Human Vision

P. Fua
(Taught by M. Salzmann)
IC-CVLab
EPFL

Human Vision

It Works!!

-->Proof of existence.

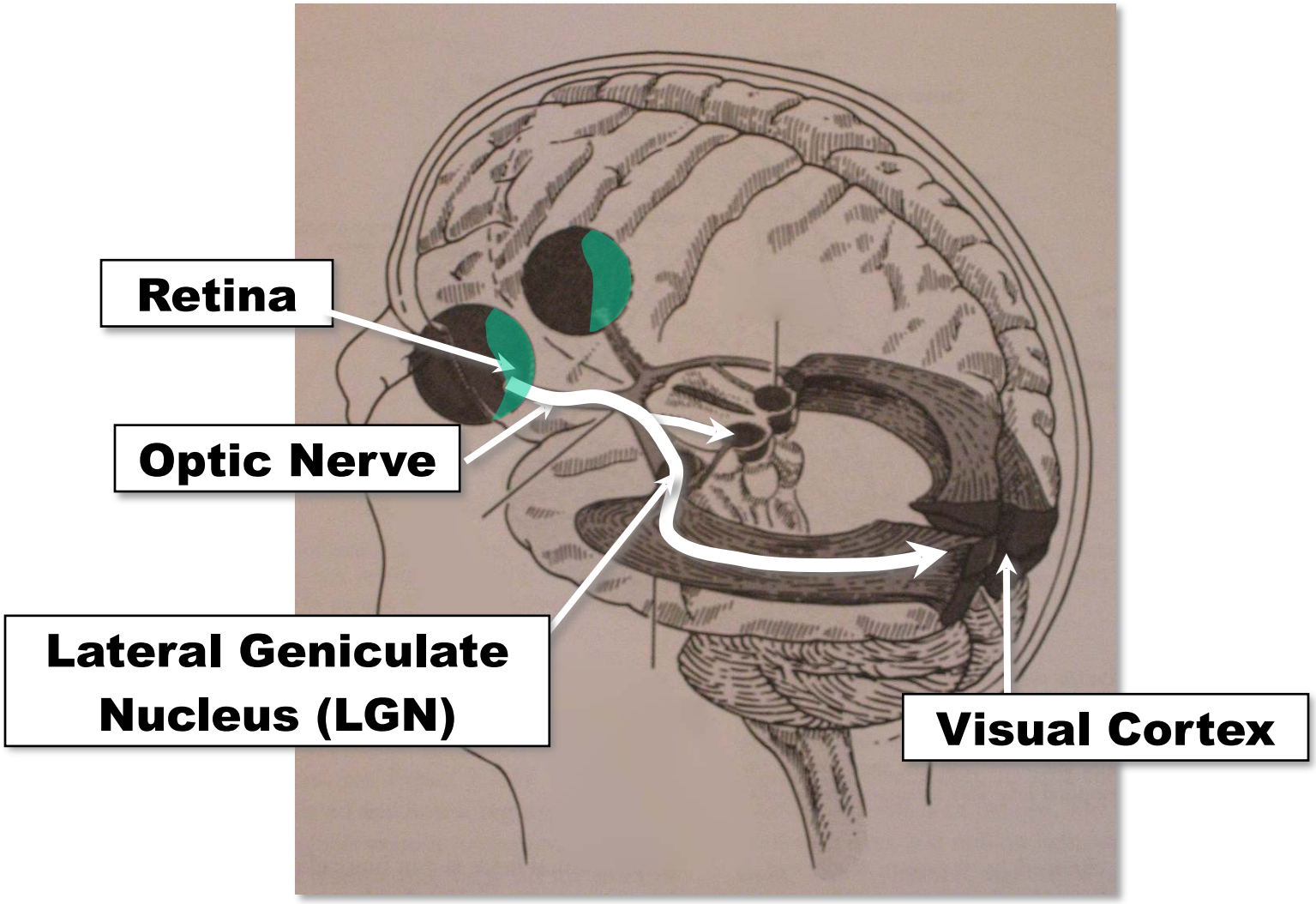
not easy

- The image formation process is well understood
- The image understanding one remains mysterious

How eyes act as cameras?

CV → image understanding, not acquisition

Pathways To The Brain



Human Eye

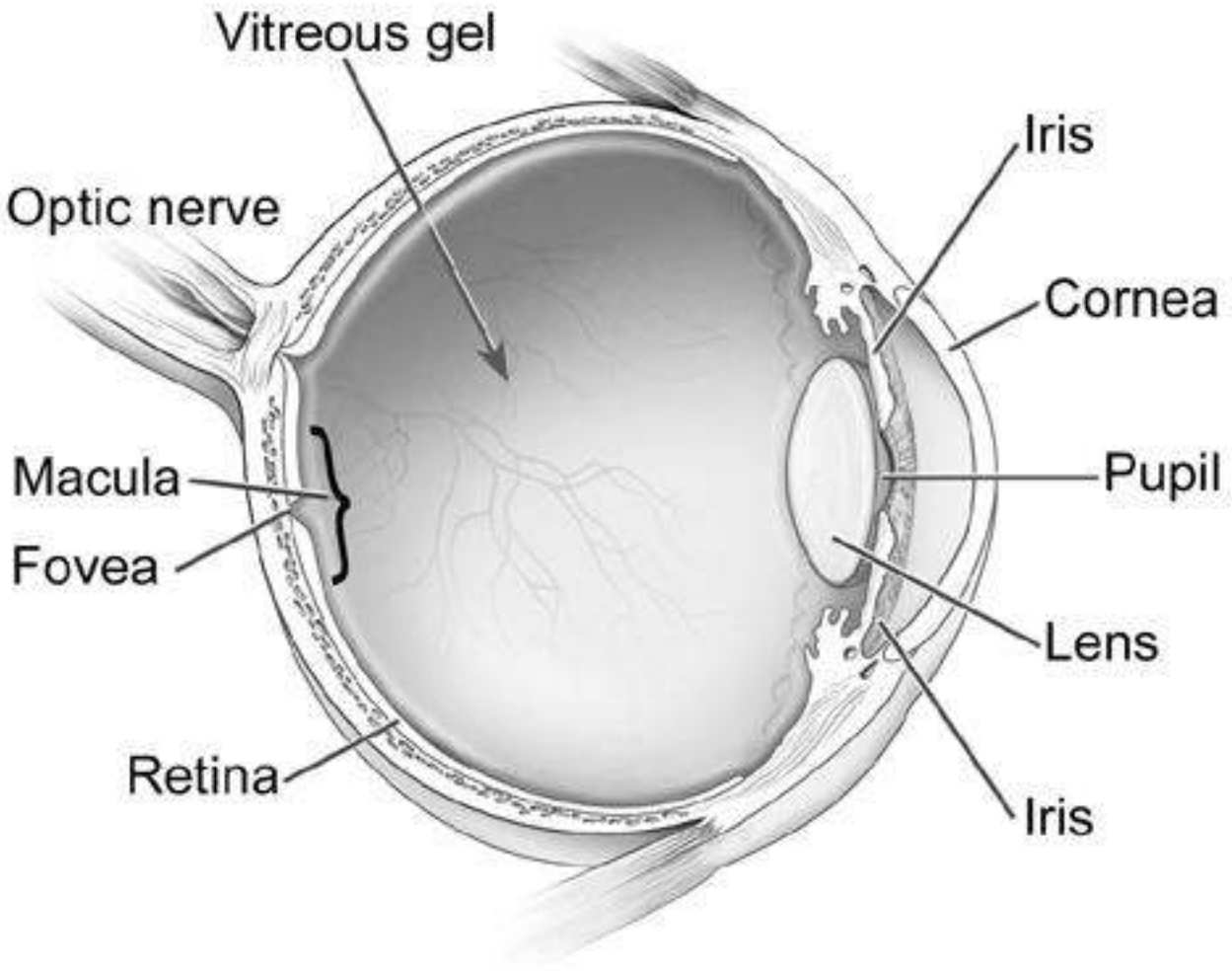
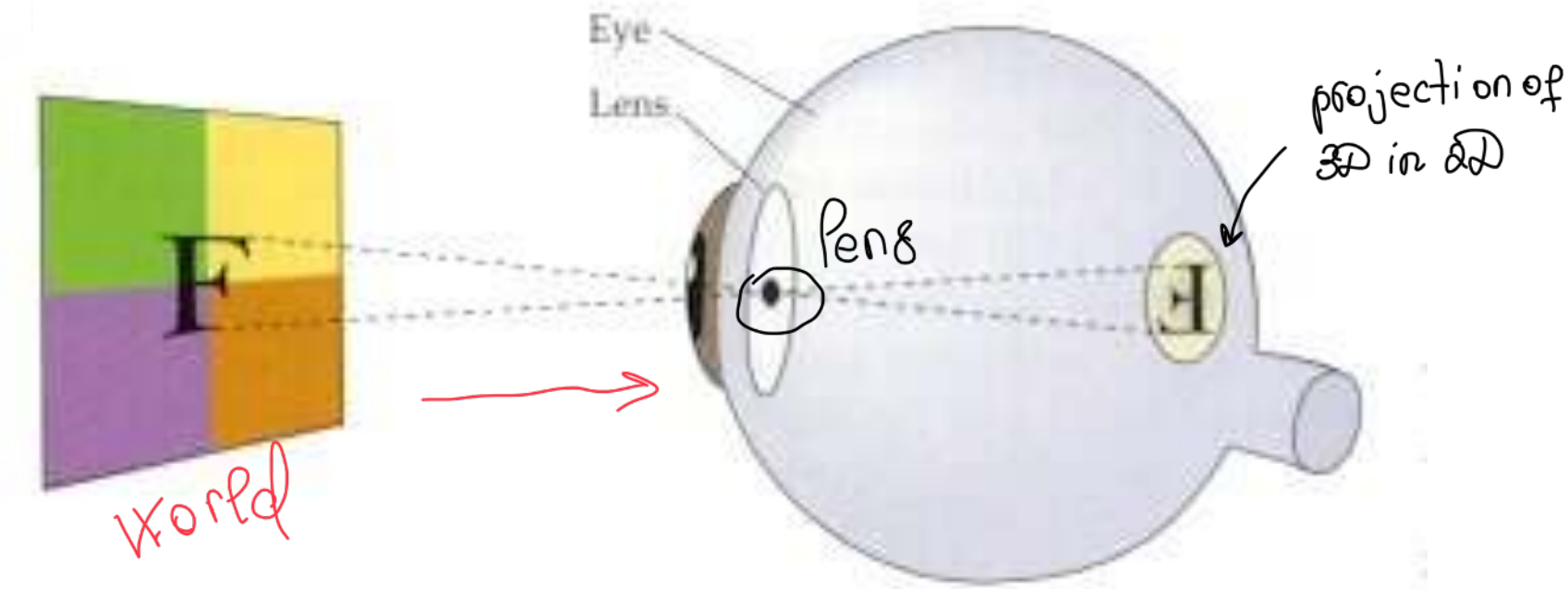
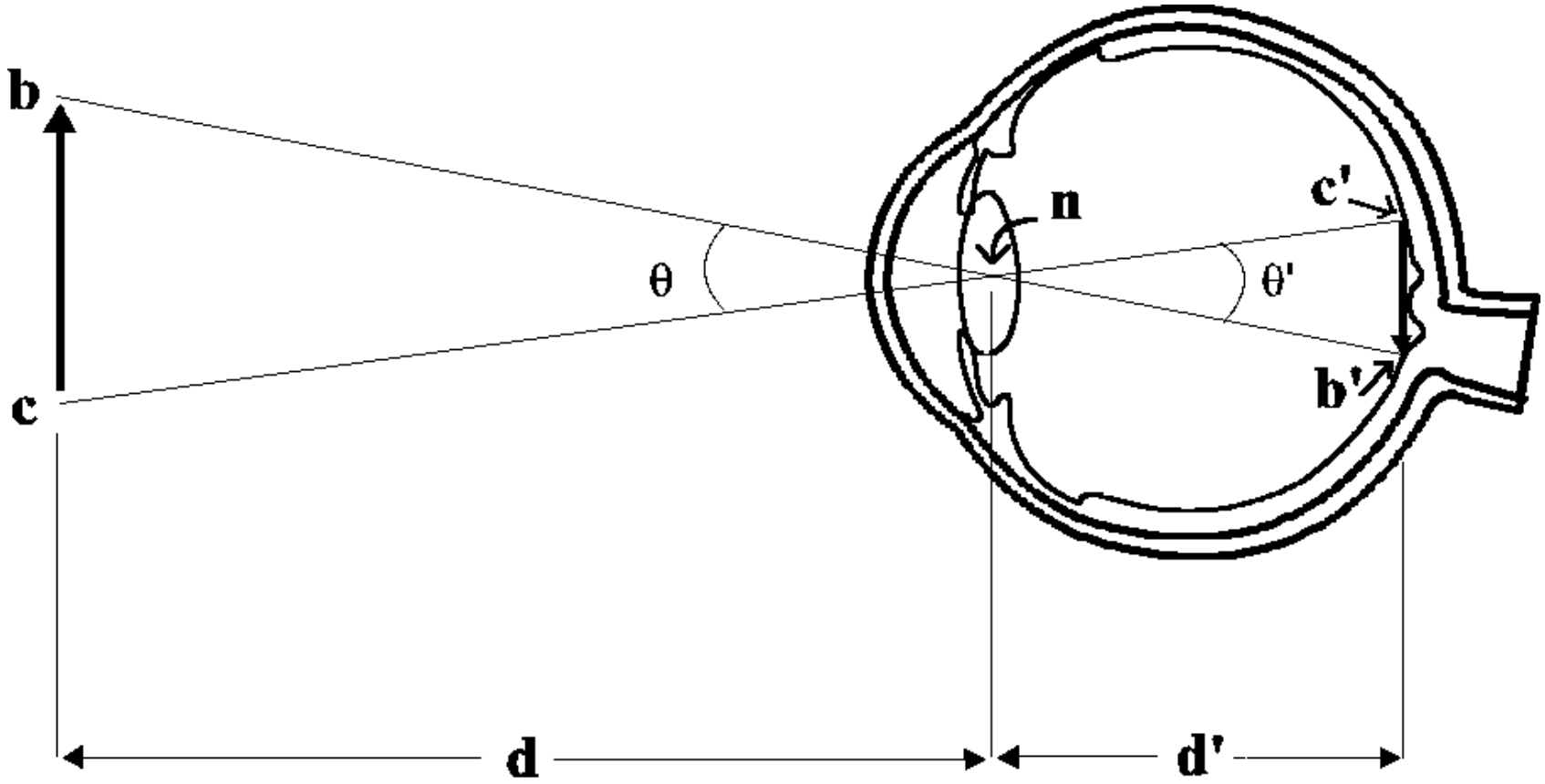

Vitreous gel
Optic nerve
Iris
Cornea
Macula
Fovea
Pupil
Lens
Retina
Iris

Image Formation



An inverted image forms on the retina.

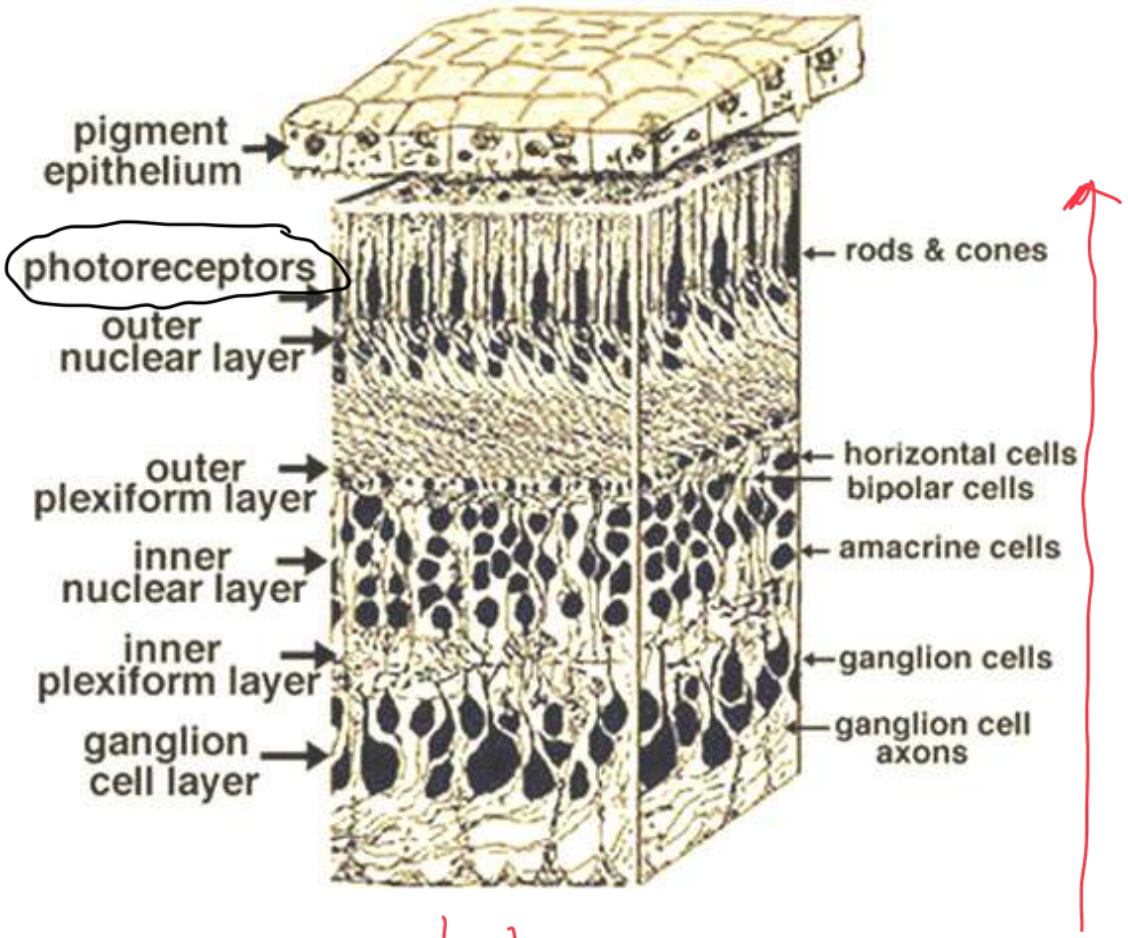
Perspective Projection

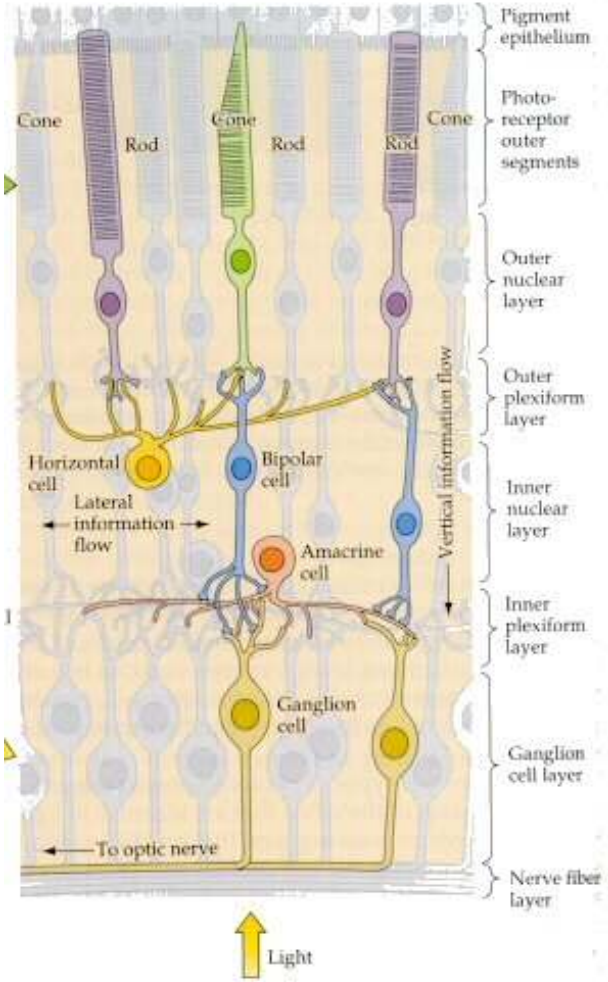


$$bc/d = b'c'/d'$$

- This is known as the pinhole camera model.
- Cameras do something similar and we will revisit it in the next lecture.

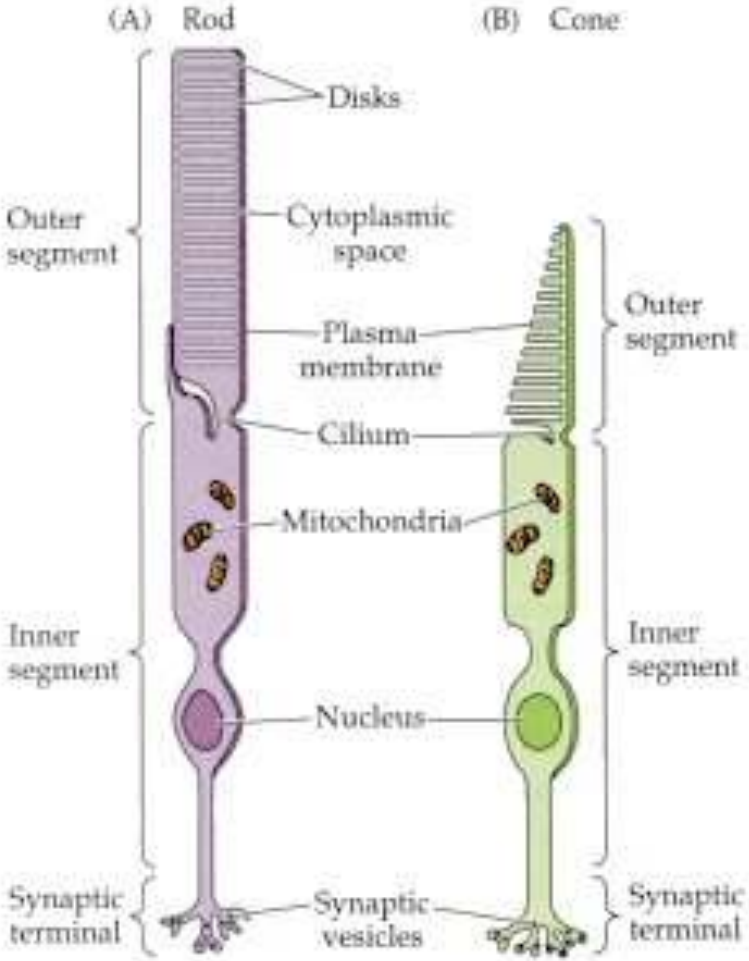
Retina





Light comes from below

Rods and Cones



Rods: Low-intensity light vision, e.g. night vision.

Cones: Color-vision with higher intensity light.

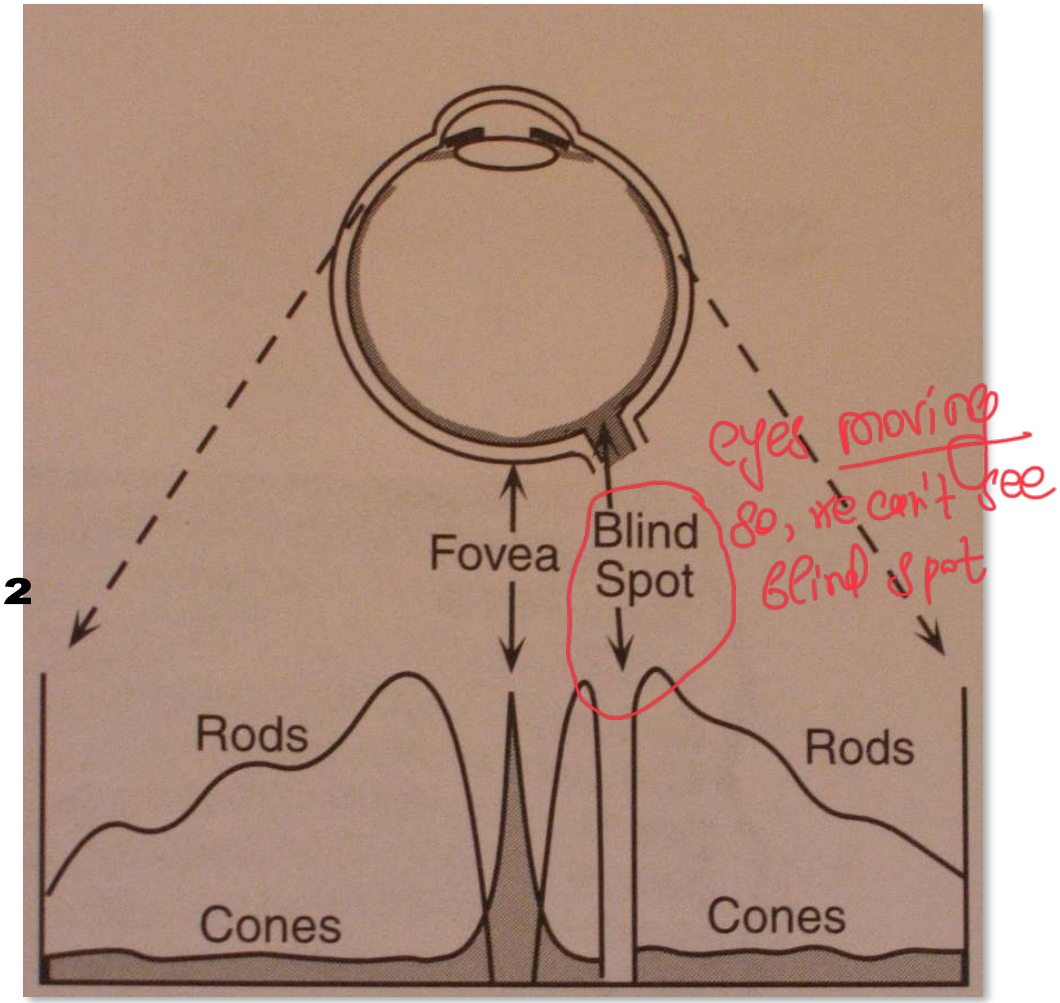
Cell Distribution

 # receptors/mm²

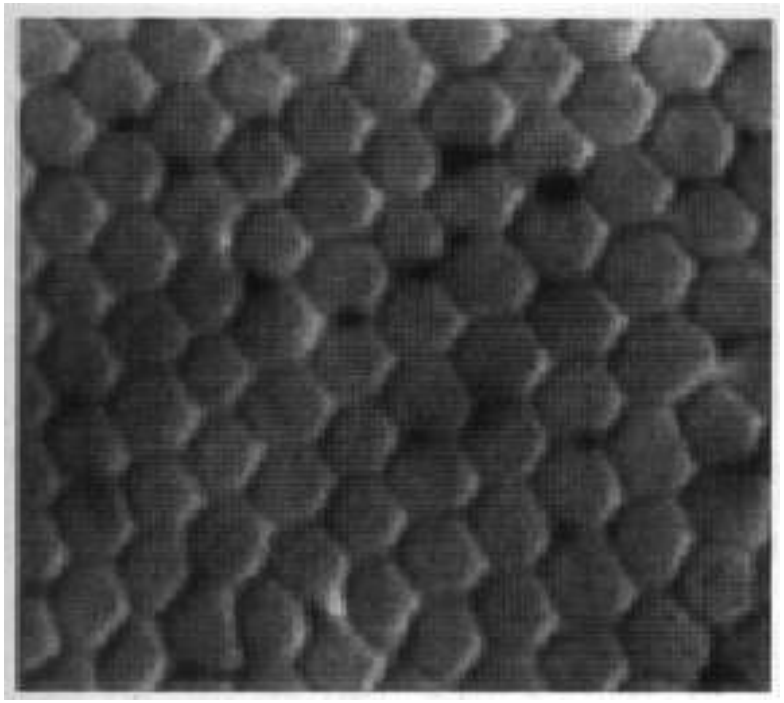
150’000

100’000

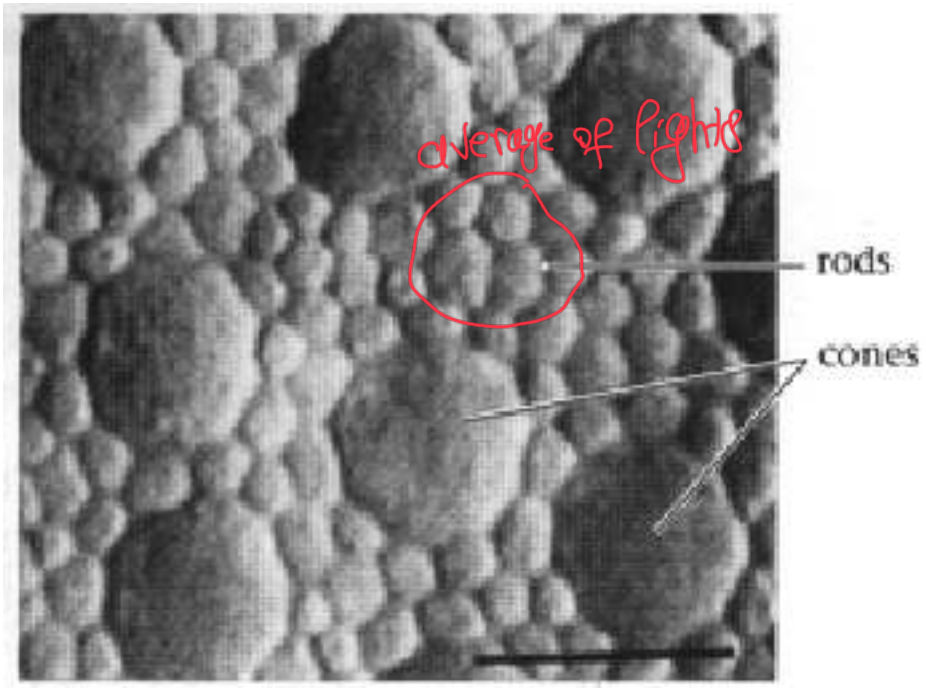
50’000



Fovea vs Periphery

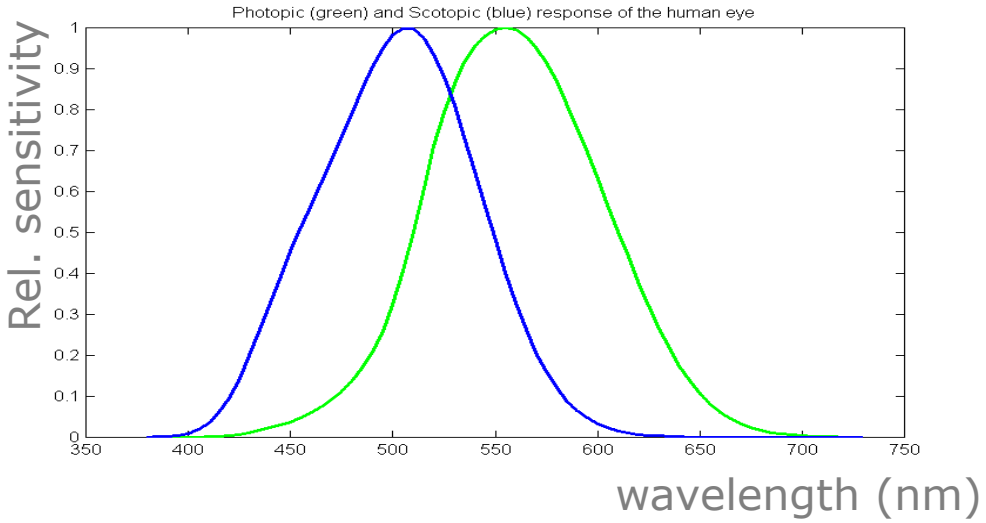


Fovea



Periphery

Scotopic vs Photopic



Low luminance (< 1 cd/m^2):

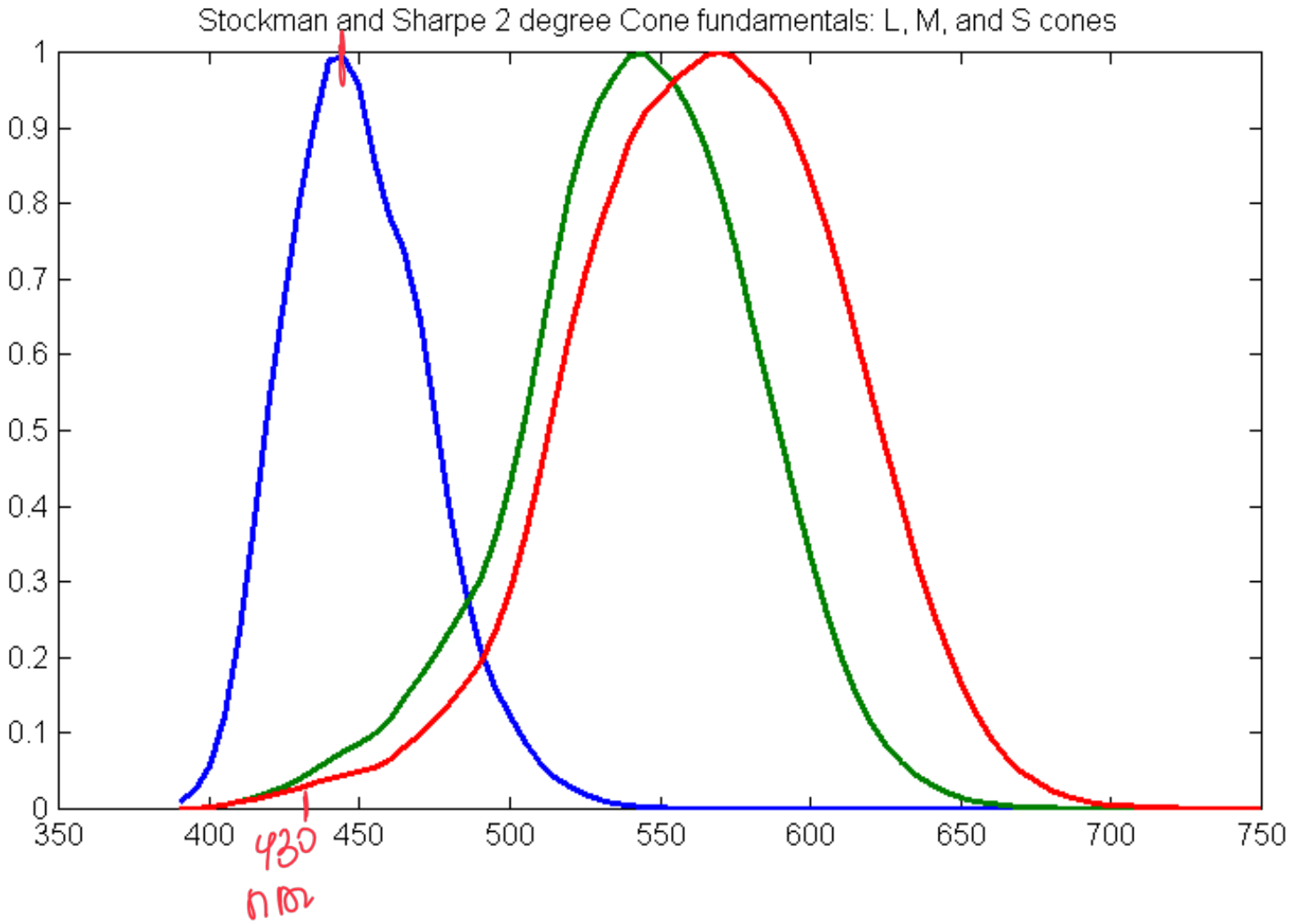
- 120 million rods with peak spectral response around 510 nm.
- Primarily located outside the fovea.

Due to rods, in periphery

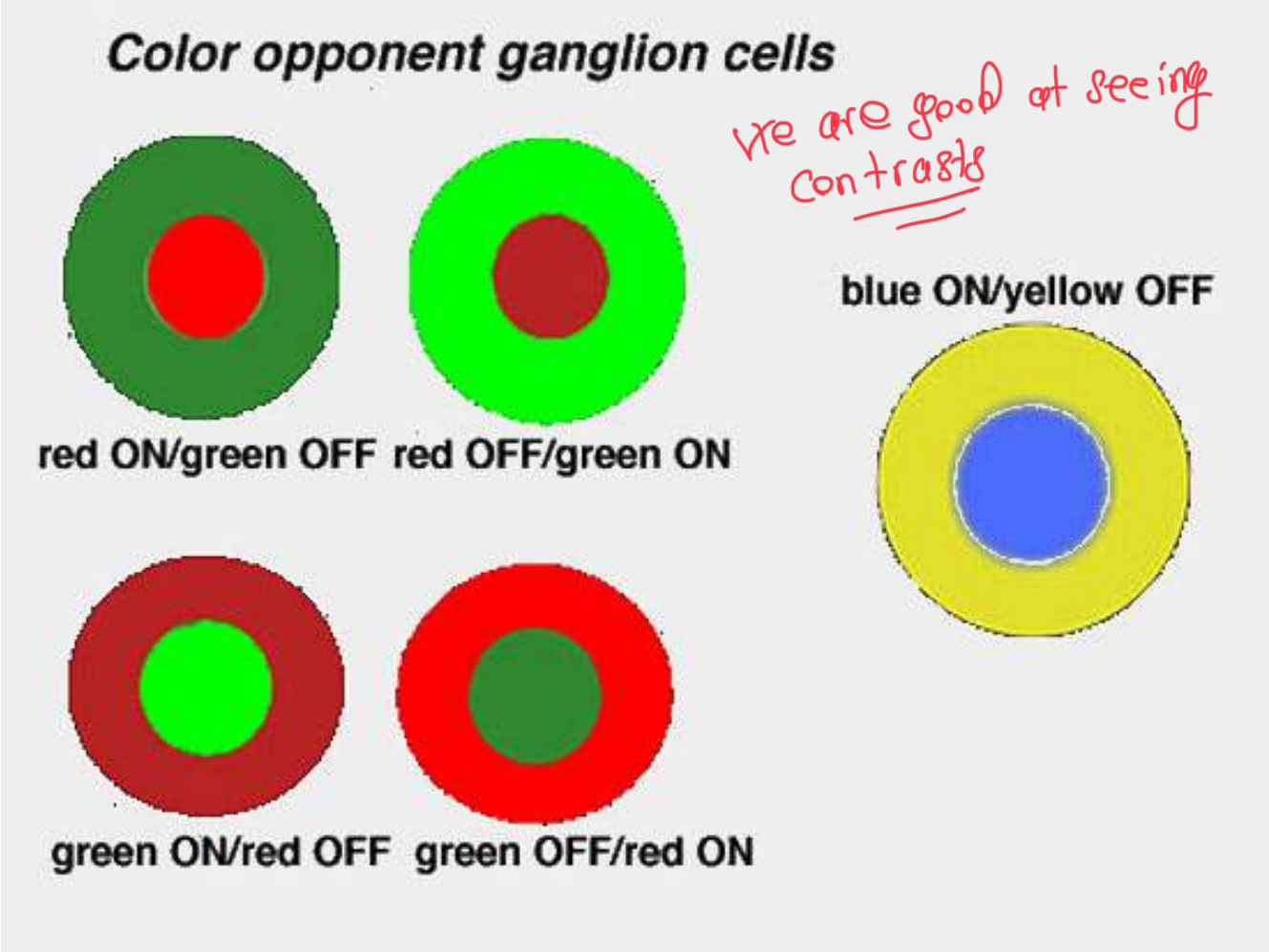
High luminance (> 100 cd/m^2):

- 7 million cones per retina.
- Primarily located in the fovea.
- Three types of cones (S, M, L) with peak spectral response at different nm.
- Ratio L:M:S ≅ 40:20:1

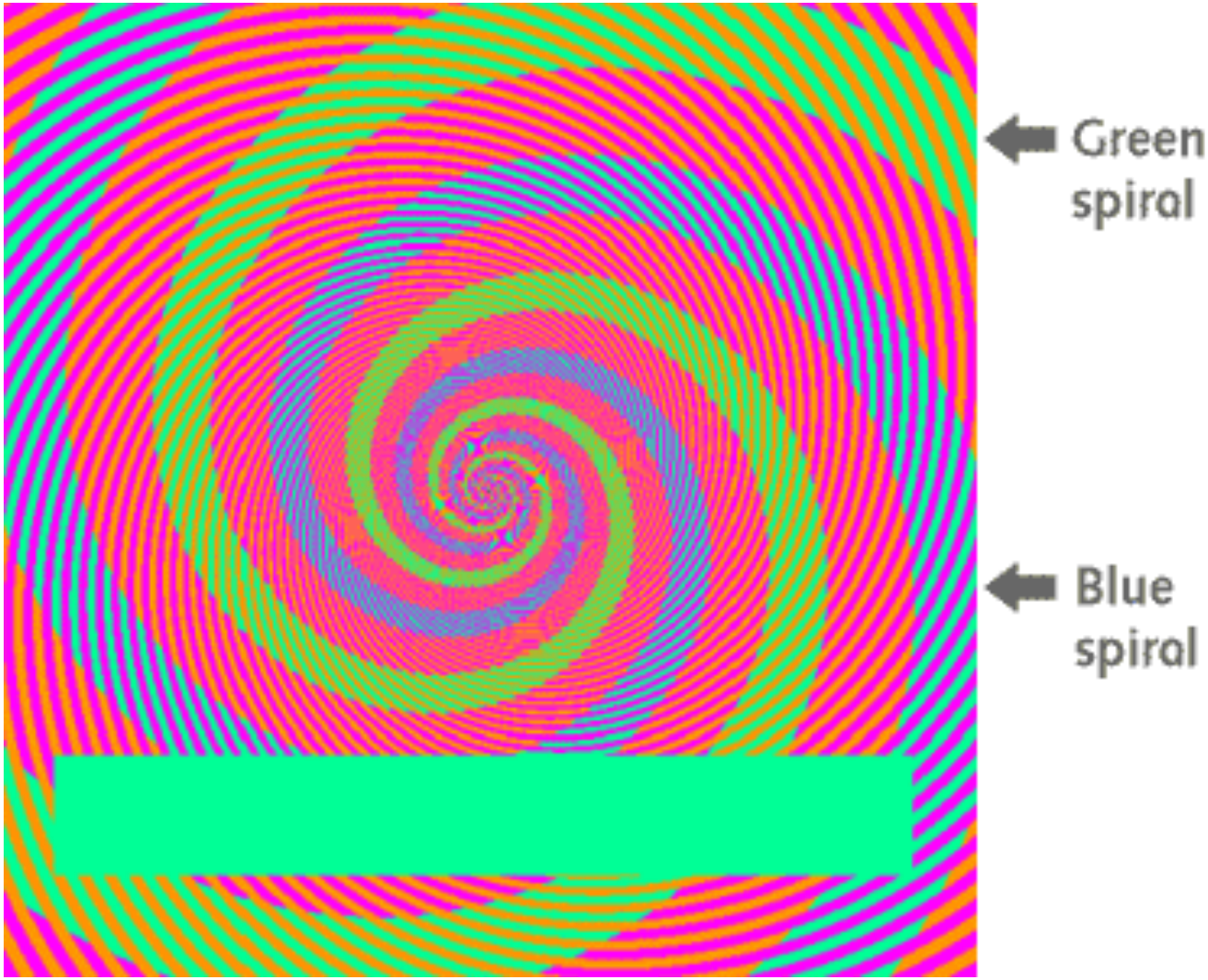
Sensitivity to Different Wavelengths



Ganglion Cells



Color Illusion



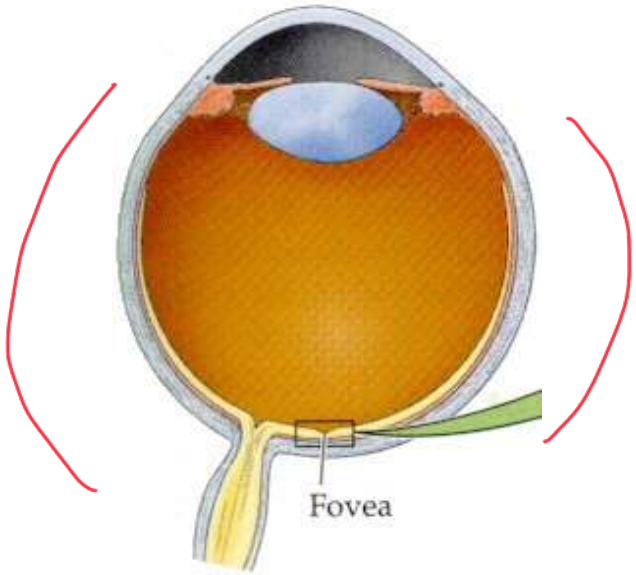
Color Balancing

- Red sand has been blown onto the slopes.
- The streak of snow in the middle should be white.

iPhone image

Color corrected image

Peripheral Vs Foveal Vision



rods → peripheral vision

Much higher concentration of cells on the Fovea

higher-resolution image

→Active vision:

- We find objects using our peripheral vision
- We concentrate our gaze on objects of interest.

using fovea

The Human Eye In Short

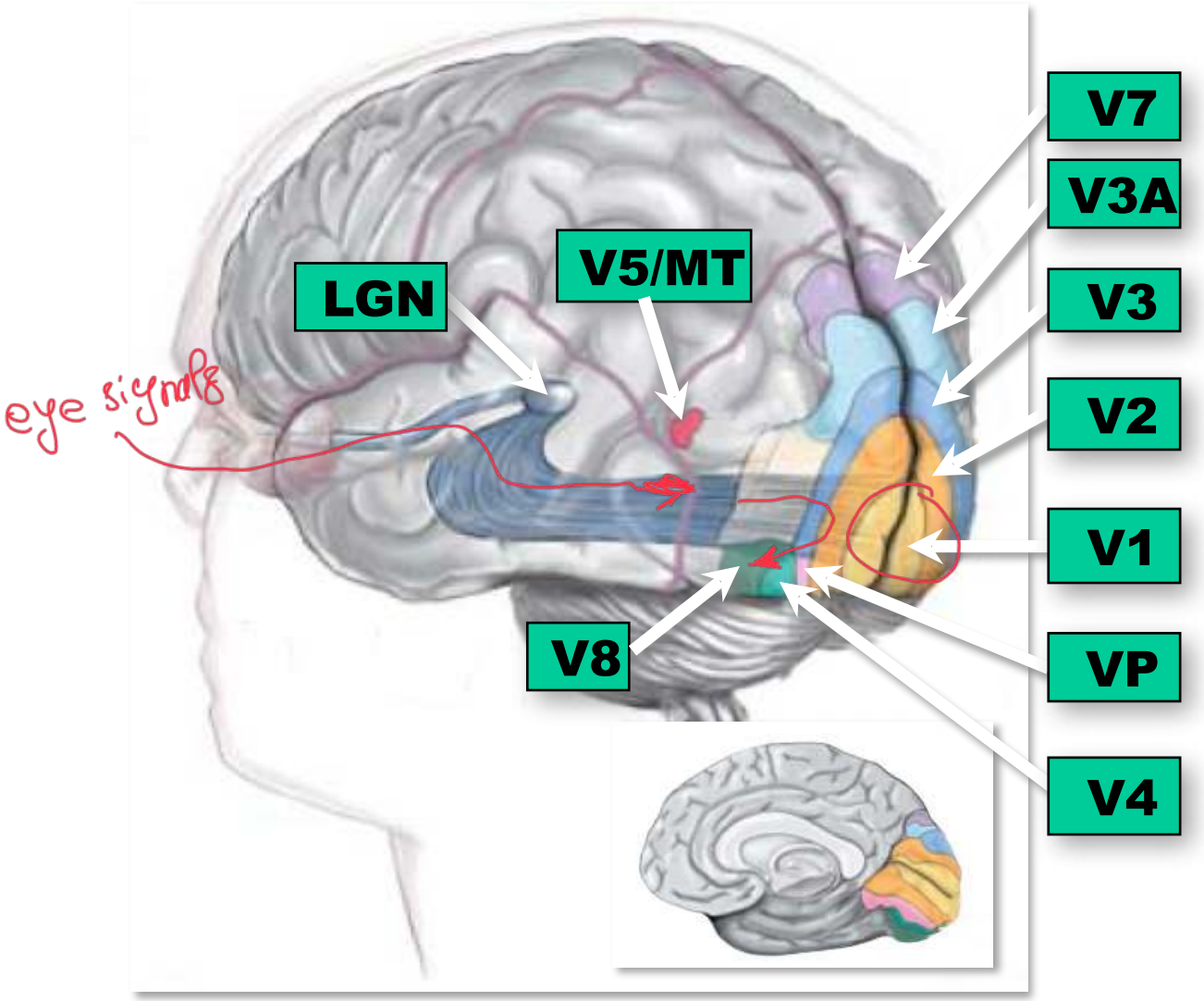
The Retina:

- rods (low-intensity light, night vision)
- cones (color-vision)
- Synapses and ganglions ⇝ circuitry
- Optic nerve fibers

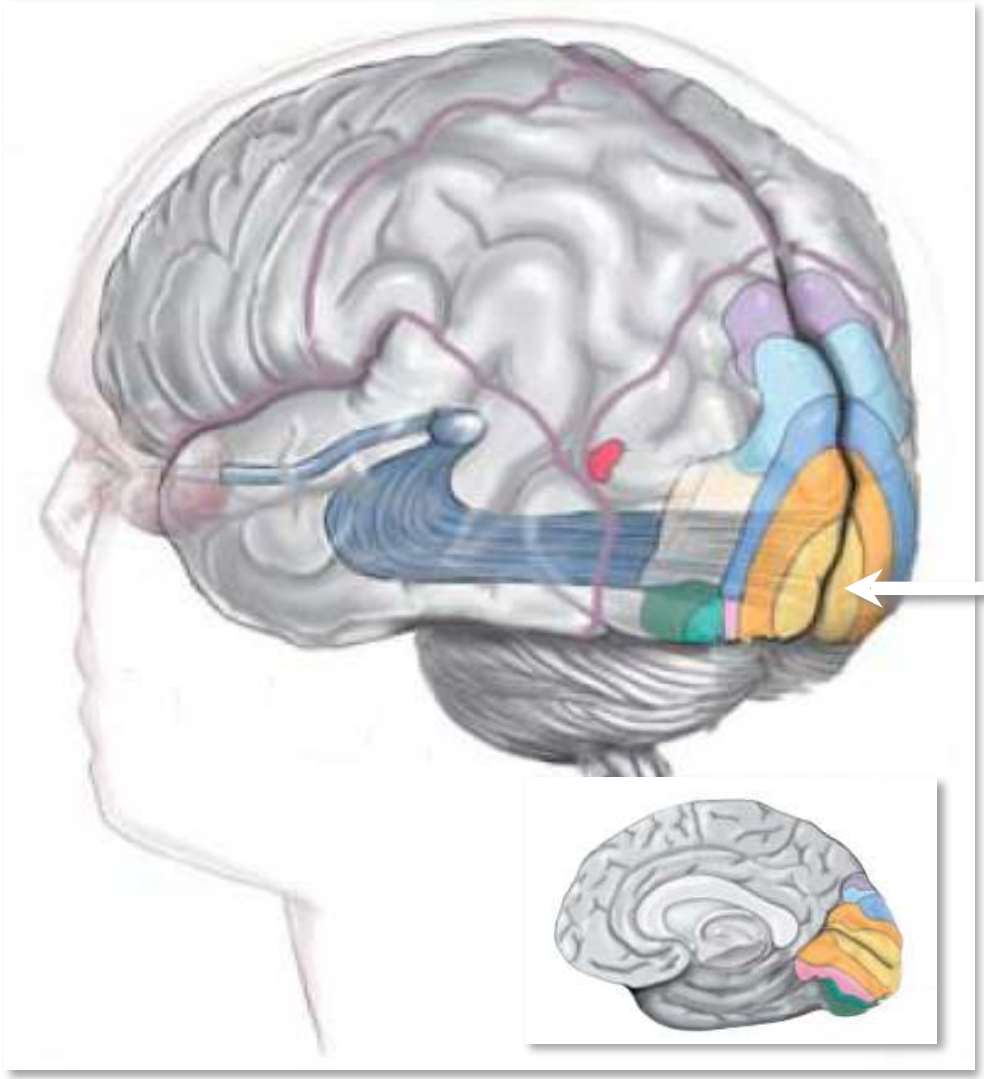
Sensing and low-level processing layer:

- 125 millions rods and cones feed into 1 million nerve fibers

Visual Cortex



Primary Visual Cortex (V1)



- Largest area in the visual cortex.
- 100 times as many neurons as retinal ganglion cells

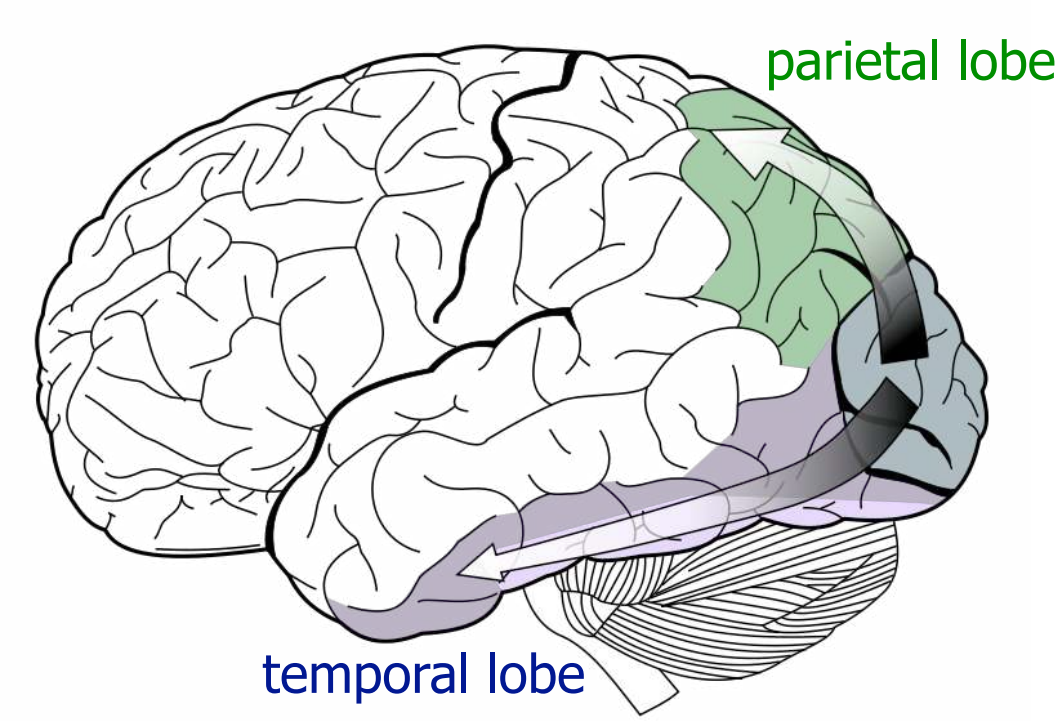
→ Overcomplete representation.

From V1 to the Others

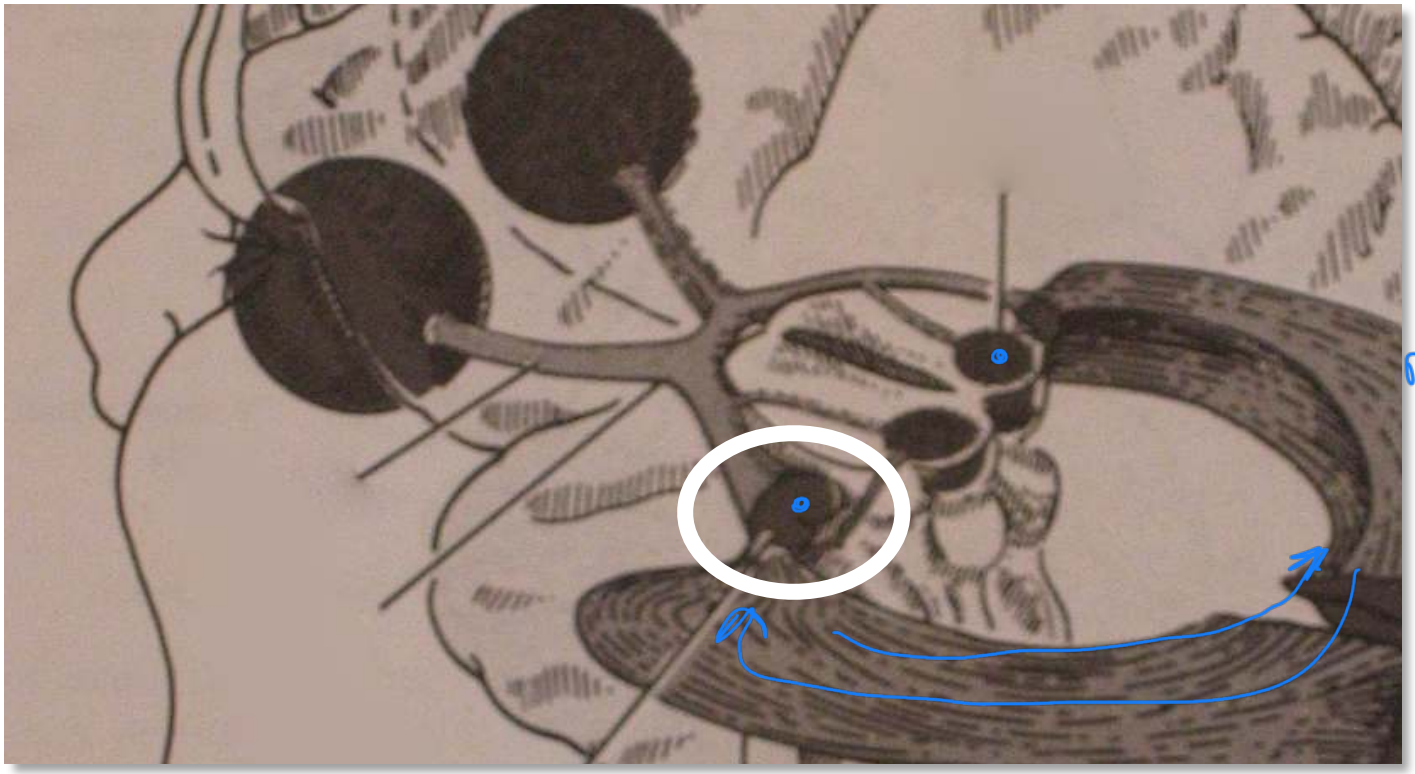
To pathways originate from V1:

- The "where" pathway: V1→V2→V5→parietal lobe.
- The "what" pathway: V1→V2→V3→V4→temporal lobe.

⇒ Motion Detection and Object Recognition are mostly performed in parallel but interconnections exist.



Lateral Geniculate Nucleus (LGN)



Receives feedbacks from V1 and V2. There is ten times more feedback than feedforward sent to V1.

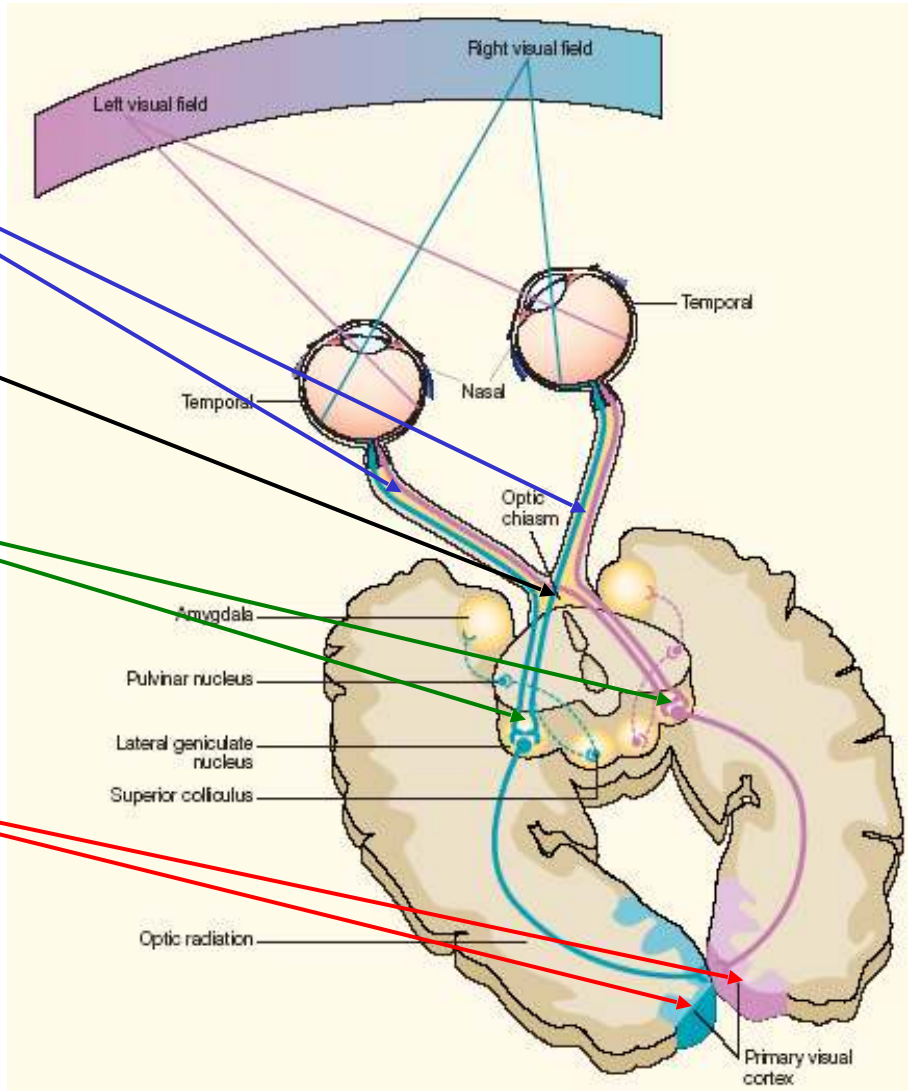
Hemispherical Vision

- Optical Nerves

Optic Chiasm

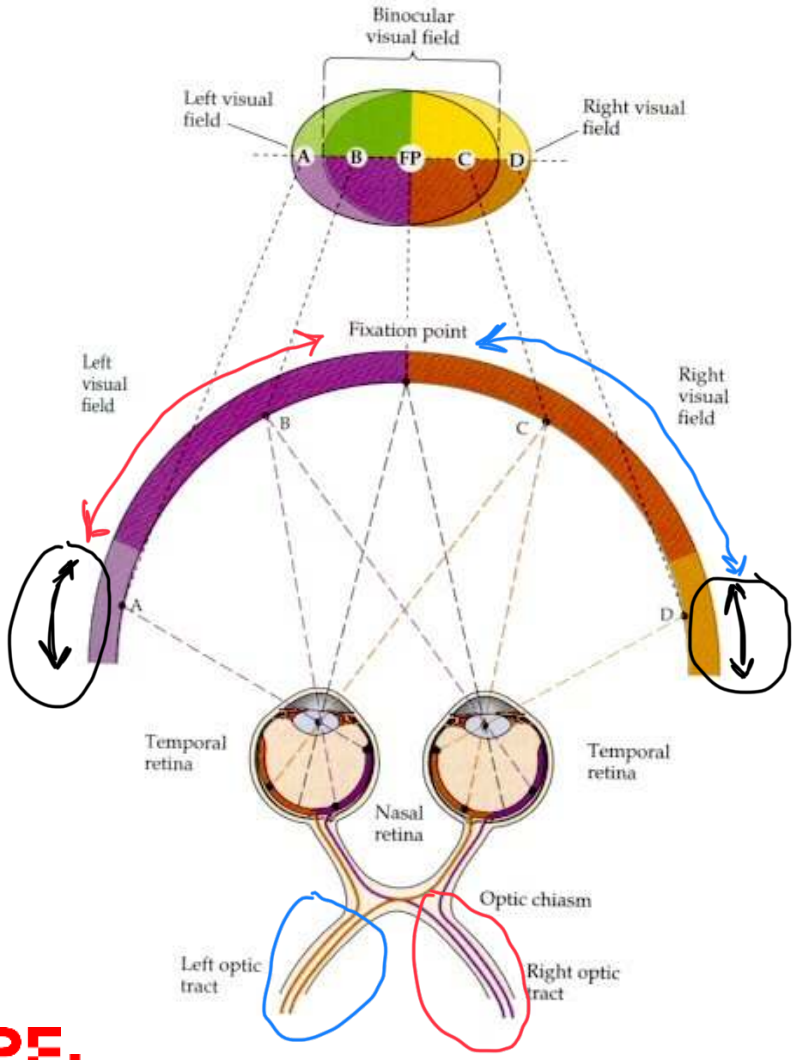
Lateral Geniculate Nucleus

Primary Visual Cortex (V1)



Stereoscopical Vision

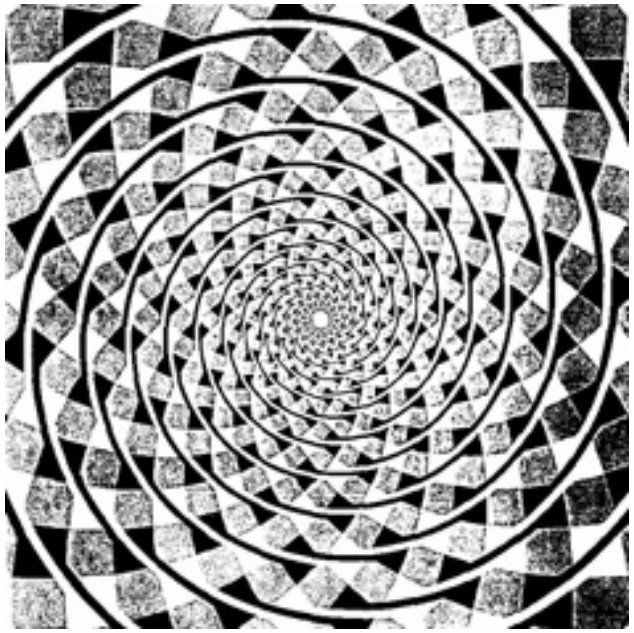
Our vision is based on predator, not prey



Our brain is wired for stereo vision:

- Redundancy
- Depth perception

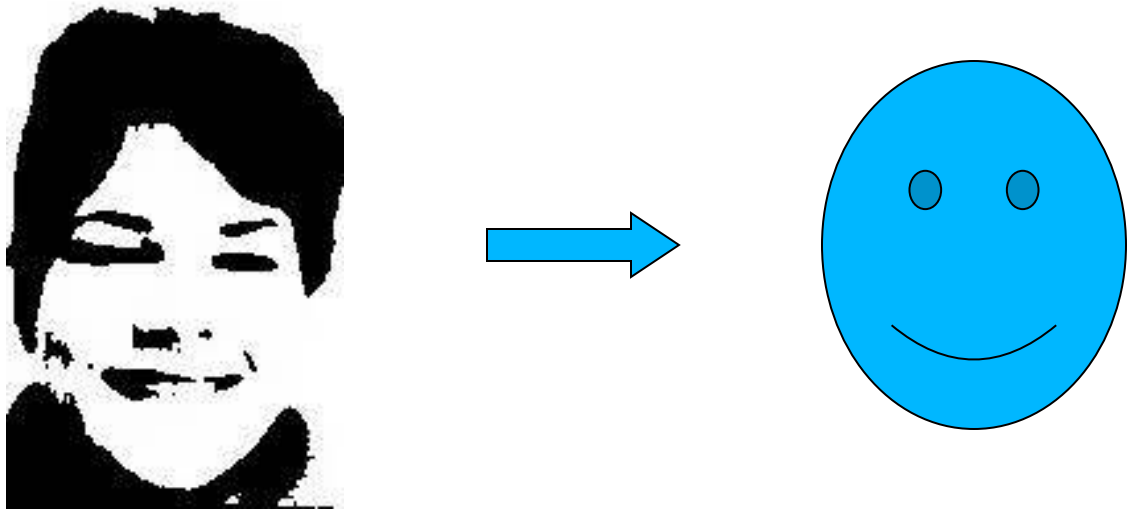
Optical Illusions



Every image is the image of thing merely to him who knows how to read it, and who is enabled by the aid of the image to form an idea of the thing.

Handbook of Physiological Optics
H. von Helmholtz

Controlled Hallucination?



Perhaps, but very cleverly implemented in “wetware”.

→ How can we emulate it in hardware?

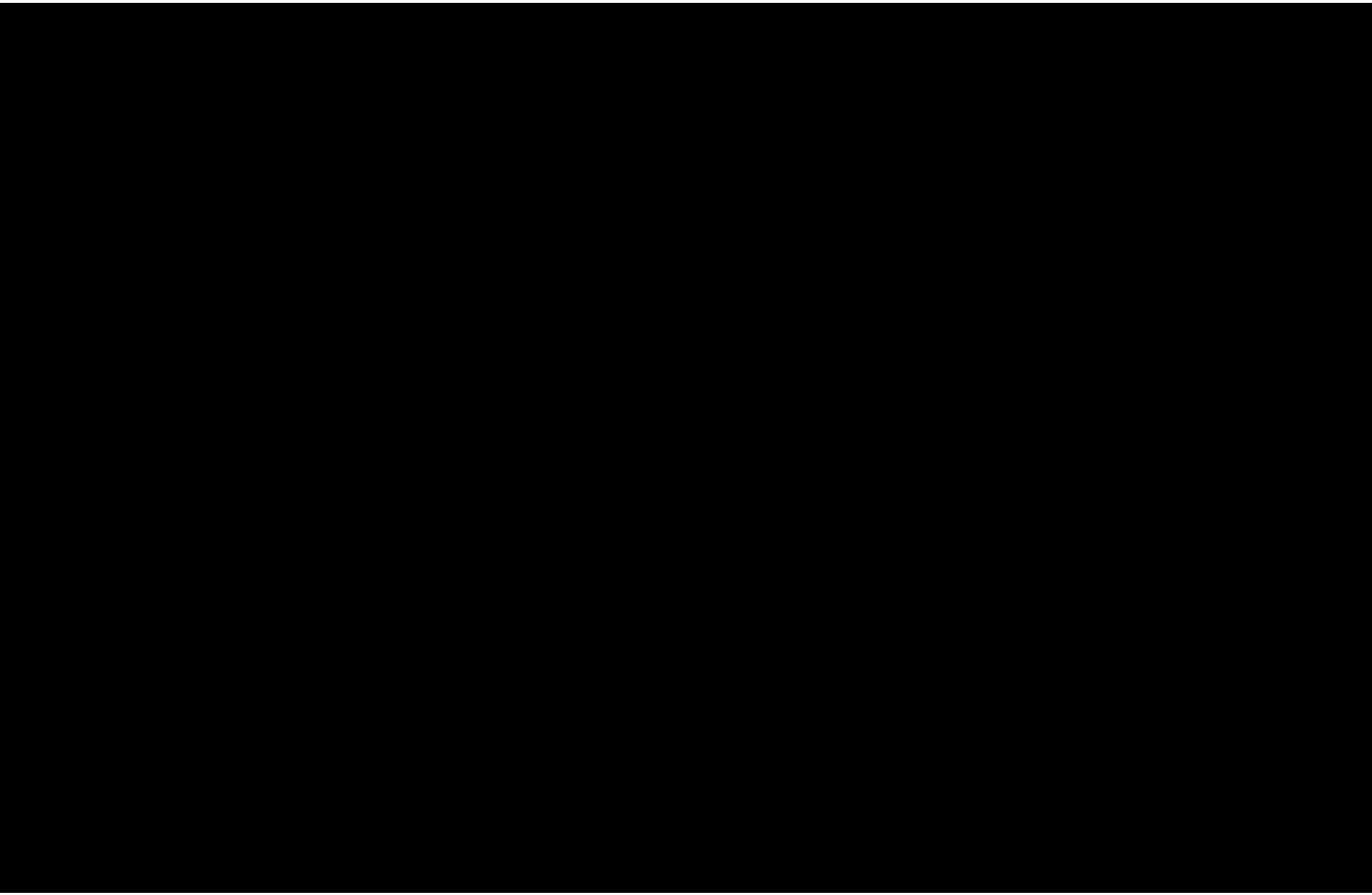
Recognize And Classify Animal vs No Animal

Subjects must raise their hand if they see an animal:

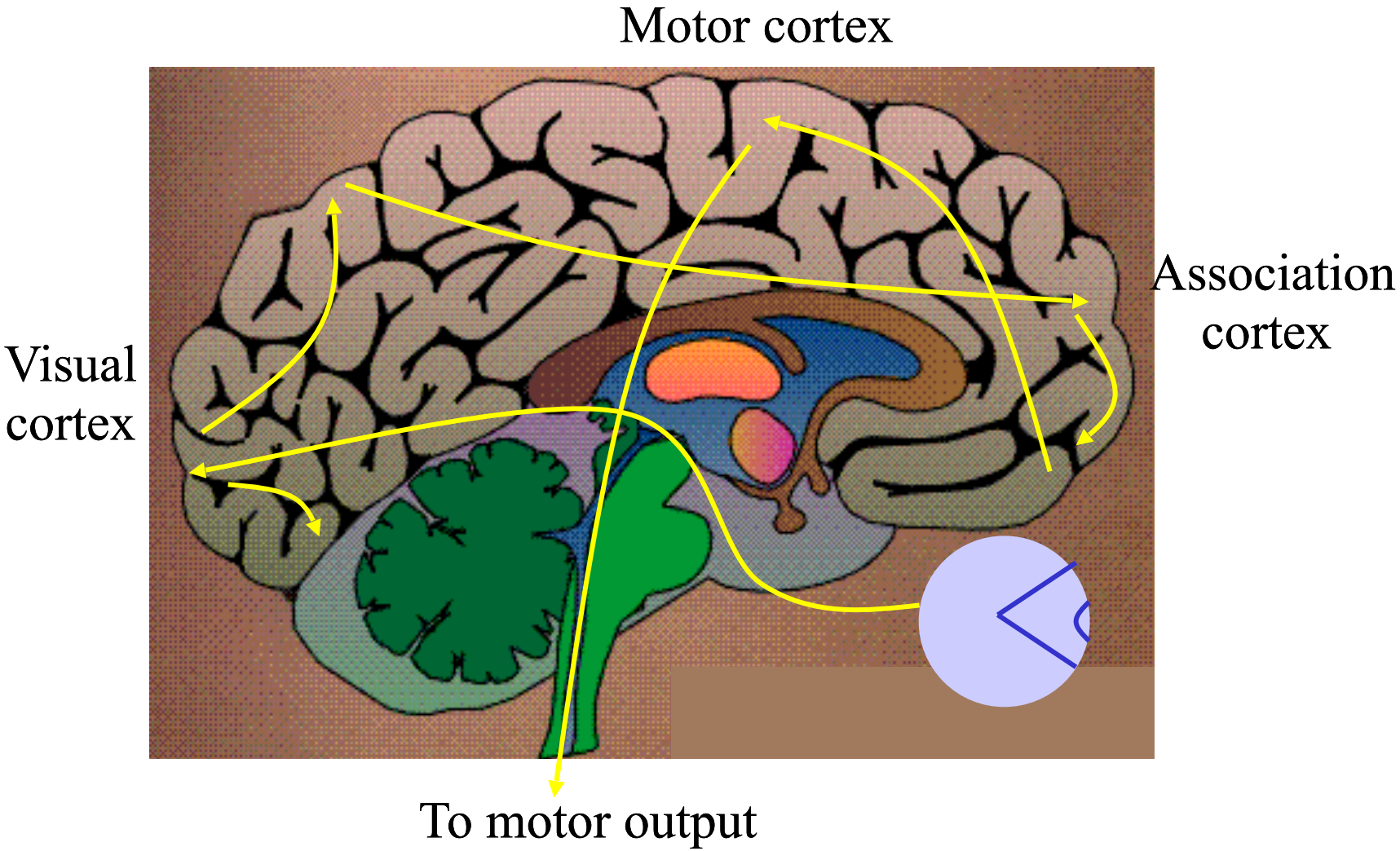
- 60 images
- 1 image per second

→ Measure their reaction time.

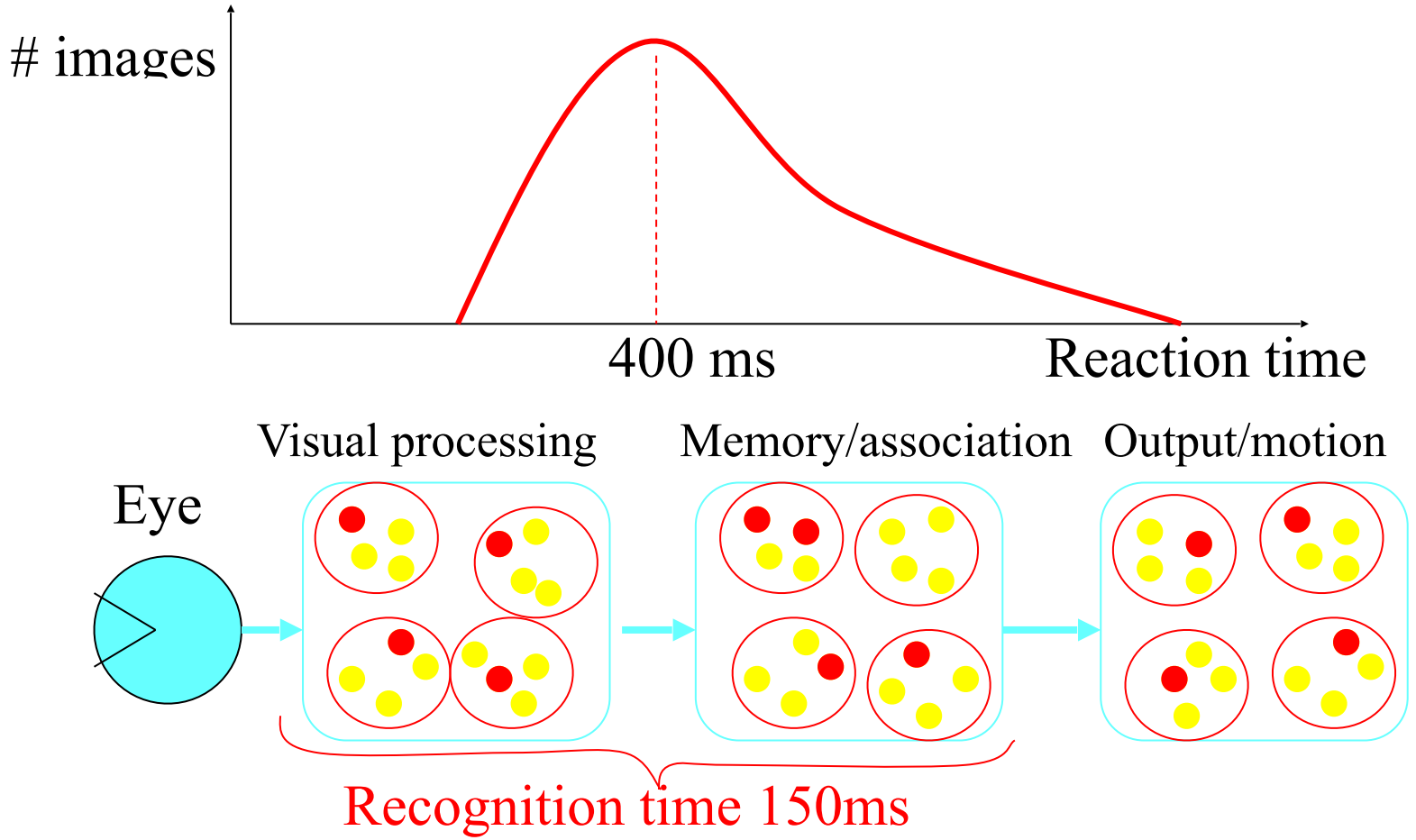
Simon Thorpe, *Nature*, 1996



Brain Pathways

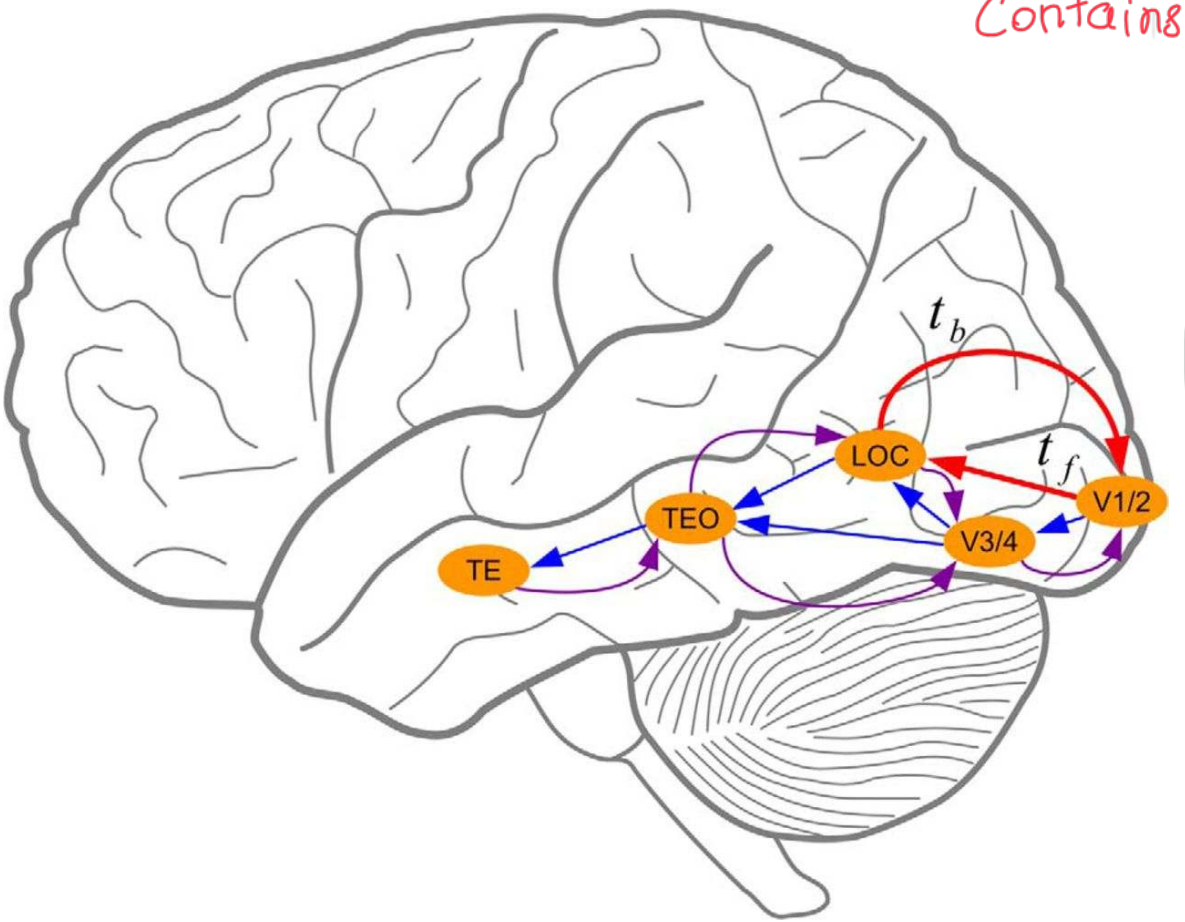


REACTION TIME



—> Suggests a purely feed-forward processing because there is not enough time for feedback loops.

Recurrent Pathways



"Shape stimuli are optimally reinforcing each other when separated in time by ~60 ms, suggesting an underlying recurrent circuit with a time constant (feedforward + feedback) of 60 ms."

Drewes et al. , *Journal of Neuroscience*, 2016

Human vs Computer Vision

The camera replaces the eye:

- Eye lens -> Camera optics
- Cones and rods -> Sensor array
- Ganglion cells -> Filter banks

The computer replaces the brain:

But how?