Johns Hopkins Engineering

Computer Vision

Course Overview

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Outline

- Course organization
- What is computer vision?
- Course overview

Class Organization

Course Information

- Introductory course
  - Audience: undergraduate and graduate students
- Required background
  - Programming fundamentals and Python experience
  - Linear algebra
  - Basic probability
  - Calculus

Class Instructional Model

- This course will aim to be consistent with asynchronous instruction with synchronous "deep dive" lectures
  - Slides will be posted
  - Lectures will be posted in short installments
  - Python notebooks will augment both

- Lectures will focus on answering questions and solving problems
  - This puts the onus on you to drive lecture content!

Zoom Intro

- We will use Zoom for all interactive classes
  - All lectures will be recorded
  - Use the chat window to ask questions
  - Inappropriate posts or other class disruptions will lead to removal from live Zoom lectures
- I'll also occasionally make use of polls
  - Let's try the first one …
- Please do not post the Zoom link or password publically

Course Logistics

- Organization
  - Three assignments, "Midterm" Exam, Final Project
  - Weekly in-class quizzes
  - Weekly python notebooks

- Office hours:
  - Hager: 5-6PM Tuesday/Wednesday (signup sheet on Piazza)
  - Chen: 5-6PM Monday/Friday and by appointment

- Course updates:
  - Announcements/lectures/support material/assignments: Piazza
  - Quizzes: JHU Blackboard

Course Information

- Grading
  - 30% Homework
  - 20% Midterm Exam
  - 30% Project
  - 10% Quizzes (Assigned weekly; drop the lowest three)
  - 10% Python exercises (Assigned weekly; drop the lowest three)

- References:
  - Optional: R. Szeliski, Computer Vision, Springer (2011).
  - Optional: Forsyth and Ponce, Computer Vision: A Modern Approach, Prentice Hall (2002).
  - Optional: Goodfellow, Bengio, Courville, *Deep Learning*, MIT Press, 2016.
  - Also online at https://www.deeplearningbook.org
  - Optional: Hartley and Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press (2010).

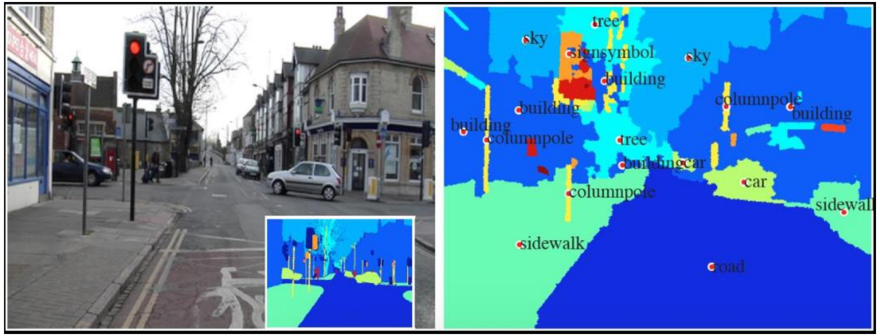
Programming Resources

- Python tools
  - Numpy
  - Pytorch
  - Matplotlib
  - OpenCV
- Google Colaboratory: https://colab.research.google.com/
- Jupyter: https://jupyter.org

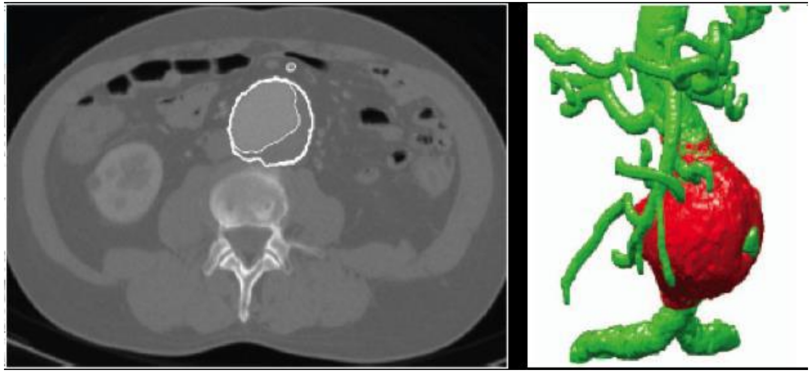
What is Computer Vision?

Computer Vision

- Recovering properties from the real-world using images
  - Examples:
    - Detecting/recognizing objects in a scene
    - Reconstructing a 3D model of the imaged buildings
- Overlaps with a few other areas from Computer Science, e.g.
  - Image processing
  - Machine learning/pattern recognition
  - Graphics



Micusik et al.



Demirci et al.

But what really is computer vision?

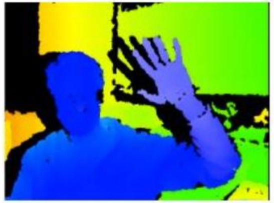
- Vision is:
  - automating human visual processes
  - an information processing task
  - inverting image formation
  - inverse graphics
  - really useful!

Image modalities: More than "just" light

Digital cameras

Infrared cameras

Time-of-flight cameras

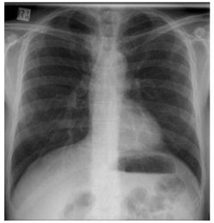


kinetic

Image modalities: More than "just" light

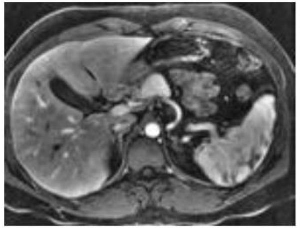

ultrasound


endoscopy

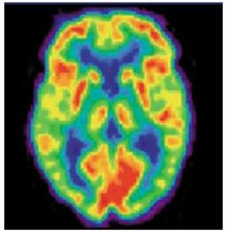



X-rays


Computed tomography

Magnetic resonance

Positron emission tomography

Application examples

Visual surveillance

Character recognition

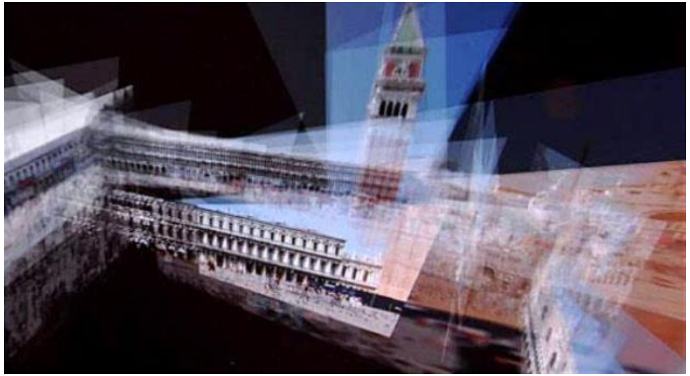
Image search

Sportvision

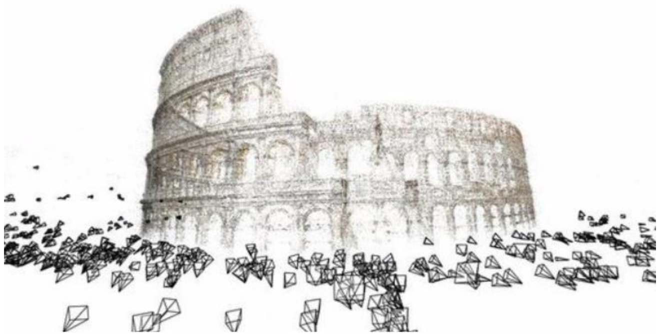
Fingerprint scanners

Robot control

Application examples



Photosynth



Multiple view reconstruction (Univ. of Washington)

Medical augmented reality (MIT)

Special effects

Vision is Easy, Right?

The Human Visual System has no problem interpreting subtle variations and correctly segmenting the object from the background

Vision is Easy, Right?

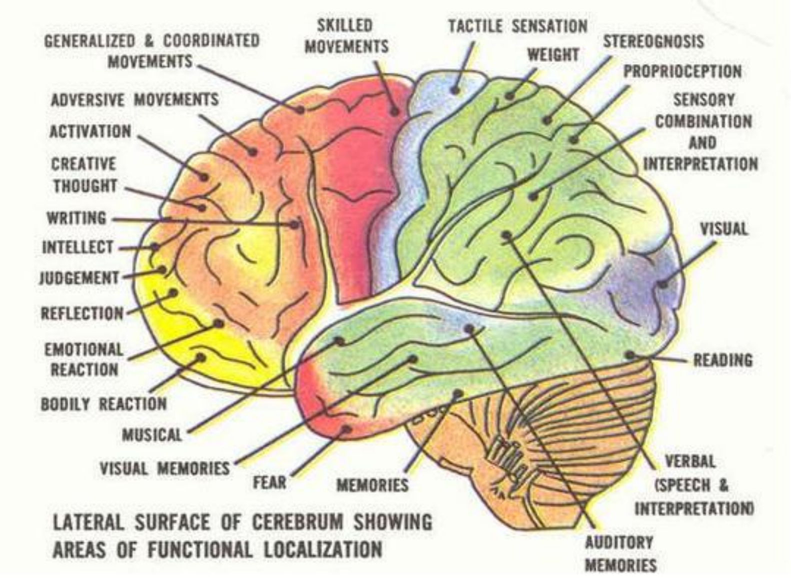
- The Human Visual System has no problem interpreting subtle variations and correctly segmenting the object from the background
- What can we infer?
  - It's a flower
  - The flower is pink
  - There are water droplets on the flower. Perhaps it rained earlier
  - It's daytime and sunny
  - ...

So why is it so hard for computers?

- The Human Visual System has no problem interpreting subtle variations and correctly segmenting the object from the background
- What can the machine infer?
  - It’s a flower.
  - …
  - That’s mostly it!

Why is vision challenging?

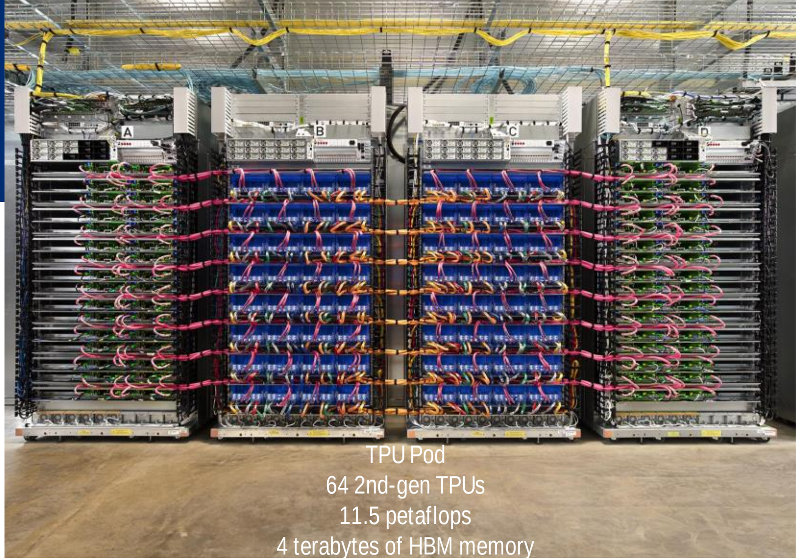
- When you see an image, you see:
  - The objects in the image
  - The context of the scene
  - Your prior knowledge and experiences automatically apply to what you see to deduce meaning
  - …and much much more…
- When a machine sees an image, it sees:
  - Numbers
  - …and…
  - …well…
  - That's it



- Approx. 1011 Neurons
- Approx. 1014 Synapses
- Firing rates 100-1000 Hz
- Asynchronous, distributed
- Consumes 20W
- Diverse (everyone is different!)

Why is vision challenging?

- When you see an image, you see:
  - The objects in the image
  - The context of the scene
  - Your prior knowledge and experiences automatically apply to what you see to deduce meaning
  - …and much much more…
- When a machine sees an image, it sees:
  - Numbers
  - …and…
  - …well…
  - That’s it

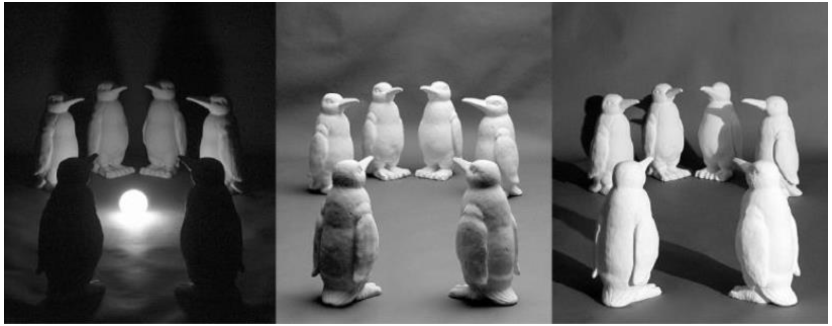


- 11.5x 1015 floating point ops
- 4 x 1012 memory locations
- Gigahertz clock
- Synchronous
- Consumes hundreds of kilowatts
- Each is the same

So why is vision challenging?

- Basically we are solving an "Inverse Problem"
  - Recover some unknowns given insufficient information to fully specify the solution
- So what do we do?
  - Use physics-based and probabilistic models to disambiguate between potential solutions
  - Let's be honest – we also use a bunch of "tricks"!

Why is vision challenging…?



Illumination and viewpoint

Intra-class variation

Scale and depth

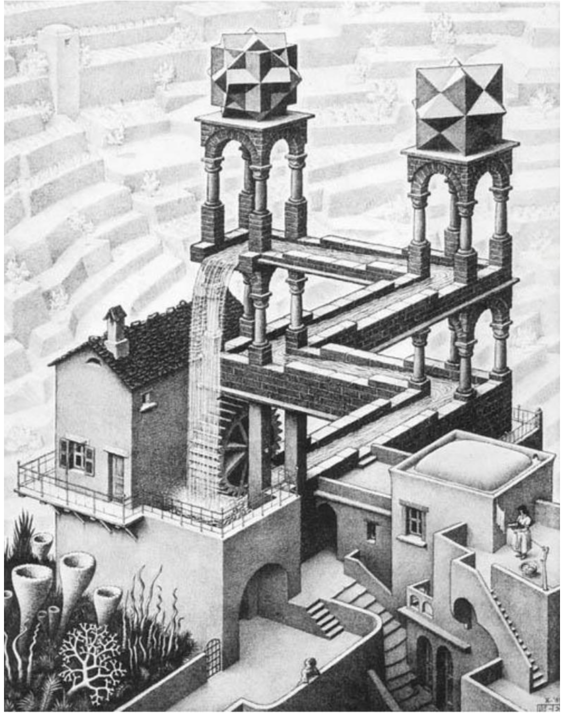
Why is vision challenging…?

Shadows

Clutter

Why is vision challenging…?

Occlusions



Illusions

Why is vision challenging…?

Why is vision challenging...?

In short…

…additional knowledge/constraints required to remove ambiguities:

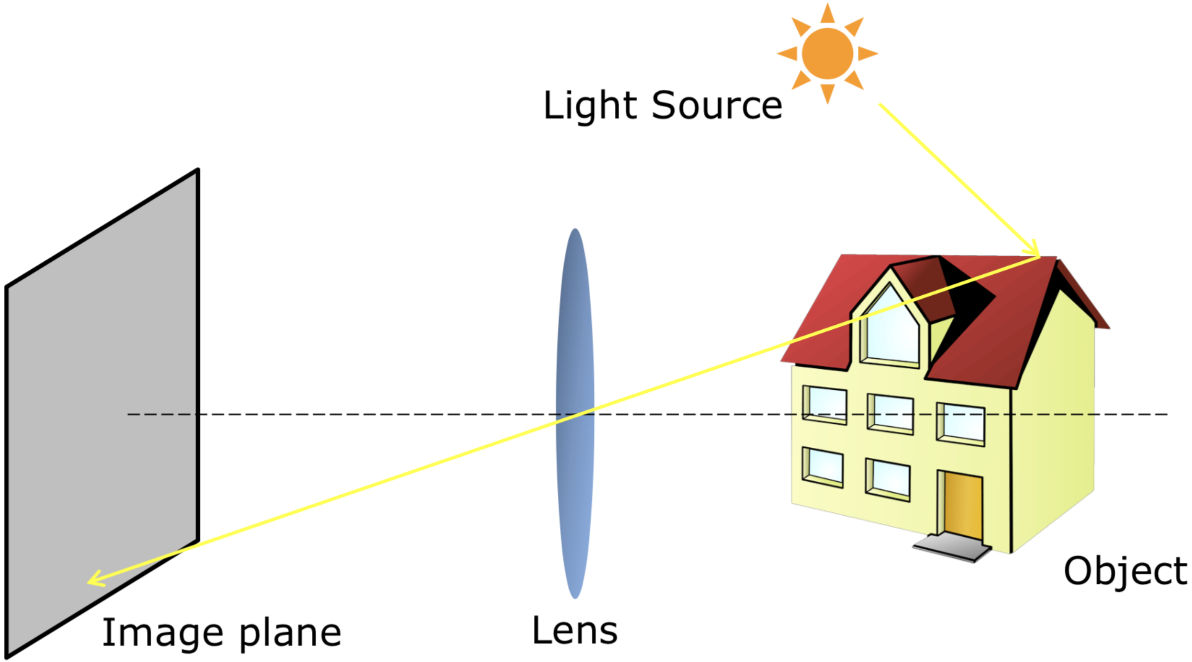
- A-prior information about the scene, context and/or the images
- Additional images from different viewpoints

This is where the truly hard aspects of vision and AI meet

Class Overview

Image formation and optics

Where do in



Projection of 3D world on a 2D Plane

Image processing

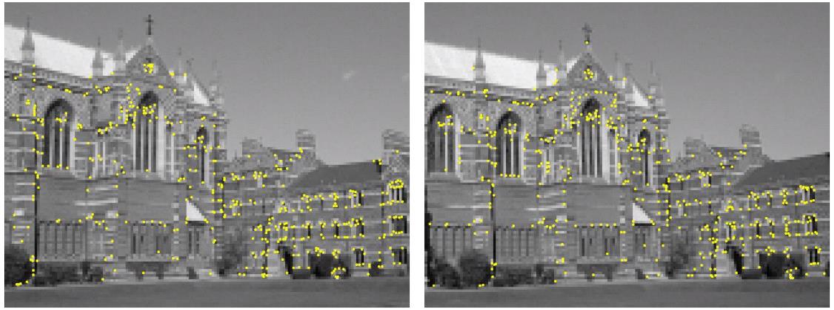
- Basic operations, linear filters, non-linear filters

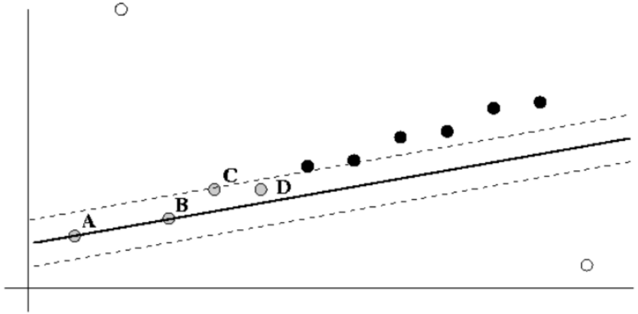




Image processing

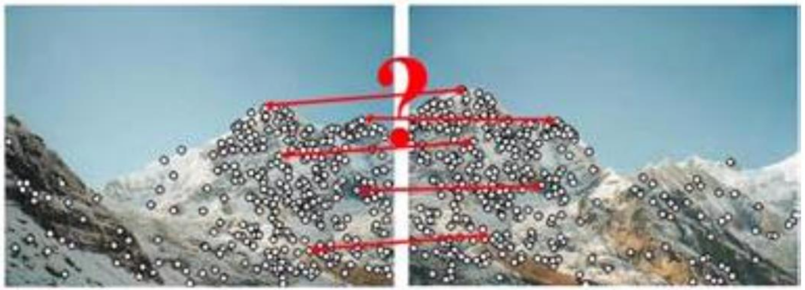
- Edge detection, corner detection, line/circle detection/fitting

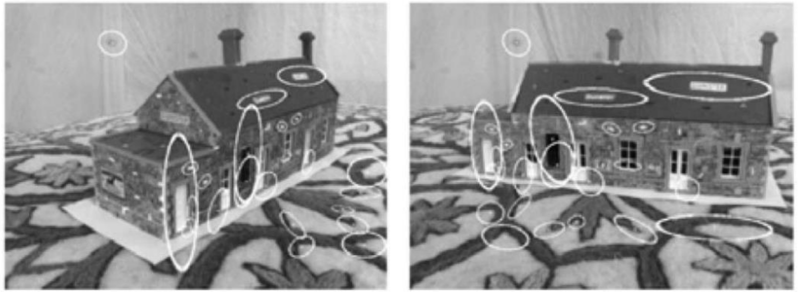




Feature descriptors

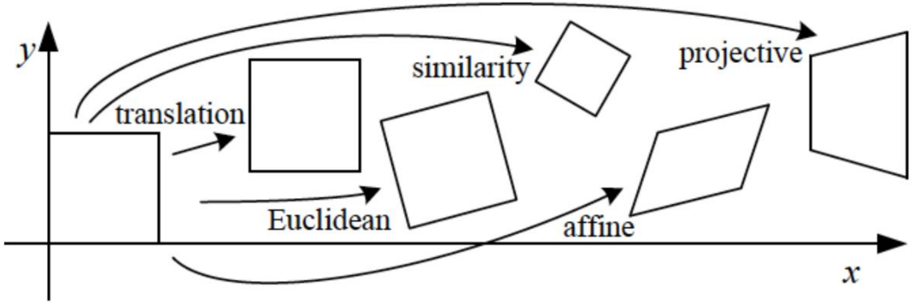
- Feature detection and matching, SIFT, SURF





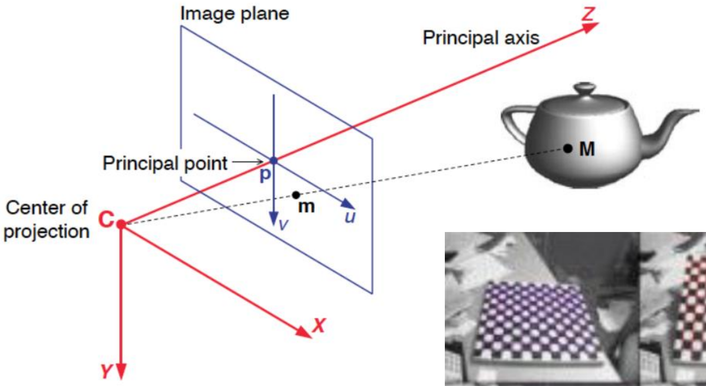
Single view geometry

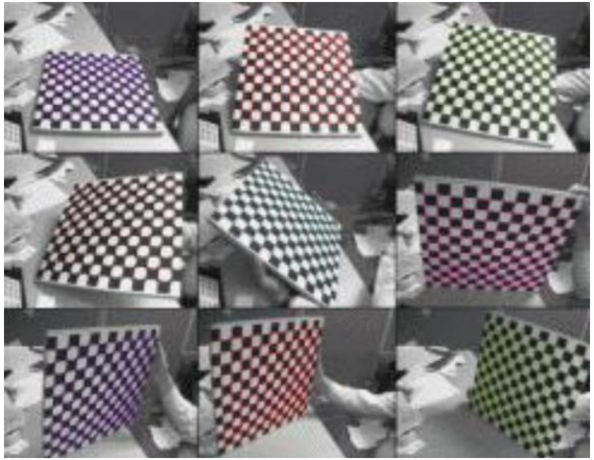
- 2D transformations, homographies, DLT, planar rectification

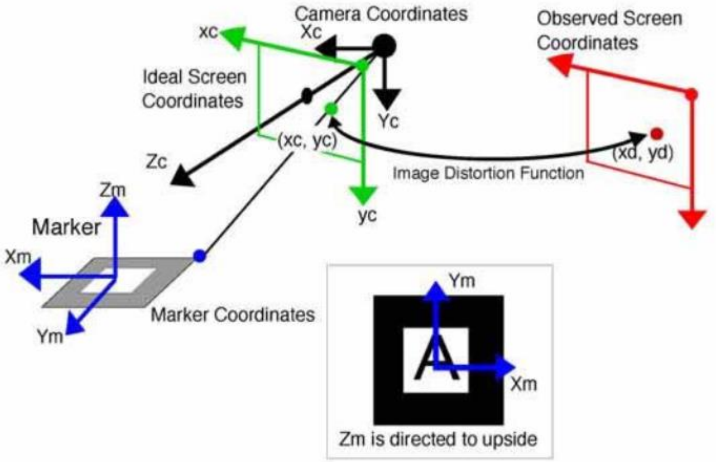


Single view geometry

- Camera models, calibration, pose estimation, mosaicing

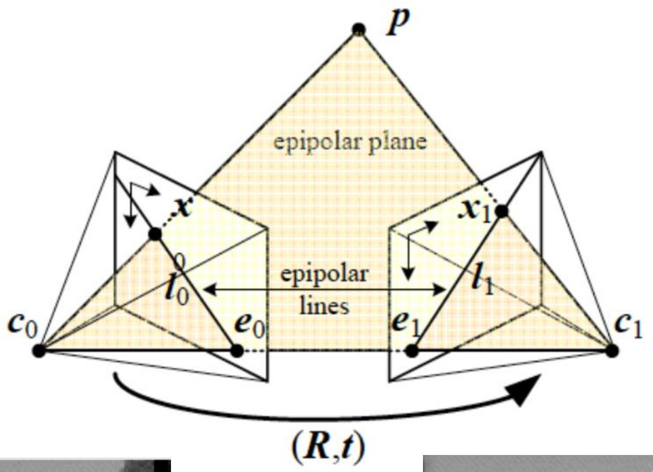


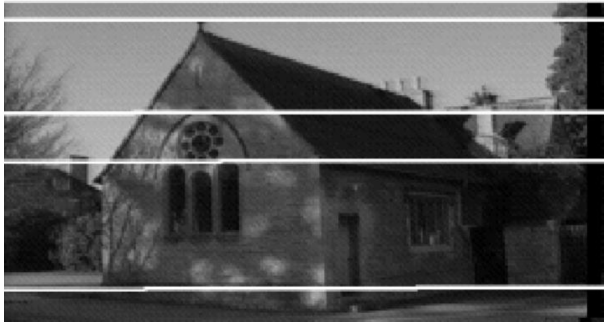


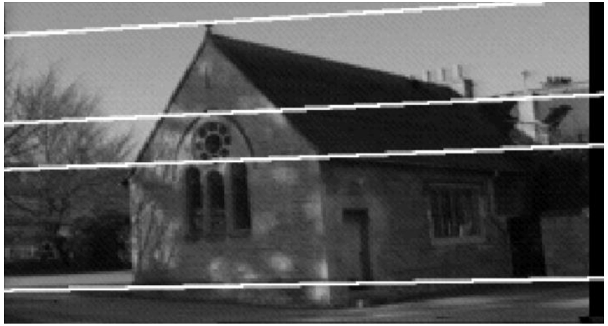


Two-view geometry

- Triangulation, stereo calibration, epipolar geometry

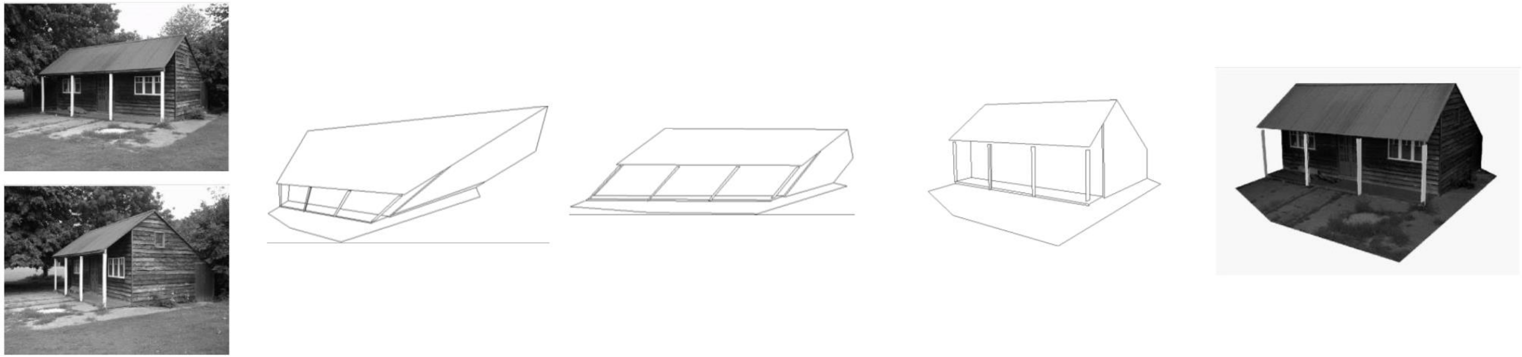






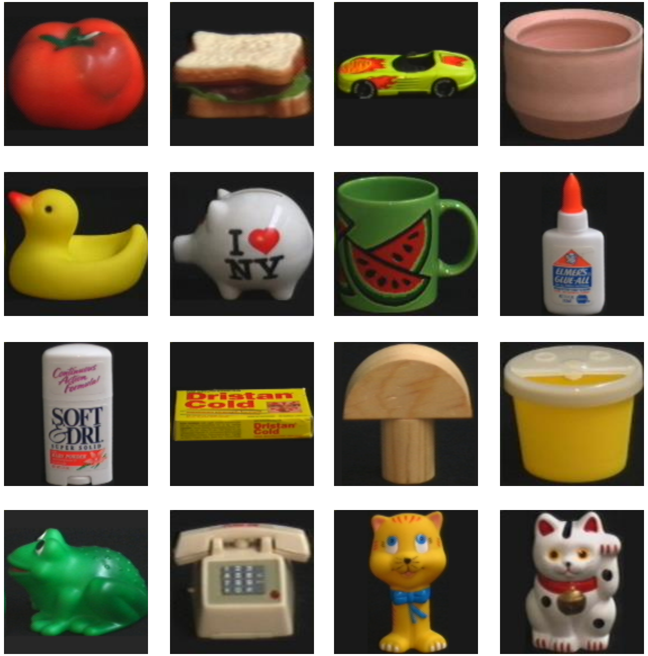
Two-view geometry

- Stratified (sparse) reconstruction, image rectification, dense reconstruction



Appearance matching

- Shape vs. Appearance, 3D object representation, Principal Components and Dimensionality Reduction



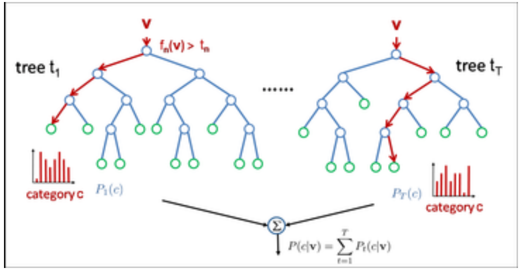
100 Objects Database

Recognition

Object Recognition: Representation and model

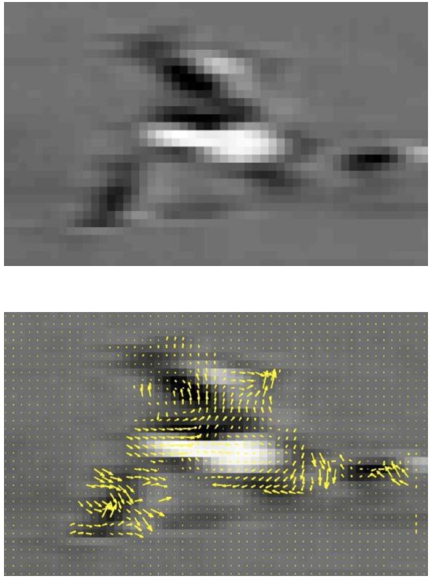
- Bag-of-words, Pedestrian Detection, and Random Forests

Object → Bag of ‘words’

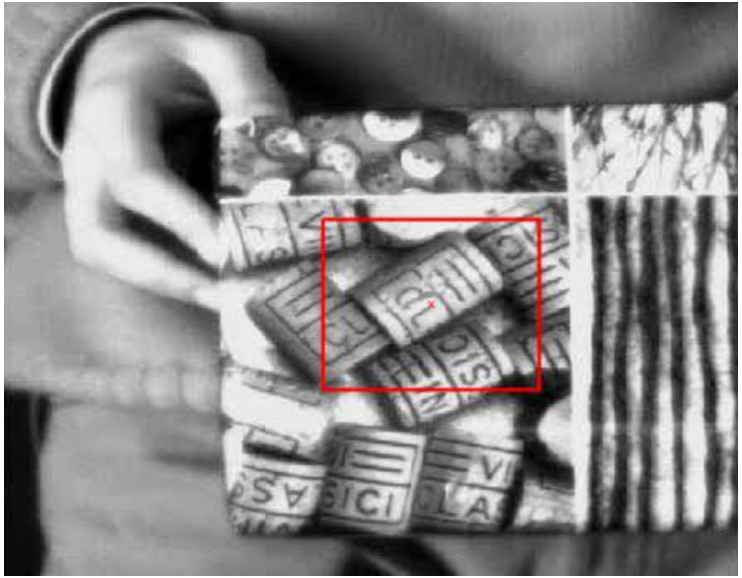


Motion and tracking

- Optical flow, feature tracking, applications, SSD template tracking

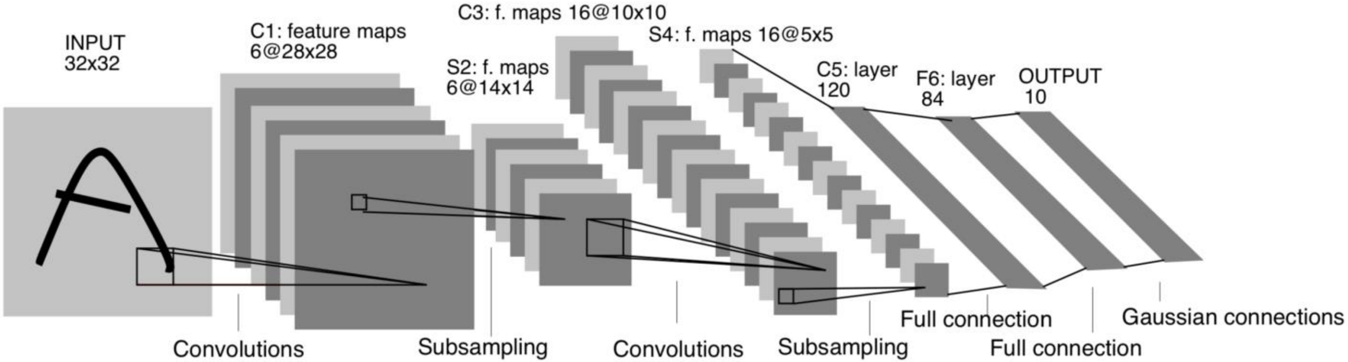


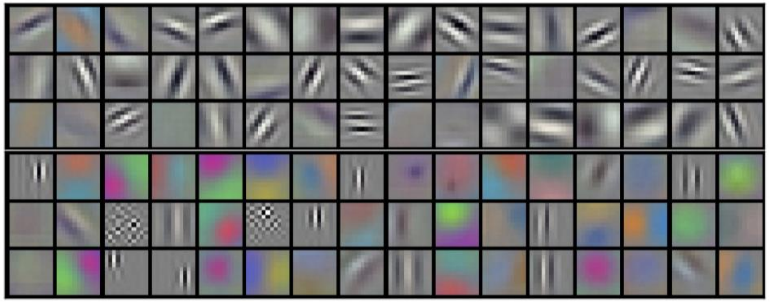
Efros et al.



INRIA

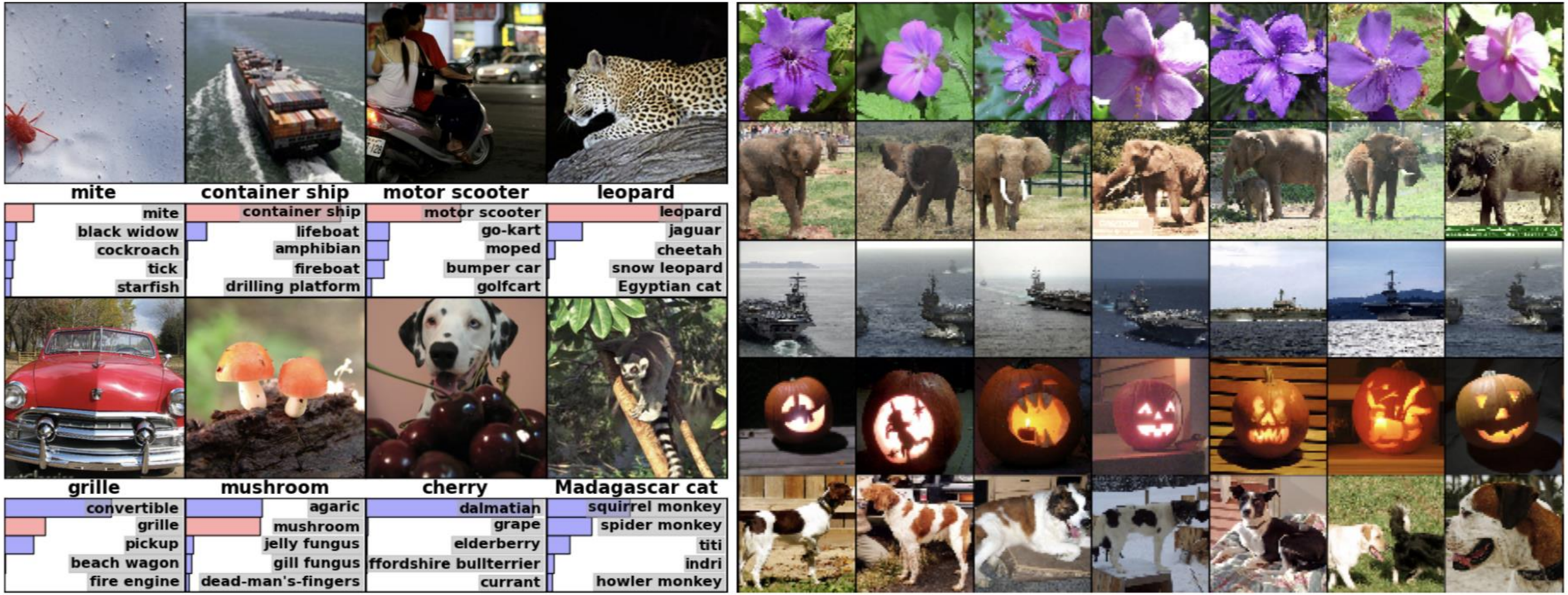
Deep learning basics





LeCun et al., Krizhevsky et al.

Image classification

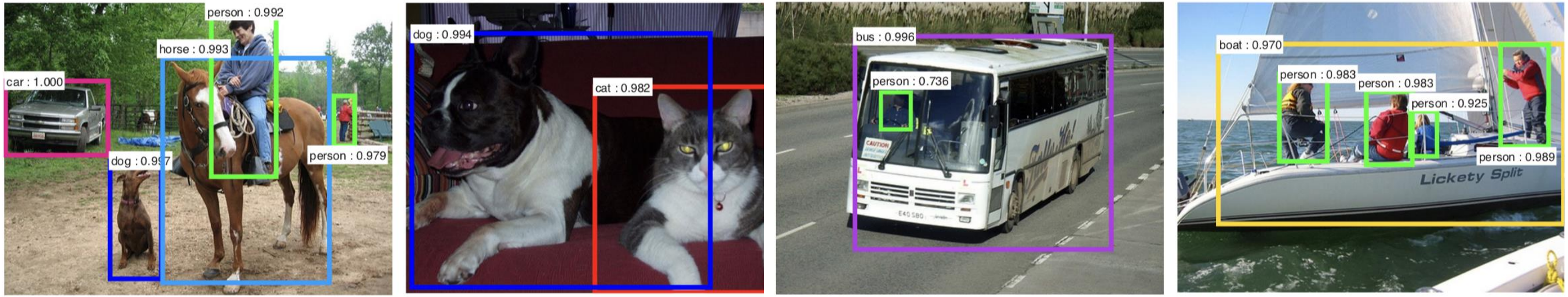


Krizhevsky et al.

Semantic segmentation

Chen et al.

Object detection

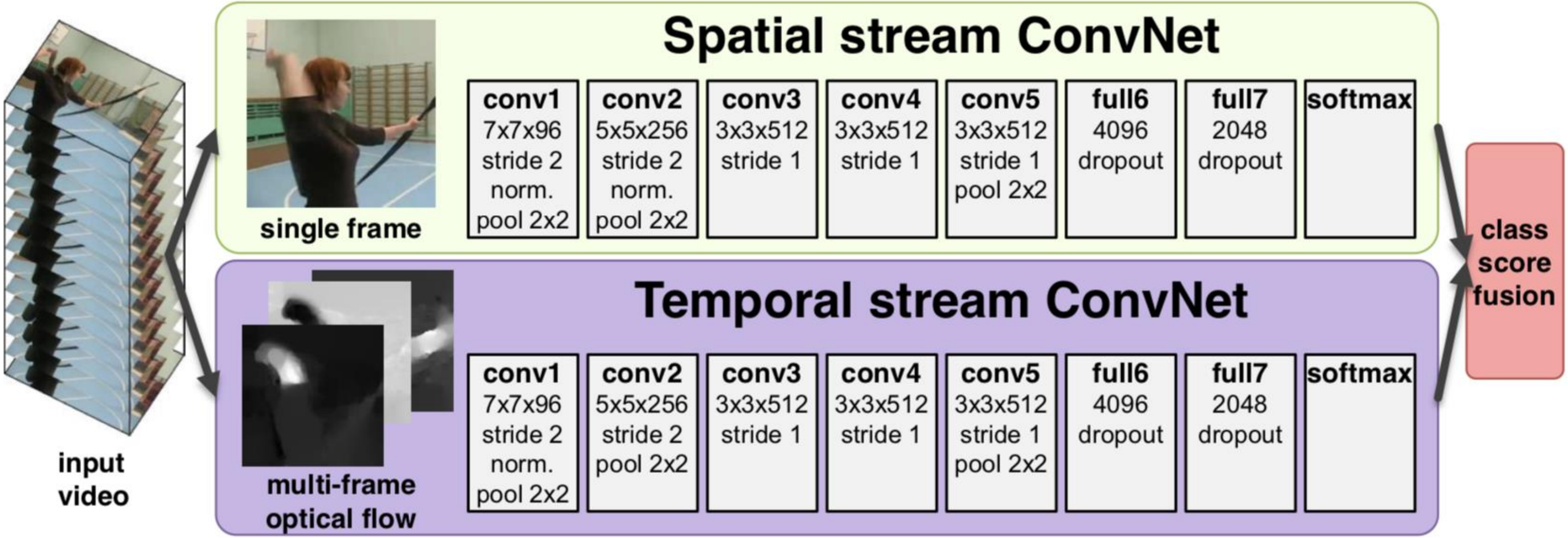


Ren et al., He et al.

Human pose estimation

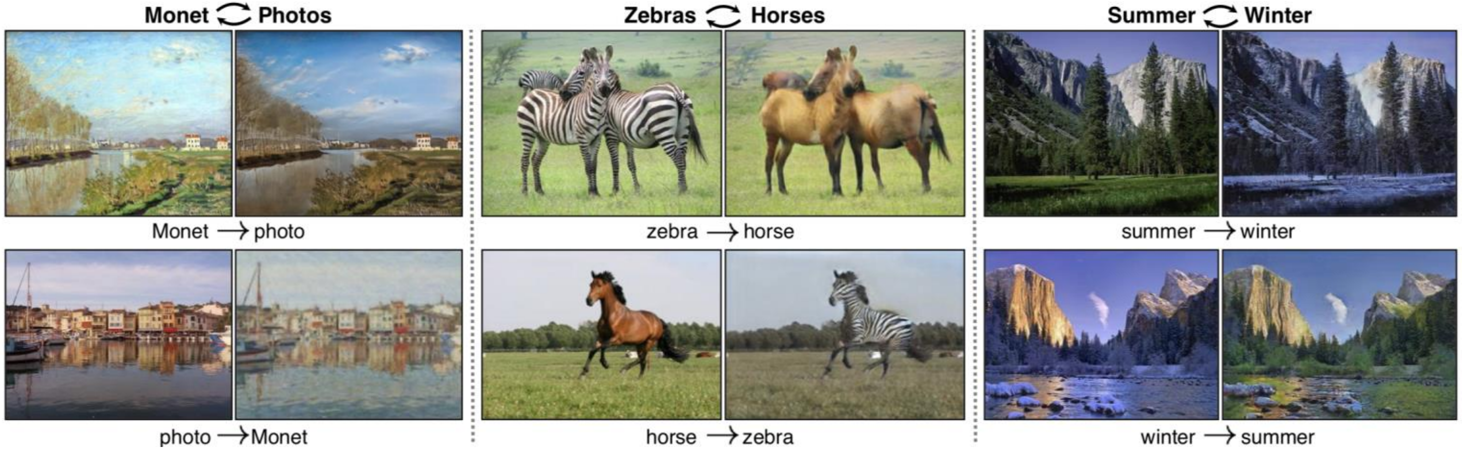
Newell et al., Cao et al.

Action recognition



Simonyan et al.

Style transfer and GAN



Gatys et al., Zhu et al.

What is next?

- This week:
  - Images and image formation (three minilectures)
- Next week
  - Basic image operations
  - Linear and non-linear filtering

- Please check:
  - Access to Piazza
  - Access to Blackboard
  - Access to Python

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING