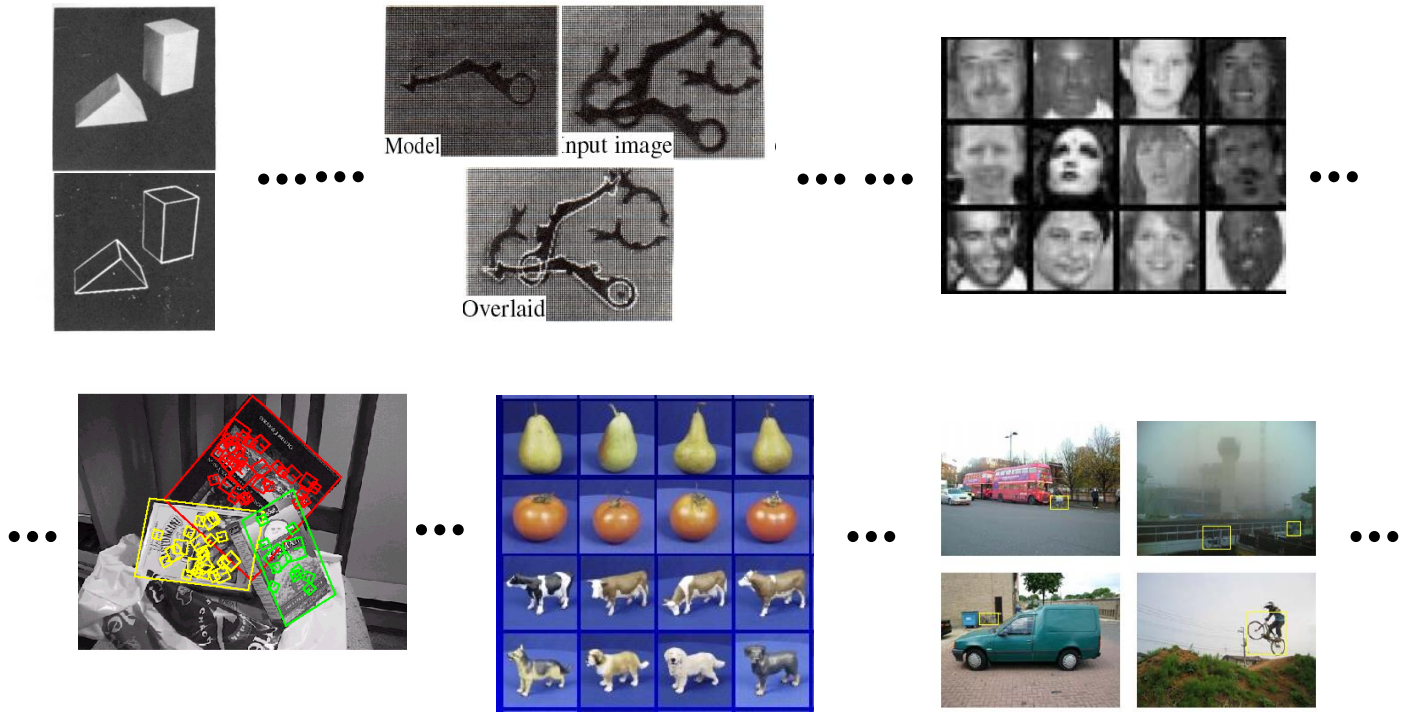
Challenges: Limited supervision



Kristen Grauman

University of Sussex

- Ok, clearly the vision problem is deep and challenging…time to give up?
- Active research area with exciting progress!



Kristen Grauman

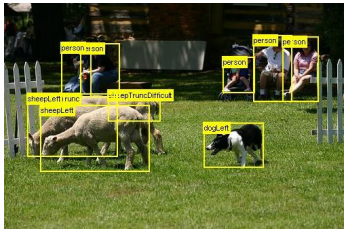
University of Sussex

Datasets today



ImageNet:
22k categories, 14mil images


Microsoft COCO:
70 categories, 300k images

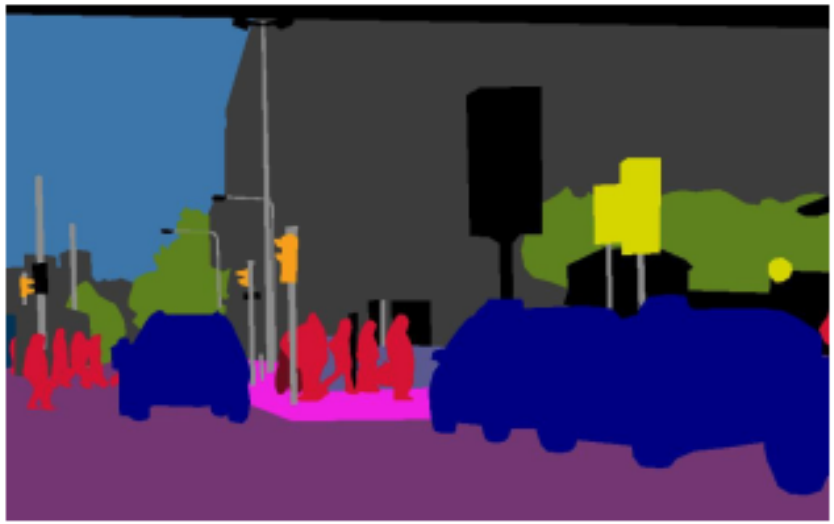

PASCAL:
20 categories, 12k images



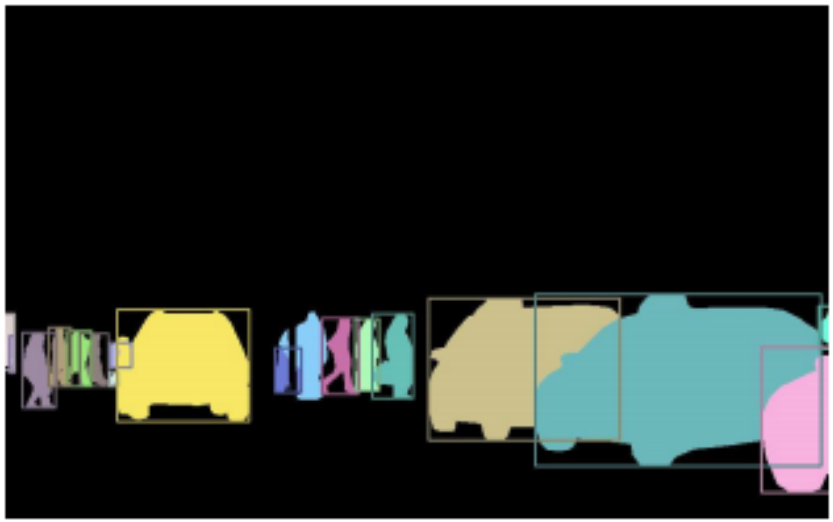
SUN:
5k categories, 130k images

New Tasks: Panoptic Segmentation

(a) image



(b) semantic segmentation

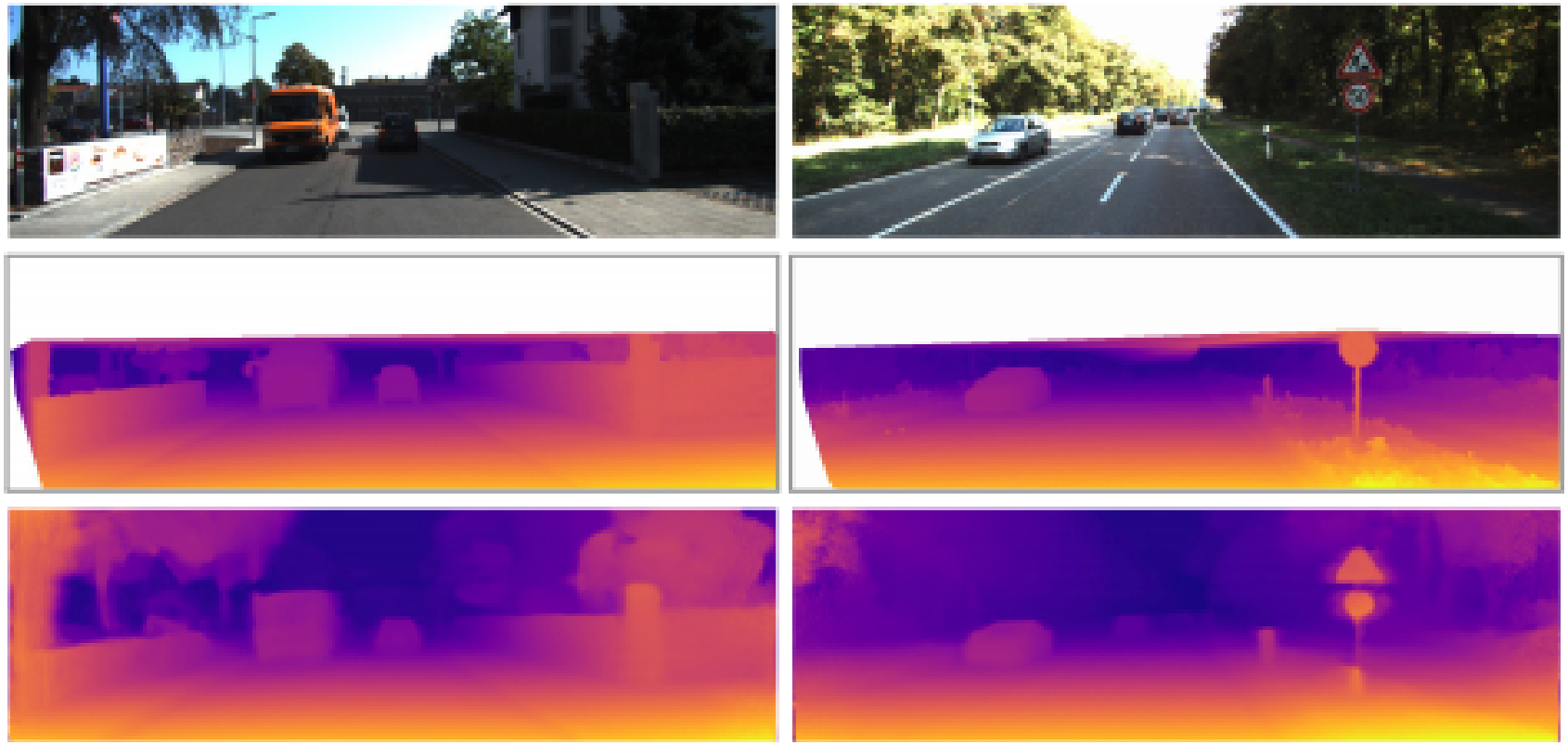


(c) instance segmentation

(d) panoptic segmentation

Kirillov et al "Panoptic segmentation." CVPR 2019.

New Tasks: Monocular Depth estimation



Godard et al "Unsupervised Monocular Depth Estimation with Left-Right Consistency." CVPR 2018.

What will be on Wednesday?

- Introduction to Python and Colaboratory (https://colab.research.google.com) for Computer Vision.
- Remember, no labs this week!

Temporary page!

LaTeX was unable to guess the total number of pages correctly. A
there was some unprocessed data that should have been added
the final page this extra page has been added to receive it.
If you rerun the document (without altering it) this surplus page
will go away, because LaTeX now knows how many pages to expe
for this document.