While we wait for everyone to join, a brief math problem:

How many possible images are there of size 100 x 100 pixels, with 256 shades of gray per pixel?
Note: no class on 09/04/2023 (Labor Day)
Class 1 [09/11/2023]. Introduction to Vision
Class 2 [09/18/2023]. The Phenomenology of Vision
Class 3 [09/25/2023]. Natural image statistics and the retina
Class 4 [10/02/2023]. Learning from Lesions
Note: no class on 10/09/2023 (Indigenous Day)
Class 5 [10/16/2023]. Primary Visual Cortex
Class 6 [10/23/2023]. Adventures into terra incognita
Class 7 [10/30/2023]. From the Highest Echelons of Visual Processing to Cognition
Class 8 [11/06/2023]. First Steps into in silico vision
Class 9 [11/13/2023]. Teaching Computers how to see
Class 10 [11/20/2023]. Computer Vision
Class 11 [11/27/2023]. Connecting Vision to the rest of Cognition [Dr. Will Xiao]
Class 12 [12/06/2023]. Visual Consciousness

Starting from the very beginning

• Let there be light, and there was light.
• Objects reflect light
• Light photons impinge on the retina (Latin: small net)
• The retina conveys visual information to the brain

An oversimplified first-order description:

*The eye functions as a very sophisticated and spectacular digital camera*
An image as a collection of pixels
Natural images are special
We only encounter a small subset of the space of possible images

Consider an image of size 100 x 100 pixels
Assume a pixel can have 256 shades of gray
How many such images are possible?

Answer
For a size of 1x1 pixel, there are 256 possible images
For a size of 1x2 pixels, there are \( 256^2 \) possible images
For a size of 100x100 pixels, there are \( 256^{10000} \) possibilities*

Yet, we only encounter a small fraction of these possibilities in natural images

*Some of those are “related” by translation, rotation or inversion, etc
Spatial regularities in natural scenes
The properties of nearby points are correlated

Simoncelli and Olshausen 2001
Images contain different spatial frequencies
Some are more prevalent than others

Low frequencies  Middle frequencies  High frequencies

Natural image statistics
Power spectrum $\sim 1/f^2$

$log(f(w)) = \alpha \log(w) + c$

Note: Scale invariance

$w' \rightarrow aw$

$log(f(w')) = \beta \log(w) + d$

There are multiple examples of power law distributions in physics, biology and social sciences

Figure 4  Power spectrum of a natural image (solid line) averaged over all orientations, compared with $1/f^2$ (dashed line).

Simoncelli and Olshausen 2001
Sizes are measured in degrees of visual angle

\[
\frac{\alpha}{2} = \arctan \left( \frac{h}{2d} \right)
\]

Size of the moon ~ 0.5 degrees
Thumb at arms length ~ 2 degrees
Natural image statistics

There are also strong correlations in time

The visual input is largely static, except for:

• External object movements
• Head movements
• Eye movements

The visual image is largely static over hundreds of milliseconds

Silent Reading       225-250 ms fixation, 2 degrees saccade size (8-9 letters)
Scene Perception    260-330 ms fixation, 4 degrees saccade

“Slowness” has been proposed as a constraint for learning about objects (Földiak 1991, Stringer et al 2006, Wiskott et al 2002, Li et al 2008)
Saccadic eye movements
Example pattern of fixations while examining an image
Microsaccades are important for perception

The eye constantly makes very small “fixational” movements, a.k.a. microsaccades.

When visual stimulation is held fixed relative to the retina, visual perception rapidly fades (Ratliff & Riggs, 1950).

These microsaccades may be critical for counteracting perceptual fading (McCamy, 2012)

Martinez-Conde, Macknik, & Hubel, 2004

Early retinal stabilization apparatus by Yarbus
The image is focused onto the retina

Images are inverted by the eye

Object  lens  Focal plane
Why don’t we see everything upside down?

Perception can adapt to reversing the image.
Some preliminary experiments on vision without inversion of the retinal image. Stratton, 1896
In octopi, light reaches directly to the photoreceptors before reaching the ganglion cells.
“Seeing” the blind spot

Cover your left eye
Fixate on the red circle and do not move your eyes!
Evolution of the retina

Vertebrate

1. In octopi, light reaches directly the photoreceptors before reaching the ganglion cells

4. Blind spot is only present in vertebrates

Octopus

1. In octopi, light reaches directly the photoreceptors before reaching the ganglion cells
“To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different mounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest degree.”

Charles Darwin, The origin of species, 1859
The retina: An amazingly beautiful circuitry composed of many different cell types

- ~0.5 mm thick
- 5 x 5 cm retinal area
- Three cellular layers
- Rods (low-illumination conditions, \(~10^8\))
- Cones (color sensitivity, \(~10^6\))
- Blind spot
- Fovea (rod free, ~0.5 mm, ~1.7 deg)
- Midget ganglion cells (small dendritic arbors)
- Parasol ganglion cells (large dendritic arbors)

Dowling (2007), Scholarpedia, 2:3487
The retina: Some cells fire action potentials whereas other cells show graded responses

- Photoreceptors transduce incoming light input into electrical signals
- Rod to bipolar convergence increases rod-pathway sensitivity
- Cones, rods, horizontal and bipolar cells are non-spiking neurons
- Many different types of amacrine cells (some of which fire action potentials)
- Retinal ganglion cells fire action potentials and carry the output signals

Rods see largely in grayscale
Ishihara tests for color blindness
Non-uniform sampling of the visual field

There is much more detail at the fovea
To be, or not to be, that is the question:
Whether 'tis nobler in the mind to suffer
The slings and arrows of outrageous fortune,
Or to take Arms against a Sea of troubles,
And by opposing end them: to die, to sleep
No more; and by a sleep, to say we end
The heart-ache, and the thousand natural shocks
That Flesh is heir to? 'Tis a consummation
Devoutly to be wished. To die, to sleep,
perchance to Dream; aye, there's the rub,
For in that sleep of death, what dreams may come,
When we have shuffled off this mortal coil,
Must give us pause. There's the respect
That makes Calamity of so long life:
For who would bear the Whips and Scorns of time,
The Oppressor's wrong, the proud man's Contumely,
The pangs of dispised Love, the Law's delay,
The insolence of Office, and the spurns
There is much more detail at the **fovea**

Try to read outside of your fixation spot
There is much more detail at the fovea
Example psychophysics task

Indistinguishable images due to reduced acuity in the periphery

Picture 1 is flashed
Picture 2 is flashed
Picture 3 is flashed

Was there an L or a T in the center?

Was picture 3 more similar to picture 1 or picture 2?

Sample 1 or 2?
The retina has a huge dynamic range especially at high background brightnesses. Likewise, it is uncertain that all summation has ceased for the stimulus of 19 degrees² (4.90 diameter) and 1 sec duration, used for measuring threshold intensity. For this reason, thresholds were determined at intermediate durations and areas at the various background intensities on one subject. It was hoped that these results would also show whether the changes in total summation involved changes in both temporal and spatial summation, and whether the changes took place in the upper limits of complete summation (Bunsen-Roscoe and Ricco ranges), or whether the amounts, or ranges, of partial summation were involved.

Vision works well in moonless nights and in bright sunlight (~$10^9$ more light intensity).

There are several mechanisms that support light adaptation:
- Changing pupil size
- Reliance on rods vs. cones
- Photopigment bleaching
- Feedback from horizontal cells to photoreceptors

These mechanisms are relatively slow.

Barlow, 1958
The receptive field
Neurons throughout the visual system are very picky about the stimulus location

This cartoon neuron responds only when a flash of light appears in the periphery, in the lower left quadrant

Blumberg and Kreiman, 2010
Physiology of retinal ganglion cells

The receptive field of most RGCs have a center-surround structure

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

About 1.2 million projections from each retina towards the brain

Diversity of retinal ganglion cells

A minority of RGCs have more complex response properties:

- Phasic cells respond briefly to stimulus onset, offset, or both
- Some phasic cells respond selectively to edge orientation
- Suppressed-by-contrast cells fire except when an edge is present in receptive field
- Bistratified RGCs lack surrounds and are color-sensitive
- Color-opponent cells have centers and surrounds with opposing color preferences
- Intrinsically photosensitive RGCs contain photoreceptors and project to regions controlling pupil size, circadian rhythm, etc.
- Direction-sensitive cells respond to direction of motion of light or dark spots

These cells likely account for approximately 10% of RGCs
Unclear to what extent, these cells contribute to visual object recognition

Stone and Fukuda, *Journal of Neurophysiology* 1974
Cleland and Levick, *Journal of Neurophysiology* 1974
Berson et al., *Science* 2002
# Eye versus digital cameras

<table>
<thead>
<tr>
<th></th>
<th>One eye</th>
<th>Camera (phone)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Angle of view</strong></td>
<td>~ 120 degrees</td>
<td>~ 70 degrees</td>
</tr>
<tr>
<td><strong>Sampling</strong></td>
<td>Eccentricity dependent</td>
<td>Uniform</td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td>~ 10 “MP” in the fovea</td>
<td>~ 20 MP (million pixels)</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>High sensitivity in low light conditions</td>
<td>Struggle in low light and requires very long integration time</td>
</tr>
<tr>
<td><strong>Energy efficiency</strong></td>
<td>Small fraction of ~3 meals a day, say 14 hours of continuous use. (2500 calories / day ~ 0.12 Watts)</td>
<td>Continuous video recording: ~ 3 hours on one battery life (~5 Watts, hence &gt; 12 watts a day)</td>
</tr>
<tr>
<td><strong>Malleability</strong></td>
<td>Hard to change (without additional devices like glasses, microscopes, telescopes)</td>
<td>Can expand in many ways</td>
</tr>
<tr>
<td><strong>Post-processing</strong></td>
<td>The human brain!</td>
<td>Minimal computer vision*</td>
</tr>
</tbody>
</table>
## Eyes in non-human animals

<table>
<thead>
<tr>
<th></th>
<th>Humans</th>
<th>Other animals</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wavelength</strong></td>
<td>“Visible” spectrum</td>
<td>Also UV, IR</td>
<td>mice, dogs, birds: UV snakes: IR</td>
</tr>
<tr>
<td><strong>Cones</strong></td>
<td>3</td>
<td>1-30</td>
<td>bat, racoon: 1 cats, dogs: 2 some dragon flies: 30</td>
</tr>
<tr>
<td><strong>Polarization</strong></td>
<td>No</td>
<td>Some animals</td>
<td>cuttlefish</td>
</tr>
<tr>
<td><strong>Number of eyes</strong></td>
<td>2</td>
<td>Up to 10</td>
<td>spiders: 8-12 horseshoe crabs: 10</td>
</tr>
<tr>
<td><strong>Angle</strong></td>
<td>Binocular: ~120°</td>
<td>Up to 300°</td>
<td>rabbits: no binocular zone in the center cows, horses: 300°</td>
</tr>
<tr>
<td></td>
<td>Up to 180°</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Eyes</strong></td>
<td>Correlated</td>
<td>Sometimes independent</td>
<td>Chameleon: can move eyes independently</td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td>~100M receptors</td>
<td>Wide variation</td>
<td>Star fish: 200 receptors Birds: higher acuity than humans</td>
</tr>
</tbody>
</table>
To cortex, through the thalamus:
The Lateral Geniculate Nucleus (LGN)

- Layers 2, 3 and 5 receive *ipsilateral* eye’s input
- Layers 1, 4 and 6 receive *contralateral* eye’s input
- Layers 1-2: *magnocellular* layers, input from parasol retinal ganglion cells (~ motion)
- Layers 3-6: *parvocellular* layers, input from midget ganglion cells (~ color)
- On and Off center cells are present in all layers
- LGN does not project back to the retina

NOTE: Most of the input to the LGN comes from visual cortex and not from the retina! (e.g. Douglas and Martin 2004)
Modeling receptive fields: difference of Gaussians
The center-surround structure can be described by a difference of gaussians (Mexican-hat)

\[
D(x, y) = \pm \left( \frac{1}{2\pi\sigma_{cen}^2} \exp\left[ -\frac{x^2 + y^2}{2\sigma_{cen}^2} \right] - \frac{B}{2\pi\sigma_{sur}^2} \exp\left[ -\frac{x^2 + y^2}{2\sigma_{sur}^2} \right] \right)
\]

Neurons respond with transient bursts of activity

Dynamic receptive fields in the retina/LGN

\[
D(x,y,t) = \pm \left( \frac{D_{cen}(t)}{2\pi \sigma_{cen}^2} \exp \left[ -\frac{x^2 + y^2}{2\sigma_{cen}^2} \right] - \frac{BD_{sur}(t)}{2\pi \sigma_{sur}^2} \exp \left[ -\frac{x^2 + y^2}{2\sigma_{sur}^2} \right] \right)
\]

\[
D_{cen}(t) = \alpha_{cen}^2 t \exp[-\alpha_{cen} t] - \beta_{cen}^2 t \exp[-\beta_{cen} t]
\]

\[
D_{sur}(t) = \alpha_{sur}^2 t \exp[-\alpha_{sur} t] - \beta_{sur}^2 t \exp[-\beta_{sur} t]
\]

Difference of Gaussians in space and time
The center-surround structure can also be seen in receptive field dynamics

Information from the retina goes to many places
Subcortical visual pathways

<table>
<thead>
<tr>
<th>Retinal projections</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lateral geniculate nucleus (LGN)</strong></td>
</tr>
<tr>
<td>Superior Colliculi</td>
</tr>
<tr>
<td>Main visual pathway in birds, reptiles, fish</td>
</tr>
<tr>
<td>Implicated in saccade generation in mammals</td>
</tr>
<tr>
<td>Suprachiasmatic Nucleus</td>
</tr>
<tr>
<td>Hypothalamus: involved in circadian rhythms</td>
</tr>
<tr>
<td>Pretectum</td>
</tr>
<tr>
<td>Pregeniculate</td>
</tr>
<tr>
<td>Accessory optic system</td>
</tr>
<tr>
<td>Primates can recognize objects after lesions to the Superior Colliculus or Suprachiasmatic Nucleus, but not after lesions to V1 (Gross 1994).</td>
</tr>
</tbody>
</table>
Visual system circuitry

Felleman and Van Essen. *Cerebral Cortex* 1991


