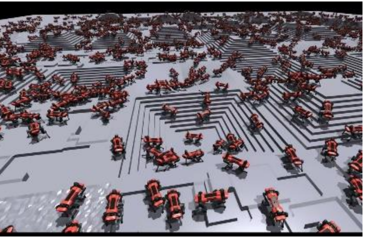
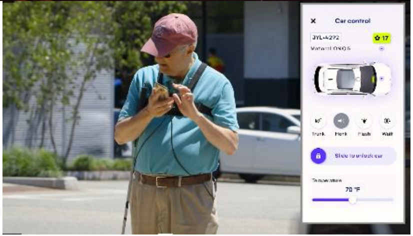
EC500: Robot Learning and Vision for Navigation



Eshed Ohn-Bar

January 23, 2023



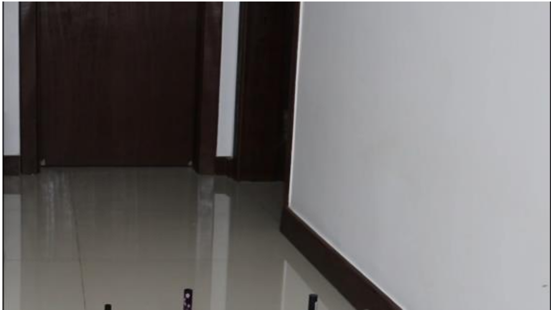
<u>Overarching Objective for the Class</u>

- Holistic understanding of ***<u>fundamental paradigms and challenges</u>*** in learning-based autonomous systems.

- Formulation of in-depth, critical analysis and communication of key concepts, terminology, ***<u>state-of-the-art research approaches</u>*** robotics

We must perceive in order to move, but we must also move in order to perceive
- Gibson (1979)

Why Navigation?



“**Come to Massachusetts** to test your cars, we have bad roads, worse weather, even worse drivers,”

- Transportation Secretary Stephanie Pollack

GO BOSTON 2030

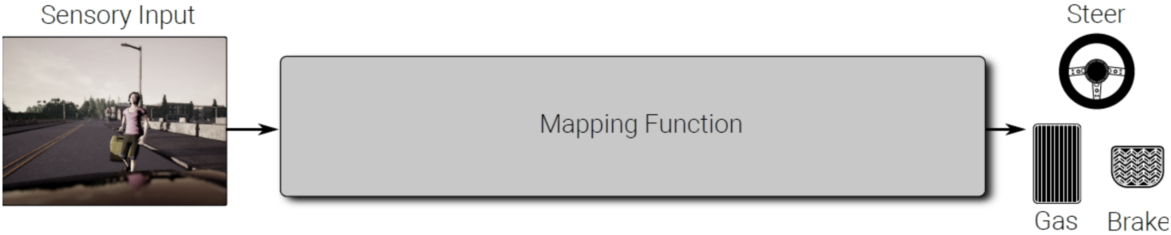
Imagining Our Transportation Future

World Health Organization

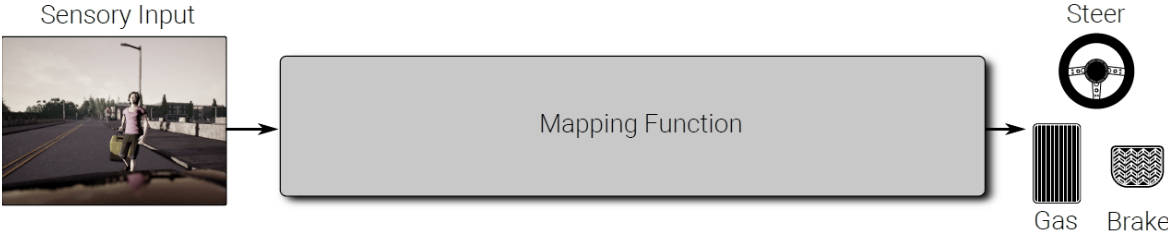
Death from car accident

~ one person every 25 seconds

Robot Learning and Vision ***for*** ***Navigation***



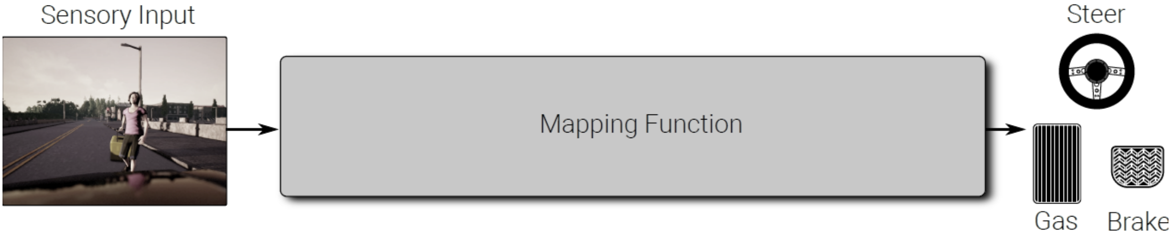
Robot Learning and Vision ***for Navigation***



- Massive technological challenge, ***real-world applications*** in safety and assistive technologies

- Embodied intelligence, system can ***act and move***

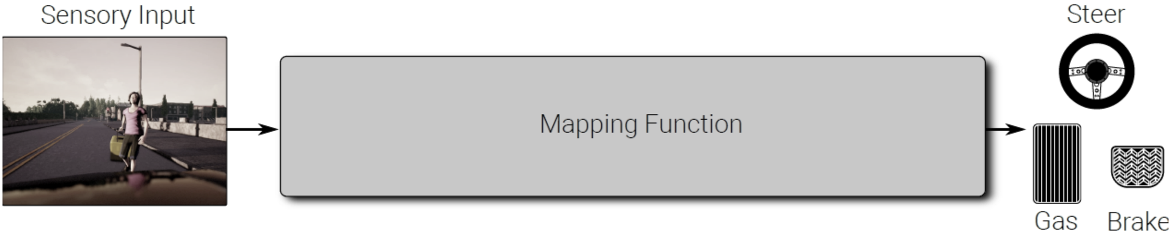
Robot Learning and ***Vision*** for Navigation



- How to code this?

If Pedestrian then
Slow down
If Traffic Light is Red then Stop
…

Robot Learning and ***Vision*** for Navigation



- How to code this?

If **Pedestrian** then

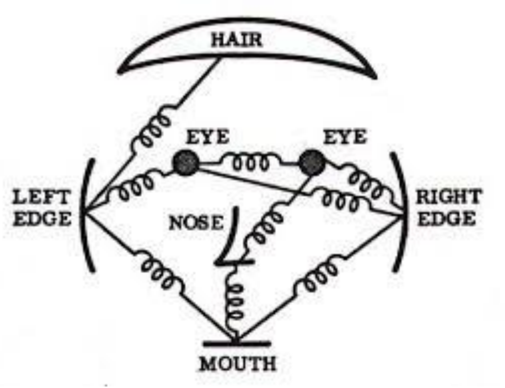
Slow down

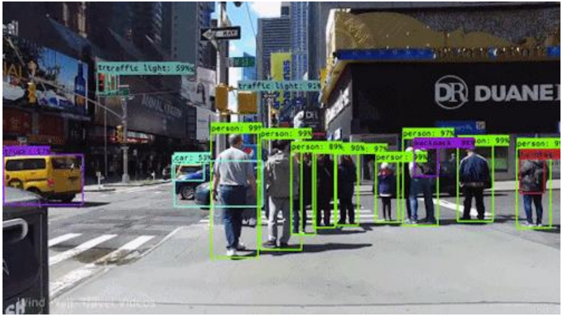
If Traffic Light is Red then Stop

…

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Computer Vision – Then and Now



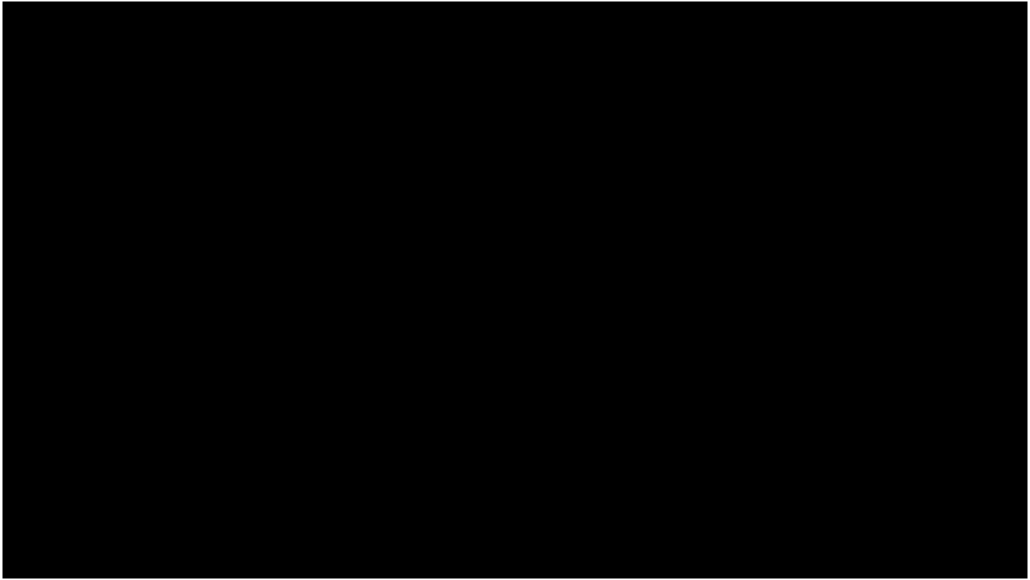


Geometry and deformable parts, object detection

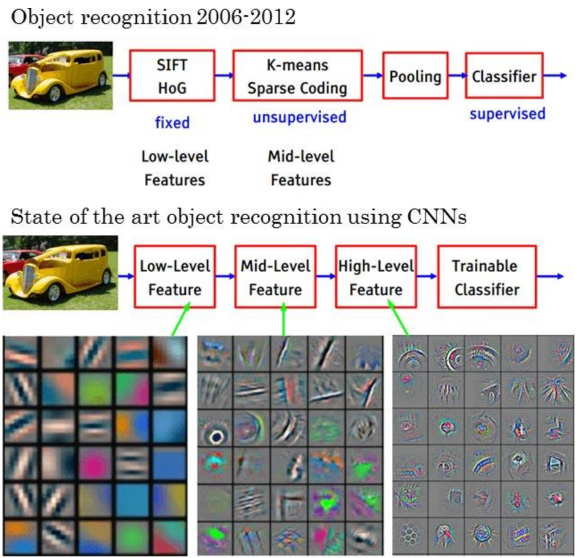
Otherwise, Could Miss (Rare) but Important Objects!

Otherwise, Could Miss (Rare) but Important Objects!

Accident with Uber Vehicle

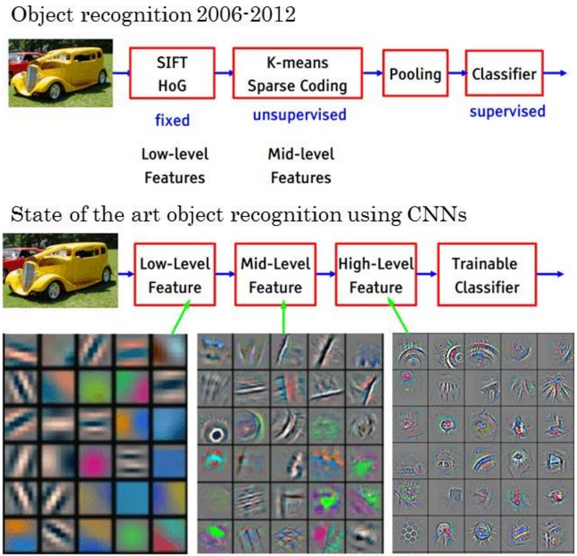


Computer Vision – Then and Now



Taken from Y. LeCunn

Computer Vision – Then and Now



- What should be the ***output***?
- What supervision?
- Complex tasks?
- Agent is passive

Taken from Y. LeCunn

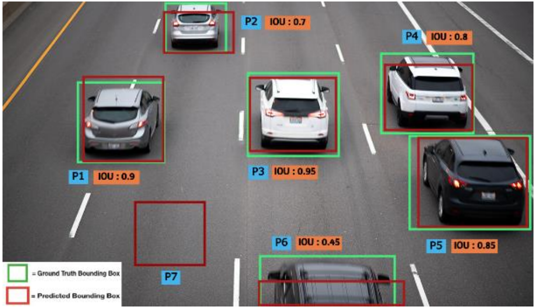
On the Importance of Embodied Intelligence

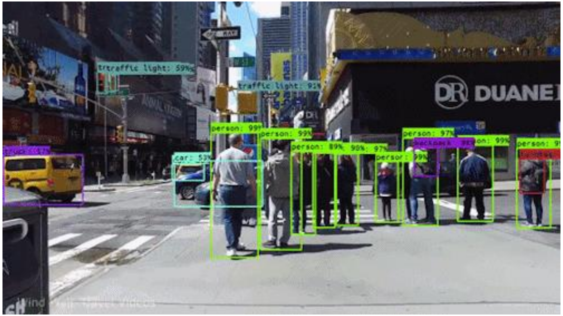
Which of the Surrounding Agents are More Important?

Which of the Surrounding Agents are More Important?

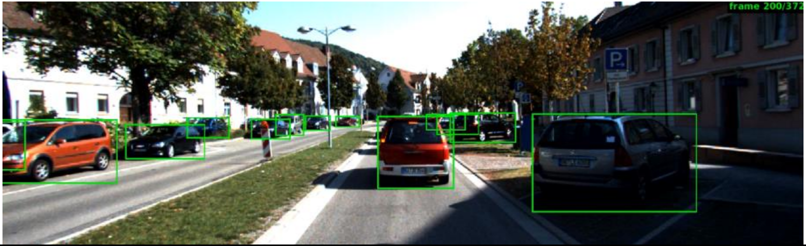
Which of the Surrounding Agents are More Important?

Computer Vision – Not Optimizing for the Right Thing!

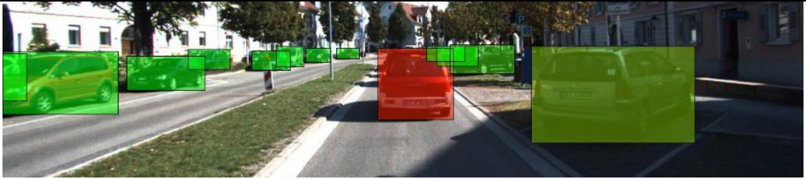




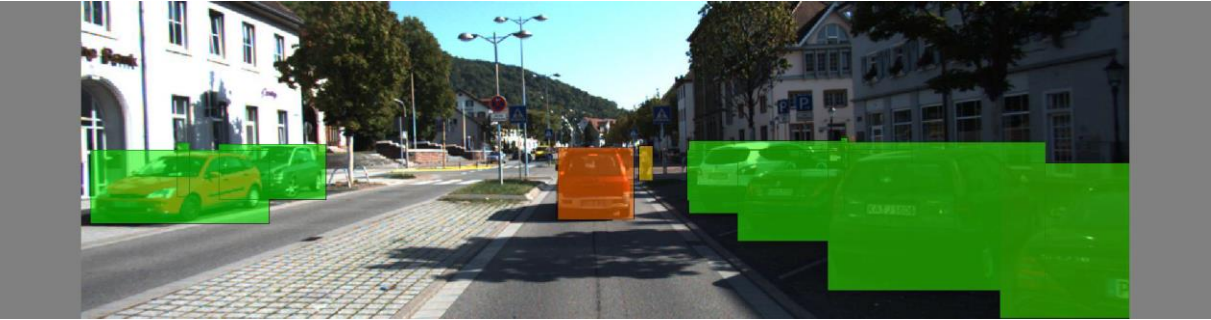
Modeling, Training, Evaluation Should Consider Navigation Task



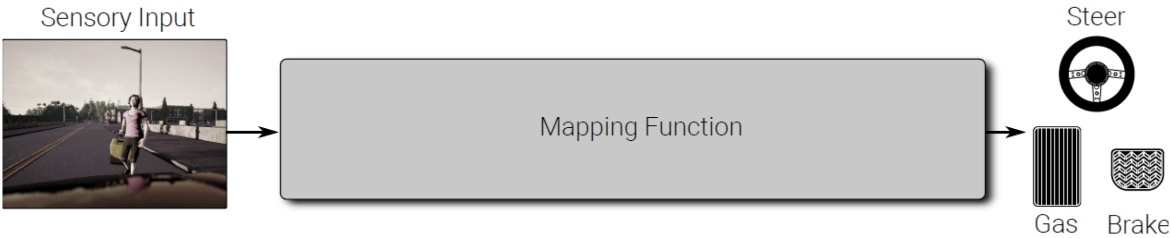
Color coded from high to moderate to low importance



Ohn-Bar et. al, ICPR 2016, Pattern Recognition 2017



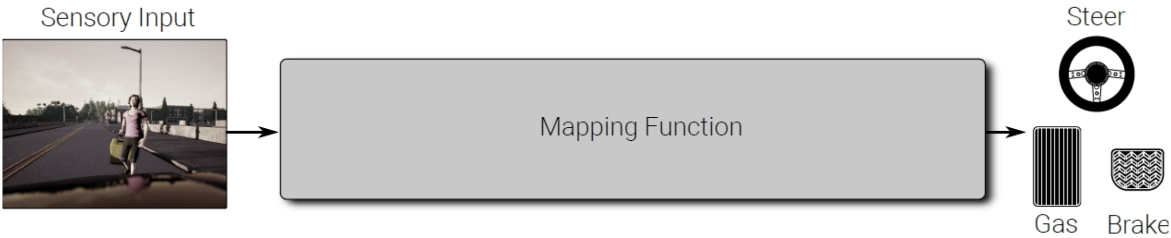
Robot Learning and Vision ***for Navigation***



- Massive technological challenge, ***real-world applications*** in safety and assistive technologies

- Embodied intelligence, system can ***act and move***

Robot Learning and Vision for Navigation



- How to code this?

  If Pedestrian then

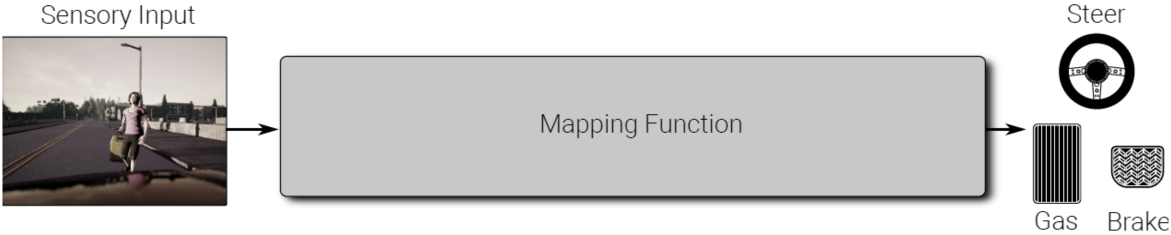
  Slow down

  If Traffic Light is Red then Stop

  …

Learning for Autonomous Navigation

Robot Learning and Vision for Navigation

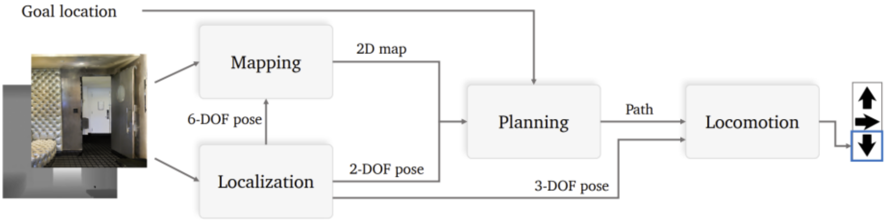


- In robotics research, vision is ***assumed a solved task***.
- Perfect knowledge of the state
- ***Decoupled*** perception from action

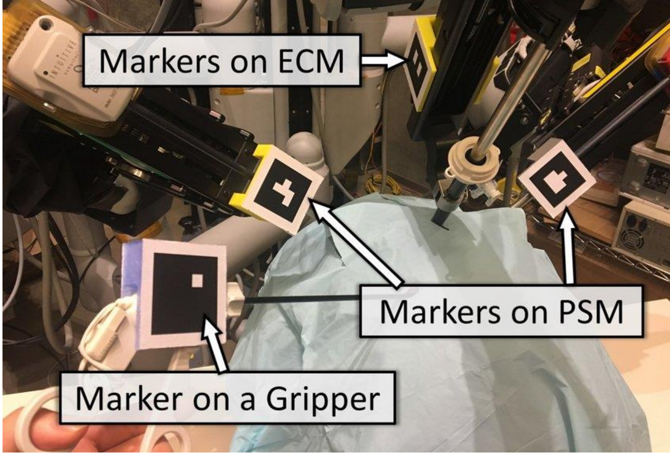
Classic sensorimotor control pipeline

- Perception system constructs a map of the environment
- Planner generates waypoints
- Continuous controller (e.g., proportional-derivative) actuates

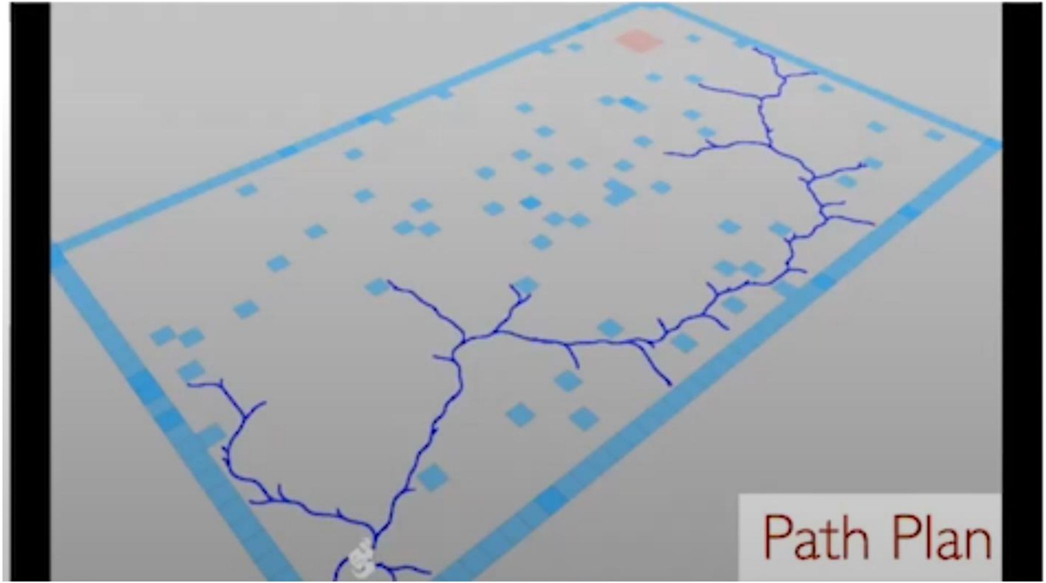
- **What matters is action**
- SLAM (Simultaneous Localization and Mapping) is solving a problem that may be both unnecessarily hard and incomplete
- Biological motivation suggests alternative approaches



Mishkin et al, Benchmarking Classic and Learned Navigation in Complex 3D Environments, 2019


Markers on ECM
Markers on PSM
Marker on a Gripper


Boston Dynamics



Rapidly exploring random trees, Lavalle, Kuffner

The planner will take care of it

- SLAM does not tell us how to explore, navigate, accomplish goals
- But planners assume that the map is accurate and sufficiently complete
- When this assumption breaks, the whole system can fail

Hints from Natural World

Critique of Pure Vision

Vision, like other sensory functions, has its evolutionary rationale rooted in improved motor control. Although organisms can of course see when motionless or paralyzed, the visual system of the brain has the organization, computational profile, and architecture it has in order to facilitate the organism's thriving at the four Fs: feeding fleeing, fighting, and reproduction. By contrast, a pure visionary would say that the visual system creates a fully elaborated model of the world in the brain, and that the visual system can be studied and modeled without worrying too much about the nonvisual influences on vision.
- Churchland, Ramachandran, and Sejnowski (1994)

https://papers.cnl.salk.edu/PDFs/A%20Critique%20of%20Pure%20Vision%201994-2933.pdf

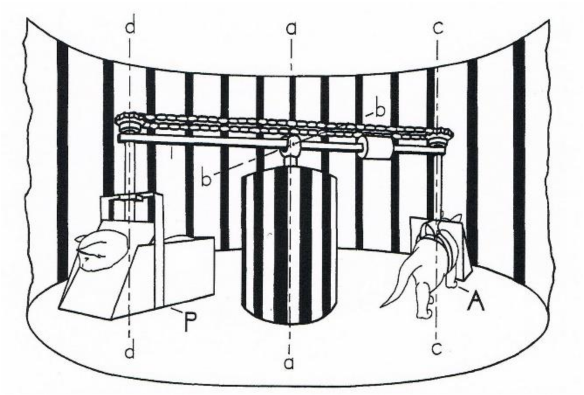
Critique of Pure Vision

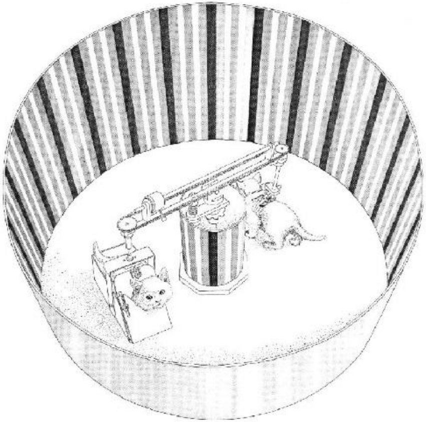
Vision, like other sensory functions, has its evolutionary rationale rooted in improved motor control. Although organisms can of course see when motionless or paralyzed, the visual system of the brain has the organization, computational profile, and architecture it has in order to facilitate the organism's thriving at the four Fs: feeding fleeing, fighting, and reproduction. By contrast, a pure visionary would say that the visual system creates a fully elaborated model of the world in the brain, and that the visual system can be studied and modeled without worrying too much about the nonvisual influences on vision.
- Churchland, Ramachandran, and Sejnowski (1994)

Critique of Pure Action!

https://papers.cnl.salk.edu/PDFs/A%20Critique%20of%20Pure%20Vision%201994-2933.pdf

On the Importance of Interactive Intelligence

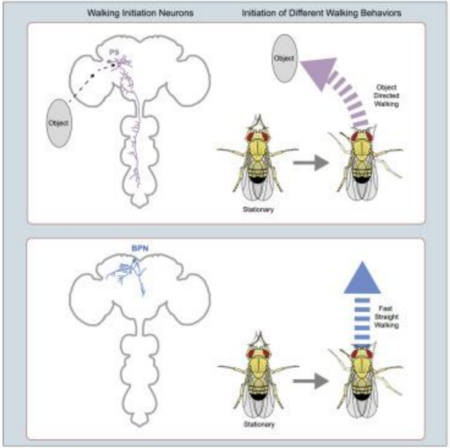




Held and Hein: Movement-produced stimulation in the development of visually guided behavior. Journal of Comparative and Physiological Psychology, 1963.

On the Importance of Interactive Intelligence

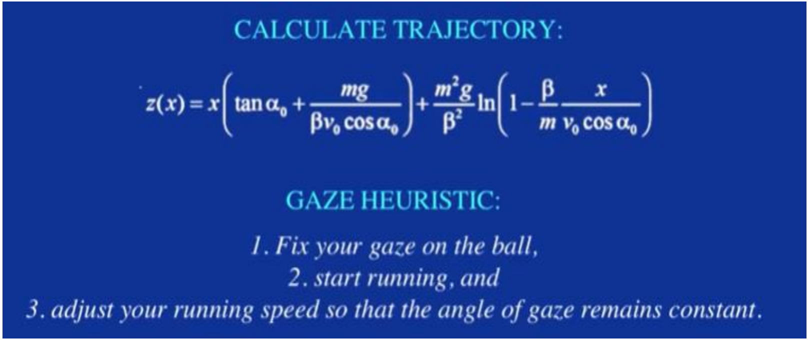
the brain switches perception and response to sensory stimuli depending on internal state and needs



Two Brain Pathways Initiate Distinct Forward Walking Programs in *Drosophila, Neuron, 2020*
Sten et al., An arousal-gated visual circuit controls pursuit during *Drosophila* courtship, bioRxiv, 2020

On the Importance of Interactive Intelligence

The ***Gaze Heuristic*** – how to catch a flying ball?



Gigerenzer, "Homo Heuristicus: Why Biased Minds Make Better Inferences“, 2009


ViralHog

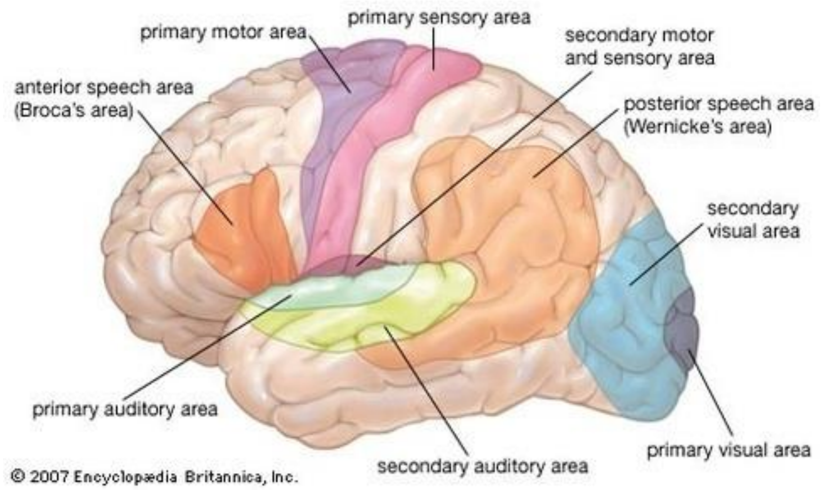
expression of 24 olfactory receptor genes in the AWA olfactory neuron is influenced by a wide array of states and 25 stimuli, including feeding state, physiological stress, and recent sensory cues

Diverse states and stimuli tune olfactory receptor expression levels to modulate food-seeking behavior, bioRxiv, 2022

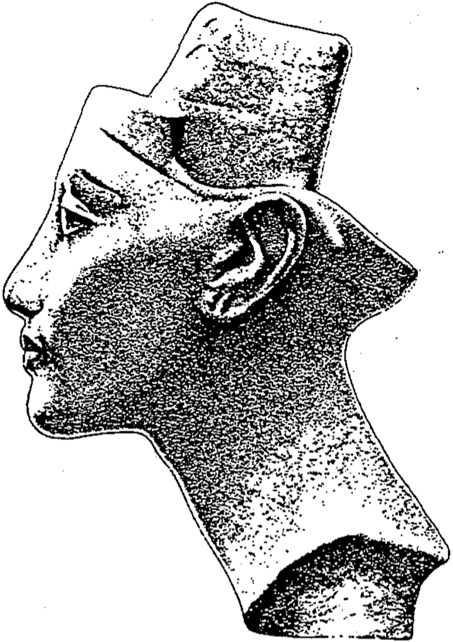
https://www.biorxiv.org/content/10.1101/2022.04.27.489714v1.full.pdf

Acoustic motion of a spider orb-web

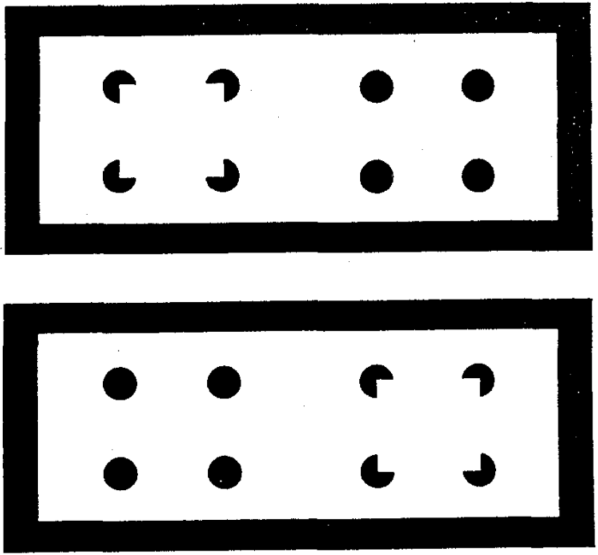
Outsourced hearing in an orb-weaving spider that uses its web as an auditory sensor, PNAS, 2022



Keller GB, et al (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. *Neuron*.




TikTok
@freighttrainofficial



Contents

Goal: How can we design robotic systems that robustly operate in complex, dynamic, and uncertain real-world environments?

- How to formulate learning algorithms for general robotics and sensori-motor navigation
- Explore challenges in perception-based control pipelines
- Affordance-based and goal-directed representation of semantic scene, motion, and geometry
- Limitations (and complementarity) of reinforcement, inverse reinforcement, and imitation learning approaches
- Policy optimization (e.g., model-based vs. model-free) and sample efficiency analysis
- Multi-modal algorithms for multi-sensor (camera, LIDAR and radar perception) and multi-task learning
- Human-in-the-loop and social learning
- Reading and presenting recent research papers
- Participation in a competition to put these ideas to the test (for awards and fame)

Tentative Agenda

Date	Topic	Notes	
1/23	Introduction: Why Navigation?	HW0: Due 1/30	
1/25	Deep Imitation Learning		
1/30	Deep Imitation Learning	HW1: Due in 2 weeks	
2/1	Affordances and Direct Perception		
2/6	Dynamics and Localization		
2/8	Semantic Scene Understanding		
2/13	Semantic Scene Understanding	HW2: Due in 2 weeks	
2/15	Object Detection and Tracking		
2/21	Object Detection and Tracking	Project Proposal Reports Due	
2/22	Markov Decision Process	HW3: Due in 2 weeks	
2/27	Q-Learning and SARSA		
3/1	Double Q-Learning and Deep RL for Robotics	Competition Presentations	
3/6	Spring Recess		
3/8	Spring Recess		
3/13	TRPO, DDPG, PPO	Project Proposal Revision Due	
3/15	LfD/Inverse Reinforcement Learning		
3/20	LfD/Inverse Reinforcement Learning		
3/22	Human-in-the-Loop Learning		
3/27	Human-in-the-Loop Learning		
3/29	Midterm Project Updates		
4/3	Model-based RL		
4/5	Model-based RL		
4/10	Sim2Real		
4/12	Bio-Inspired and Evolutionary Techniques		
4/17	Patriot's Day		
4/19	Incorporating Language		
4/24	Partial Observability		
4/26	Human-Robot Interaction		
5/1	Ethics and Social Implications		
5/3	Research Presentations		
5/10	Research Paper Submission		

Team

Zhongkai Shangguan

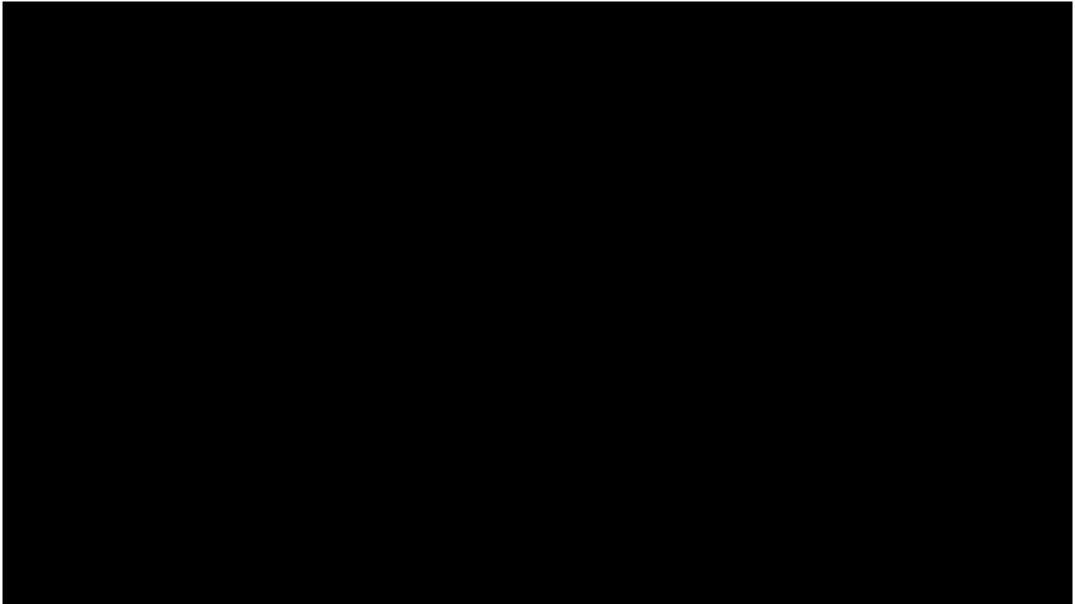
Jimuyang Zhang

Office hours: Wednesdays after class, 6:15-7:15?

Grading

- Homework: 30%
- Presentations and participation: 20%
- Final Project: 50%

HW 0 details (not graded)



HW 0 details (not graded)

Exercises

0.1 SCC Setup (0 Points)

We recommend using the SCC desktop version for a better interactive experience. For more details, you can refer to https://www.bu.edu/tech/support/research/system-usage/connect-scc/scc-ondemand/.

a) Read the cluster documentation and log in to the cluster after you get the account.

b) Apply for a desktop in Interactive Apps with 4 cores and 1 GPU, and load python3/3.8.3, libjpeg-turbo/2.0.4, and any other modules which you think is useful.

c) Connect to your desktop and you are now having a remote computer with powerful computing resources!

0.2 Anaconda Installation (0 Points)

To ensure a working environment, we ask you to install and create an Anaconda environment on your machine:

a) It is recommended to manage your Python environment with Anaconda. Please install Anaconda following instructions at https://www.anaconda.com/download/.

b) Create and activate your own environment by *conda create -n < your_env_name > python=3.7*. You can choose any other python version > 3.7.

c) Install PyTorch following instructions at https://pytorch.org/get-started/locally/ in your environment.

d) We ask you to run the toy regression code `exercise_pytorch.py` we provided. Please save the output log in the terminal as a `.txt` file.

0.3 CARLA Simulator Installation (0 Points)

In most of our homework, we use CARLA as our environment:

a) Read documents and get to know what CARLA is and what can it do https://carla.org/.

b) Download ubuntu version CARLA 0.9.10.1 (development) from https://github.com/carla-simulator/carla/releases. In order to avoid bugs caused by version inconsistencies, we restrict the CARLA version to 0.9.10.1.

c) Unzip your CARLA, cd to your CARLA directory, start your carla-server by running *DISPLAY= ./CarlaUE4.sh -opengl*. Note: it is okay to have error messages like "Disabling core dumps." or "error: XDG_RUNTIME_DIR not set in the environment." as long as it is still running.

d) Open a new terminal and import CARLA by :

```
export CARLA\_ROOT=YOUR_CARLA_DRECTORY
export PYTHONPATH=$PYTHONPATH:$CARLA_ROOT/PythonAPI/carla
export PYTHONPATH=$PYTHONPATH:$CARLA_ROOT/PythonAPI/
    ↳ carla/dist/carla-0.9.10-py3.7-linux-x86_64.egg
```

then run the toy manual control code `manual_control.py` under "YOUR_CARLA_DRECTORY/PythonAPI/examples". You are supposed to use ↑↓←→ to control the car in a simulated driving environment. Please make screenshot and compress to your submission .zip.

Today's Readings

- Critique of pure vision,
https://papers.cnl.salk.edu/PDFs/A%20Critique%20of%20Pure%20Vision%201994-2933.pdf

Patricia S. Churchland, V. S. Ramachandran, and, Terrence J. Sejnowsk, 1994