Overview of Robot Perception

Prof. Roberto Martín-Martín

Fall 2022

Logistics

Office Hours

Instructor: 5-6pm Tuesdays (in person) or by appointment (in person or Zoom)

TAs:

Yifeng: Thursdays 10:30-11:30 am

Jeff: Thursdays 5:30-6:30 pm

Presentation Sign-Up: Deadline Tomorrow (EOD)

First three abstracts due: Monday 9:59pm (AlexNet, Mask-RCNN, YOLO)

On the abstracts

- What should be contained in the abstract?
  - 1-2 sentences describing the problem
  - 1-2 sentences explaining why the state-of-the-art is not enough for this, why it fails
  - 1-2 sentences explaining the clever idea of this paper
  - 1-2 sentences explaining how the idea is implemented
  - 1 sentence about the experimental evaluation
- The entire abstract should be 5-7 sentences long. Be concise.
- We will run a plagiarism software on the abstracts to compare to the original abstract.

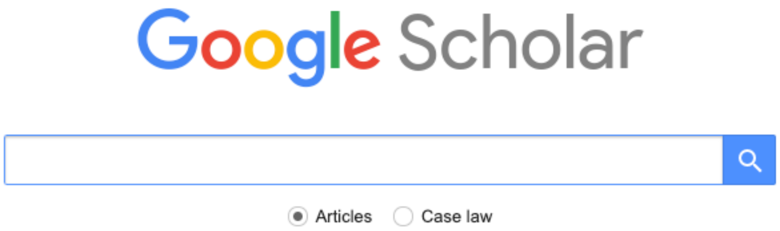
On the abstracts

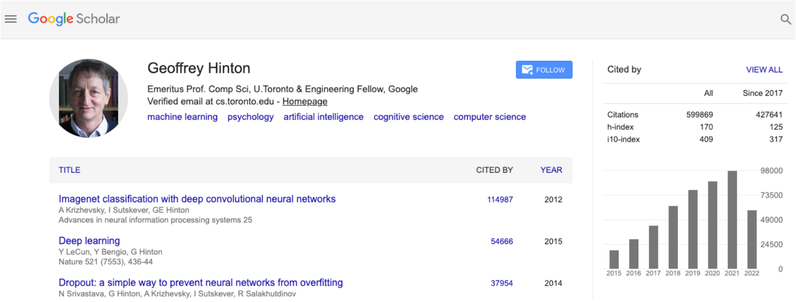
AlexNet:

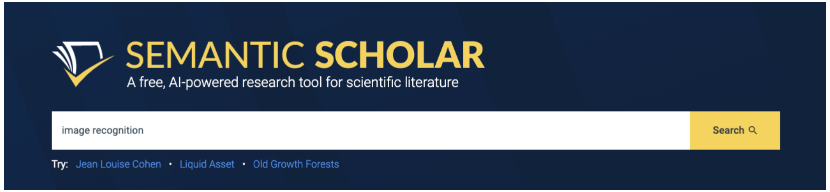
The paper presents a learning algorithm for image classification: assigning semantic labels to images based on their content.
Prior work failed because they applied hardcoded features to represent each image, and these features were not optimal, and thus the methods were not able to improve performance, or they used learned features but trained on few images, because the models didn't have enough capacity (trainable parameters).
In this paper, the authors proposed to map directly input images to the right label with a very large learnable model. They do not hardcode features, instead they propose to learn the most optimal features to represent the image by learning a large capacity (many learnable parameters) model based on a large amount of training data.
They implement their idea with a deep artificial neural network, trained with backpropagation using labeled images in ImageNet.
Their results were a large leap over all previous image recognition methods and started the Deep Learning era.

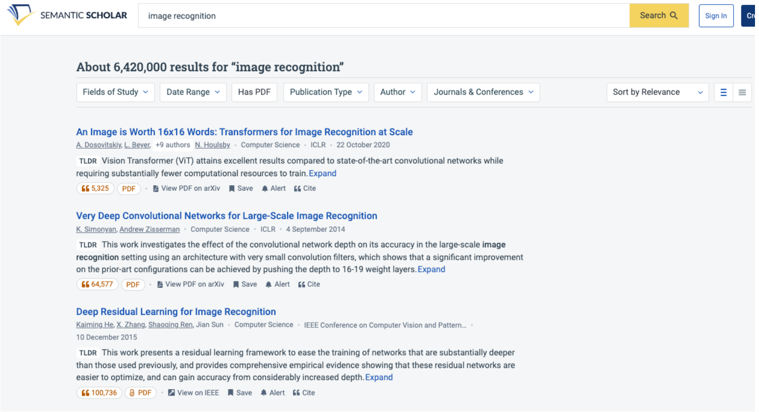
- 1-2 sentences describing the problem
- 1-2 sentences explaining why the state-of-the-art is not enough for this, why it fails
- 1-2 sentences explaining the clever idea of this paper
- 1-2 sentences explaining how the idea is implemented
- 1 sentence about the experimental evaluation

On reading scientific text









What is your background?

- Machine learning?
- Deep Neural Networks?
- Computer Vision?

Today’s Agenda

- What is Robot Perception?
- Robot Vision vs. Computer Vision
- Landscape of Robot Perception
- Quick Review
  - Deep Learning (if time permits)
  - Image formation and projective geometry (if time permits)

What is Robot Perception?

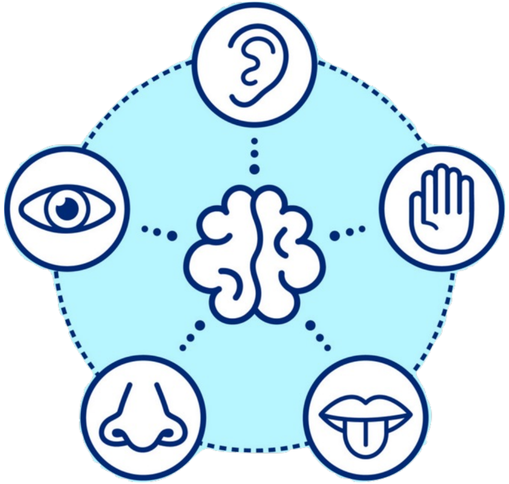
Techniques that allow robots to make sense of the unstructured real world extracting information from noisy sensor signals

A robot may need to perceive…

- Task-relevant information of objects and scene
- The progress and result of its own actions, that may lead to failure
- Environment dynamics and other agents

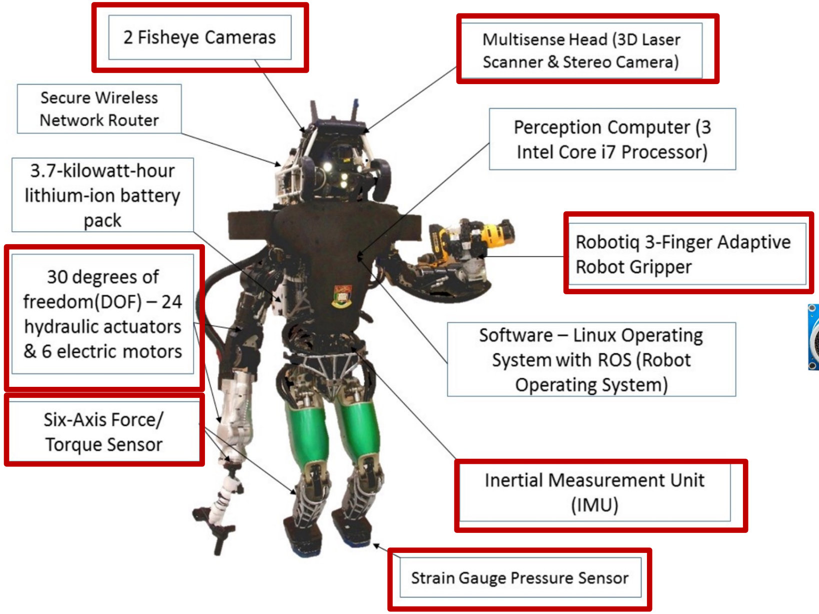
Robotic Sensors

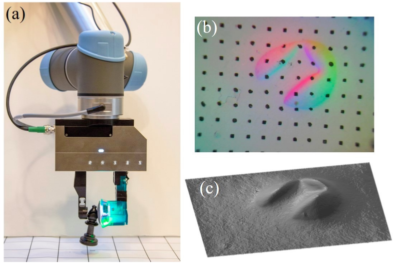
Observing the physical world through multimodal senses



Robotic Sensors

Observing the physical world through multimodal senses

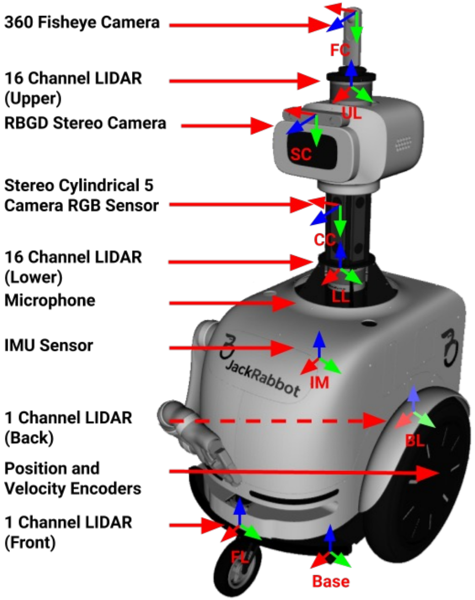




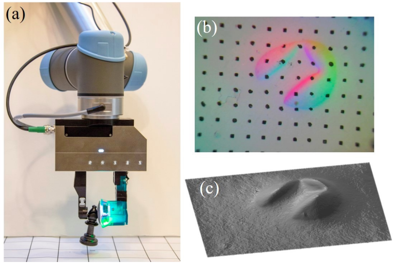
[Source: HKU Advanced Robotics Laboratory]

Robotic Sensors

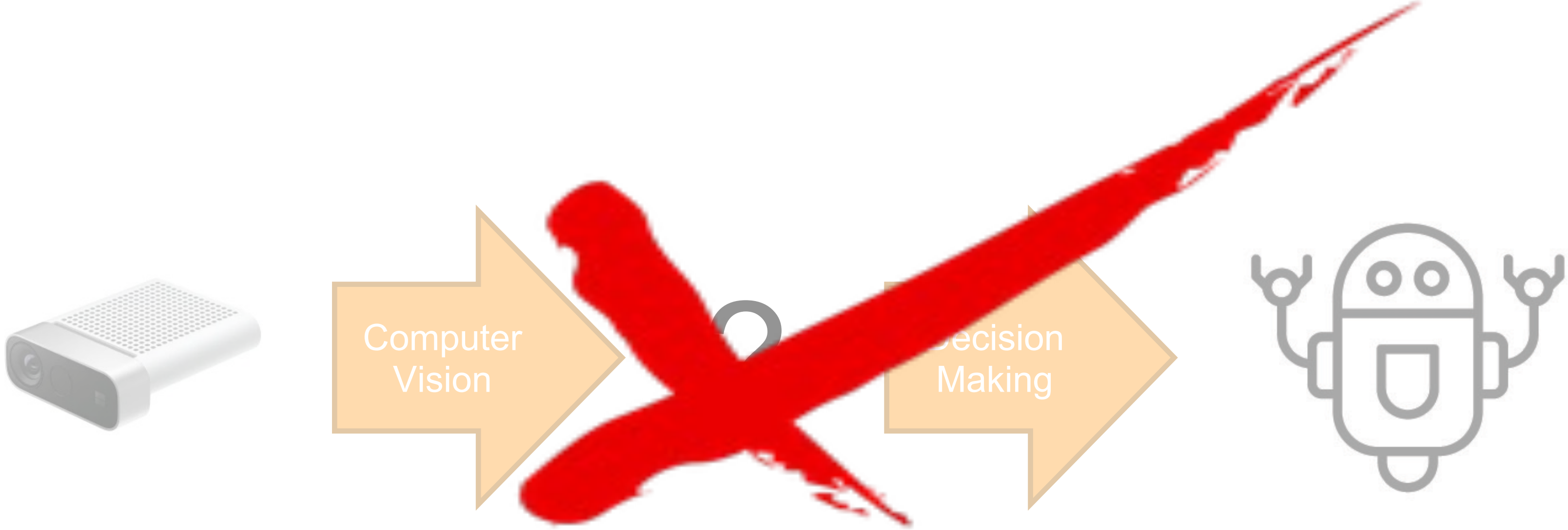
Observing the physical world through multimodal senses



[Source: JackRabbot, Stanford]



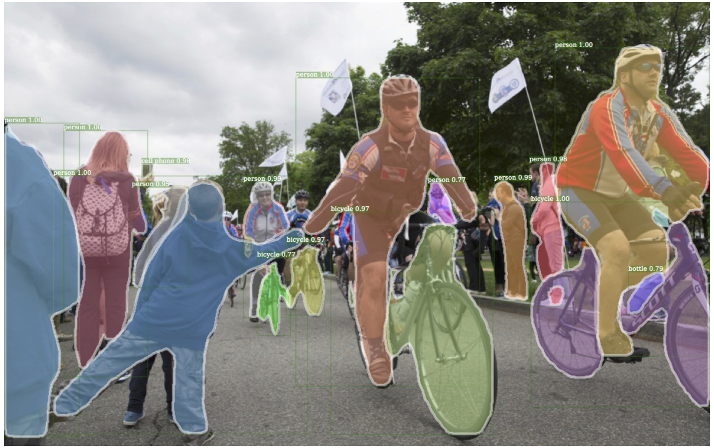
Robot Vision?



Robot Vision vs. Computer Vision

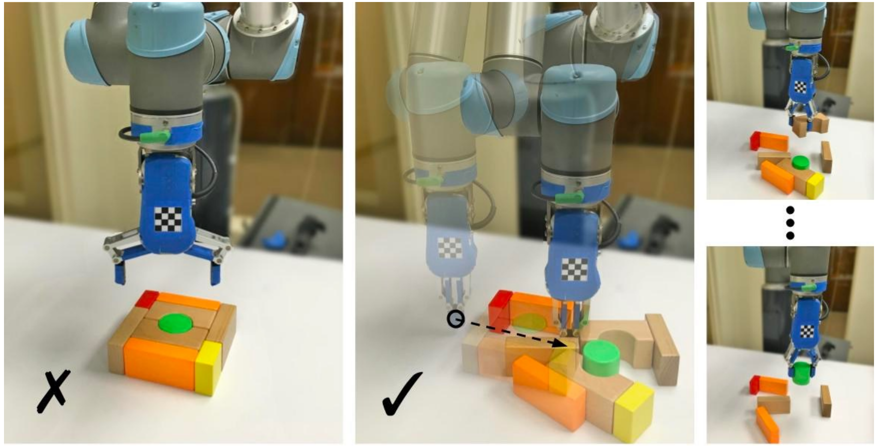
- The Limits and Potentials of Deep Learning for Robotics. Niko Sünderhauf, Oliver Brock, Walter Scheirer, Raia Hadsell, Dieter Fox, Jürgen Leitner, Ben Upcroft, Pieter Abbeel, Wolfram Burgard, Michael Milford, Peter Corke (2018)
  - A Sensorimotor Account of Vision and Visual Consciousness. Kevin O'Regan and Alva Noë (2001)

MIND THE GAP



[Detectron - Facebook AI Research]

metrics: pixel accuracy, FP/FN…



[Zeng et al., IROS 2018]

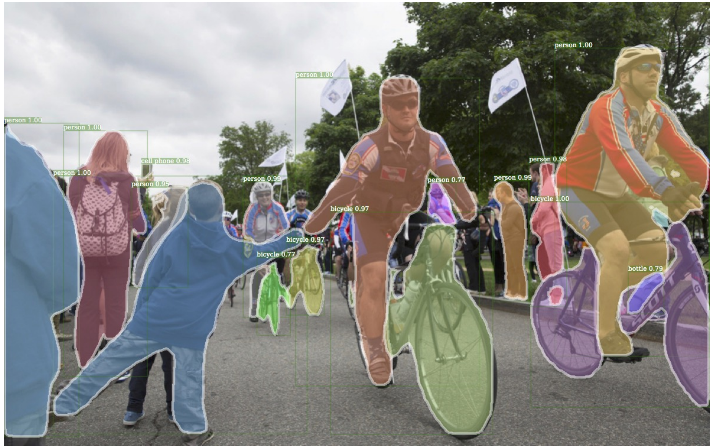
performance in a robotic task

1. Robot vision is **task-oriented**

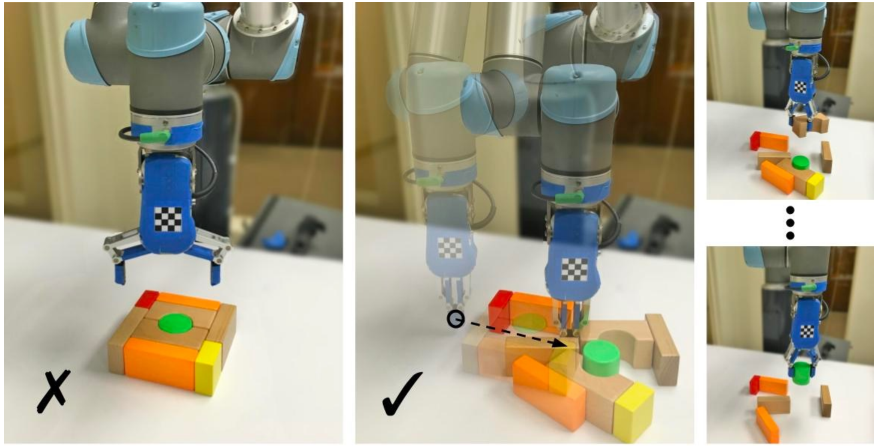
- The Limits and Potentials of Deep Learning for Robotics. Niko Sünderhauf, Oliver Brock, Walter Scheirer, Raia Hadsell, Dieter Fox, Jürgen Leitner, Ben Upcroft, Pieter Abbeel, Wolfram Burgard, Michael Milford, Peter Corke (2018)
  - A Sensorimotor Account of Vision and Visual Consciousness. Kevin O'Regan and Alva Noë (2001)

Robot Vision vs. Computer Vision

MIND THE GAP



[Detectron - Facebook AI Research]



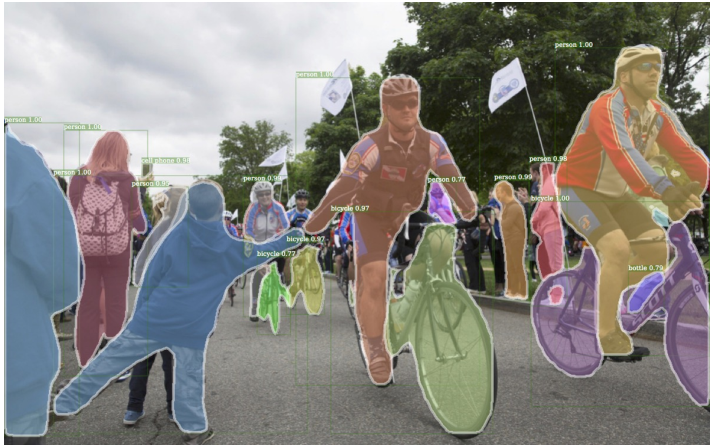
[Zeng et al., IROS 2018]

2. Robot vision is **embodied**, **active**, and **environmentally situated**

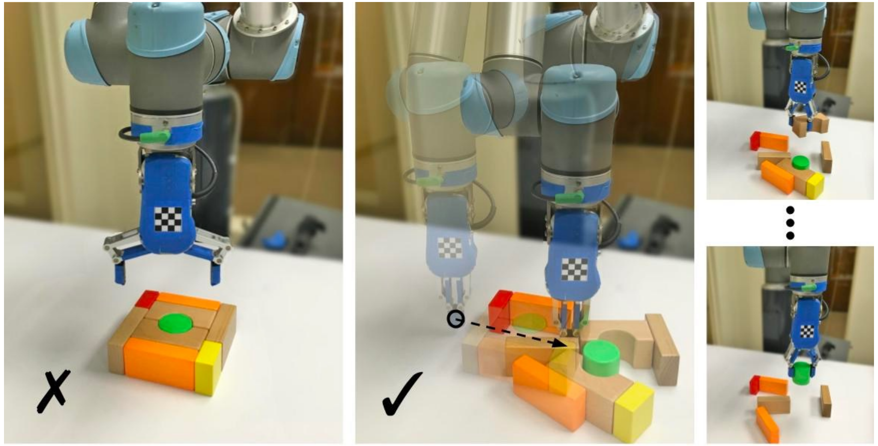
Robot Vision vs. Computer Vision

- The Limits and Potentials of Deep Learning for Robotics. Niko Sünderhauf, Oliver Brock, Walter Scheirer, Raia Hadsell, Dieter Fox, Jürgen Leitner, Ben Upcroft, Pieter Abbeel, Wolfram Burgard, Michael Milford, Peter Corke (2018)
  - A Sensorimotor Account of Vision and Visual Consciousness. Kevin O'Regan and Alva Noë (2001)

MIND THE GAP



[Detectron - Facebook AI Research]



[Zeng et al., IROS 2018]

2. Robot vision is **embodied**, **active**, and **environmentally situated**

Robot Vision vs. Computer Vision

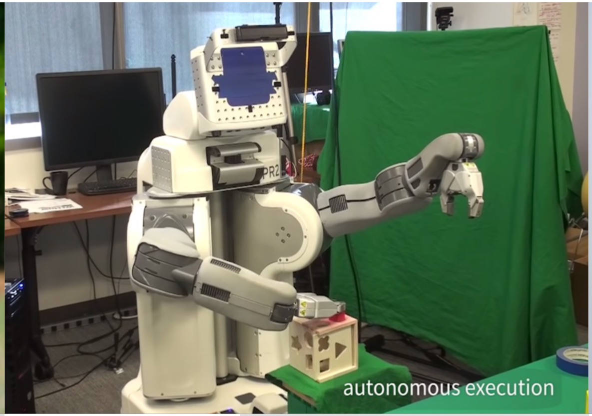
Robot vision is **embodied**, **active**, and **environmentally situated**.

- **Embodied**: Robots have physical bodies and experience the world directly. Their actions are part of a dynamical system together with the environment and have immediate feedback on their own sensation.
- **Active**: Robots are active perceivers. They should know what and why it wishes to sense, and they should be able to choose what to perceive, and determine how, when and where to achieve that perception.
- **Situated**: Robots are situated in the world. They do not deal with abstract descriptions, but with the here and now of the world directly influencing the behavior of the system.

[Brooks 1991; Bajcsy 2018]

The Perception-Action Loop

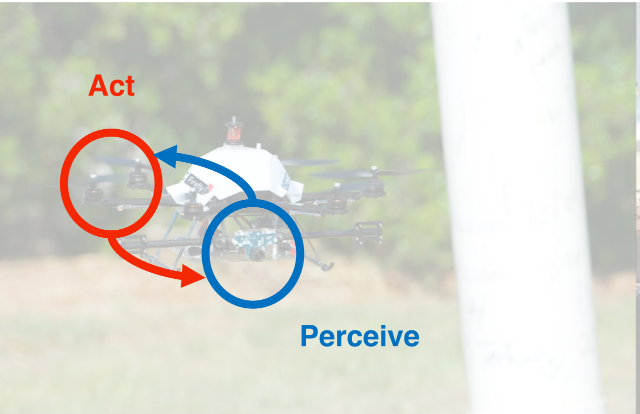
[Sa et al. IROS 2014]



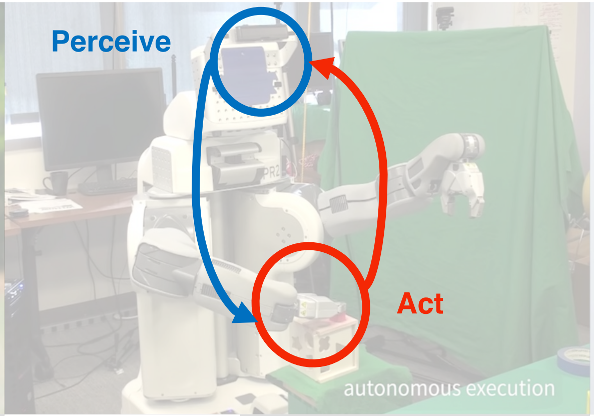
[Levine et al. JMLR 2016]

[Bohg et al. ICRA 2018]

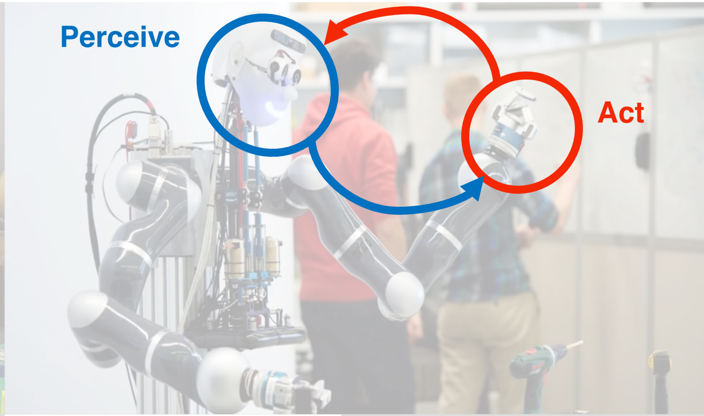
The Perception-Action Loop



[Sa et al. IROS 2014]



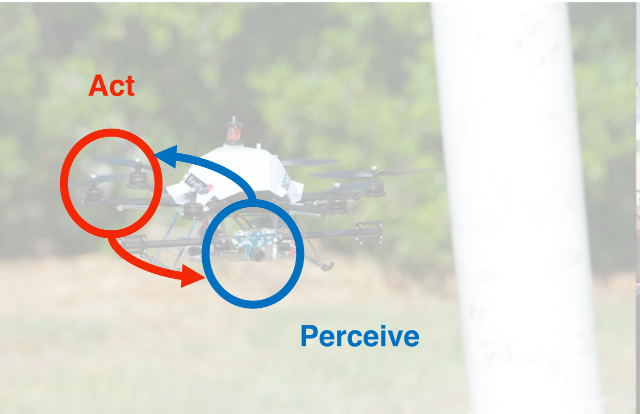
[Levine et al. JMLR 2016]



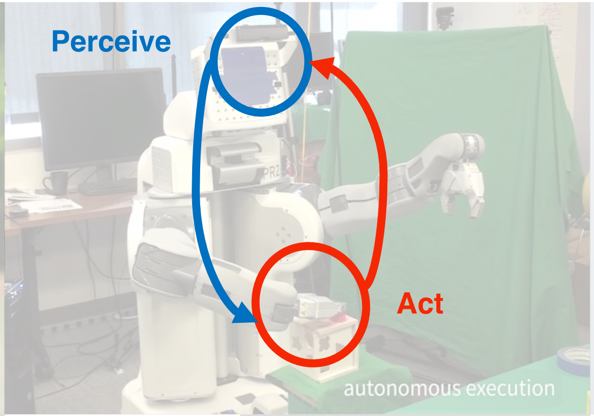
[Bohg et al. ICRA 2018]

The Perception-Action Loop

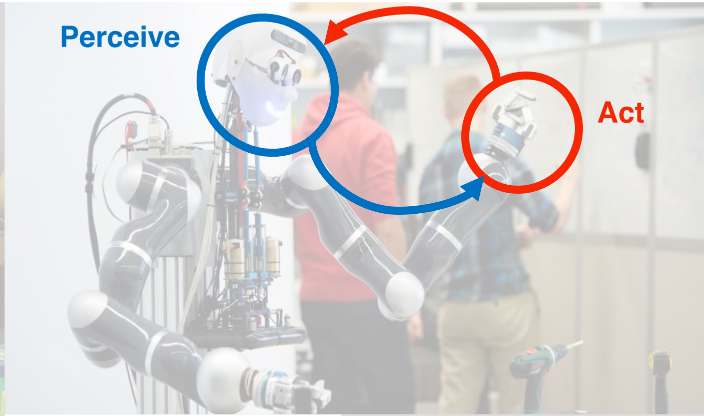
A key challenge in **Robot Learning** is to close the **perception**-action loop.



[Sa et al. IROS 2014]



[Levine et al. JMLR 2016]



[Bohg et al. ICRA 2018]

Today’s Agenda

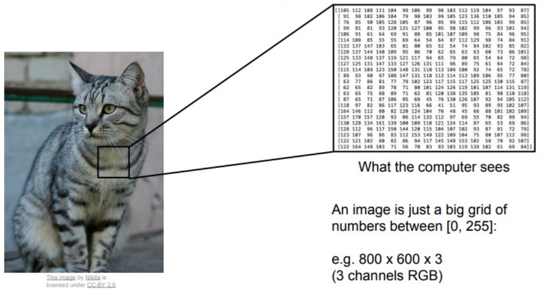
- What is Robot Perception?
- Robot Vision vs. Computer Vision
- Landscape of Robot Perception
- Quick Review
    - Deep Learning (if time permits)
    - Image formation and projective geometry (if time permits)

Robot Perception: Landscape

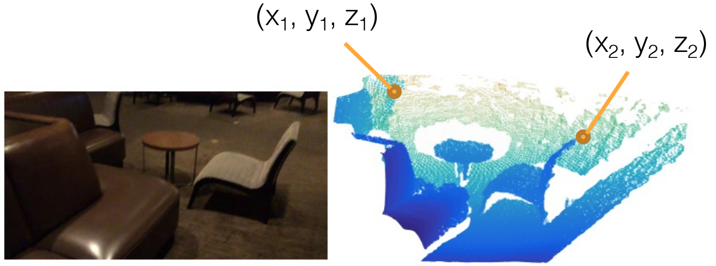
What you will learn in the chapter of Robot Perception

1. **Modalities**: neural network architectures designed for different sensory modalities
2. **Representation & Attention**: representation learning algorithms and latest attention mechanisms
3. **Temporal Integration**: state estimation tasks for robot navigation and manipulation
4. **Interactive Perception**: embodied perceptual learning, Embodied AI

Robot Perception: Modalities

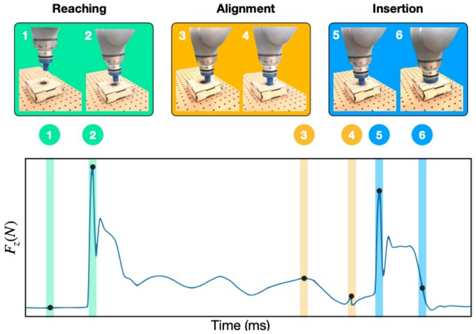


Pixels (from RGB cameras)



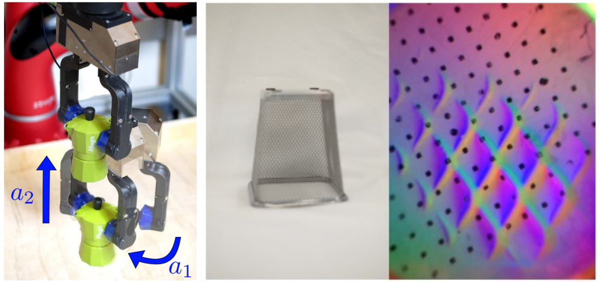
[Source: PointNet++; Qi et al. 2016]

Point cloud (from structure sensors)



[Source: Lee*, Zhu*, et al. 2018]

Time series (from F/T sensors)

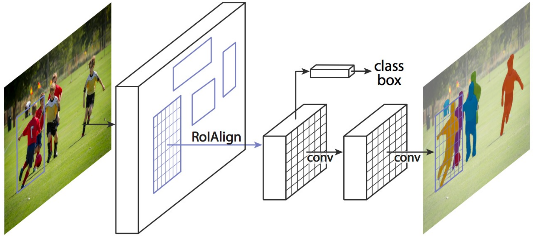


[Source: Calandra et al. 2018]

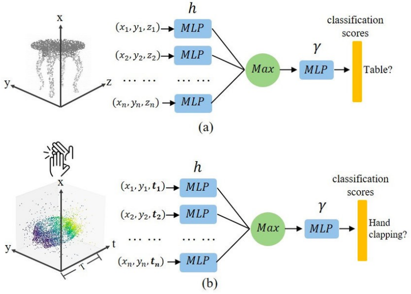
Tactile data (from the GelSights sensors)

Robot Perception: Modalities

How can we design the **neural network architectures** that can effectively process raw sensory data in vastly different forms?



Week 2: 2D Object Detection

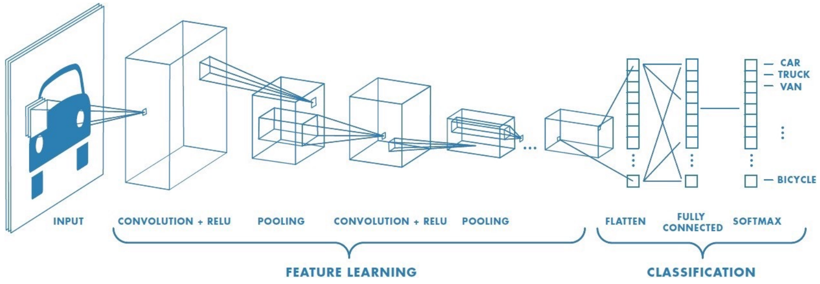


Week 3: 3D Data Processing

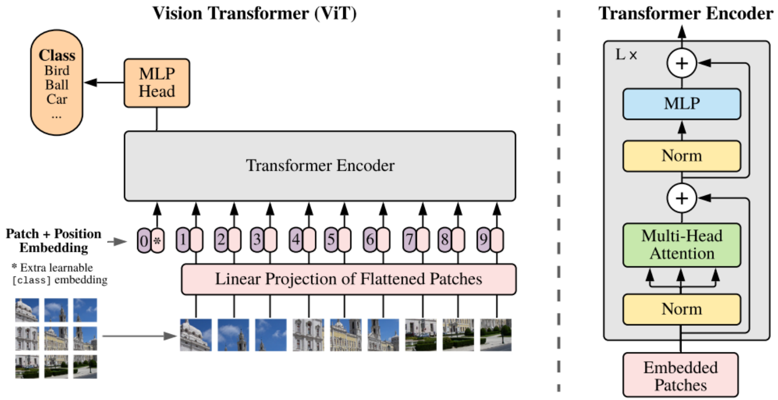
More sensory modalities in later weeks…

Robot Perception: Modalities

How can we design the **neural network architectures** that can effectively process raw sensory data in vastly different forms?



[Credit: Sumit Saha]



Week 4: Attention Architectures

Robot Perception: Multimodality

How can we learn to fuse **multiple sensory modalities** together?

Is seeing believing?

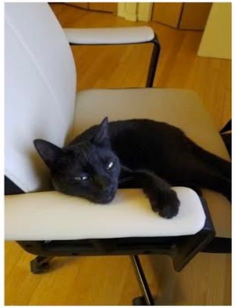
[The McGurk Effect, BBC]

https://www.youtube.com/watch?v=2k8fHR9jKVM

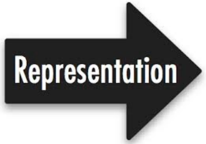
Robot Perception: Representations

A fundamental problem in robot perception is to learn the proper **representations** of the unstructured world.

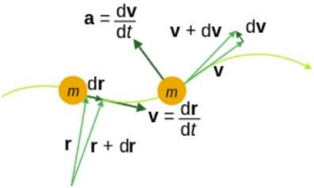
Things...

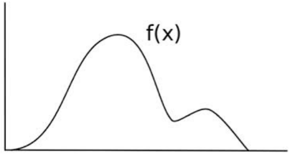


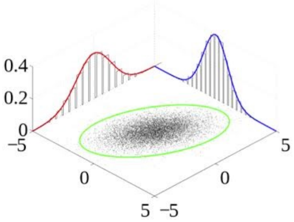
My heart beats as if the world is dropping,
you may not feel the love but i do its a heart
breaking moment of your life. enjoy the times
that we have, it might not sound good but
one thing it rhymes it might not be romantic
but i think it is great,the best rhyme i've ever
heard.

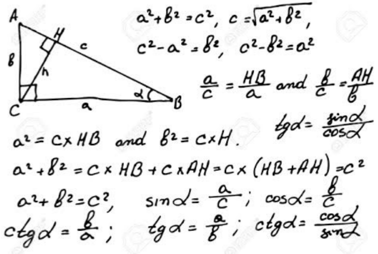


Engineering Knowledge...









[Source: Stanford CS331b]

Robot Perception: Representations

"Solving a problem simply means representing it so as to make the solution transparent."

Herbert A. Simon, Sciences of the Artificial

Our secret weapon? **Learning**

ICLR | 2023
Eleventh International Conference on Learning Representations

ICLR

Robot Perception: Representations

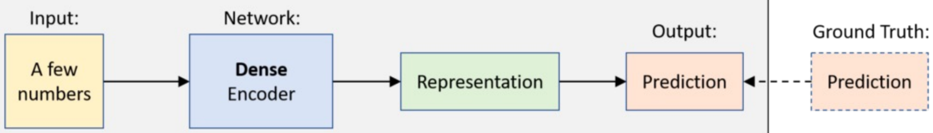
"Solving a problem simply means representing it so as to make the solution transparent."

Herbert A. Simon, Sciences of the Artificial

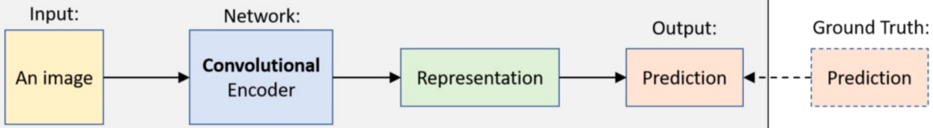
What representations to learn? How to learn them?

Supervised Learning

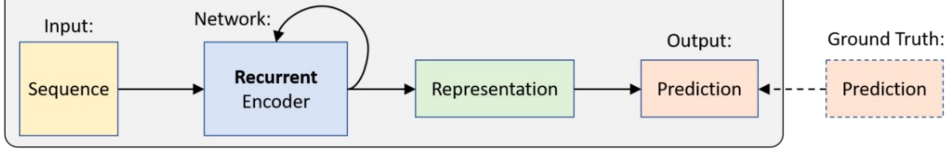
1. Feed Forward Neural Networks



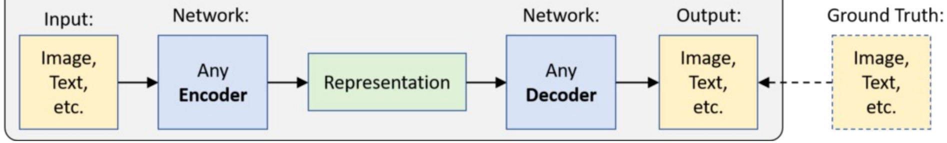
2. Convolutional Neural Networks



3. Recurrent Neural Networks

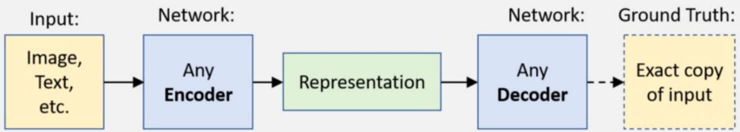


4. Encoder-Decoder Architectures

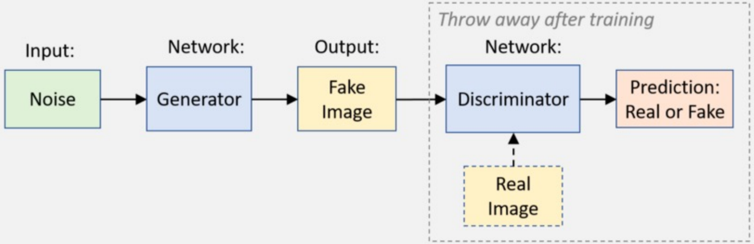


Unsupervised Learning

5. Autoencoder

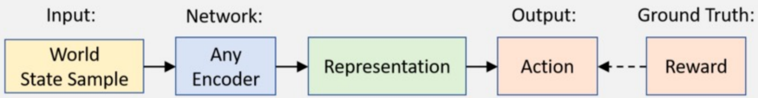


6. Generative Adversarial Networks



Reinforcement Learning

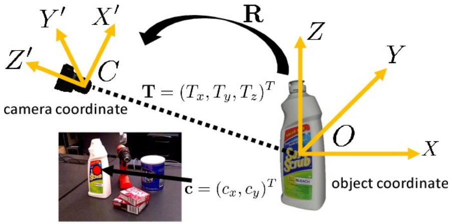
7. Networks for Actions, Values, Policies, and Models



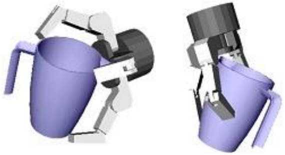
[6.S094, MIT]

Robot Perception: Temporal Integration

Noisy Sensory Data

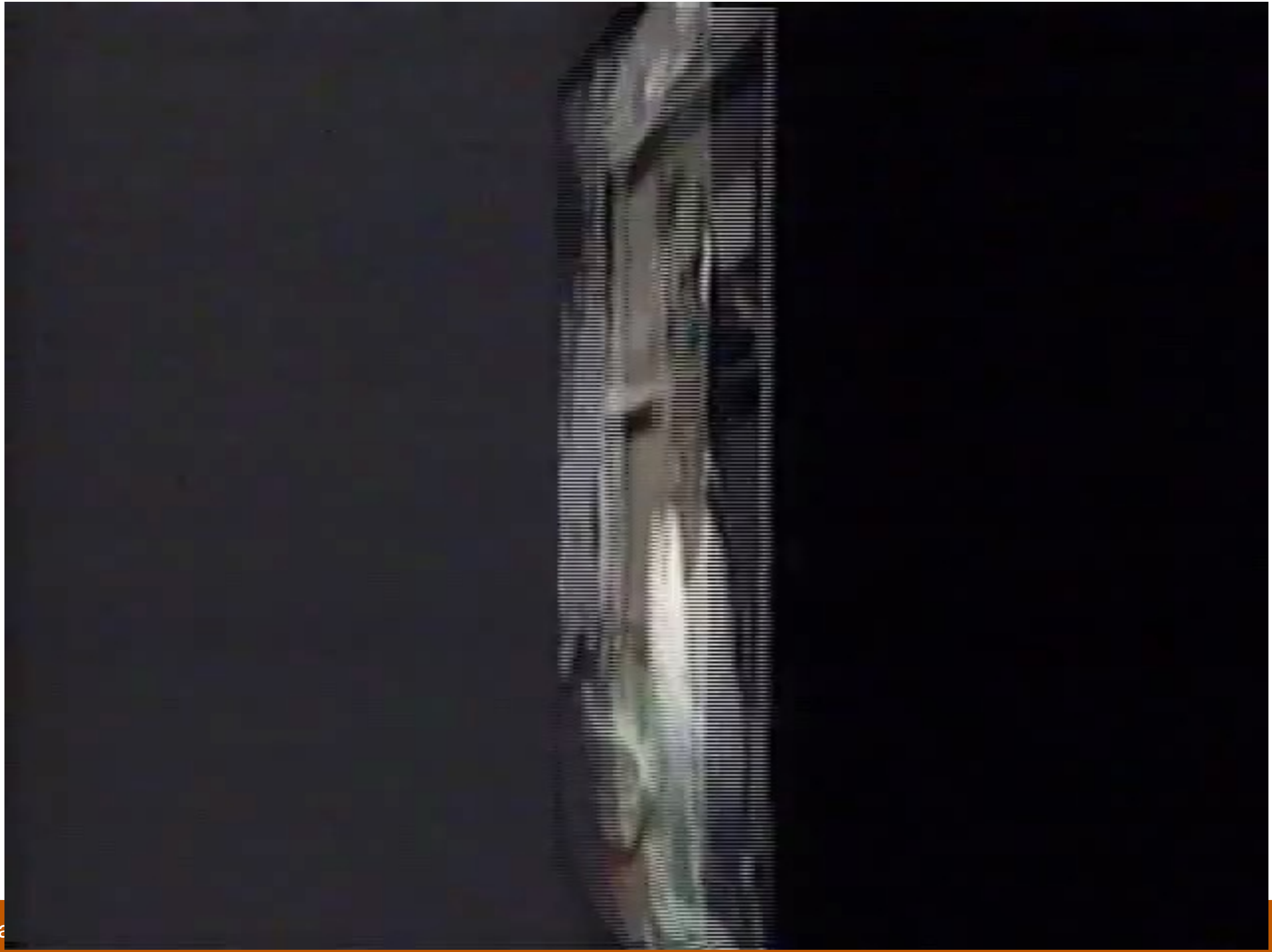


Physical State



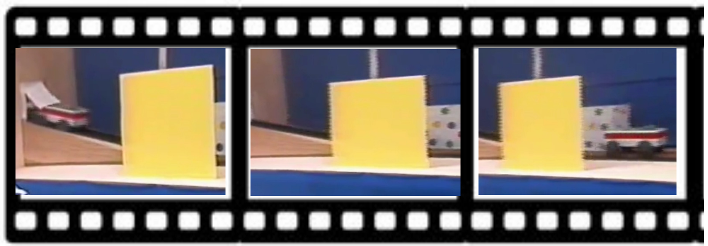
Perception & Computer Vision

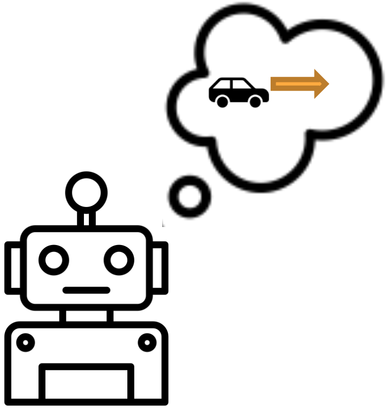
Robot Control & Decision Making



Robot Perception: Temporal Integration

sensor signals

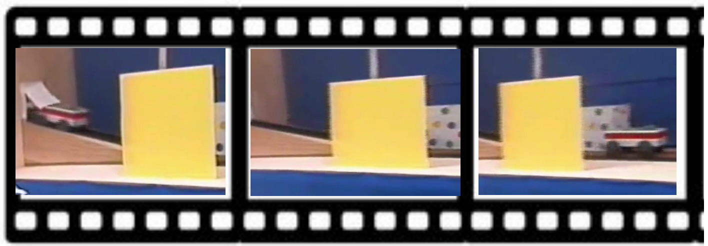


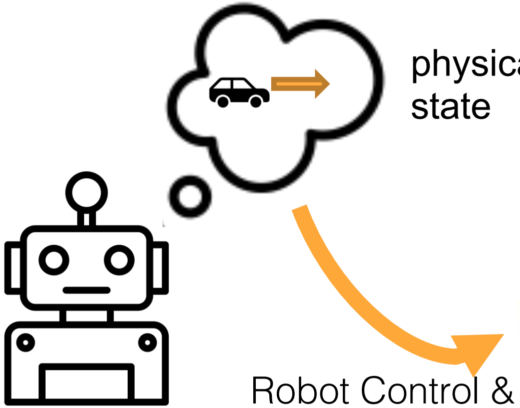


physical state

Robot Perception: Temporal Integration

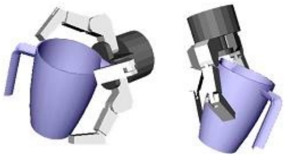
sensor signals





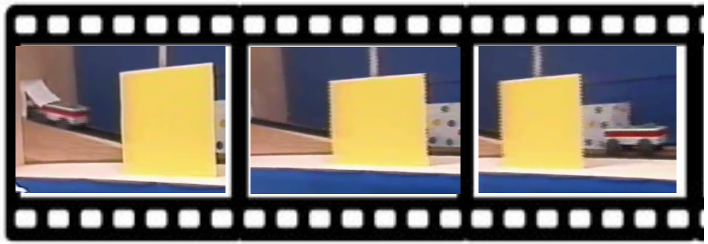
physical state

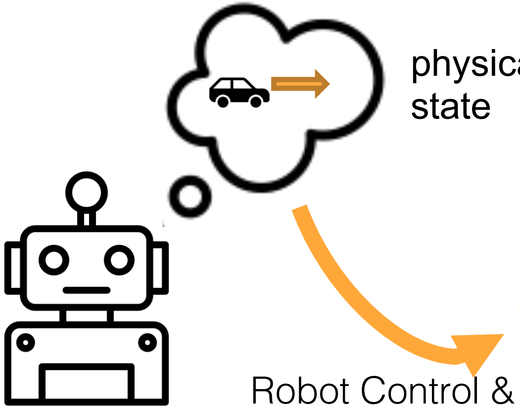
Robot Control & Decision Making



Robot Perception: Temporal Integration

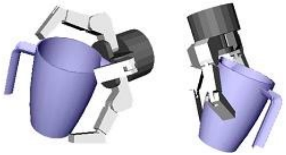
sensor signals



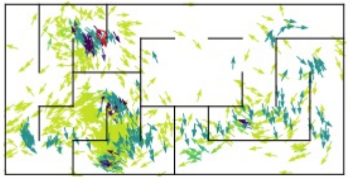


physical
state

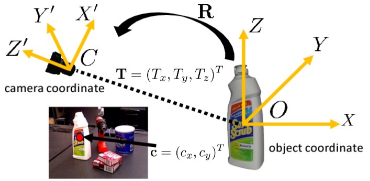
Robot Control &
Decision Making



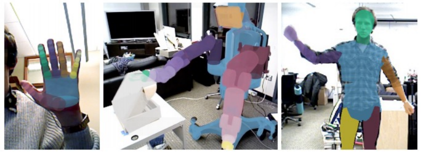
Localization



Pose Tracking

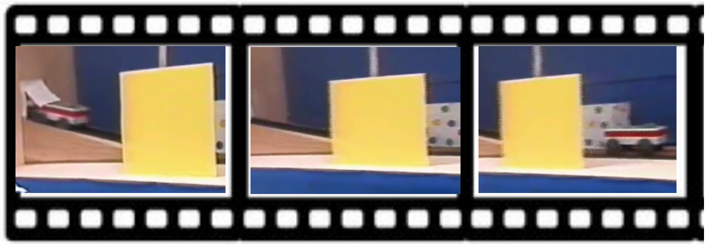


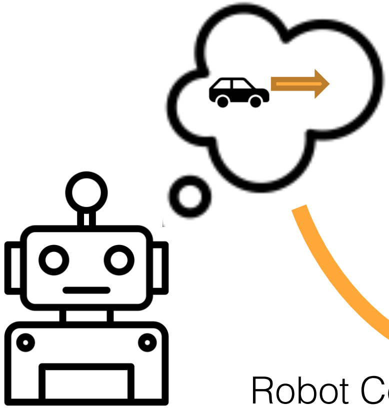
Visual Tracking



Robot Perception: Temporal Integration

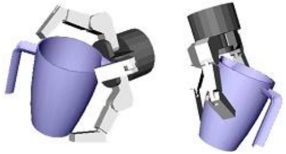
sensor signals

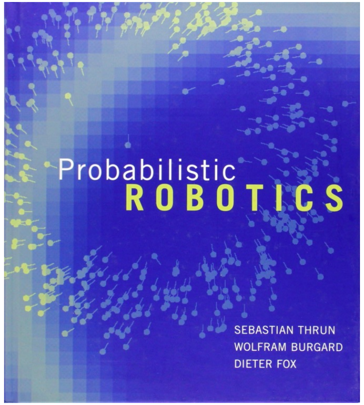




physical state

Robot Control & Decision Making





http://www.probabilistic-robotics.org/

Robot Perception: Temporal Integration

State estimation methods: **Bayes Filtering**

Algorithm 1 The general algorithm for Bayes filtering

1: **for each** x_t **do**
2: $\quad \overline{bel}(x_t) = \int p(x_t|u_t, x_{t-1})\, bel(x_{t-1})\, dx_{t-1}$ $\quad \triangleright$ transition update
3: $\quad bel(x_t) = \eta\, p(z_t|x_t)\, \overline{bel}(x_t)$ $\quad \triangleright$ measurement update
4: **end for each**

x_t: state $\quad z_t$: observation $\quad u_t$: action $\quad bel(x_t)$: belief

$p(x_t|u_t, x_{t-1})$: transition model (motion model)

$p(z_t|x_t)$: measurement model (observation model)

Week 4 Thu, Sept 15 — Recursive State Estimation

- Differentiable Particle Filters: End-to-End Learning with Algorithmic Priors. Rico Jonschkowski, Divyam Rastogi, Oliver Brock (2018)
  - Presenter:
- Online Interactive Perception of Articulated Objects with Multi-Level Recursive Estimation Based on Task-Specific Priors. Roberto Martín-Martín, Oliver Brock (2014)
  - Presenter:
- Multimodal sensor fusion with differentiable filters. Michelle A Lee, Brent Yi, Roberto Martín-Martín, Silvio Savarese, Jeannette Bohg (2020)
  - Presenter:

  - A Brief Tutorial On Recursive Estimation With Examples From Intelligent Vehicle Applications. Hao Li

Robot Perception: Temporal Integration

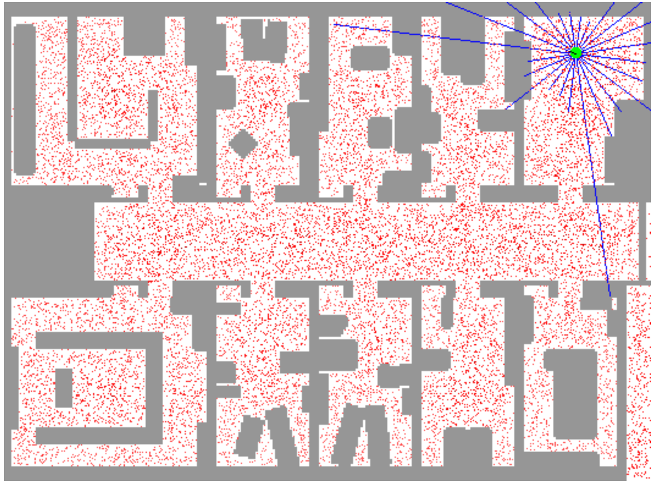
State estimation methods: **Bayes Filtering**

x_t: state z_t: observation u_t: action $bel(x_t)$: belief

$p(x_t|u_t, x_{t-1})$: transition model (motion model)

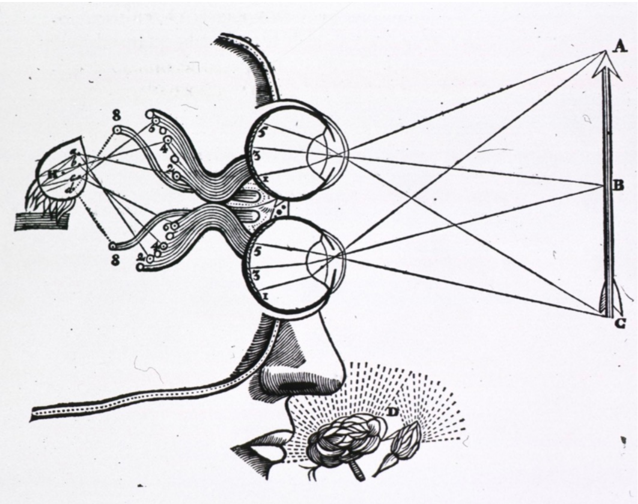
$p(z_t|x_t)$: measurement model (observation model)

What if models are hard to specify? **Learning**



Example: Particle Filter Localization

Robot Perception: Embodied AI



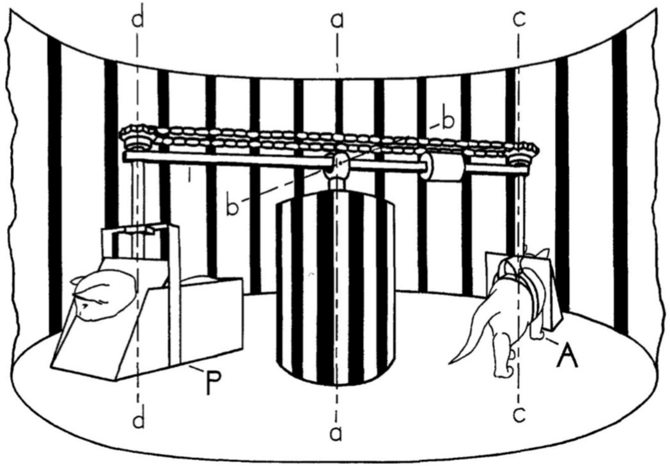
Input-Output Picture (Susan Hurley, 1998)

Classical View of Perception

- Perception is the process of building an internal representation of the environment
- Perception is input from world to mind, and action is output from mind to world, thought is the mediating process.

[Action in Perception, Alva Noë 2004]

Robot Perception: Embodied AI

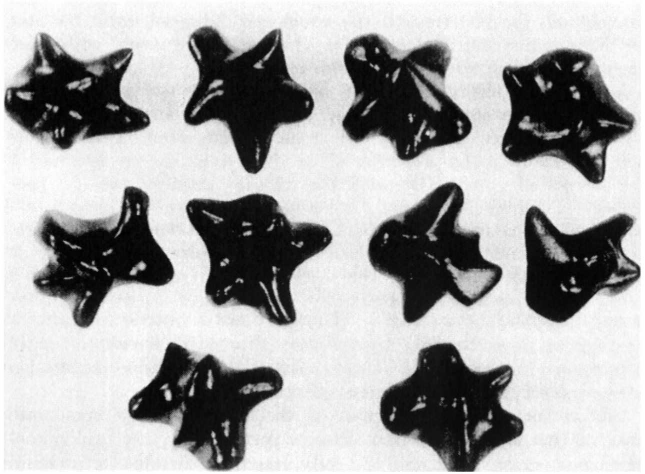


Kitten Carousel (Held and Hein, 1963)

Embodied View of Perception

- As the active cat (A) walks, the other cat (P) moves and perceives the environment passively.
- Only the active cat develops normal perception through *self-actuated* movement.
- The passive cat suffers from perception problems, such as 1) not blinking when objects approach, and 2) hitting the walls.

Robot Perception: Embodied AI



Pebbles (James J. Gibson 1966)

Embodied View of Perception

- Subjects asked to find a reference object among a set of irregularly-shaped objects
- Three groups
  a. Passive observers of one static image (49%)
  b. Observers of moving shapes (72%)
  c. Interactive observers (99%)
- The ability to condition input signals with actions is crucial to perception.

Robot Perception: Embodied AI

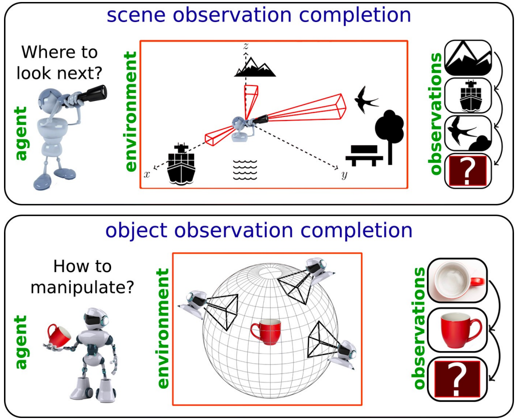
Take-home messages

- Perceptual experiences do not present the sense in the way that a photograph does.
- Perception is developed by an embodied agent through actively exploring in the physical world.
- "We see in order to move; we move in order to see." – William Gibson

Robot Perception: Embodied AI

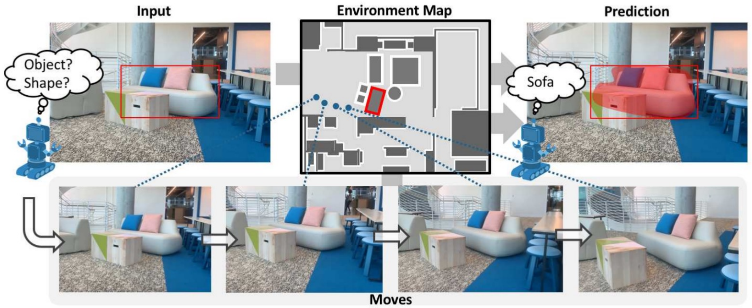
Week 5 (Tue) – Active & Interactive Perception: How can embodied agents (robots) improve perception based on visual experiences through (inter)active exploration?

View Selection



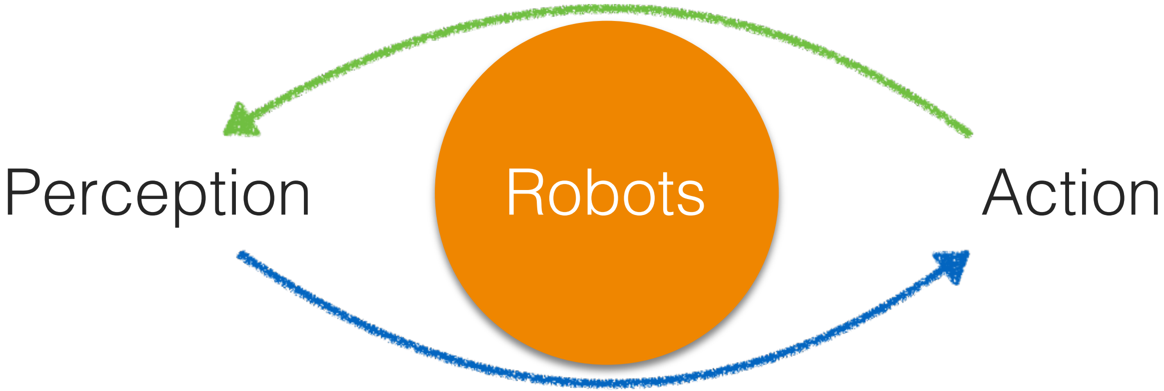
[Jayaraman and Grauman 2017]

Amodal Recognition



[Yang et al. 2019]

Research Frontier: Closing the Perception-Action Loop

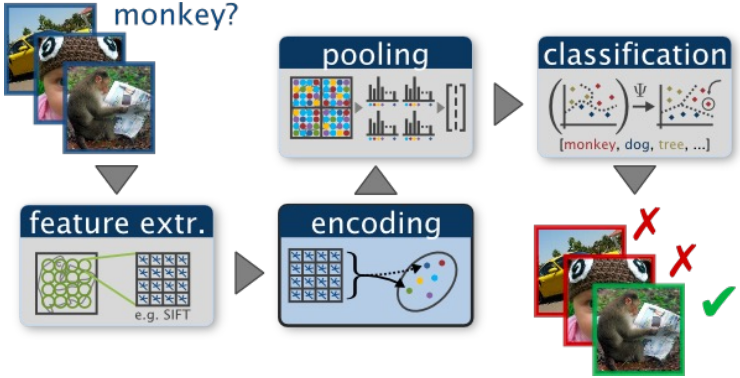


How robots develop better understanding of their surroundings from embodied sensorimotor experiences

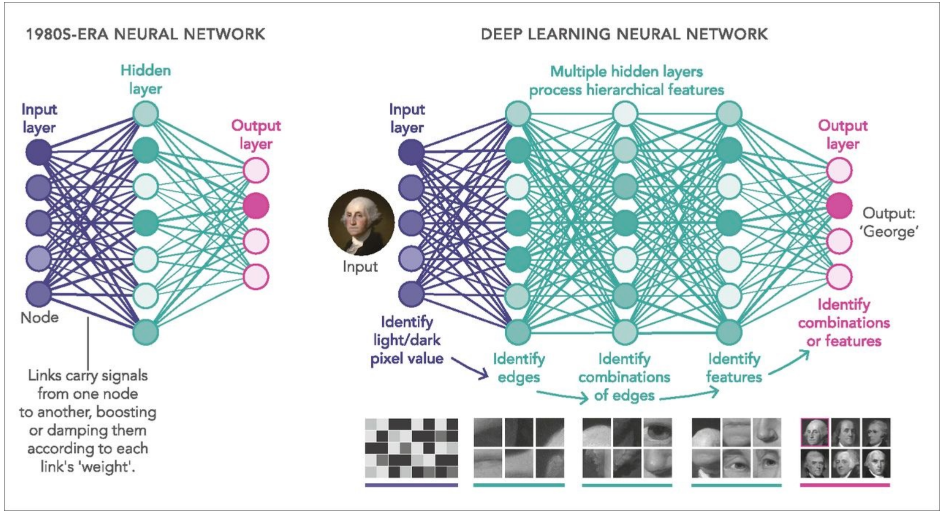
How robots' intelligent behaviors are guided by their interactive perception

Visual Processing Methods

What is new since 1980s?

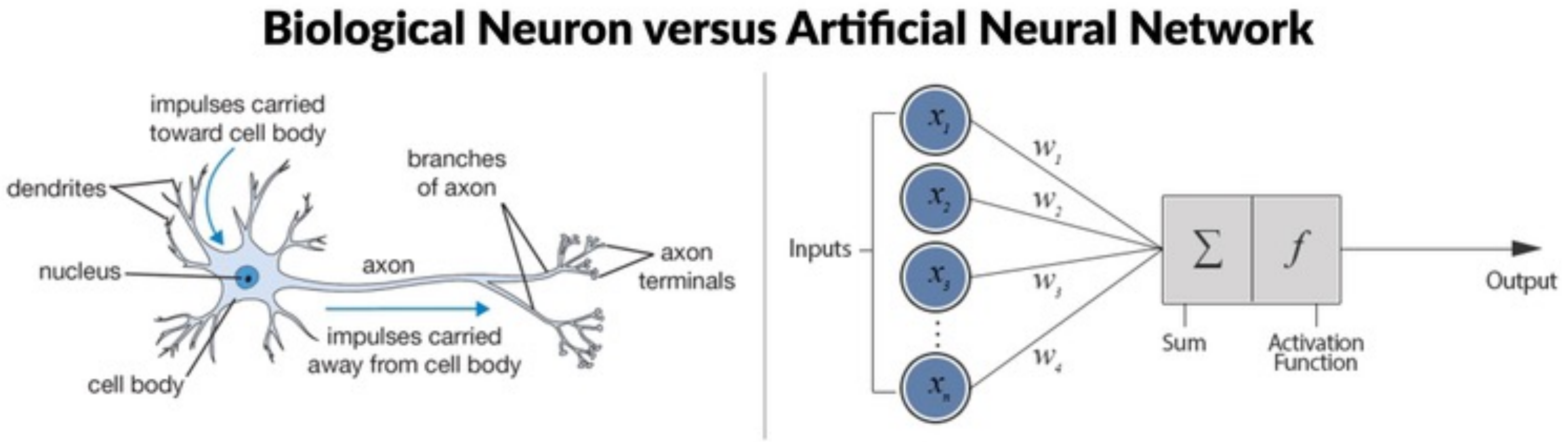


Staged Visual Recognition Pipeline



End-to-end Deep Learning

Quick Review of Deep Learning: Artificial Neurons



Biological Neuron

Computational building block for the brain

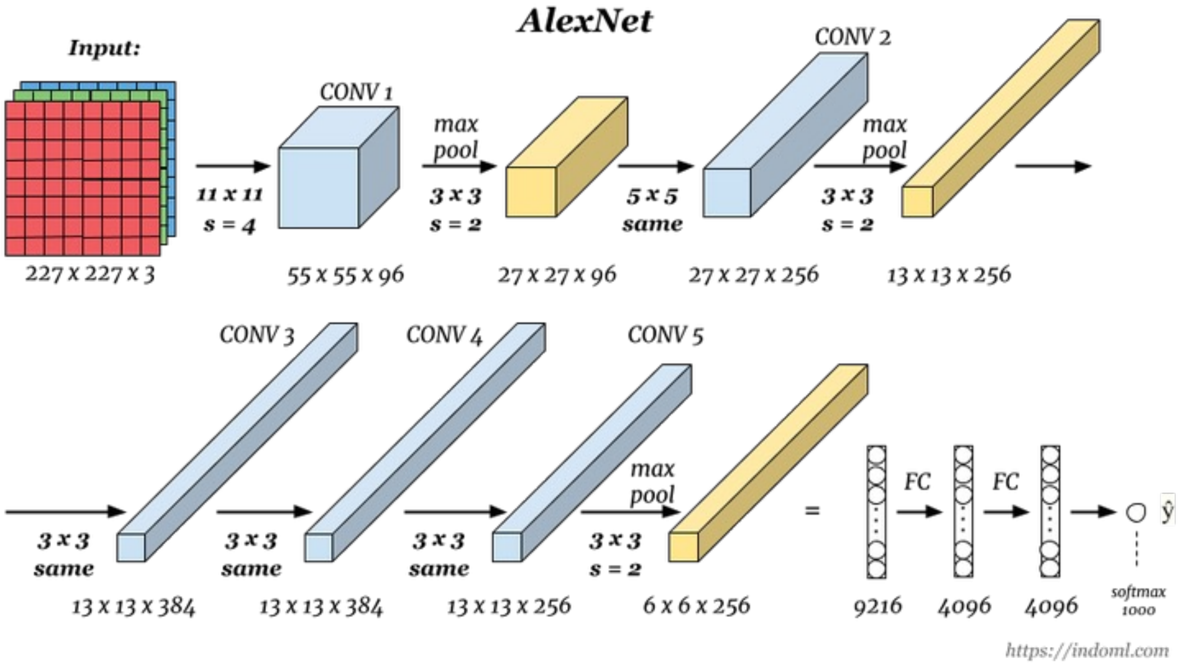
Artificial Neuron

Computational building block for the neural network

Note: Many differences exist – be careful with the brain analogies!

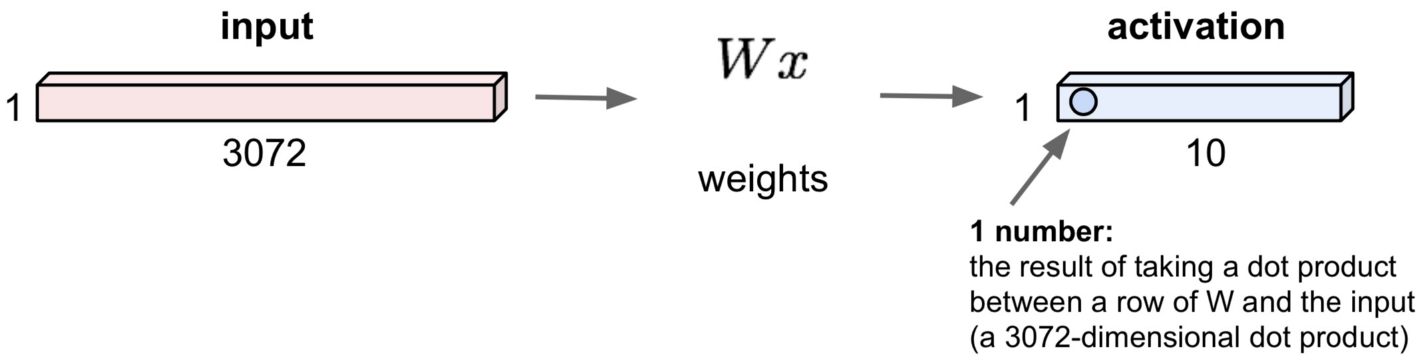
[Dendritic Computation, Michael London and Michael Hausser 2015]

Quick Review of Deep Learning: Convolutional Networks



Quick Review of Deep Learning: Fully-Connected Layers

32x32x3 image -> stretch to 3072 x 1

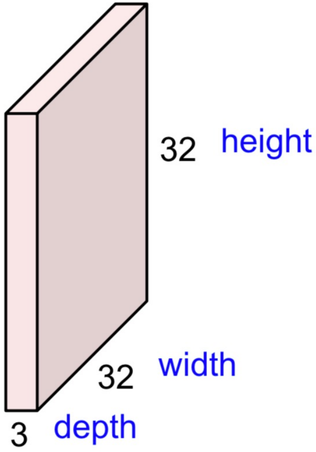


What is the dimension of W?

[Source: Stanford CS231N]

Quick Review of Deep Learning: Convolutional Layers

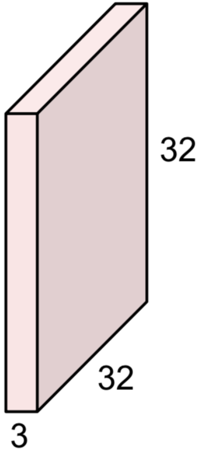
32x32x3 image -> preserve spatial structure



[Source: Stanford CS231N]

Quick Review of Deep Learning: Convolutional Layers

32x32x3 image

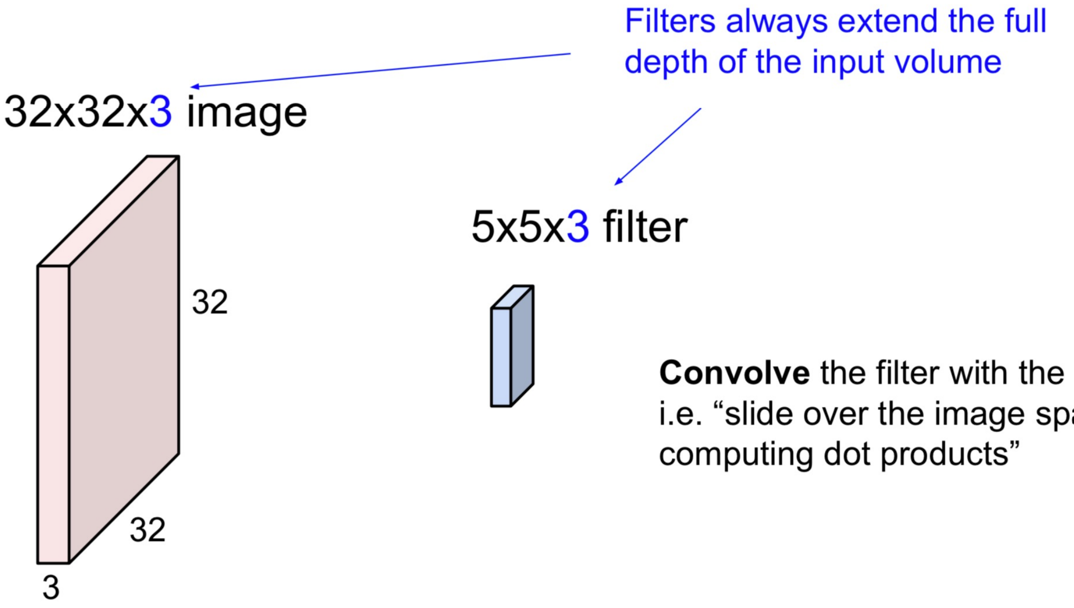


5x5x3 filter

Convolve the filter with the image
i.e. "slide over the image spatially,
computing dot products"

[Source: Stanford CS231N]

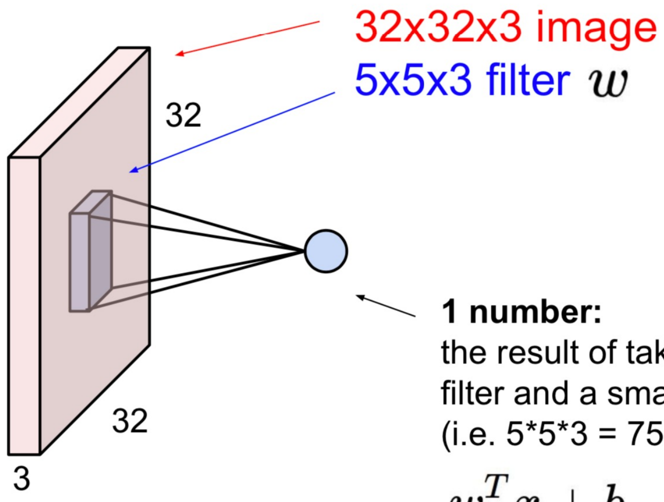
Quick Review of Deep Learning: Convolutional Layers



Convolve the filter with the image i.e. "slide over the image spatially, computing dot products"

[Source: Stanford CS231N]

Quick Review of Deep Learning: Convolutional Layers

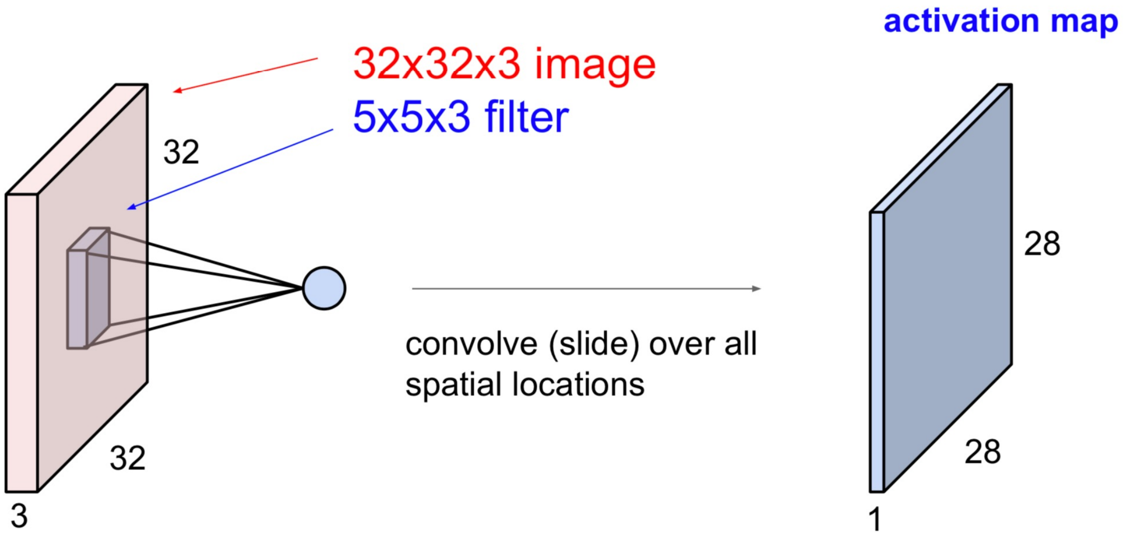


1 number:
the result of taking a dot product between the filter and a small 5x5x3 chunk of the image
(i.e. 5*5*3 = 75-dimensional dot product + bias)

$$w^T x + b$$

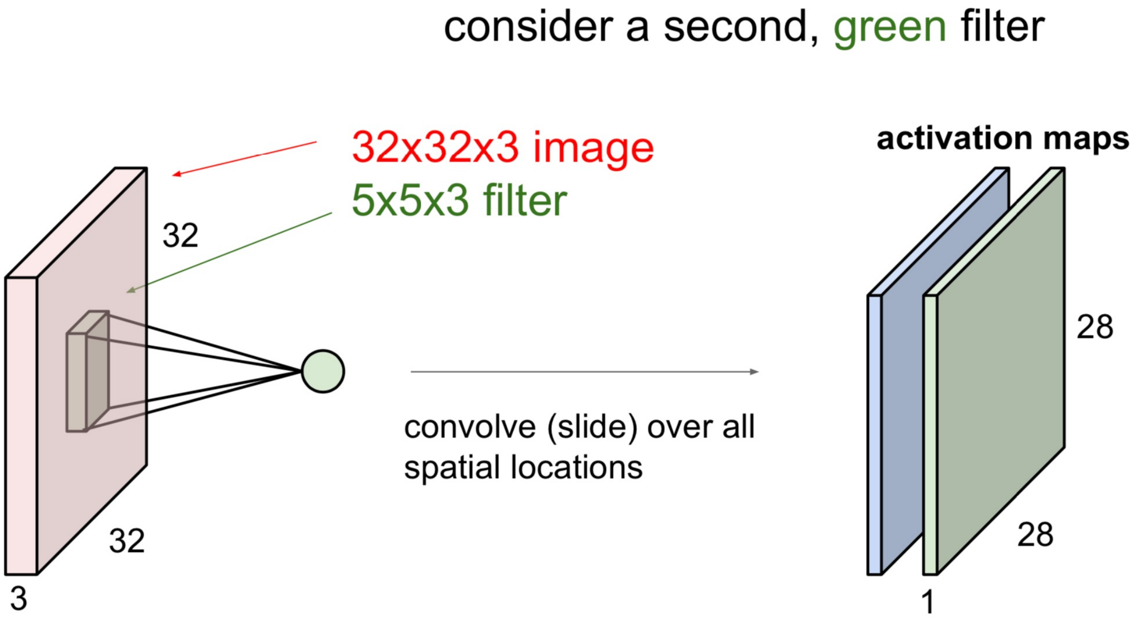
[Source: Stanford CS231N]

Quick Review of Deep Learning: Convolutional Layers



[Source: Stanford CS231N]

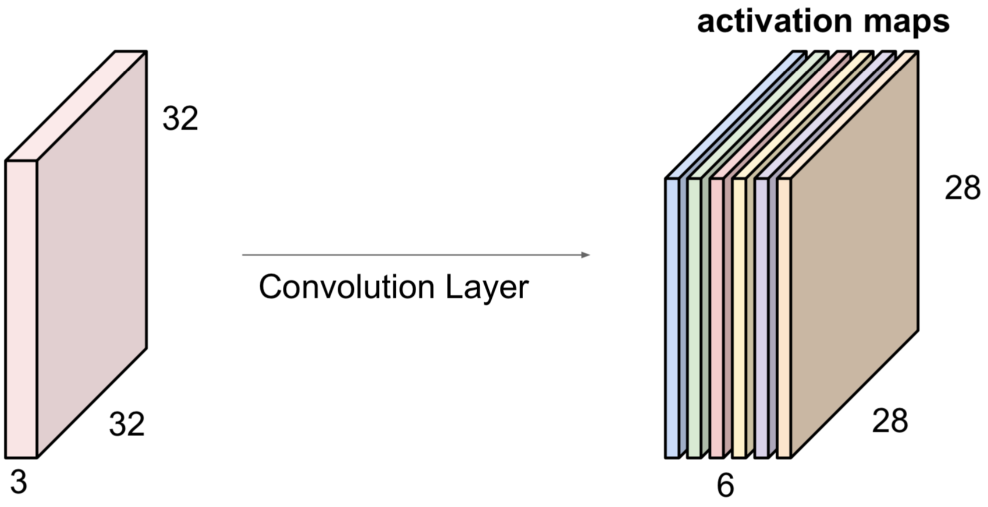
Quick Review of Deep Learning: Convolutional Layers



[Source: Stanford CS231N]

Quick Review of Deep Learning: Convolutional Layers

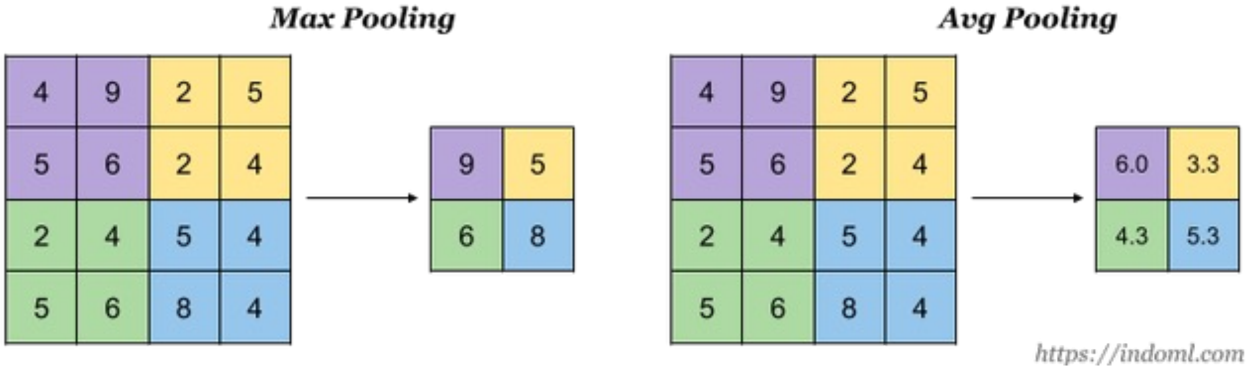
For example, if we had 6 5x5 filters, we'll get 6 separate activation maps:



We stack these up to get a "new image" of size 28x28x6!

[Source: Stanford CS231N]

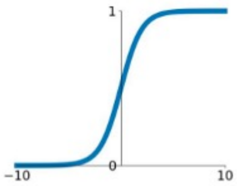
Quick Review of Deep Learning: Pooling Operations



Quick Review of Deep Learning: Activation Functions

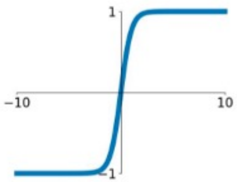
Sigmoid

$\sigma(x) = \frac{1}{1+e^{-x}}$



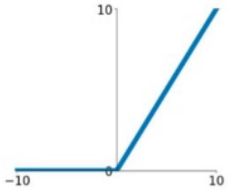
tanh

$\tanh(x)$



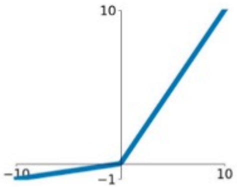
ReLU

$\max(0, x)$



Leaky ReLU

$\max(0.1x, x)$

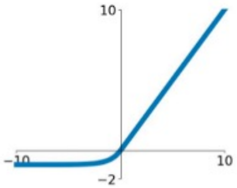


Maxout

$\max(w_1^T x + b_1, w_2^T x + b_2)$

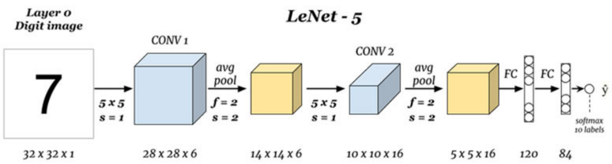
ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

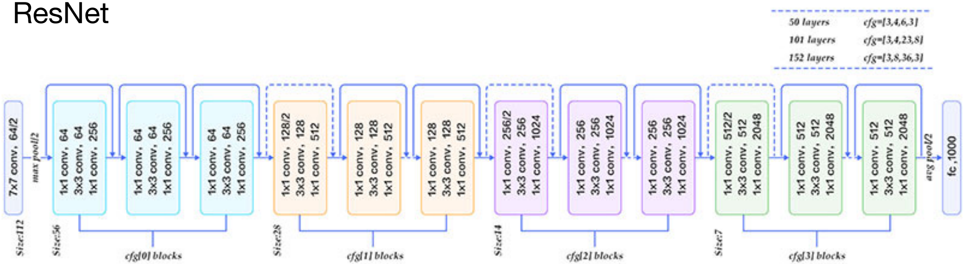


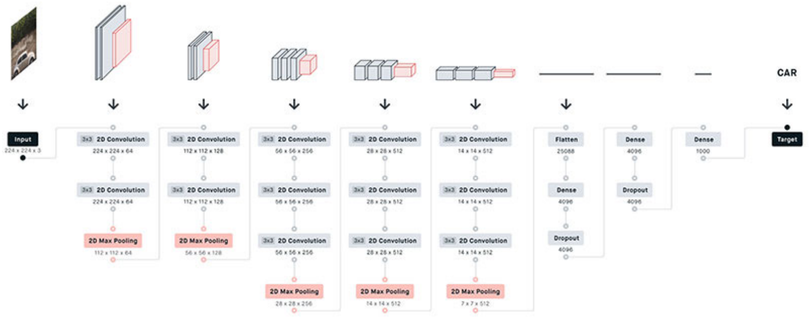
Quick Review of Deep Learning: CNN Architectures

LeNet

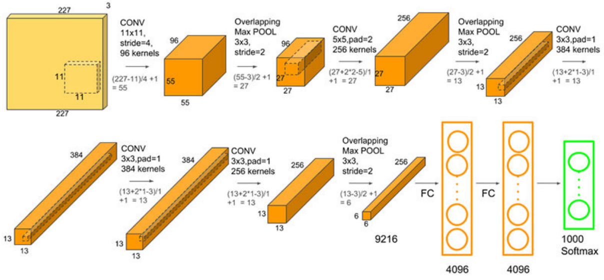


ResNet





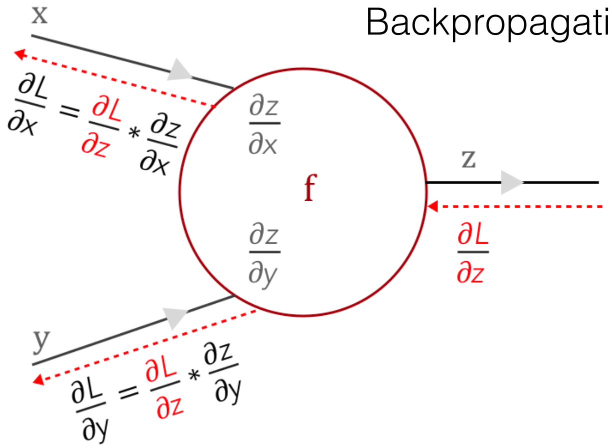
VGG-16



AlexNet

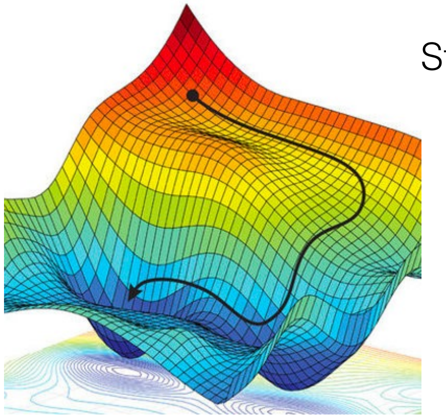
Quick Review of Deep Learning: Optimization

Backpropagation



$\frac{\partial z}{\partial x}$ & $\frac{\partial z}{\partial y}$ *are local gradients*

$\frac{\partial L}{\partial z}$ *is the loss from the previous layer which has to be backpropagated to other layers*



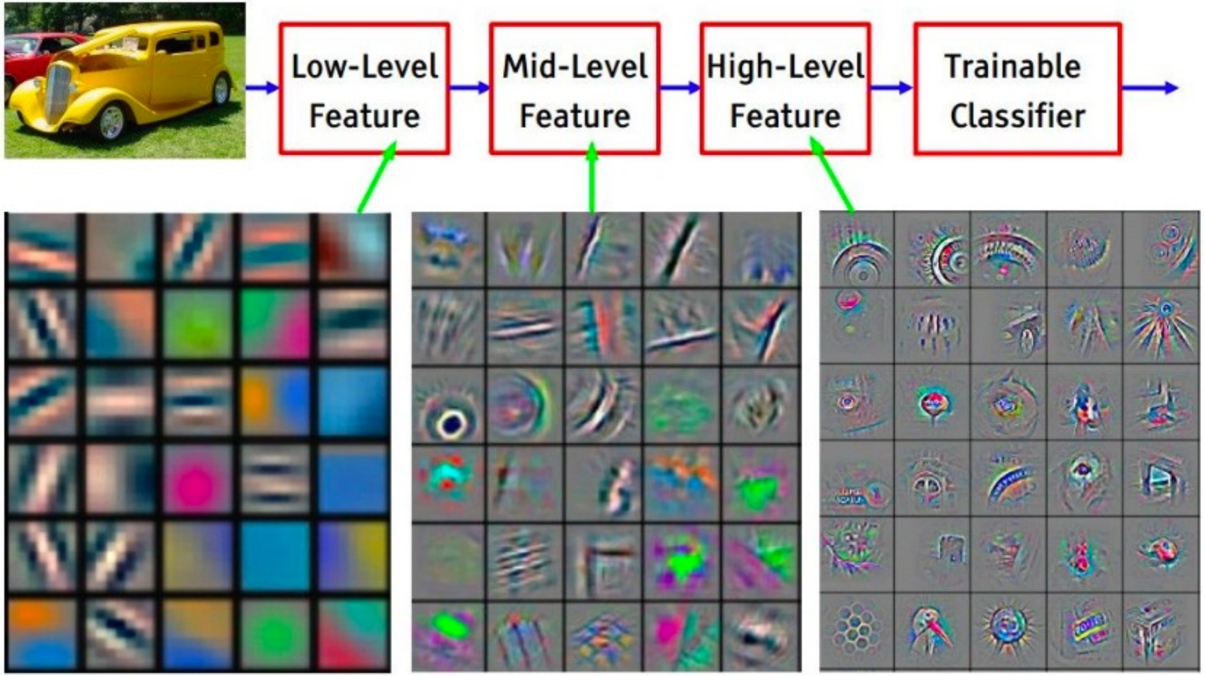
Stochastic Gradient Descent (SGD)

learning rate

$$\underbrace{\theta}_{\text{weights}} = \theta - \overbrace{\eta}^{\text{learning rate}} \nabla_\theta J(\theta; \underbrace{x^{(i)}}_{\text{input}}; \underbrace{y^{(i)}}_{\text{label}})$$

weights input label

Quick Review of Deep Learning: Features



[Source: Stanford CS231N]

Quick Review of Deep Learning: Implementation

PyTorch

TensorFlow

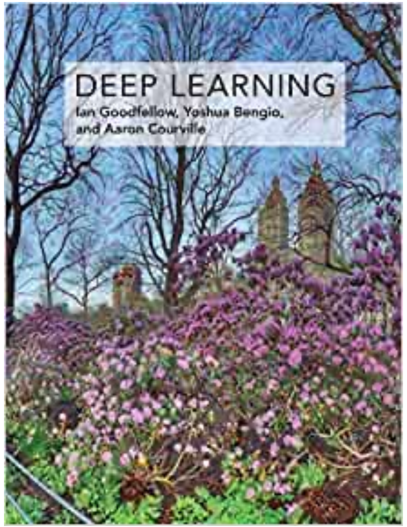
PyTorch tutorial by Yifeng

```
[ ] import torch
    from torch import nn

    class MNISTClassifier(nn.Module):

      def __init__(self):
        super(MNISTClassifier, self).__init__()

        # mnist images are (1, 28, 28) (channels, width, heig
        self.layer_1 = torch.nn.Linear(28 * 28, 128)
        self.layer_2 = torch.nn.Linear(128, 256)
        self.layer_3 = torch.nn.Linear(256, 10)

      def forward(self, x):
        batch_size, channels, width, height = x.size()

        # (b, 1, 28, 28) -> (b, 1*28*28)
        x = x.view(batch_size, -1)

        # layer 1
        x = self.layer_1(x)
        x = torch.relu(x)

        # layer 2
        x = self.layer_2(x)
        x = torch.relu(x)

        # layer 3
        x = self.layer_3(x)

        # probability distribution over labels
        x = torch.log_softmax(x, dim=1)

        return x
```

Quick Review of Deep Learning: Resources

Online Courses

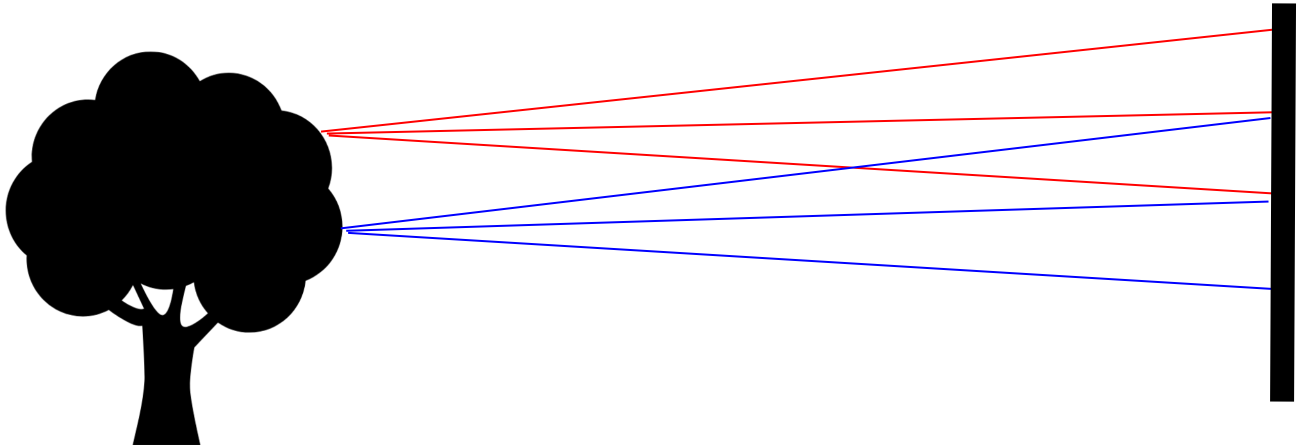
- CS231N: Convolutional Neural Networks for Visual Recognition
  http://cs231n.stanford.edu/
- MIT 6.S191: Introduction to Deep Learning
  http://introtodeeplearning.com/



Textbooks:

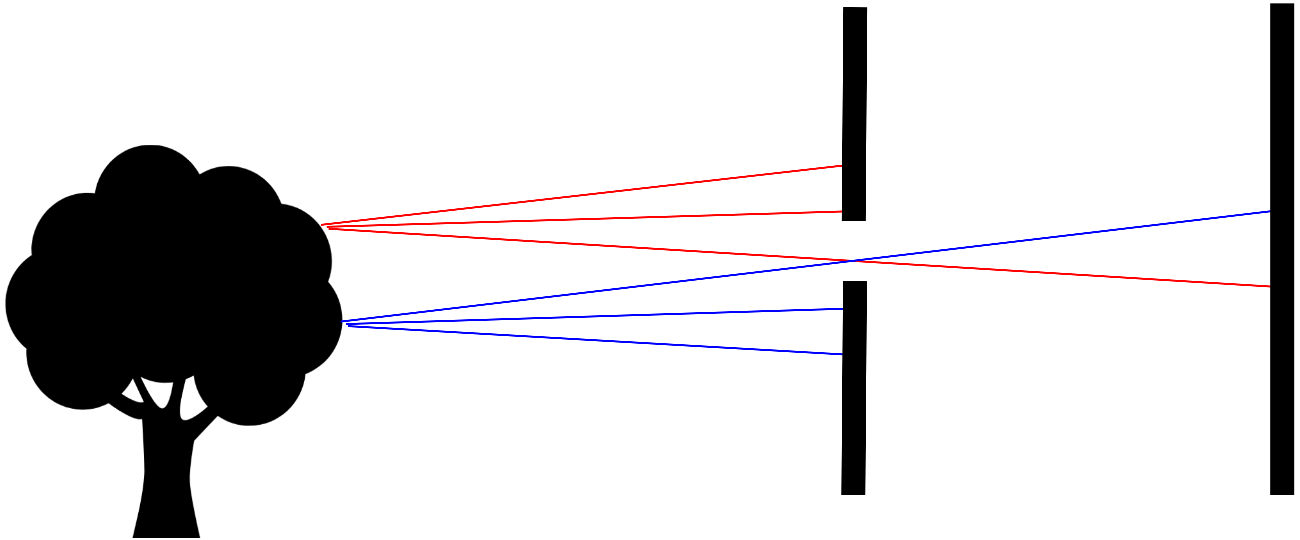
- Deep Learning. Ian Goodfellow, Yoshua Bengio, Aaron Courville
  http://www.deeplearningbook.org/

Quick Review: Image Formation - Pinhole Camera Model



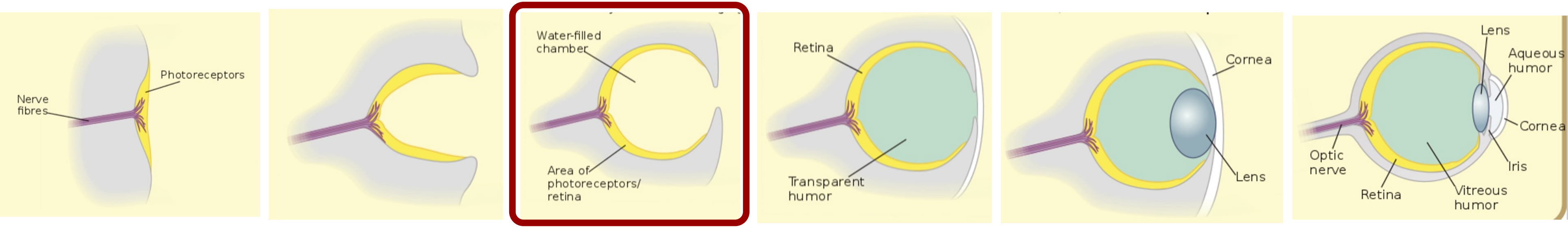
slides credit to Prof. Savarese

Quick Review: Image Formation - Pinhole Camera Model



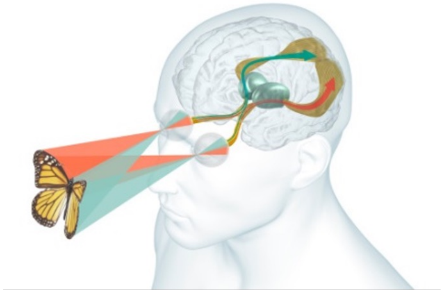
slides credit to Prof. Savarese

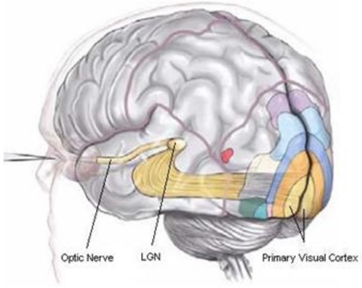
Evolution of the Eye



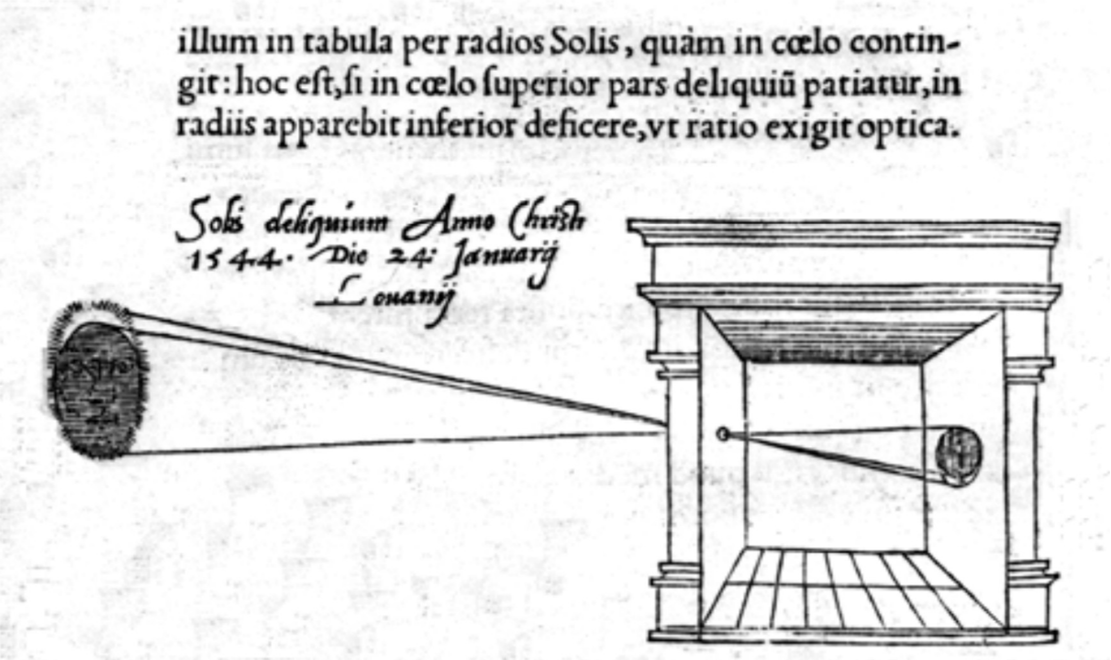
Pin Hole Model

+ More than 50% of the human cortex “involved” in vision!



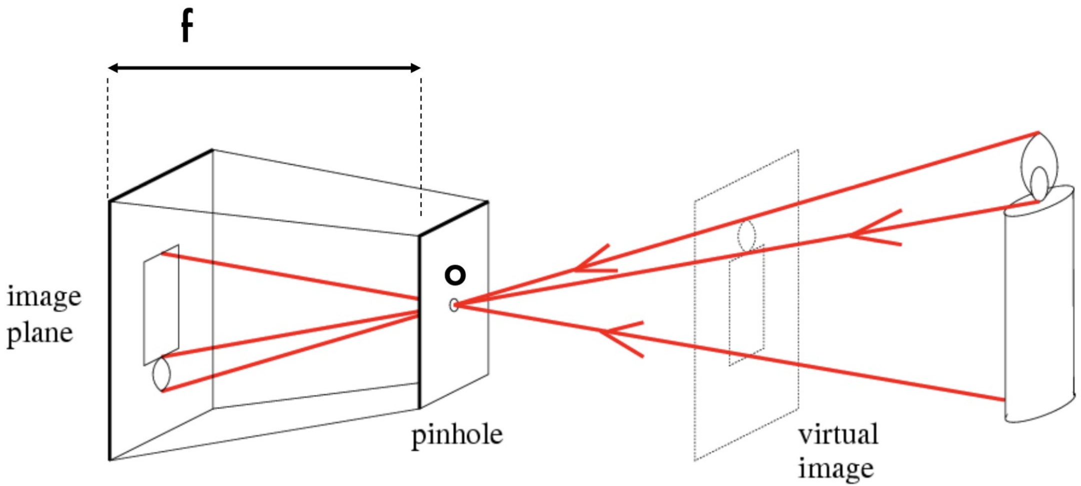


First one to do it (that we know about…)



Leonardo da Vinci (1452-1519)

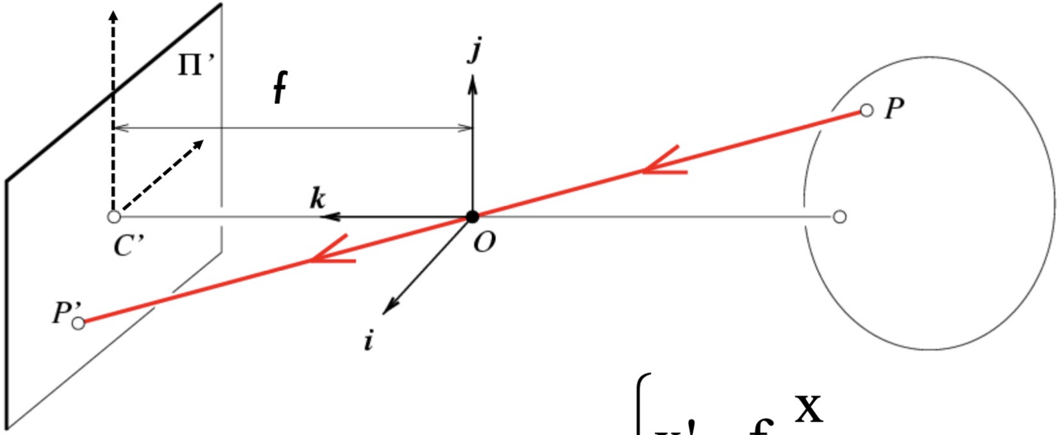
Pinhole Camera Model



slides credit to Prof. Savarese

f = focal length
o = aperture = pinhole = center of the camera

Pinhole Camera Model

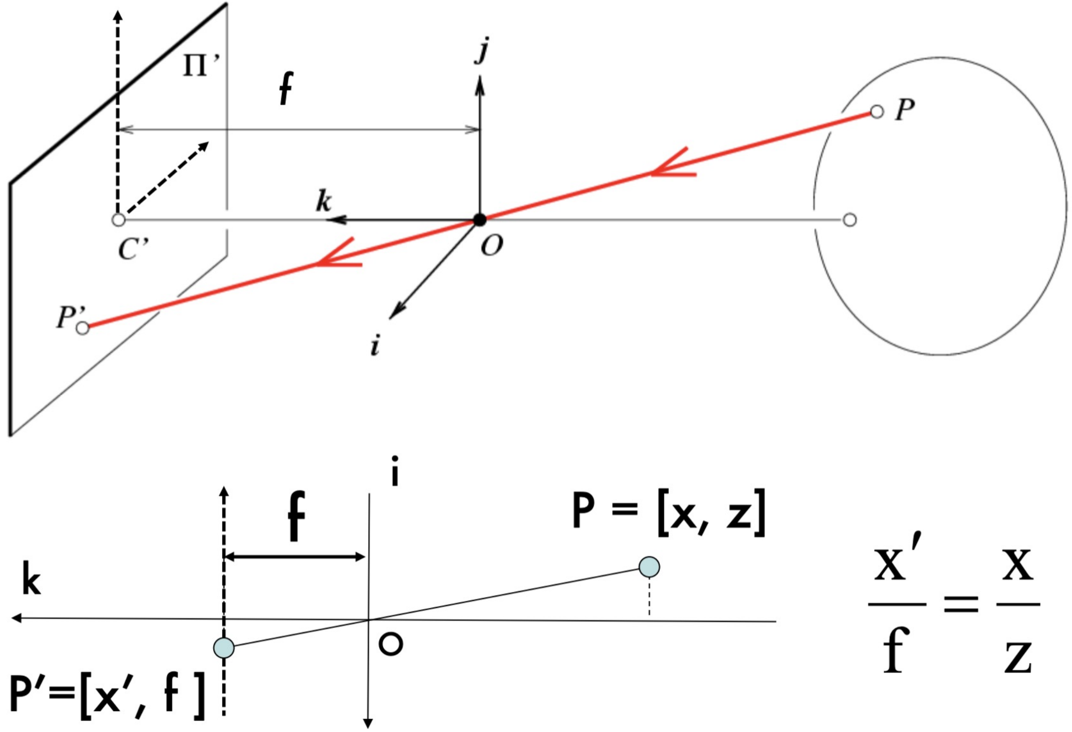


$$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

$$\begin{cases} x' = f\,\dfrac{x}{z} \\ y' = f\,\dfrac{y}{z} \end{cases}$$

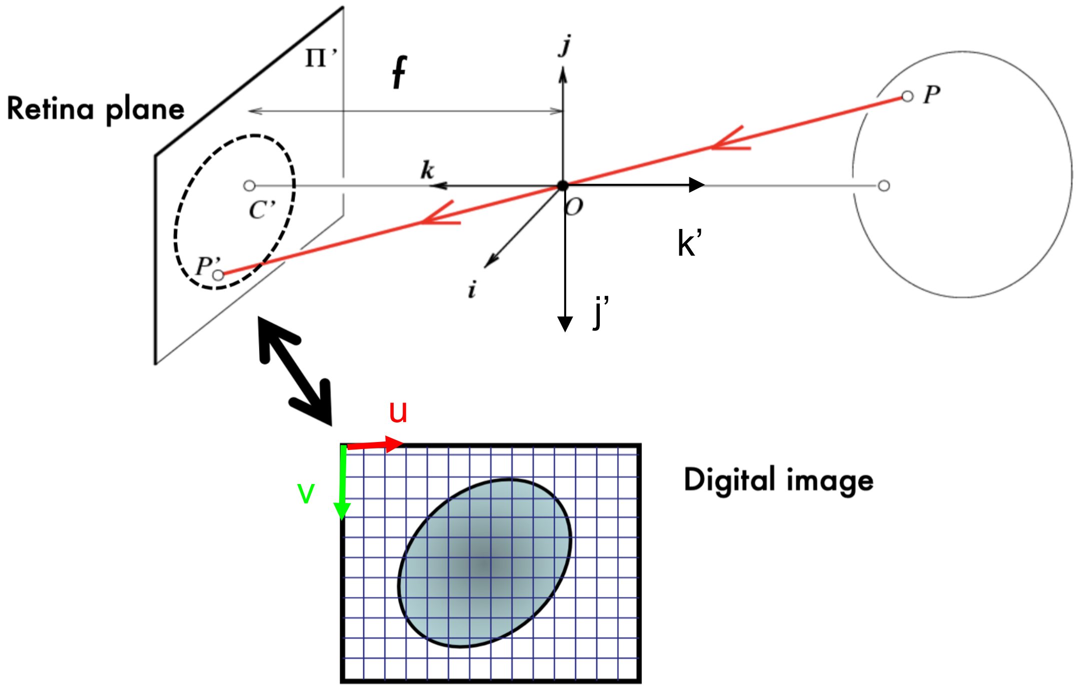
Derived using similar triangles

Pinhole Camera Model

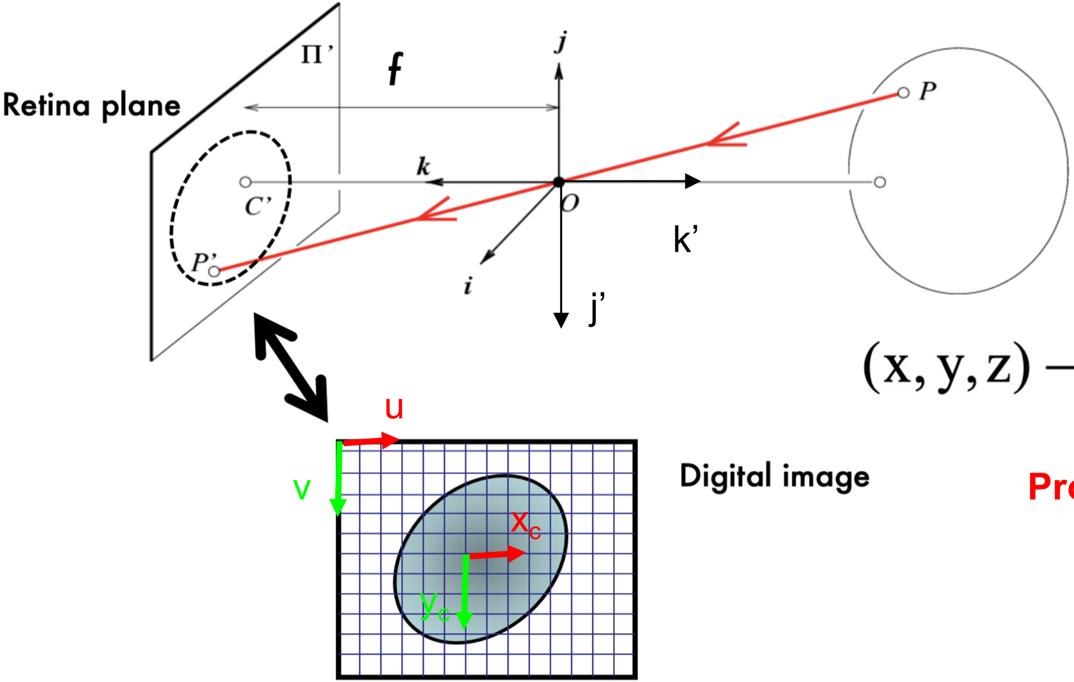


$$\frac{x'}{f} = \frac{x}{z}$$

Digital Image



Offset to Image Center



$$(x, y, z) \rightarrow (f\frac{x}{z} + c_x, f\frac{y}{z} + c_y)$$

Projective Transformation

Homogeneous Coordinates

E→H

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene coordinates

- **Converting back *from* homogeneous coordinates**

H→E

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Projective Transformation with Homogeneous Coordinates

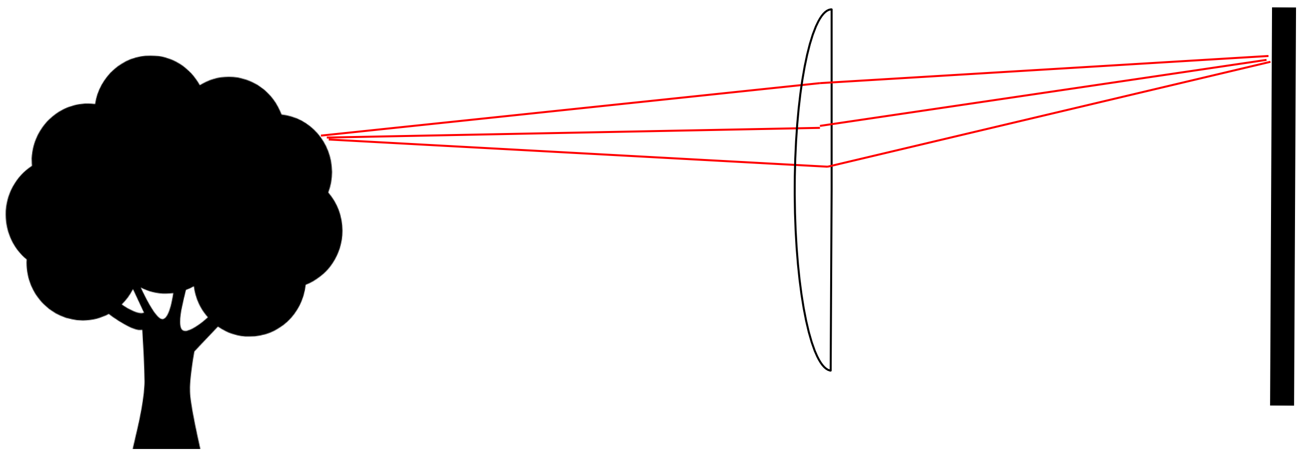
$$P_h' = \begin{bmatrix} f_x x + c_x z \\ f_y y + c_y z \\ z \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}}_{P_h}$$

[Eq.8]

$$K = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

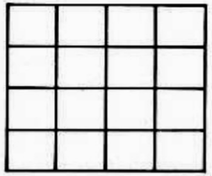
$$\overbrace{P_h'}^{\text{Homogenous}} \rightarrow \overbrace{P' = \left(f_x \frac{x}{z} + c_x, f_y \frac{y}{z} + c_y\right)}^{\text{Euclidian}}$$

Camera Lenses



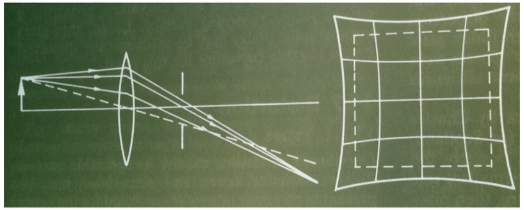
slides credit to Prof. Savarese

Problem: Radial Distortion



No distortion

Pin cushion



Barrel (fisheye lens)

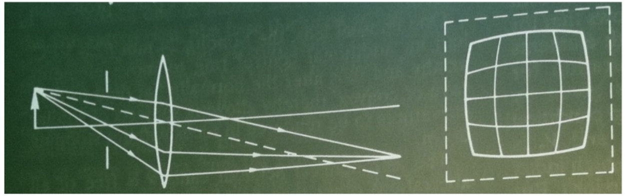
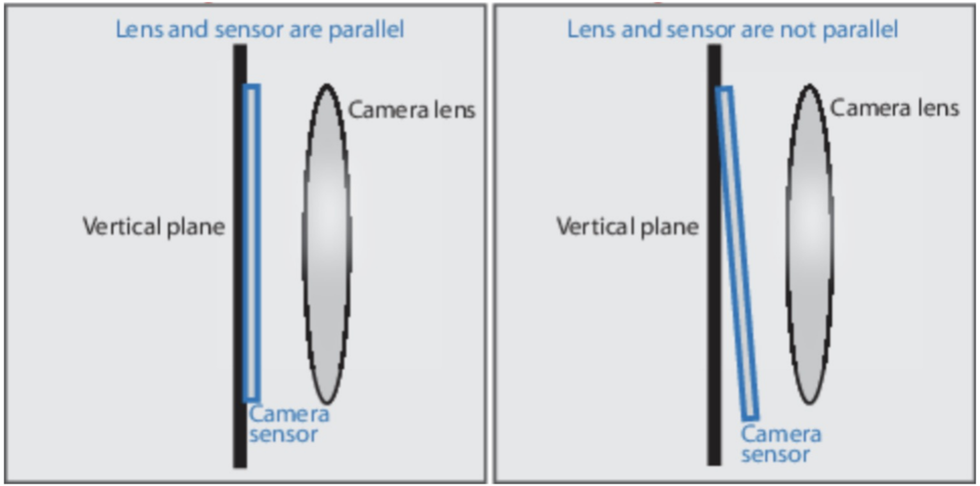


Image magnification decreases
with distance from the optical axis

Problem: Tangential Distortion



Modeling Distortion: Plumb Bob Model

$$p_c = \begin{pmatrix} x_c \\ y_c \end{pmatrix} = \begin{pmatrix} f_x \frac{x}{z} \\ f_y \frac{y}{z} \end{pmatrix}$$

$$p'_c = p_c \cdot (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + \begin{pmatrix} 2t_1 x_c y_c + t_2 (r^2 + 2x_c^2) \\ t_1 (r^2 + 2y_c^2) + 2k_2 x_c y_c \end{pmatrix}$$

Radial distance $r^2 = x_c^2 + y_c^2$

Distortion parameters $d = (k_1, k_2, k_3, t_1, t_2)$

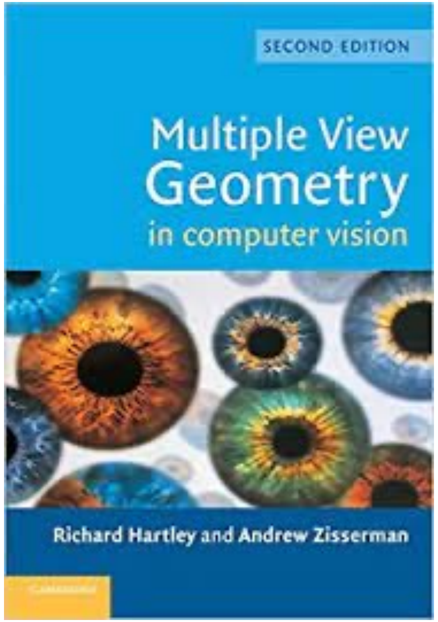
Quick Review of Projective Geometry: Resources

Online Course

- CS231A: Computer Vision
  http://vision.stanford.edu/teaching/cs231a_autumn1112/lecture/

Textbook:

- Multiple View Geometry. Richard Hartley and Andrew Zisserman (some content: https://www.robots.ox.ac.uk/~vgg/hzbook/)



Resources

Related courses at UTCS

- CS342: Neural Networks
- CS 376: Computer Vision
- CS 378 Autonomous Driving
- CS 393R: Autonomous Robots
- CS394R: Reinforcement Learning: Theory and Practice

Extended readings:

- Action-based Theories of Perception, Stanford Encyclopedia of Philosophy
- Action in Perception, Alva Noë