Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neuro 130/230. Harvard College/GSAS 78454
Class 1 [09/01/2021]. Introduction to Vision
Note: no class on 09/06/2021
Class 2 [09/13/2021]. Natural image statistics and the retina
Class 3 [09/20/2021]. The Phenomenology of Vision
Class 4 [09/27/2021]. Learning from Lesions
Class 5 [10/04/2021]. Primary Visual Cortex
Note: no class on 10/11/2021
Class 6 [10/18/2021]. Adventures into terra incognita
Class 7 [10/25/2021]. From the Highest Echelons of Visual Processing to Cognition
Class 8 [11/01/2021]. First Steps into in silico vision
Class 9 [11/08/2021]. Teaching Computers how to see
Class 10 [11/15/2021]. Computer Vision
Class 11 [11/22/2021]. Connecting Vision to the rest of Cognition

FINAL EXAM, PAPER DUE 12/14/2021. No extensions.
An image is worth a million words

Who is there?
What is there?
Where is this?
What are they doing?
What is their relationship?
What will happen next?
Let there be light
The light switch theory

Photosynthesis: ~ 3,500 million years ago

Trilobites, circa 500 million years ago

Parker, A. (2004). *In the blink of an eye: how vision sparked the big bang of evolution.*
Why visual recognition?
Selective advantage of visual processing

- Navigation
- Assessing danger
- Identifying food
- Social interactions
- Detecting far away signals (cf. tactile & auditory senses)
- High speeds (cf. olfactory signals)
- Detecting patterns such as constellations
- Reading/Symbols
Four fundamental properties of visual recognition

1. Selectivity
2. Invariance
3. Speed
4. Large capacity
Fundamental properties of visual recognition

Selectivity

Selectivity: discriminating among many (similar) objects
Fundamental properties of visual recognition

Invariance

Invariance: recognizing an object in spite of changes in scale, position, illumination, contrast, viewpoint, cue, clutter, background, etc.
Fundamental properties of visual recognition

Speed

~10 frames/sec

Fundamental properties of visual recognition

Capacity

Standing L (1973) Learning 10,000 pictures;
Shepard RN (1987) Toward a universal law of generalization for psychological science;
Vision is a construct

Class 3: The phenomenology of vision
In the eye brain of the beholder

What color is this dress?
Visual recognition is instantiated by the most powerful computational device on Earth.
Visual system circuitry

Class 5: Primary visual cortex

Class 2: Natural image statistics and the retina

Modified from Society for Neuroscience Brain Facts
Visual system circuitry

Number of neurons in the human brain: \( \sim 10^{11} \)

Number of synapses in the human brain: \( \sim 10^{15} \)

Earth population: \( \sim 7 \times 10^9 \)

Connections (?): \( \sim 7 \times 10^{12} \)

Felleman and Van Essen. *Cerebral Cortex* 1991
Figuring out how the brain works from lesion studies

Distribution of lesion sites in cases of face agnosia

Functional anatomy of the primate visual system

Class 6: Adventures into *terra incognita*: beyond primary visual cortex

Hubel and Wiesel (1959)
*J. Physiol.* **148**: 574-591

Kuffler, S. (1953)
*J. Neurophys.* **16**: 37-68

Class 5:
Primary visual cortex

Desimone *et al* (1984)
*J. Neurosci.* **4**:2051-2062

Class 7: From the highest echelons of vision to cognition

Newsome *et al* (1989)
*Nature* **341**:52-54
Before the removal was carried out, stimulation at points 5 and 7 produced the following experiential responses.

5. Patient did not reply.
5. Repeated. “Something.”
5. Patient did not reply.
5. Repeated. “Something.”
5. Repeated again. “People’s voices talking.” When asked, he said he could not tell what they were saying. They seemed to be far away.
5. Stimulation without warning. He said, “Now I hear them.” Then he added, “A little like in a dream.”
7. “Like footsteps walking—on the radio.”
7. Repeated. “Like company in the room.”
7. Repeated. He explained “it was like being in a dance hall, like standing in the doorway—in a gymnasium—like at the Kenwood Highschool.” He added, “If I wanted to go there it would be similar to what I heard just now.”
7. Repeated. Patient said, “Yes, yes, yes.” After withdrawal of the stimulus, he said it was “like a lady was talking to a child. It seemed like it was in a room, but it seemed as though it was by the ocean—at the seashore.”
7. Repeated. “I tried to think.” When asked whether he saw something or heard something, he said, “I saw and heard. It seemed familiar, as though I had been there.”

5. Repeated (20 minutes after last stimulation at 5). “People’s voices.” When asked, he said, “Relatives, my mother.” When asked if it was over, he said, “I do not know.” When asked if he also realized he was in the operating room, he said “Yes.” He explained it seemed like a dream.

5. Repeated. Patient said, “I am trying.” After withdrawal of the electrode he said, “It seemed as if my niece and nephew were visiting at my home. It happened like that many times. They were getting ready to go home, putting their things on—their coats and hats.” When asked where, he said, “In the dining room—the front room—they were moving about. There were three of them and my mother was talking to them. She was rushed—in a hurry. I could not see them clearly or hear them clearly.”

Penfield & Perot. *The brain’s record of auditory and visual experience.* A final summary and discussion. *Brain* (1963) **86:**595-696
A flower, as seen by a computer

Classes 8-11: Can computers see the way we do? Computer vision
Why is vision difficult?
Towards a theory of object recognition

Computational models can

- Integrate existing data
- Explain apparently disparate observations
- Quantify and formalize knowledge
- Suggest experimentally-testable predictions
- Provide a useful engineering tool
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".
A feed-forward hierarchical model of ventral cortex

Classes 9-10: Computer vision

Rapid progress in image classification tasks

Detection, segmentation, recognition

Face detection

Segmentation

Recognition

Classes 10-11: Can computers see the way we do?
Computer vision

Fig. 7: The 101 object categories and the background clutter category. Each category contains between 450 and 400 images. Two randomly chosen samples are shown for each category. The categories were selected prior to the experimentation, collected by operators not associated with the experiment, and augmented by various digital transformations. This dataset is

---

Detection, segmentation, recognition

Face detection

Segmentation

Recognition

Classes 10-11: Can computers see the way we do?
Computer vision

Fig. 7: The 101 object categories and the background clutter category. Each category contains between 450 and 400 images. Two randomly chosen samples are shown for each category. The categories were selected prior to the experimentation, collected by operators not associated with the experiment, and augmented by various digital transformations. This dataset is
Why visual recognition?

Applications

- Face recognition
- Pedestrian recognition
- Self-driving cars
- Robot navigation
- Clinical applications
- Security
- Intelligent image understanding
I think it's a group of people standing next to a man in a suit and tie.
A Turing test for vision

Can machines (be taught to) see the world the way we do?

Alan Turing, 1950. Computing Machinery and Intelligence. “Can machines think?”
Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an “intelligence explosion,” and the intelligence of man would be left far behind.

Thus the first ultraintelligent machine is the last invention that man need ever make . . .

I.J. Good “Speculations Concerning the First Ultra-intelligent Machine”, 1965
Bistable percepts and subjective perception


Class 12: Visual consciousness
Further reading


