Welcome!

COMP 546 Computational Perception

Prof: Michael Langer

See public web page for this course:

http://www.cim.mcgill.ca/~langer/546.html

What do you know about visual perception ?

- optics (glasses)

- color (color blindness)

- binocular depth perception (3D cinema)

- perspective (art)

- ....

What do you know about auditory perception ?

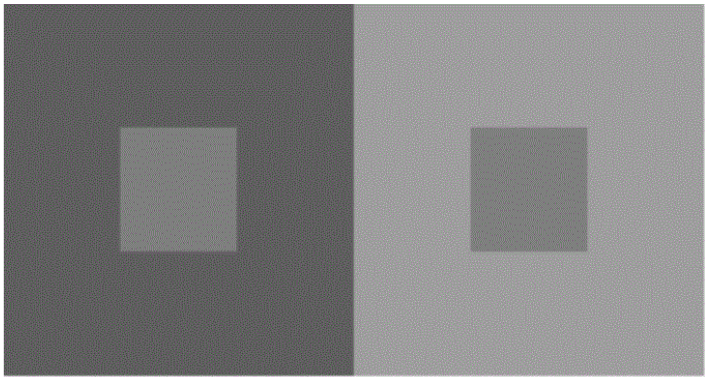
- sound (waves)

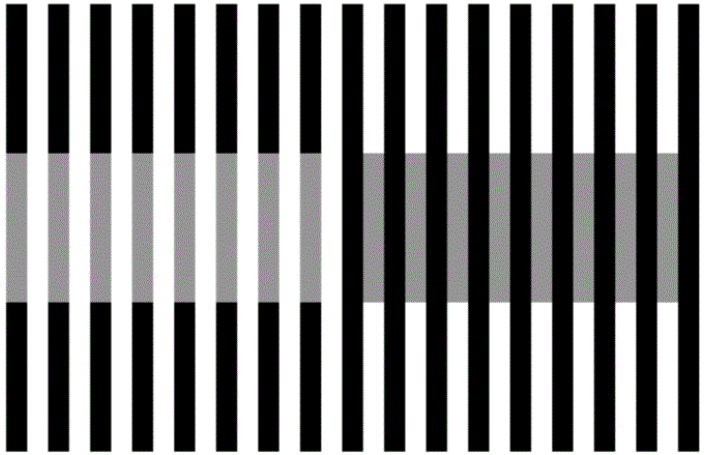
- music (tone related to frequency)

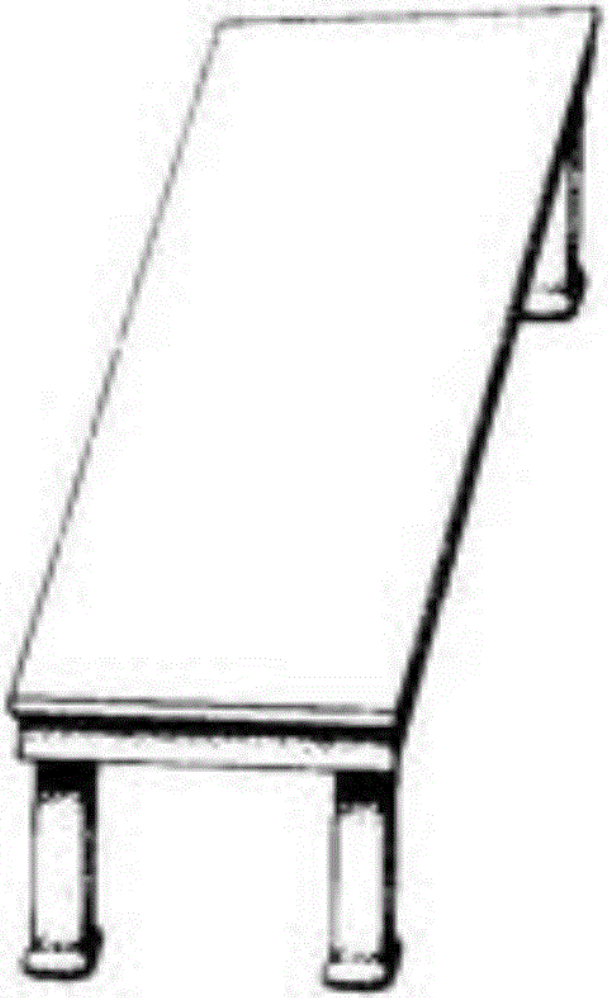
- voice (automatic speech recognition)

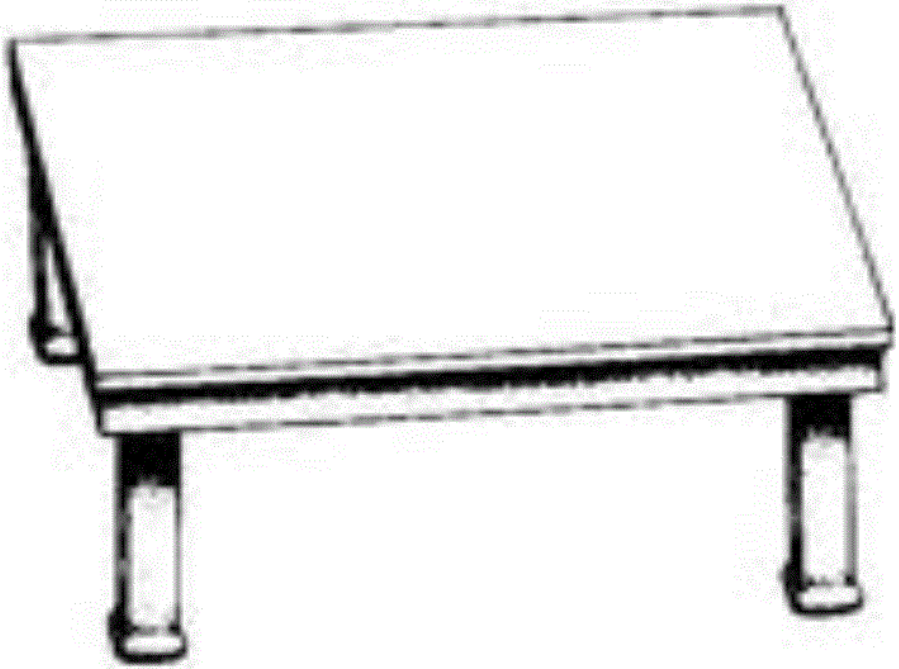
- hearing aids (external vs. cochlear implants)

Perception and Visual Illusions









Sensation and Perception

physical stimulus	sensory organ	sense
light (optics)	eye	vision (seeing)
sound (acoustics)	ear	audition (hearing)
pressure (mechanics)	skin	haptics (touch)
chemistry	mouth, nose	olfaction (taste, smell)

... + proprioception, balance, pain, temperature, nausea,....

Perception is...

... knowing *what* is *where*

(by seeing, hearing, touching, smelling ....)

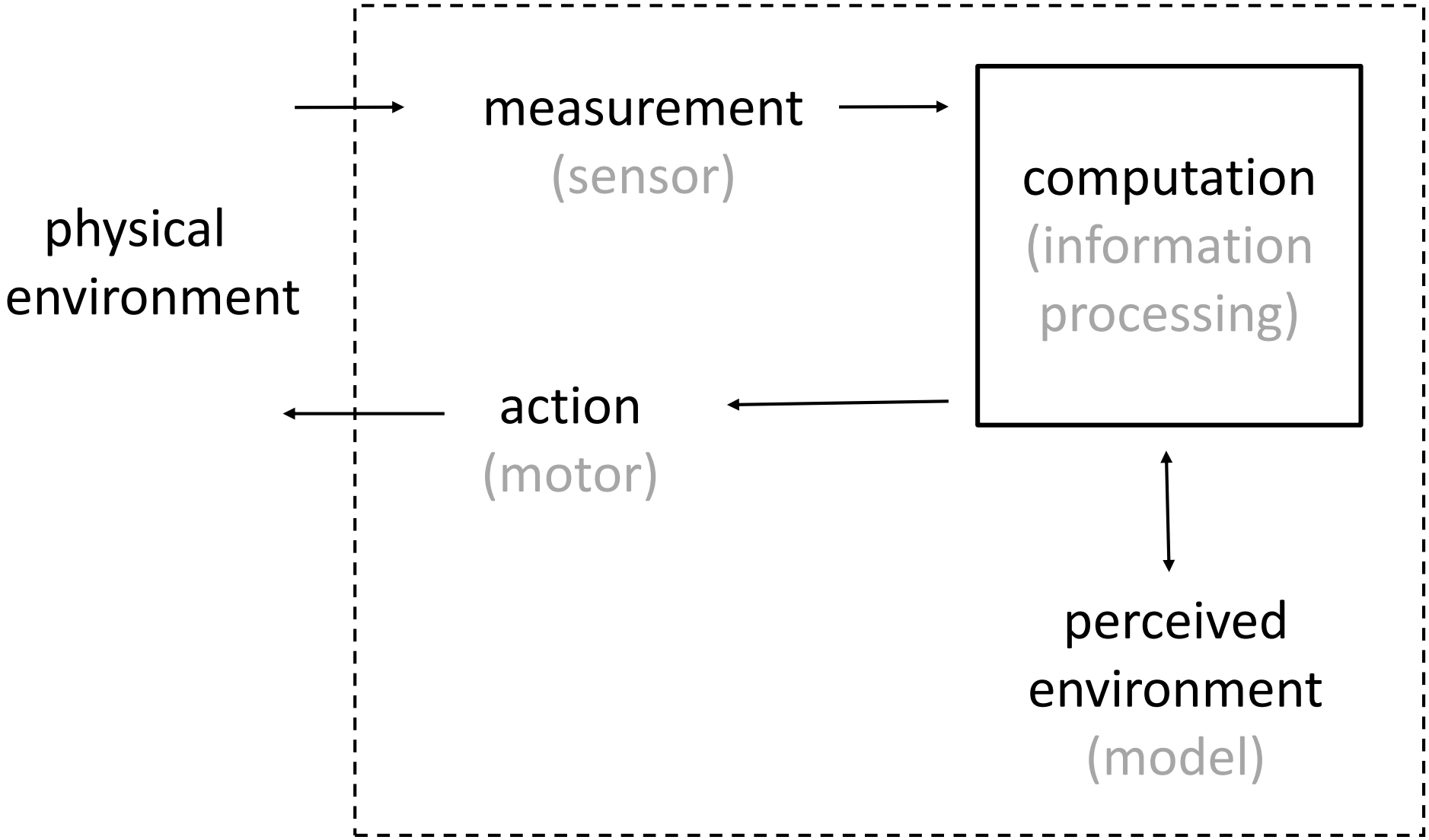
Perception is...

... knowing *what* is *where*

(by seeing, hearing, touching, smelling ....)

... a process

Perception is a *process*.



Philosophical Problems in Perception

physical environment $\neq$ perceived environment

Example: Vision

physical objects

- 3D shape
- 3D position
- material

perceived objects

- 3D shape
- 3D position
- material

Scientific Approaches to Perception

Neuroscience: Physiology, Anatomy, Biology

- Experiments measure individual or populations of neurons, or brain (imaging)

Behavioral Psychology

- experiments that measure performance in a task
  (detection and discrimination, recognition, attention, ...)

Computational Modelling

- computational neuroscience, cognitive science

As we will see, one often combines several of the above. Our emphasis will be on the last of these.

Level of Analysis in Perception

high

- behavior (task)
- brain areas and pathways
- nerve cells and coding
- neuron mechanisms

low

Behavior: What is the task ?

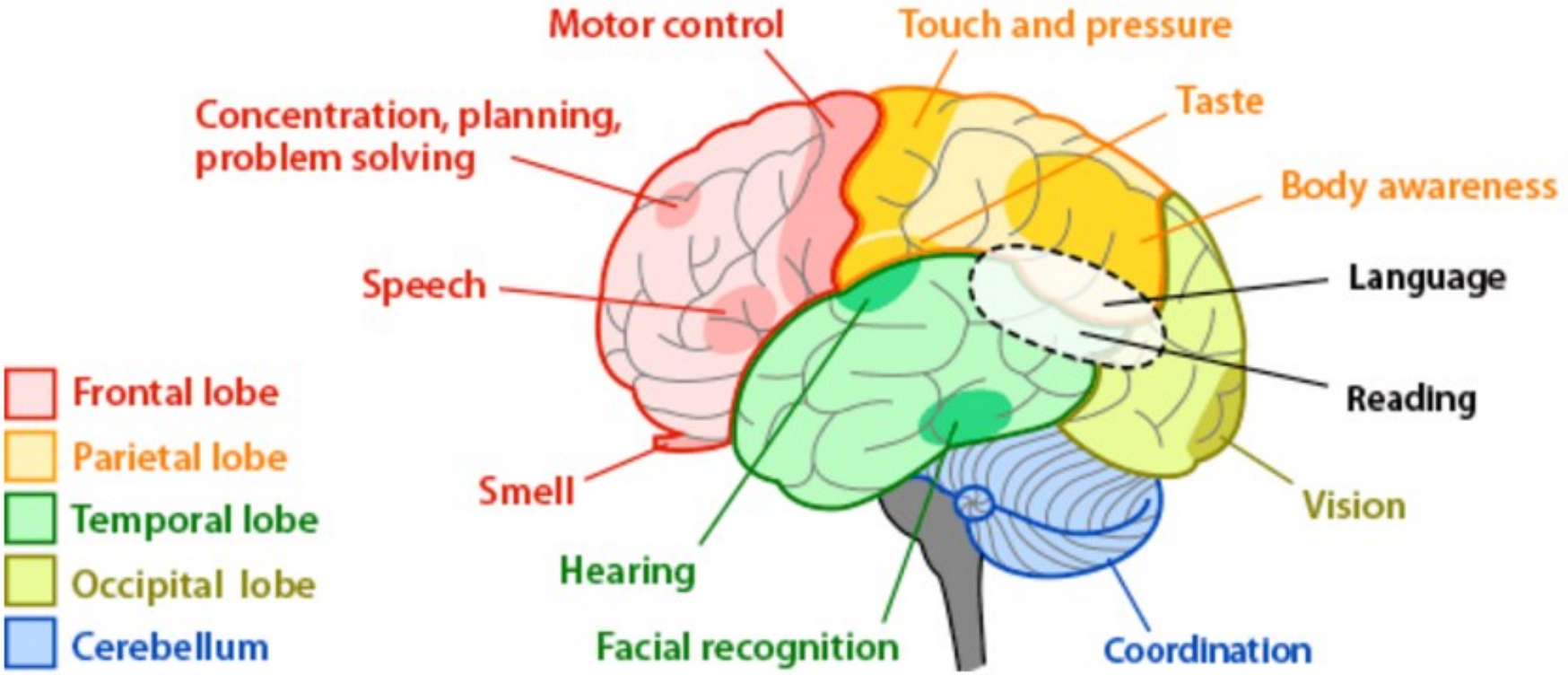
Vision

- Combine images from the two eyes to infer depth and 3D scene layout
- Estimate material and shape ("discounting the illuminant")
- Detect objects and boundaries
- Detect and recognize objects (faces, written characters, ...)
- …..

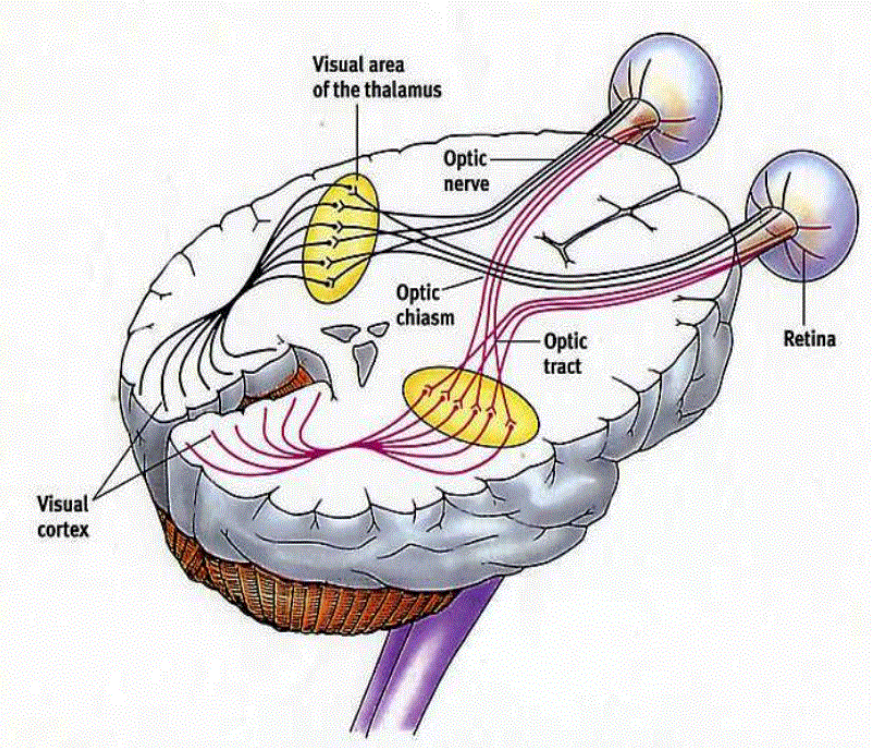
Audition

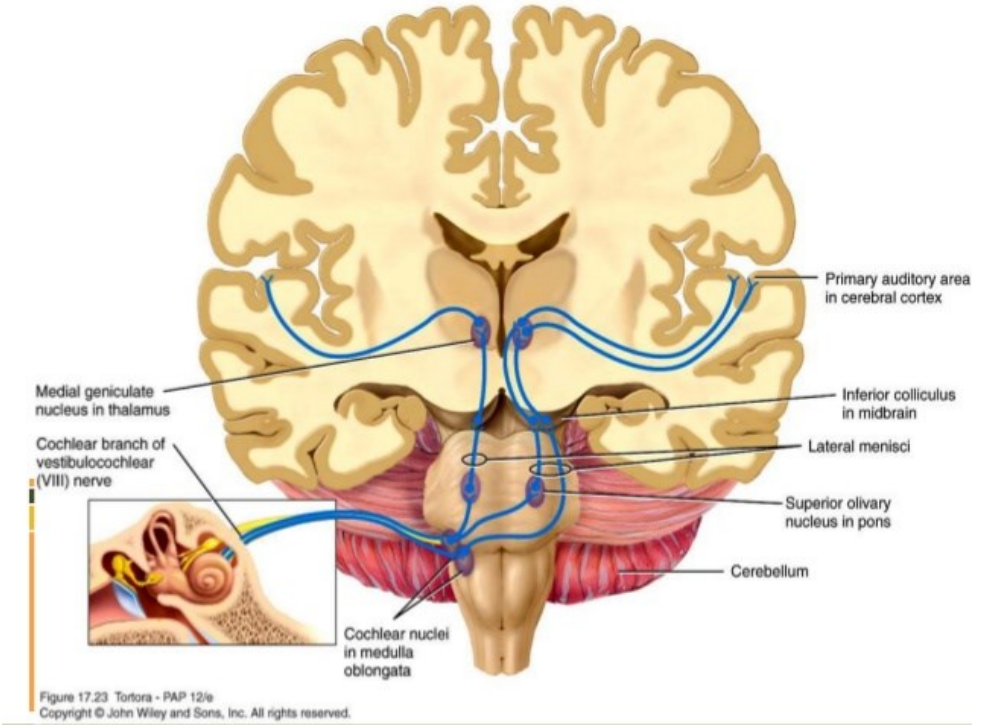
- Combine images from the two ears to infer direction of a sound source
- Estimate source (discount echos)
- Segregate sounds into distinct sources
- Detect and recognize speech sounds or other sounds (musical instruments)
- ….

Brain Areas:
functional specialization of cortex (surface)



Brain Pathways

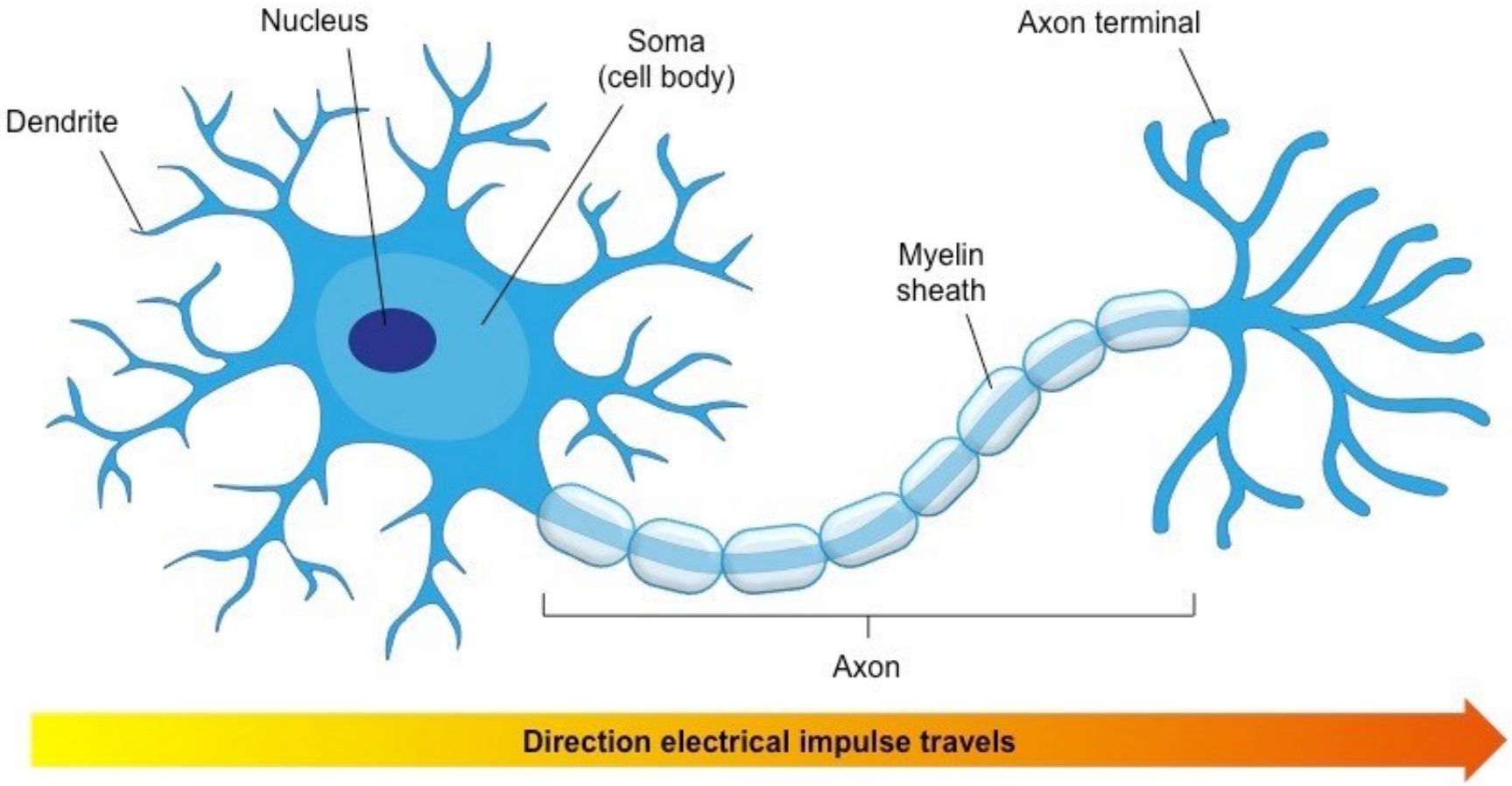




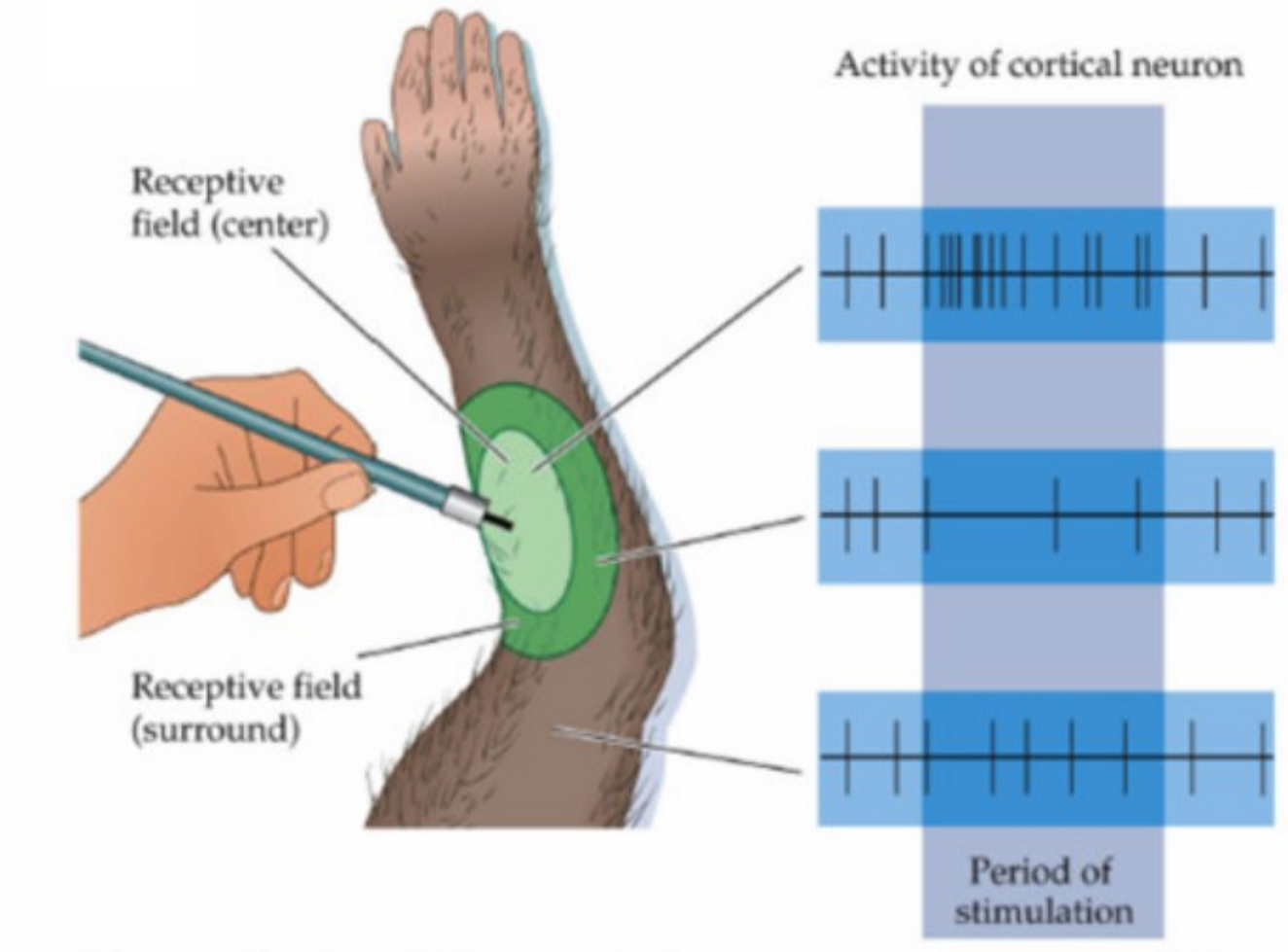
Vision

Audition

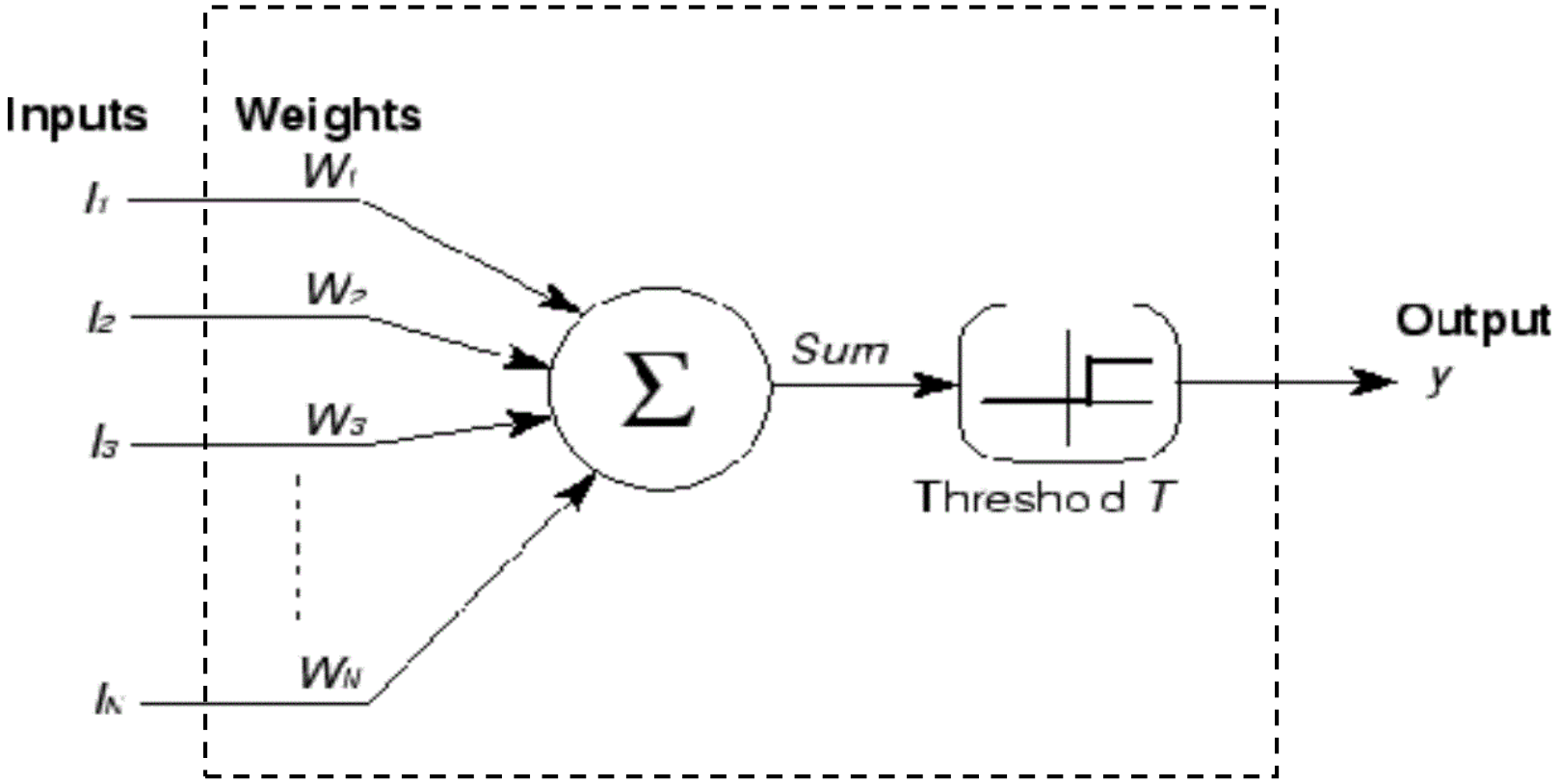
Nerve cell (neuron)



Receptive field of single sensory cell in brain e.g. touch

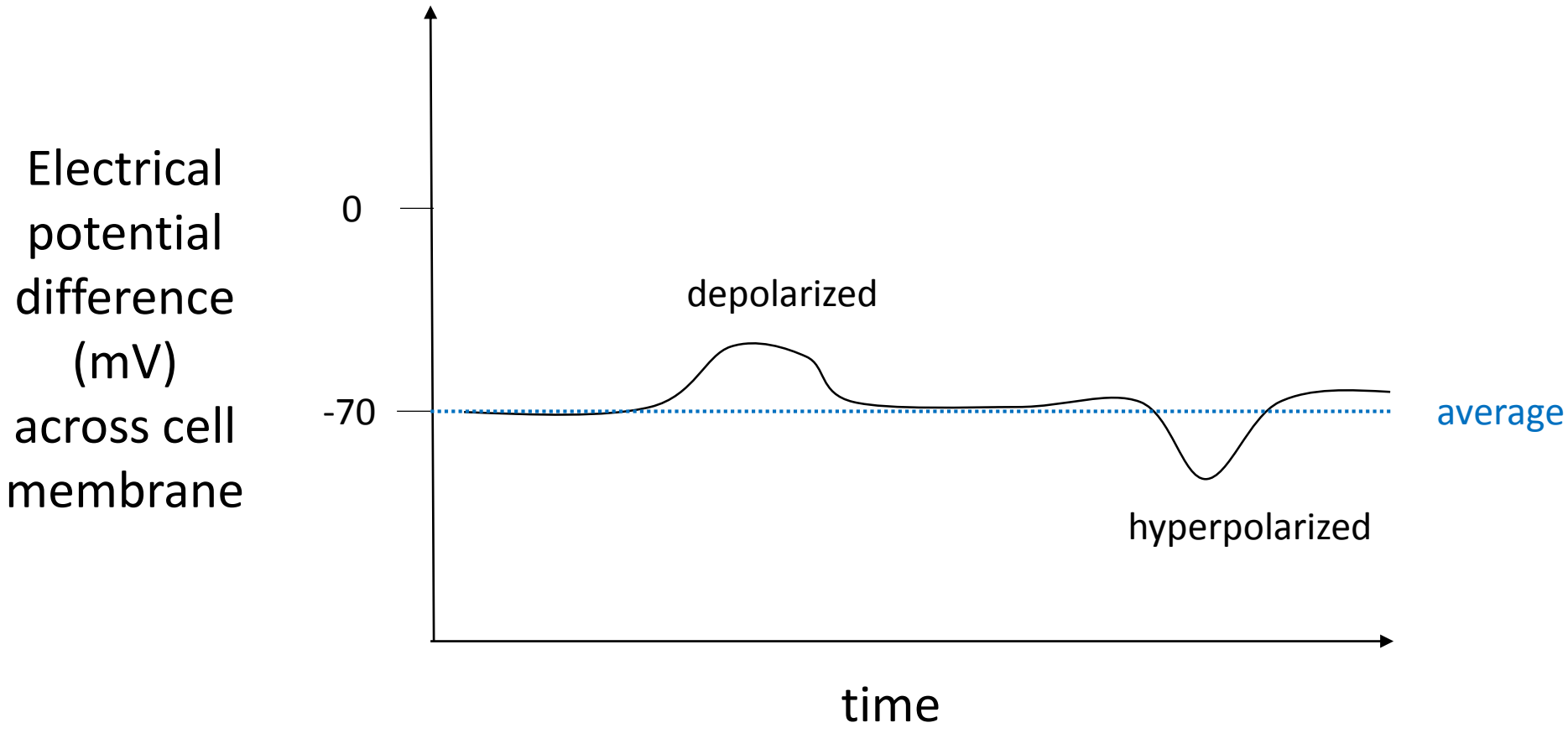


Neural Code:
Model of Neuron Response

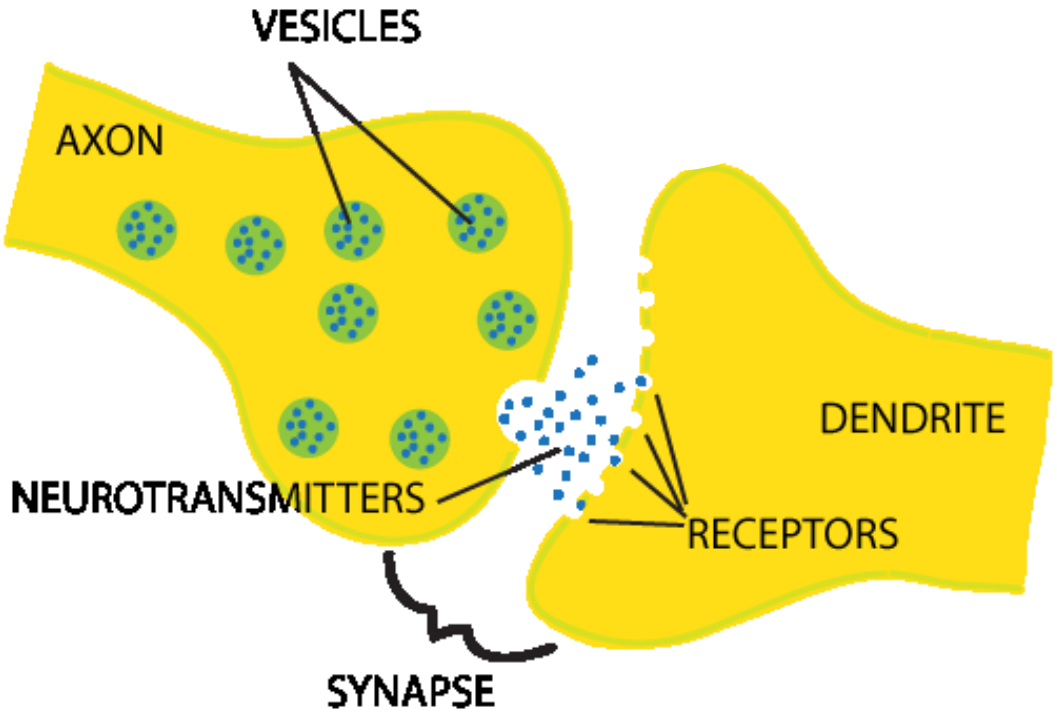


McCulloch-Pitts (1943)

Single neuron Mechanism
(activity = membrane potential)



Single neuron Mechanism
(Signalling between cells: the synapse)

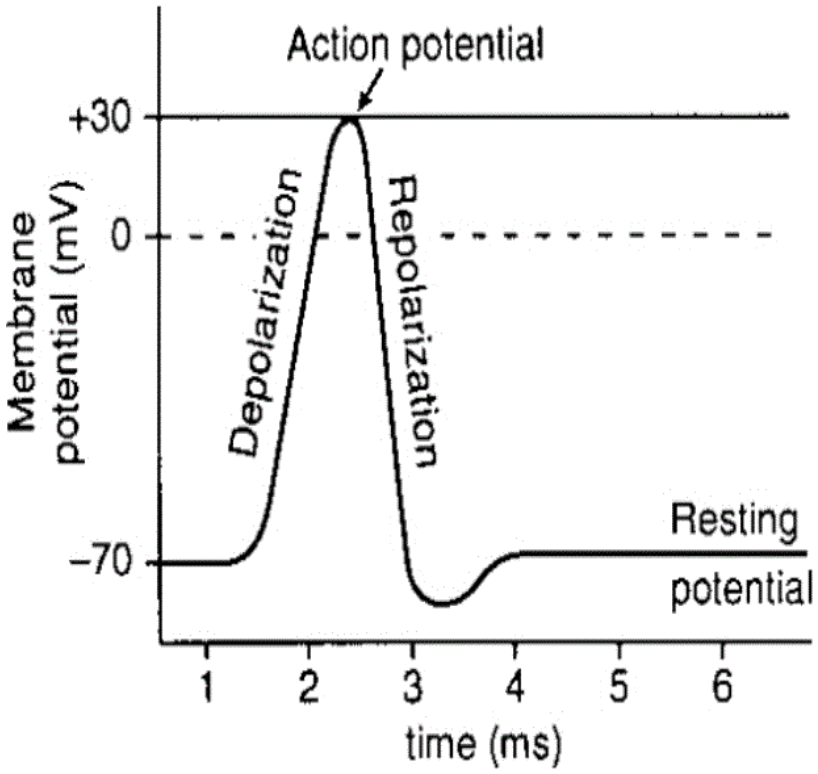


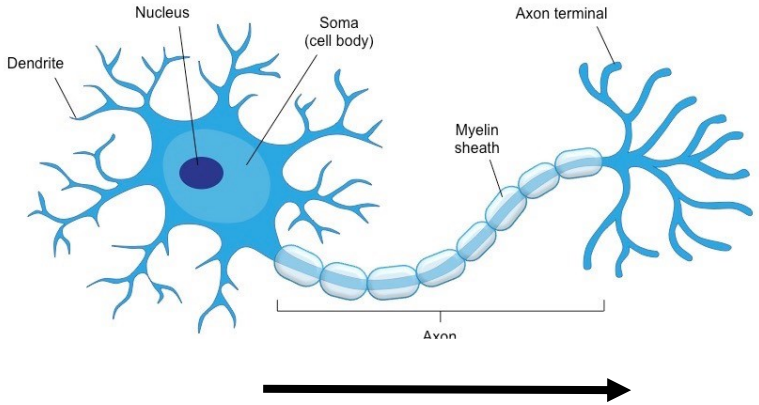
Release rate of neurotransmitters depends on the membrane potential.

Neurotransmitters can be either excitatory (depolarizing) or inhibitory (hyperpolarizing).

pre-synaptic cell ⟶ post-synaptic cell

Mechanism:
Spike (action potential)





Spike travels as an inpulse (wave) along the axon to a "terminal", which it is presynaptic to a neighboring cell.

http://www.youtube.com/watch?v=ifD1YG07fB8

Summary: Level of Analysis in Perception

high

- behavior: what is the task ? what problem is being solved?
  (how well does system solve some problem)
- brain areas and pathways
  (where in the brain do we recognize faces?)
- neural coding
  (what is a sensory cell's receptive field ? How to model responses?)
- neural mechanisms
  (membranes, synapses, spikes)

low

Analogy*: Levels of Analysis in *Computer Science*

high

- problem specification (input and output)
- algorithms
- programs in a high level language
- machine and assembly language
- gates, circuits
- transistors

low

*See book by David Marr: "Vision: A Computational Investigation into the Human Representation and Processing of Visual Information." (1982)

COMP 546 Public web page

www.cim.mcgill.ca/~langer/546.html

Apps Bookmarks Bookmarks Dropbox

Computational Perception COMP 546
Winter 2018
Tues/Thurs 8:35-9:55
ENGTR 1080

Instructor: Professor Michael Langer
School of Computer Science
Office: ENGMC 329
Tel: 514-398-3740
Email: langer [at] cim.mcgill.ca
Office Hours: Tues and Thurs 10 AM -12 PM or by appointment

Teaching Assistant (T.A.) TBD
Email: TBD [@] mail.mcgill.ca
Office: TBD
Office Hours: by appointment

Announcements

- **Winter 2017 lecture notes in one file**
- Please check mycourses for Announcements, if you are not subscribed.
- The yellow notes and slides below are from Winter 2017. I will update them gradually during the Winter 2018 semester.

Resources

- Matlab tutorials
- Official Course Outline

LECTURE SCHEDULE

0. introduction (slides) (notes)
intro, course outline, origin of eyes and spatial vision

Exercises, Exams

Course Overview (by lecture)

- Visual image formation (1-3)
  - geometry: 3D scene to 2D image
  - parallax & binocular disparity
  - focus and blur
  - color

- Early vision (4-7)
  - image coding in the retina
  - image coding in the primary visual cortex

Course Overview (by lecture)

- mid and high level vision (8-10)
  - attention
  - perceptual organization
  - object recognition
- 3D visual perception (11-13)
  - depth cues
- Cue combinations (14-16)
  - maximum likelihood and Bayesian models

Course Overview (by lecture)

- Linear system theory: frequency analysis (17,18)
  - Fourier transform, filtering
- Auditory image formation (19,20)
  - sound waves & head related effects
- 3D audition (21-23)
  - spatial hearing

Unofficial Prerequisites

- COMP 250
- multivariable Calculus (MATH 222)
- linear algebra (MATH 223)
  - vector spaces, linear operators, orthogonality, complex numbers
- probability
  - normal distributions, joint and conditional probabilities.
- waves and optics
  - PHYS 101/102

Evaluation

- Three Assignments (10% each)
  - A1 posted before last week of January
  - A2 posted in early February
  - A3 posted in late March
- Midterm Exam (20%)
  - in class on March 13 (Study Break is March 5-9)
- Final Exam (50%)

You can replace your midterm exam grade with your final exam grade, i.e. final exam would be 70%.

Who are you? (65)

- B. A. (5)
- B.A.Sc. Cog. Sci. (5)
- B.Sc. Neuroscience (15)
- B.Sc. Comp. Sci. (10)
- M.Sc. Comp. Sci (20)
- miscellaneous (10)

- U1 & U2 (10)
- U3 (30)
- MSc (25)

Who am I?

- BSc at McGill in early 1980s (Math Major, CompSci Minor)
  (interest in AI, undergrad summer research in visual neuroscience lab)

- MSc in Computer Science at U of Toronto in late 1980s
  (topic: image coding and compression)

- PhD at McGill in early 1990s
  (topic: shading, shadows, and 3D shape perception)

- postdoc at NEC in NJ, USA in mid-1990s (3 years)
  (computer vision)

- postdoc at Max Planck Inst. in Germany in late 1990s (2 years)
  (human visual perception)

- professor here since 2000
  (taught various versions of this course over 10x)

Want to get involved in research ?

Undergraduates:

- COMP 400 Project in Computer Science
- COMP 396 Undergraduate Research Project

These can be done in any semester (F, W, S).

Graduate students (M.Sc.):

- Project
- Thesis

See www.cim.mcgill.ca/~langer/resources-gradschool.html