Vision: a most remarkable sensor

random stuff Peter finds interesting and hopes you might as well

AUSTRALIAN CENTRE FOR ROBOTIC VISION SUMMER SCHOOL

Our senses

- hearing
- smell
- taste
- sight
- touch
- balance

- echo location
- electric field
- magnetism

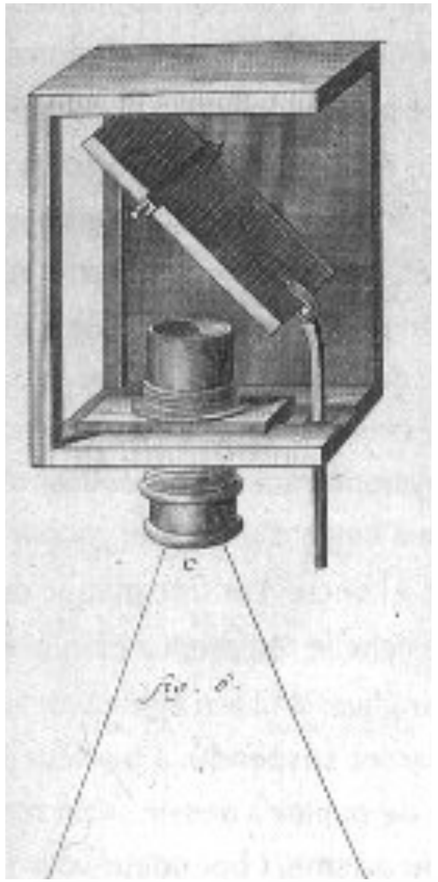
We use vision for myriad daily tasks including manipulation and navigation

- recognition of objects, people, places
- prediction
- navigation


experiment time

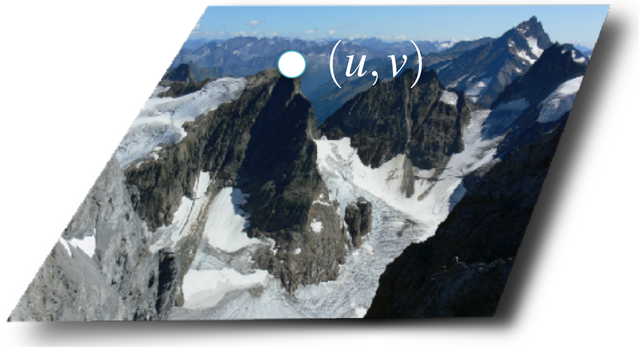

WC


A
B



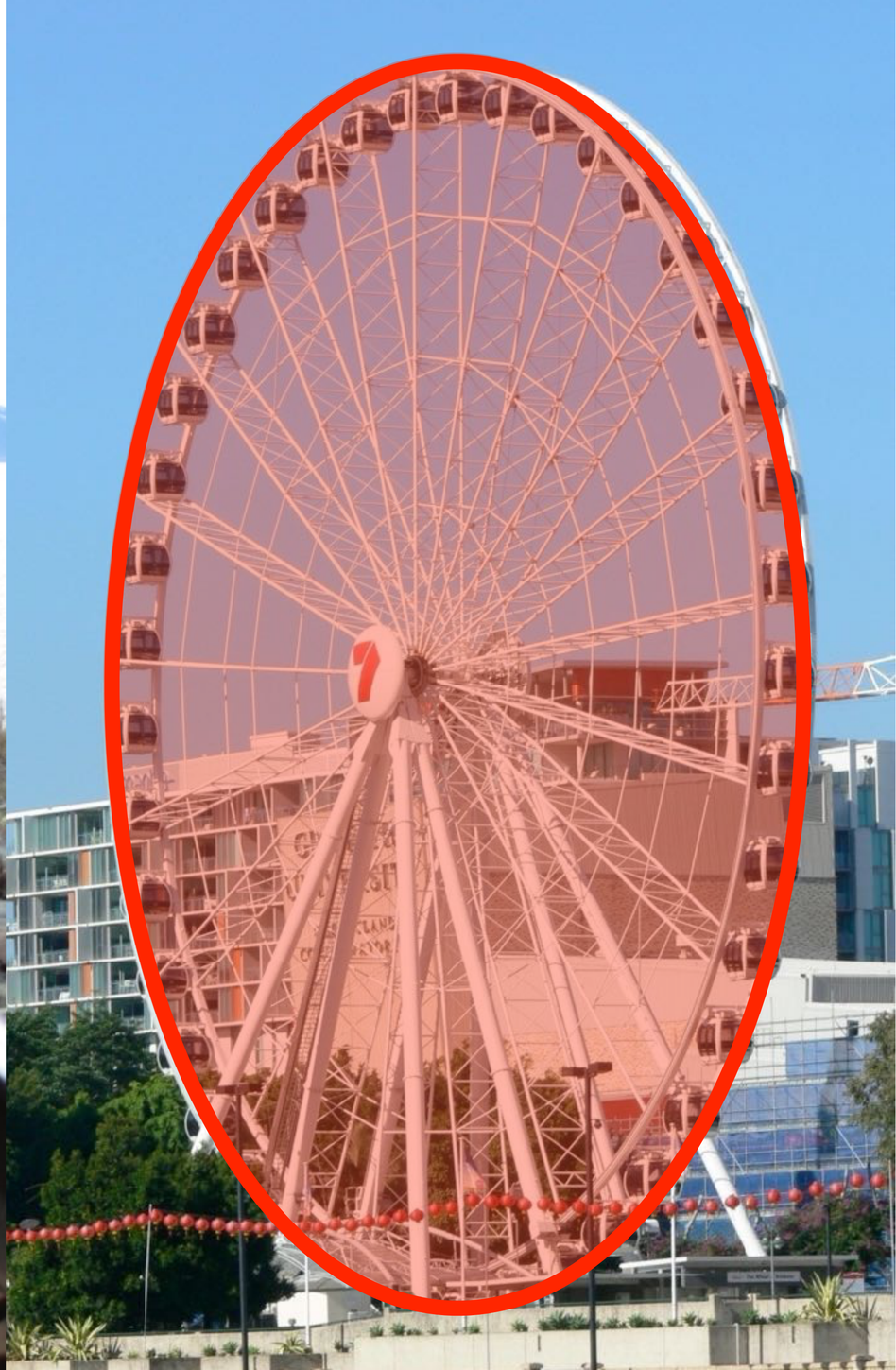
$$(X,Y,Z) \mapsto (u,v)$$

$$\mathbb{R}^3 \mapsto \mathbb{R}^2$$



$$u = \frac{fX}{Z}$$

$$v = \frac{fY}{Z}$$



Cave Paintings ~40,000 years ago

Ideal City (1470)
Piero della Francesca (1415–1492)

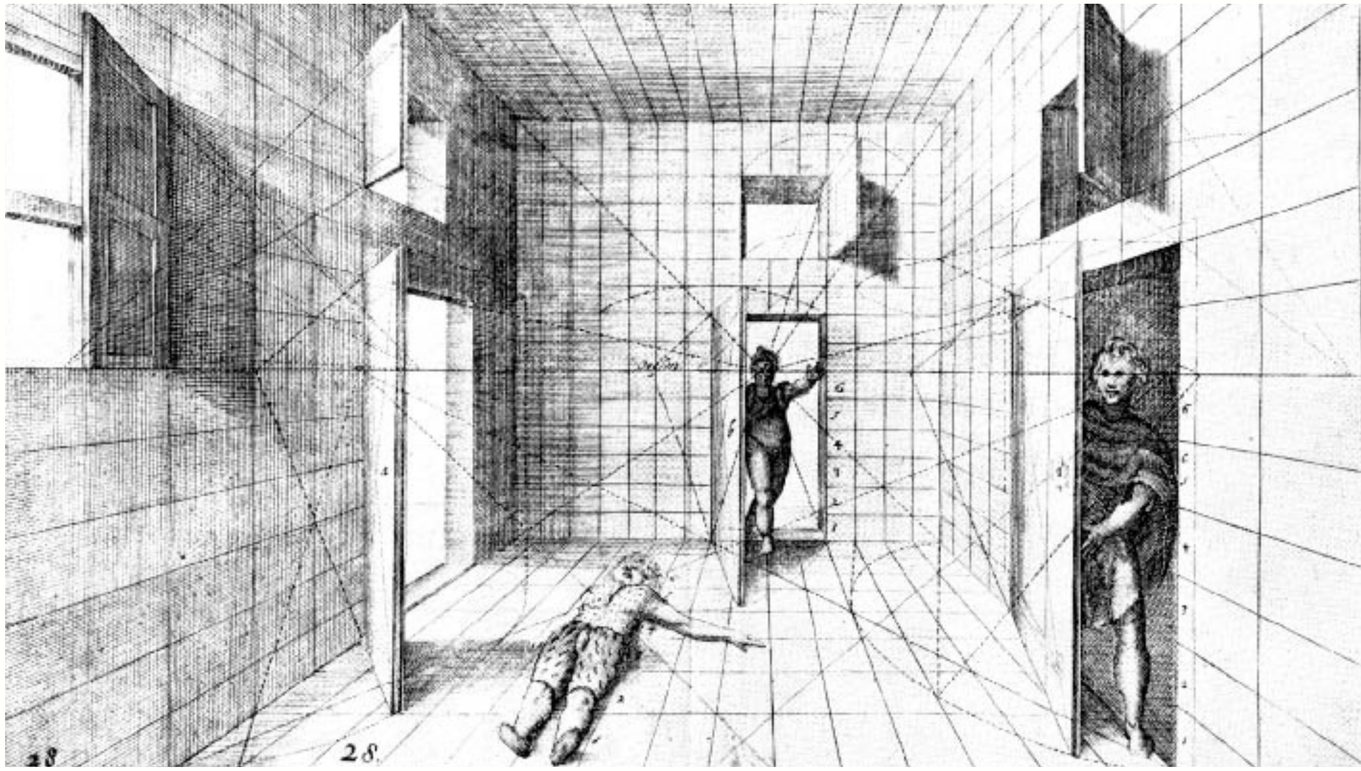
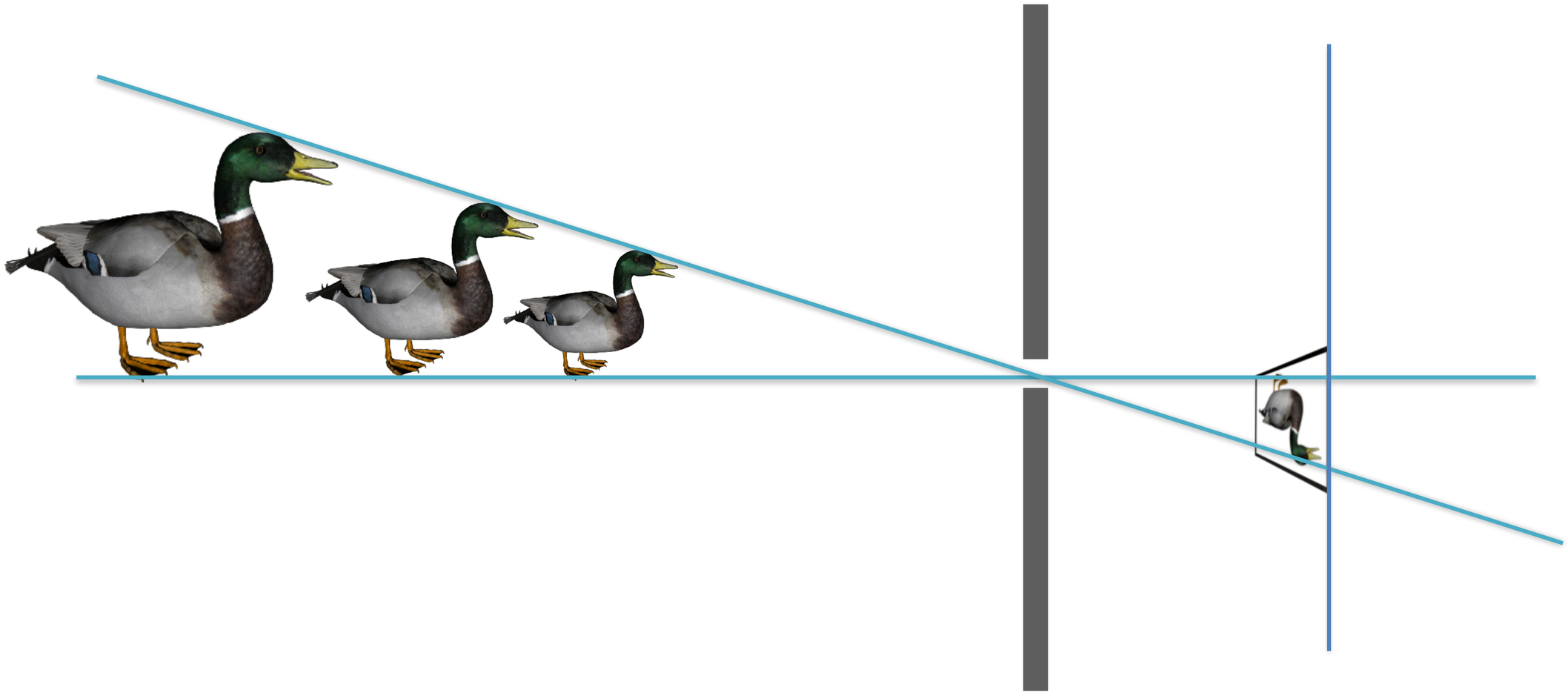


Figure 28 (Jan Vredeman de Vries, 1604).
Used with permission from Perspective, Dover Publications,1964.

People are actually avoiding walking in the "hole“ 2007
Joe Beever | CC A2.0

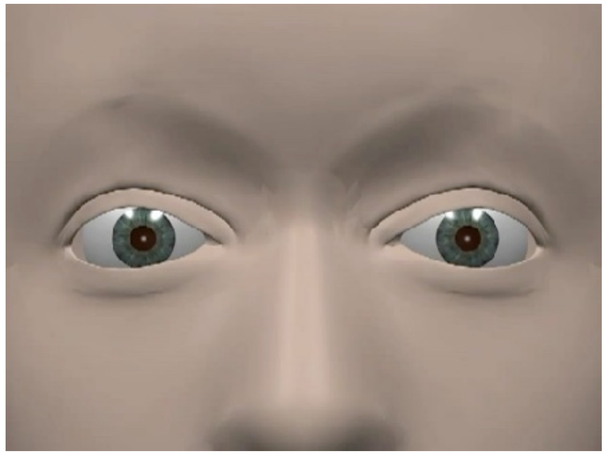
No unique inverse



Humans use eight different tricks to estimate distance

binocular stereo

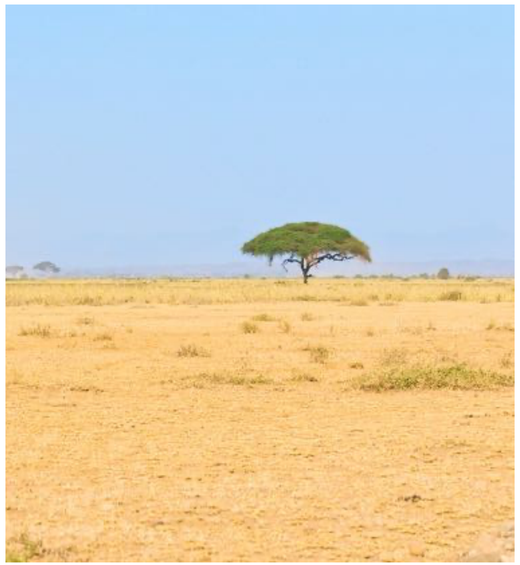
accomodation



vergence

motion perspective

occlusion

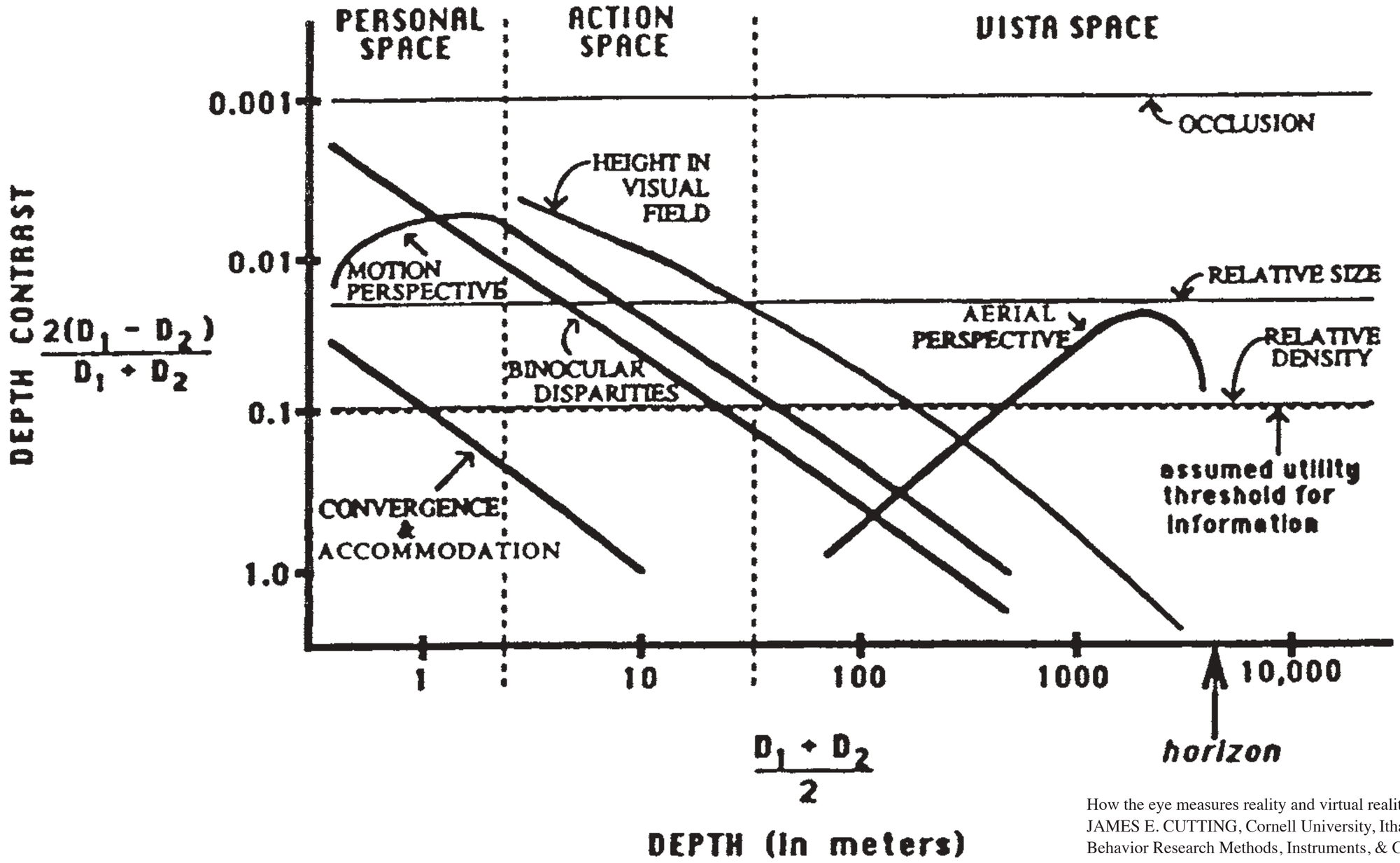


apparent height

texture density

aerial perspective

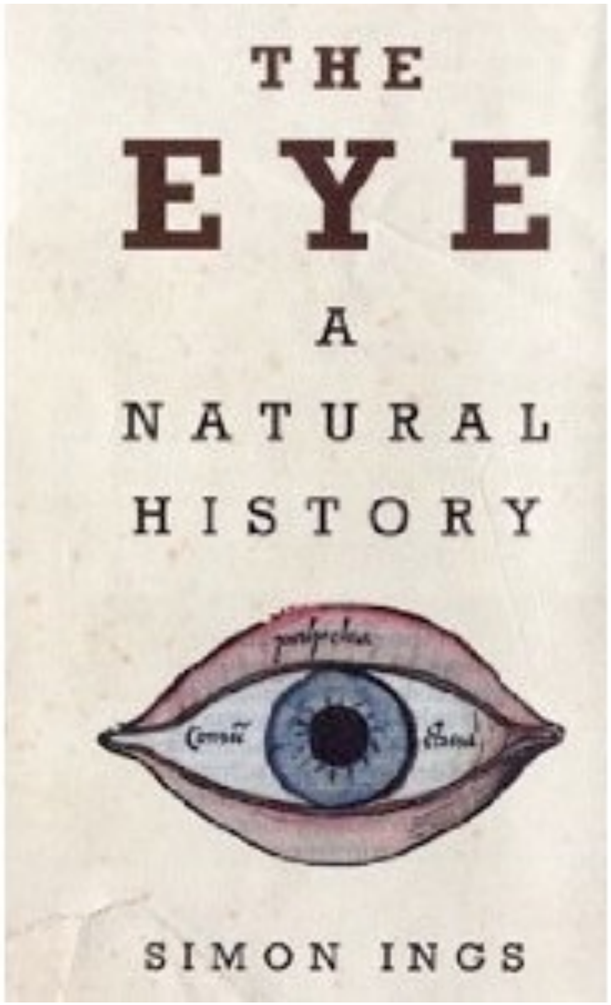
Determining distance

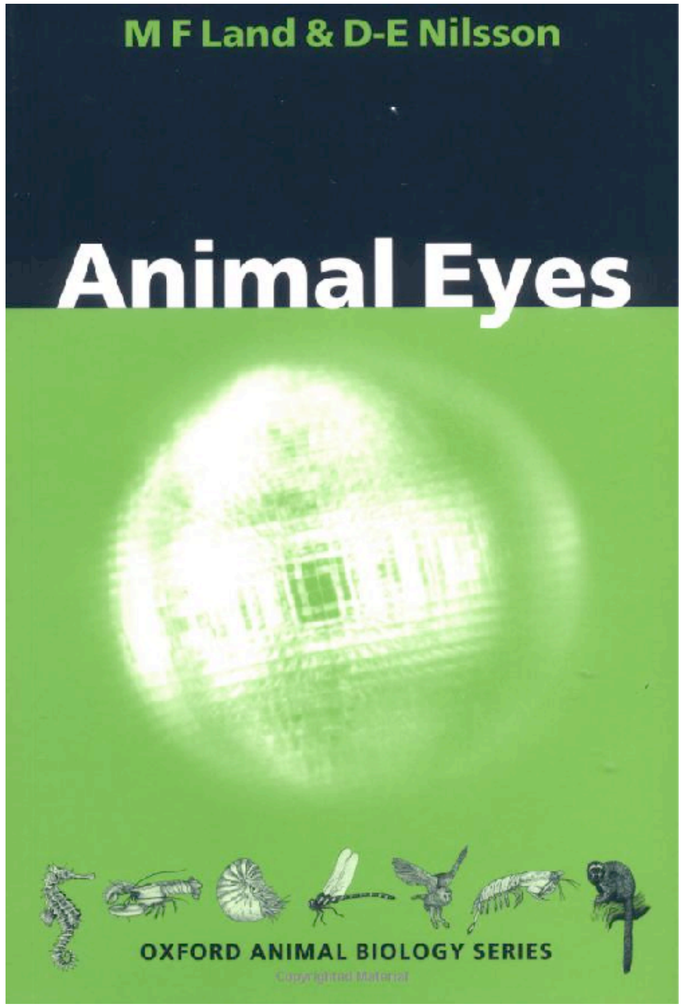


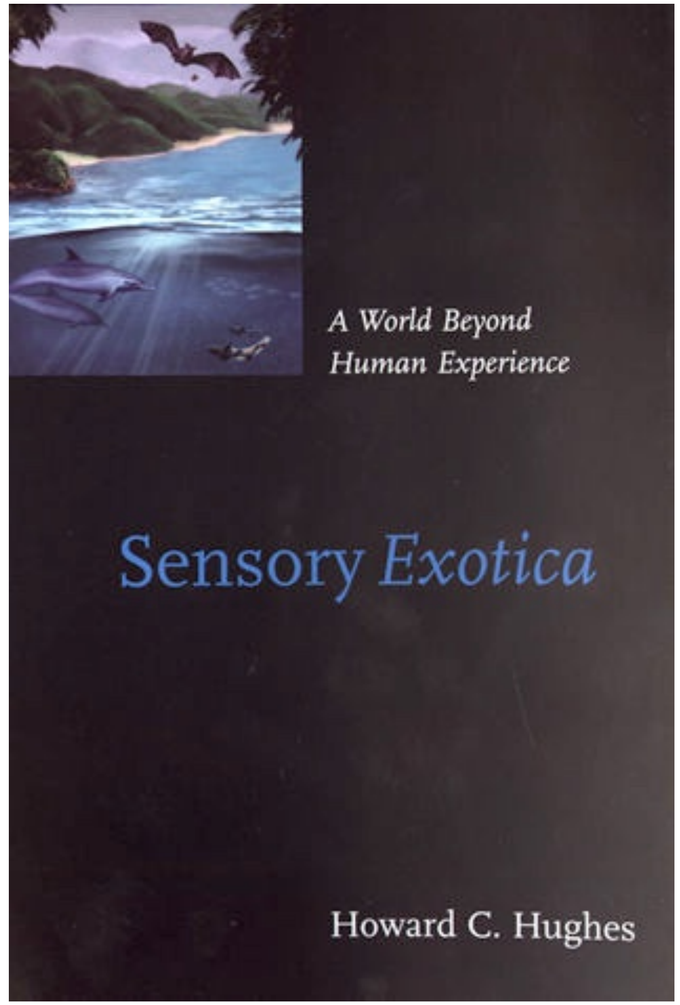
How the eye measures reality and virtual reality.
JAMES E. CUTTING, Cornell University, Ithaca, New York.
Behavior Research Methods, Instruments, & Computers 1997, 29 (1),

Puzzling World, Lake Wanaka, New Zealand

Further reading







Vision is a key sensing modality for almost all organisms: simple or complex

- Bee brain
  - 1g
  - 10^6 neurons

- Human brain
  - 1500g
  - 10^{11} neurons

Human vision data sheet

- 120Mpixel
- Total dynamic range 10^6:1 (20bit)

- Rod cells
  - ~110M
  - respond to single photons
- Cone cells
  - ~4.5M
  - high resolution
  - 2um across

×2

- 3 gyroscopes
- 2 accelerometers

×2

- Vision engine
  - $3x10^{10}$ neurons
  - 500g
  - 6W

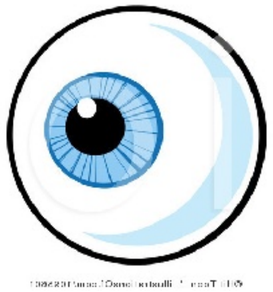

experiment time

thalmus
(day/night)

pretectum
(pupil control)

superior colliculus
(gaze control)

126M pixels



1.2M axons

10%

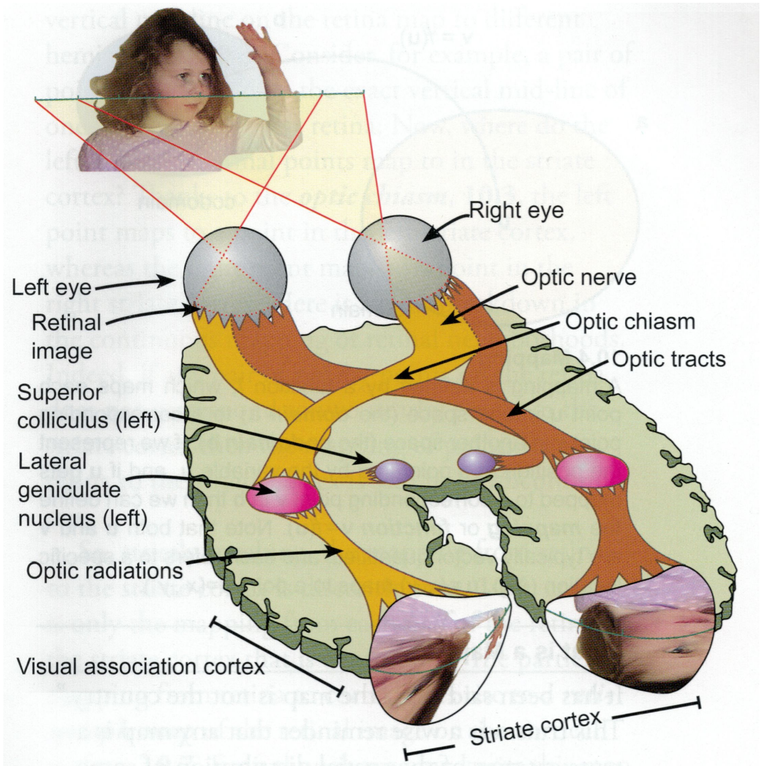
Lateral
genuculate
nucleus (LGN)

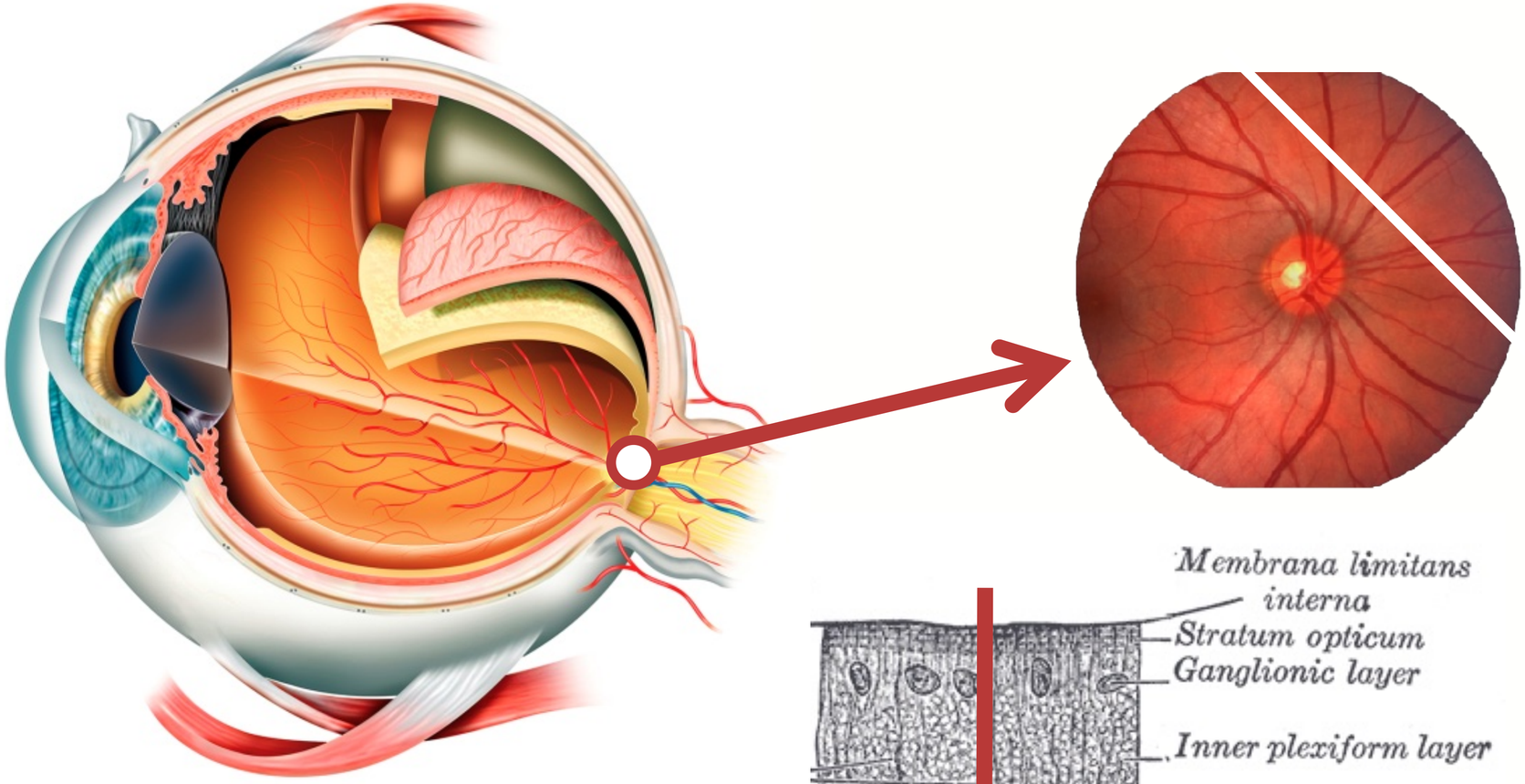
V1 (striate cortex)

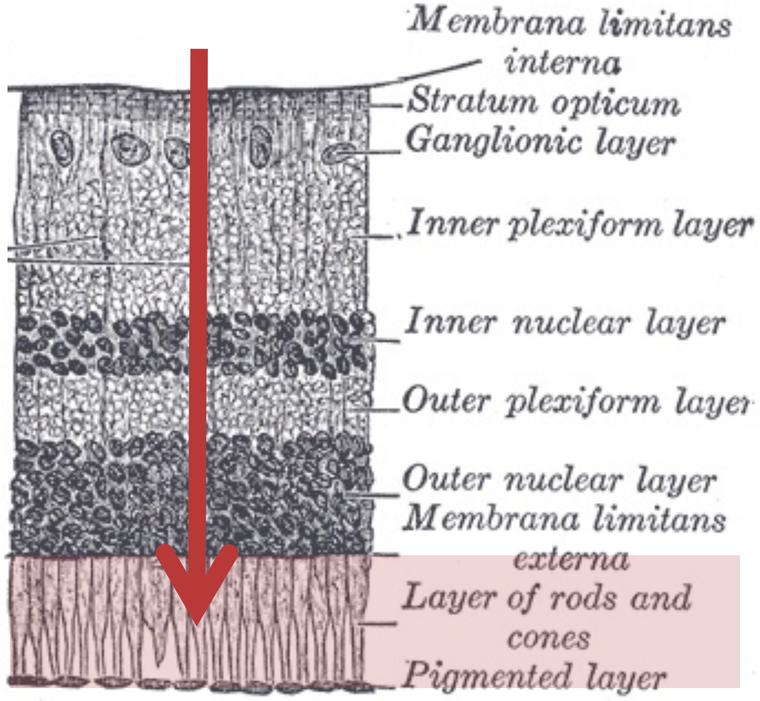
80%

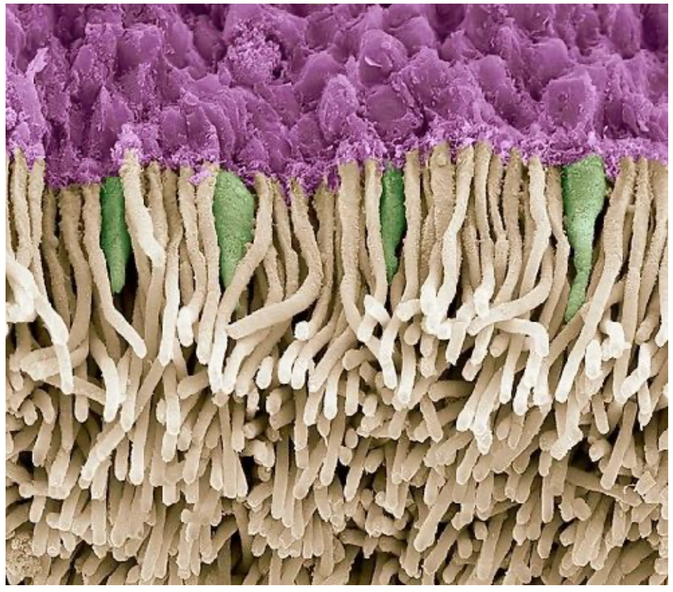
10%

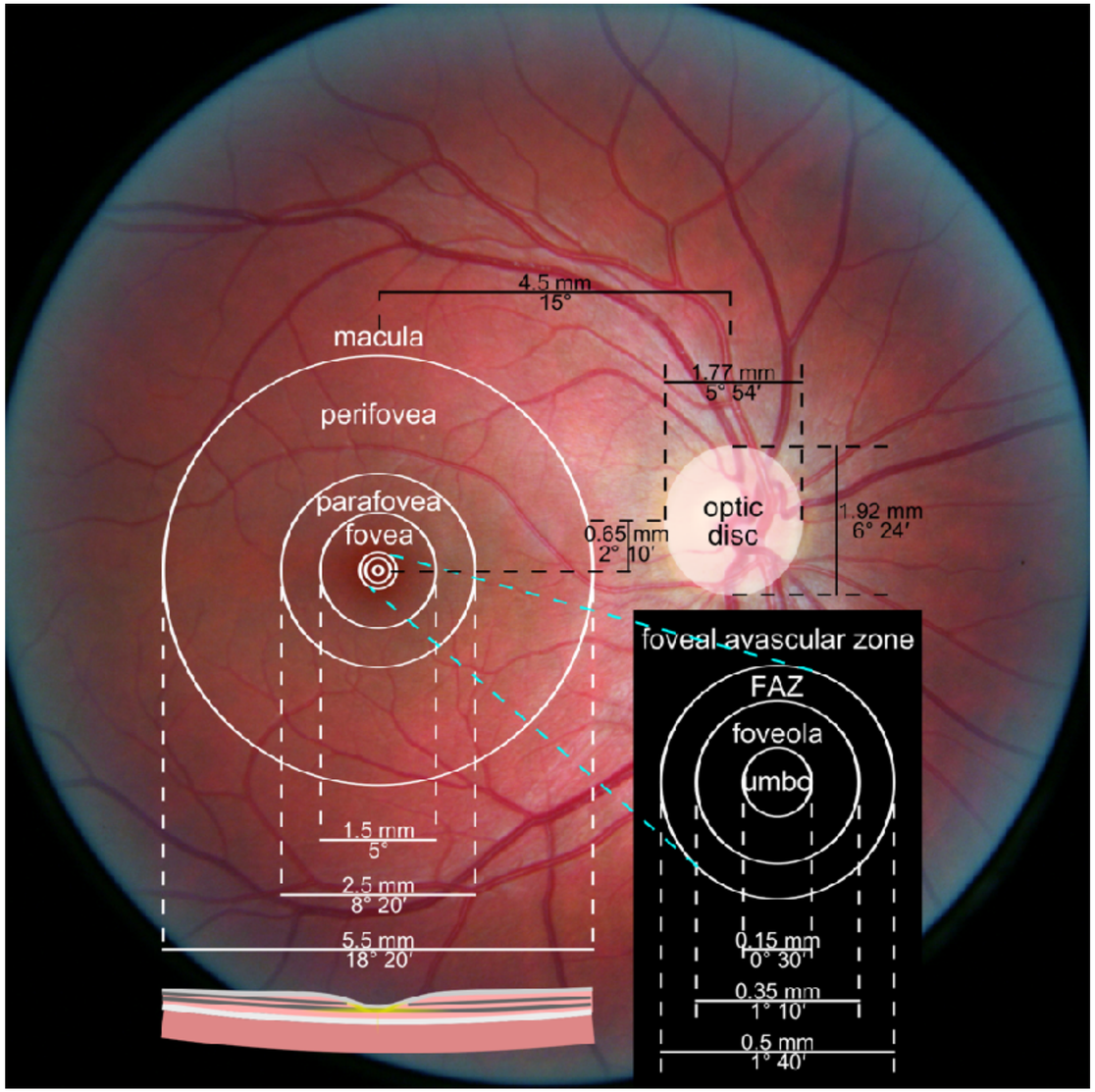
brain stem

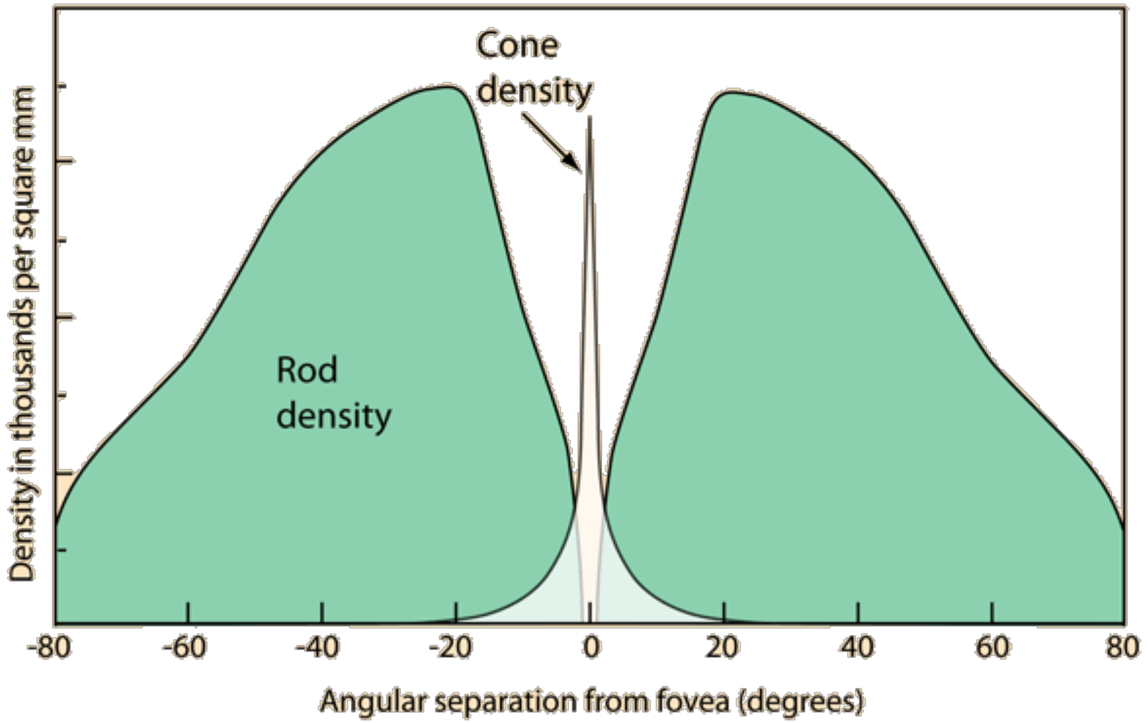





Membrana limitans interna
Stratum opticum
Ganglionic layer
Inner plexiform layer
Inner nuclear layer
Outer plexiform layer
Outer nuclear layer
Membrana limitans externa
Layer of rods and cones
Pigmented layer

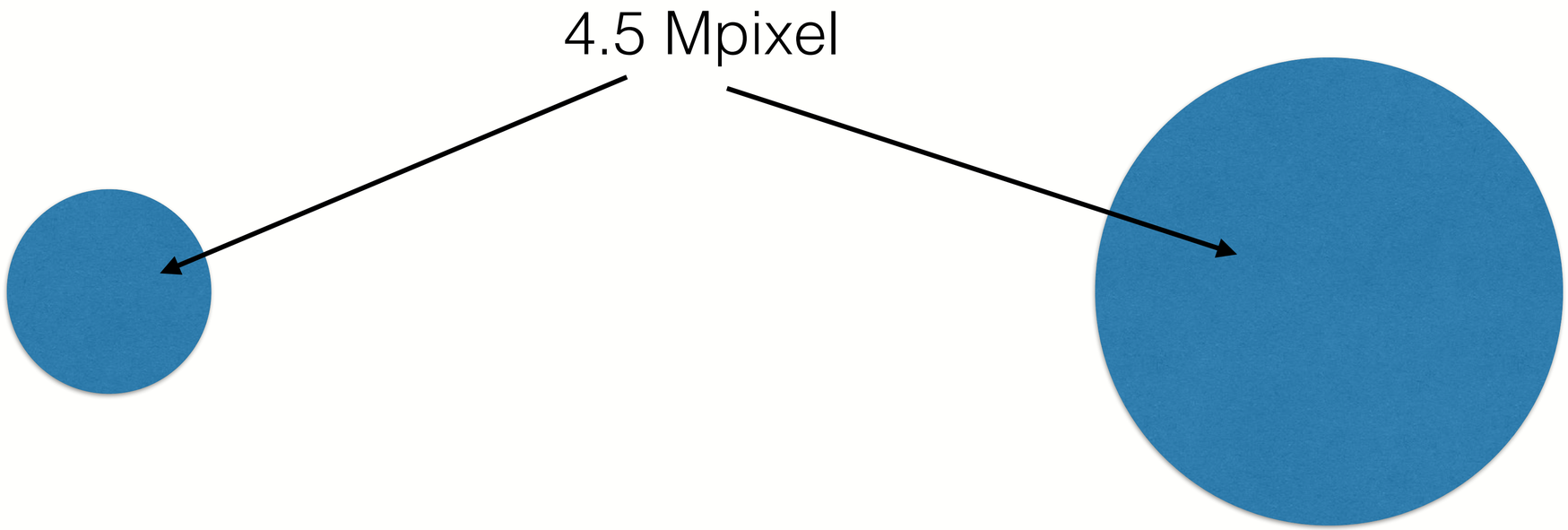






~90 million rod cells (can't see red)
~5 million cone cells (see color)

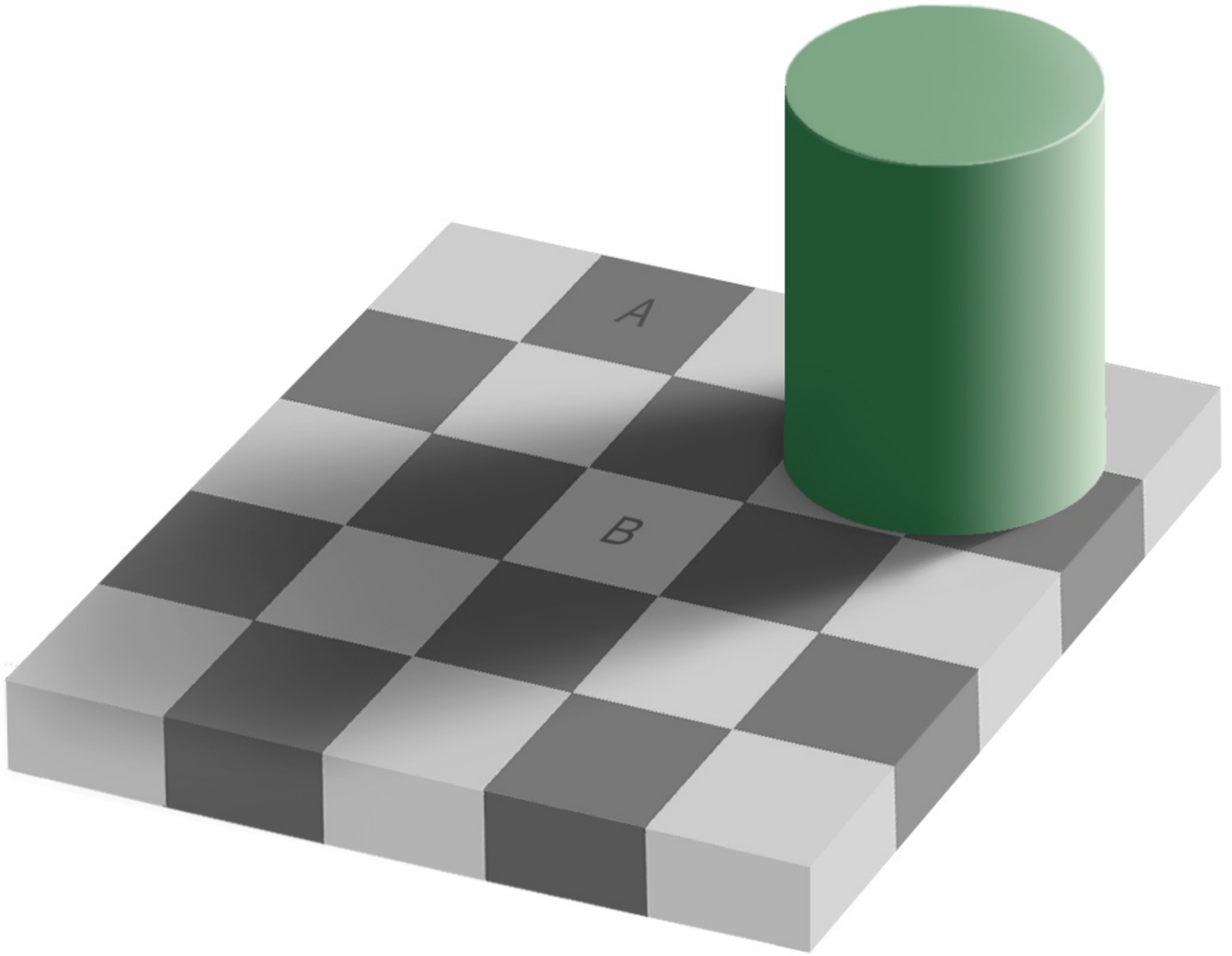
Your 5° foveal field of view



if you are 600mm in front
of a 27” monitor

if you are 5m in front
of a 2m wide
projection screen

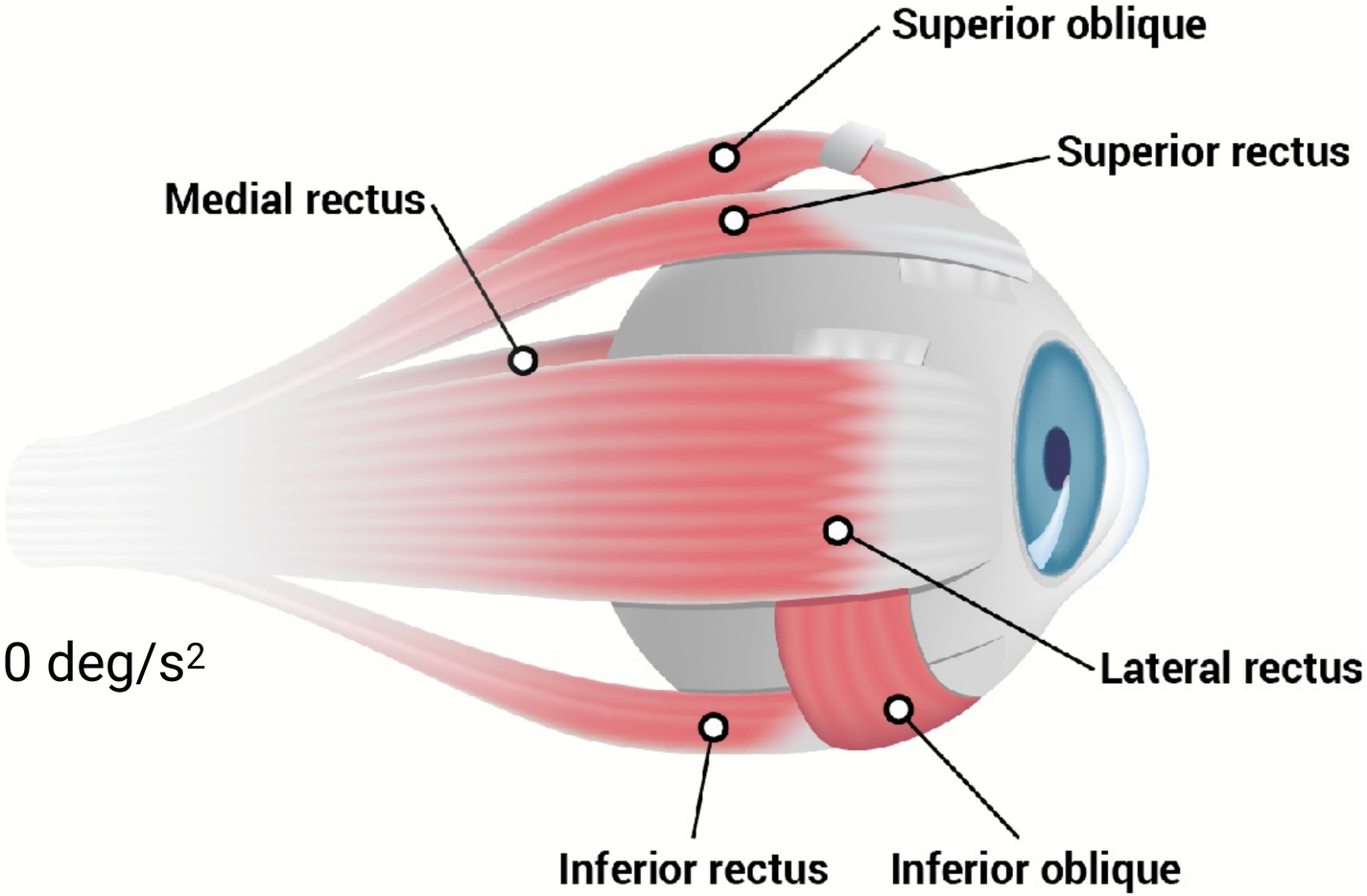
Unpacking reflectance and illumination



Why can’t you see your own nose?

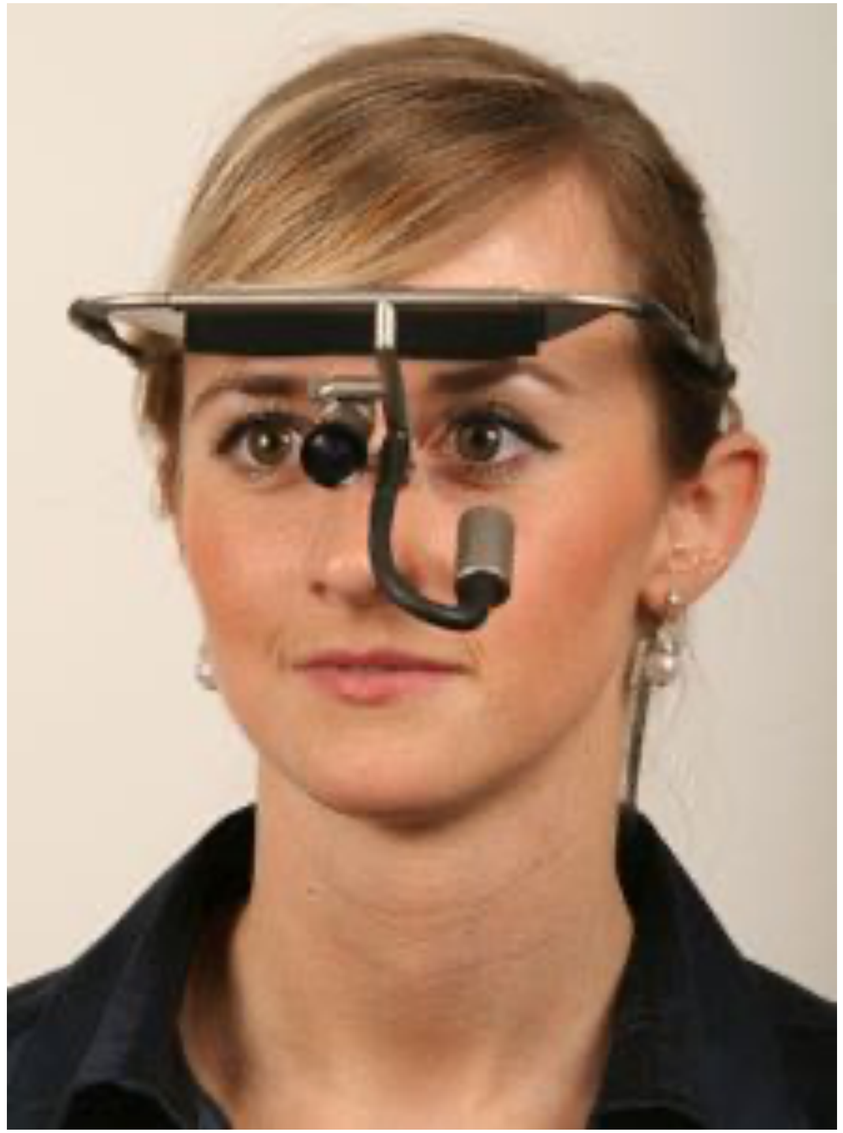
Seeing is an active process

- Fastest moving muscles in the human body
  - rotational speed: 600deg/s
  - rotational acceleration: 35,000 deg/s^2



Muscles of the eye, Lateral view

Looking at looking



Subject wearing the Ergoneers Dikablis Eye-Tracking Glasses
2012

Which question was the test subject responding to?

a) What are the material circumstances of the family?

b) What type of clothes are the family wearing?

c) What age are the figures in the painting?

Figure 109 "Seven Records of eye movements by the same subject" 1967
Yarbus, A. L.

Artwork: *Unexpected visitors*
Ilya Repin | 1884-1888

Which question was the test subject responding to?

a) What are the material circumstances of the family?

b) What type of clothes are the family wearing?

c) What age are the figures in the painting?

Figure 109 "Seven Records of eye movements by the same subject" 1967
Yarbus, A. L.

Artwork: *Unexpected visitors*
Ilya Repin | 1884-1888

Which question was the test subject responding to?

a) What are the material circumstances of the family?

b) What type of clothes are the family wearing?

c) What age are the figures in the painting?

Figure 109 "Seven Records of eye movements by the same subject" 1967
Yarbus, A. L.

Artwork: *Unexpected visitors*
Ilya Repin | 1884-1888

Which question was the test subject responding to?

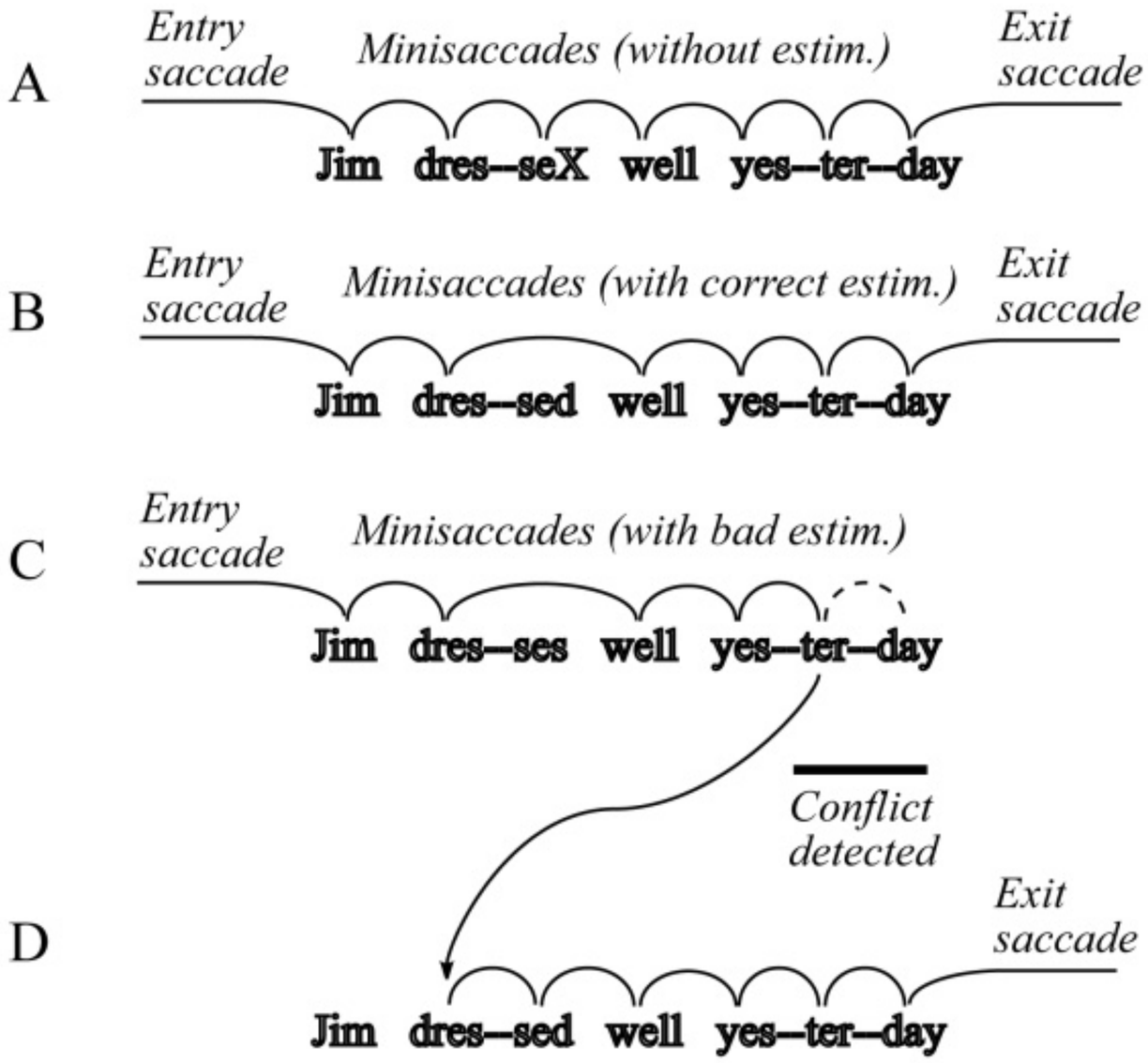
a) What are the material circumstances of the family?

b) What type of clothes are the family wearing?

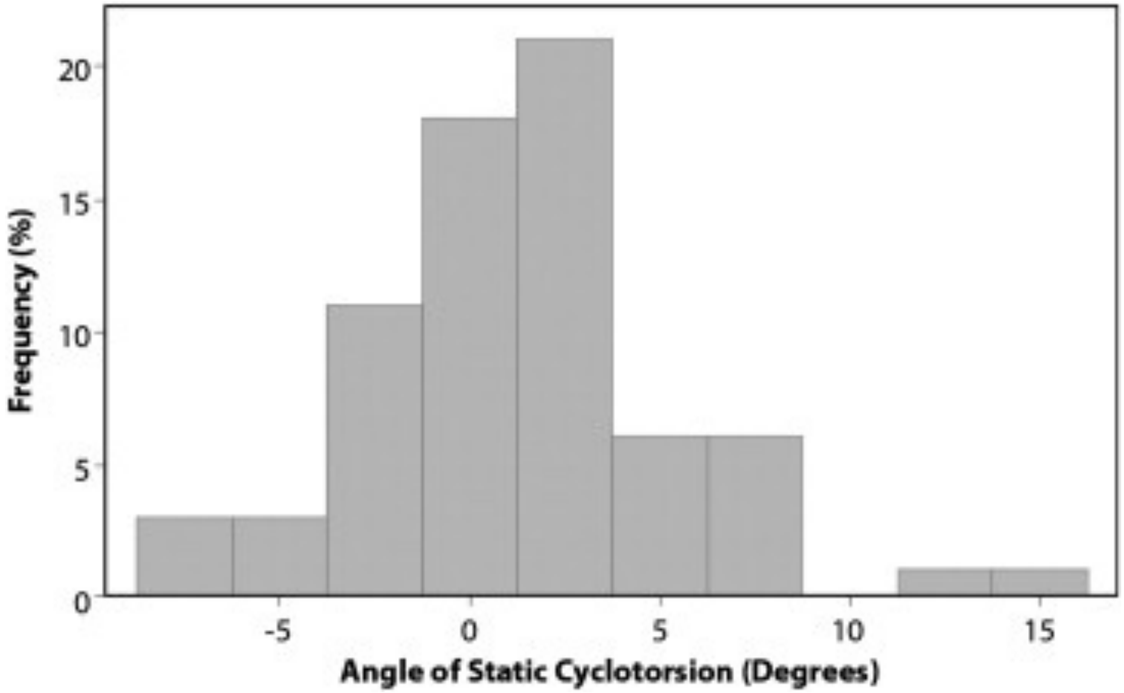
c) What age are the figures in the painting?

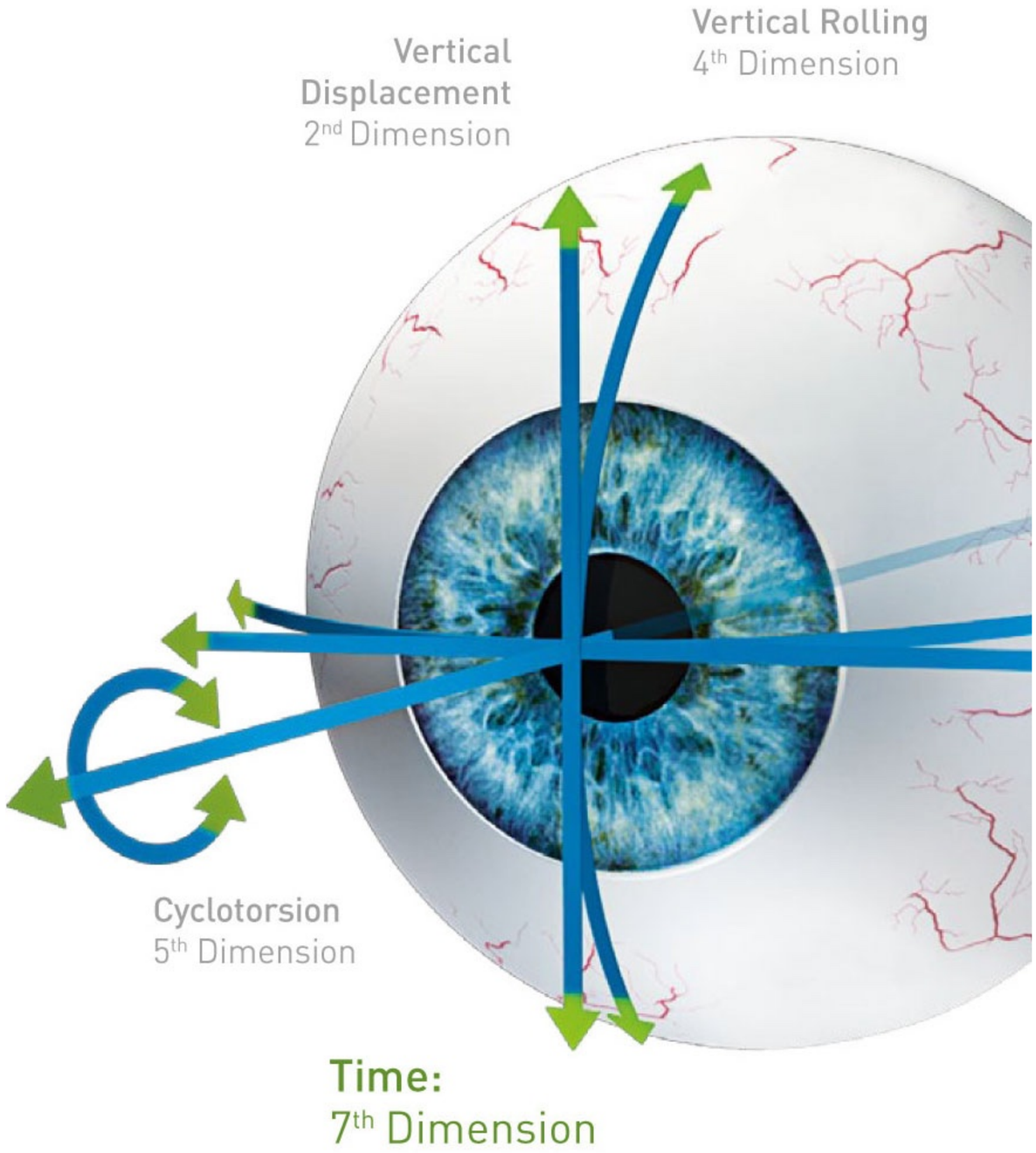
Figure 109 "Seven Records of eye movements by the same subject" 1967
Yarbus, A. L.

Artwork: *Unexpected visitors*
Ilya Repin | 1884-1888

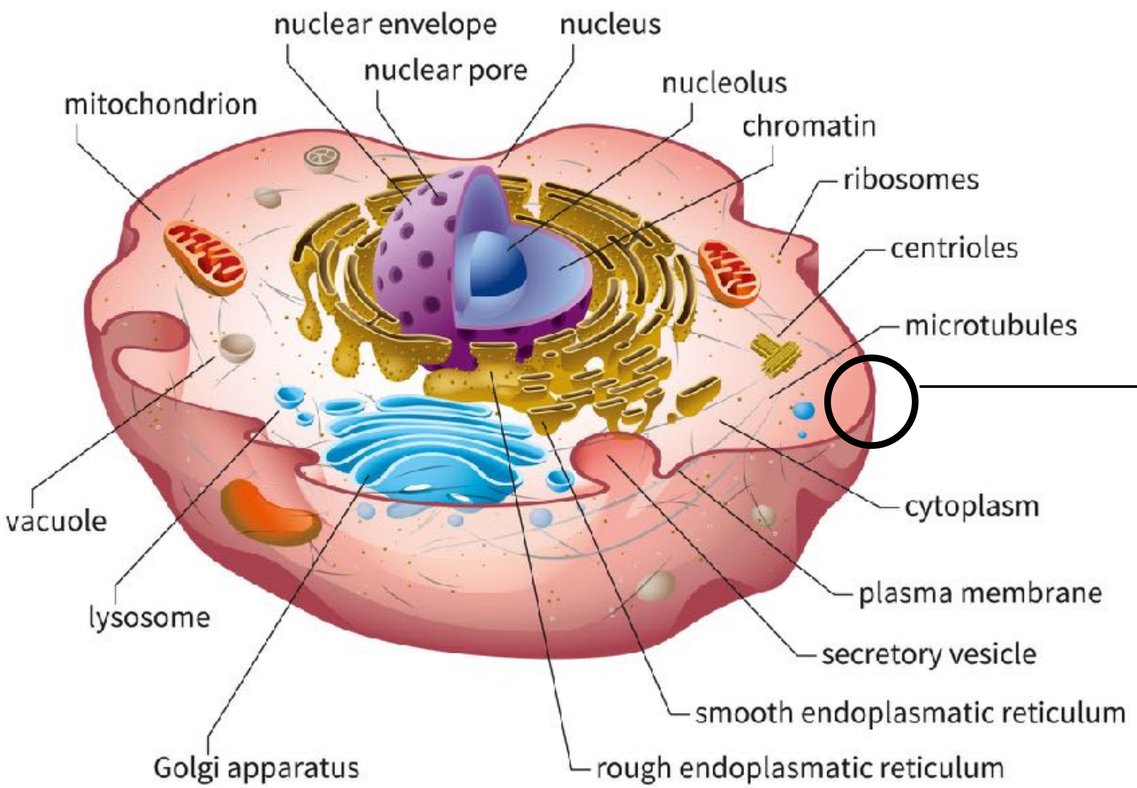

A
Entry saccade
Minisaccades (without estim.)
Exit saccade
Jim dres--seX well yes--ter--day
B
Entry saccade
Minisaccades (with correct estim.)
Exit saccade
Jim dres--sed well yes--ter--day
C
Entry saccade
Minisaccades (with bad estim.)
Jim dres--ses well yes--ter--day
Conflict detected
D
Exit saccade
Jim dres--sed well yes--ter--day

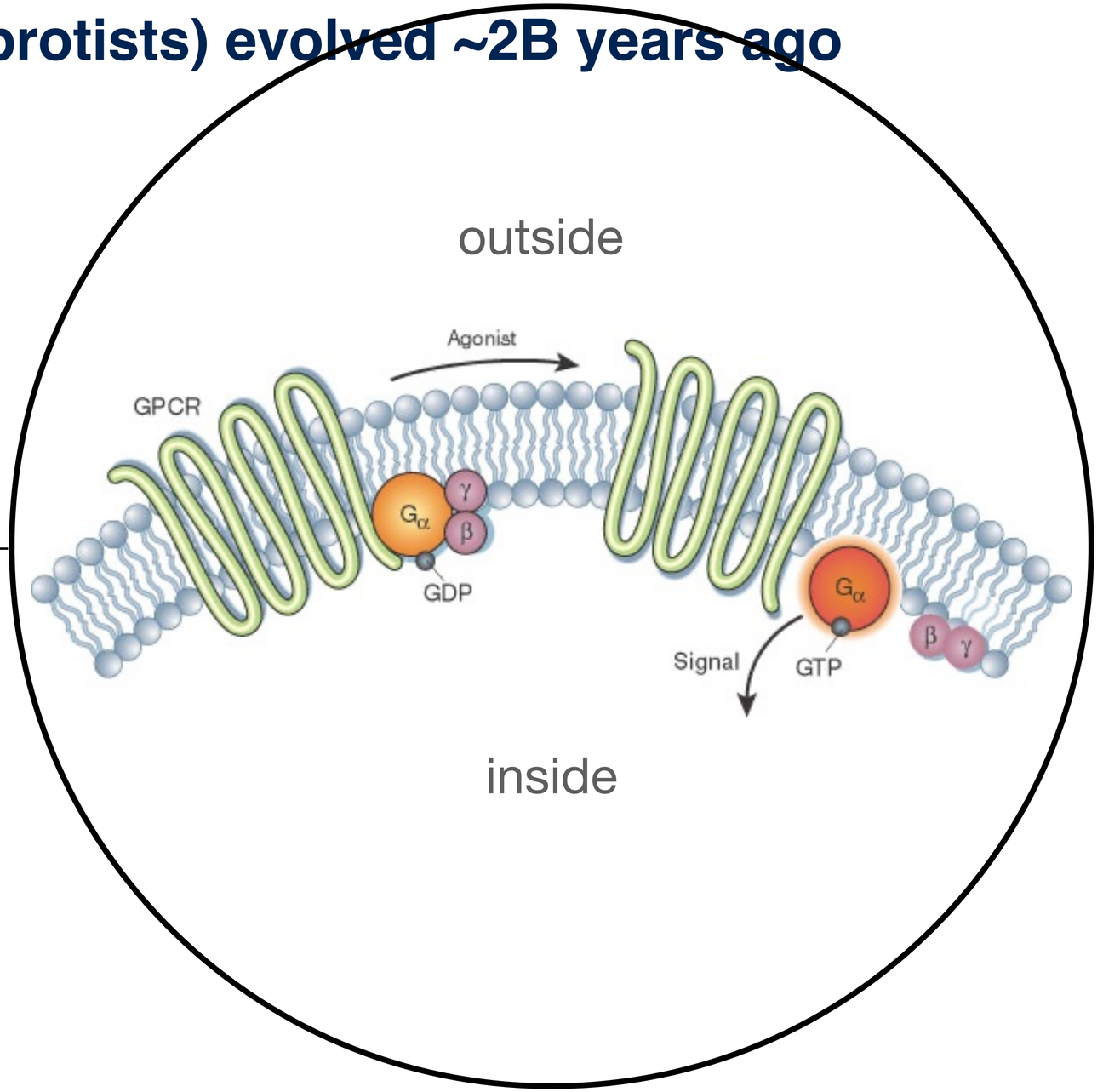
cyclotorsion





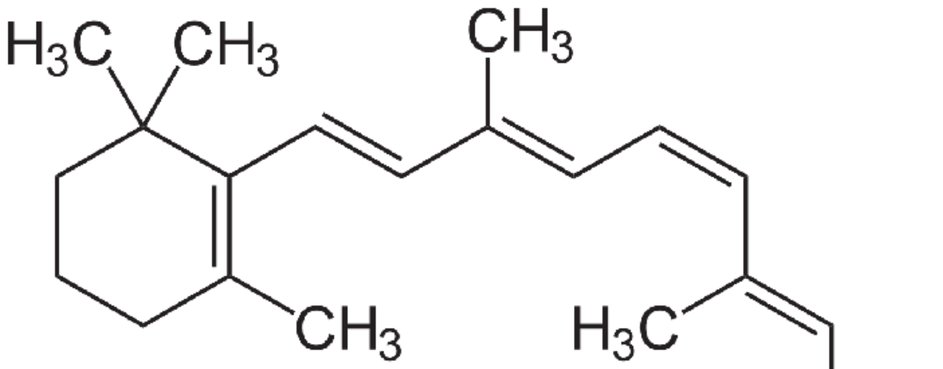
Eukaryotes (animals, plants, fungi, protists) evolved ~2B years ago





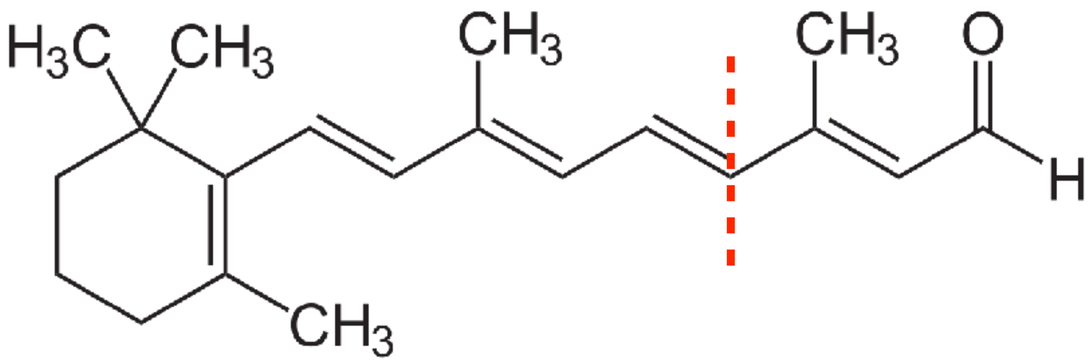
0.55B years ago a random mutation turned a chemical receptor into a light sensor. The first opsin protein.

cis-retinal

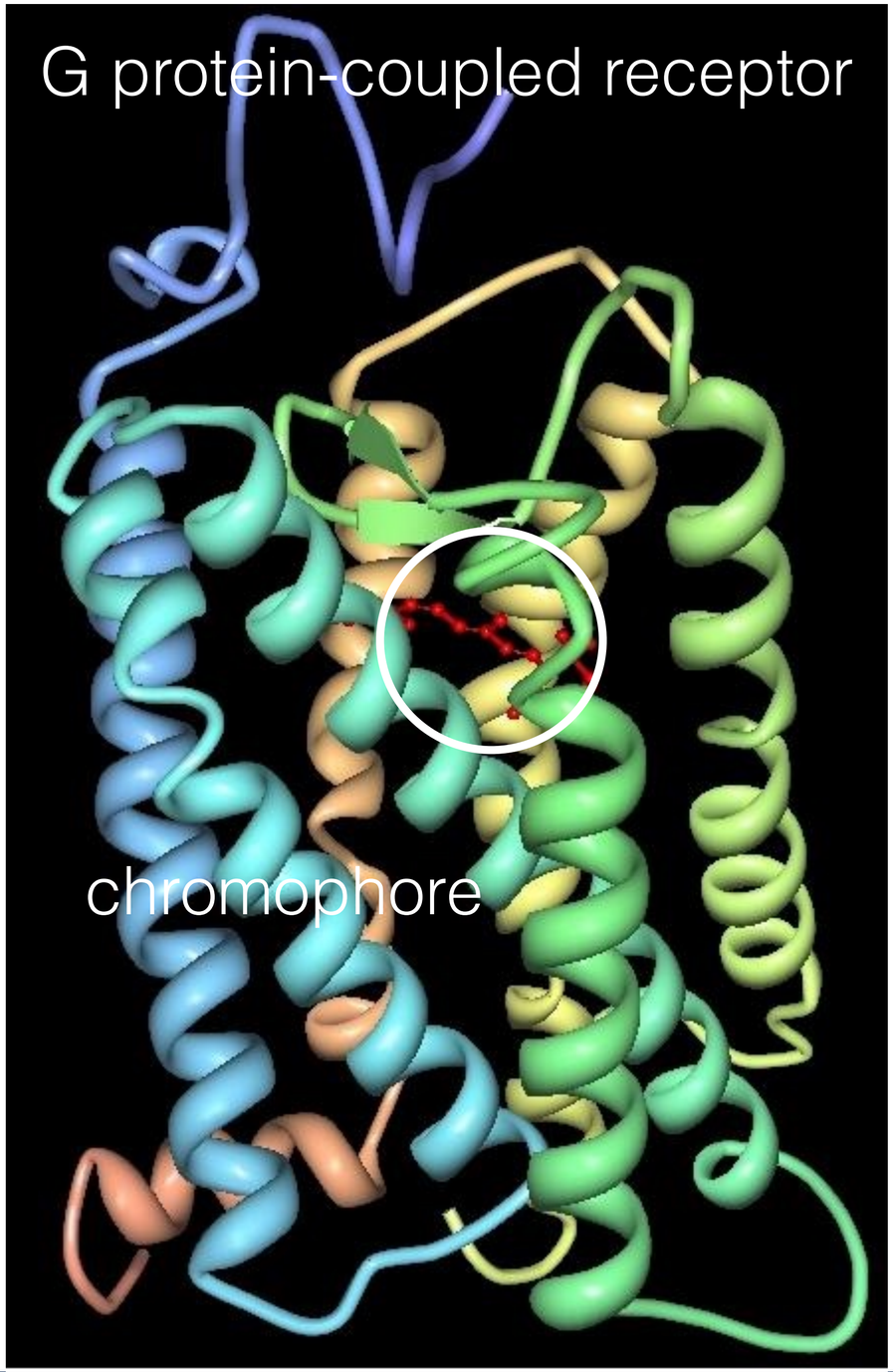


+ photon →

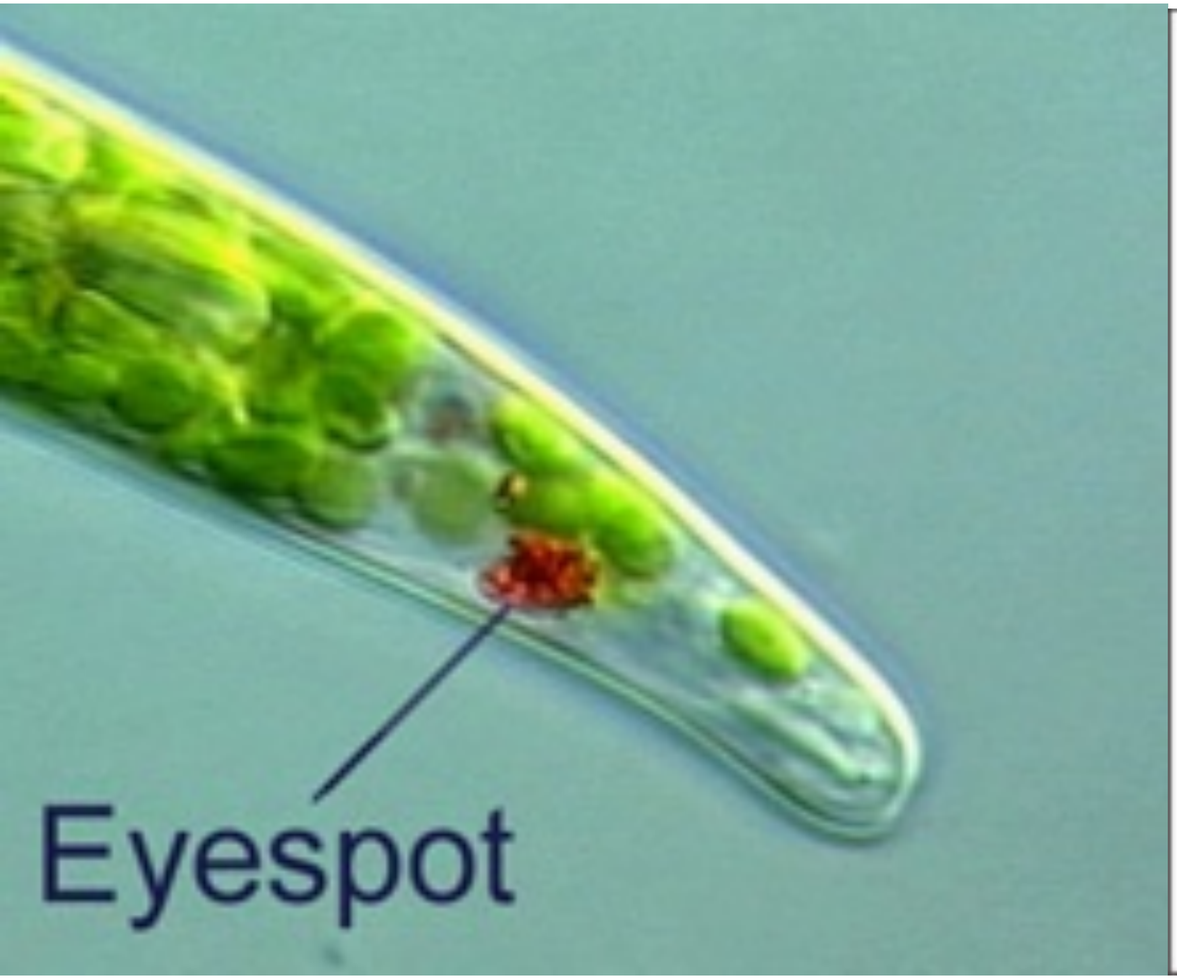
trans-retinal

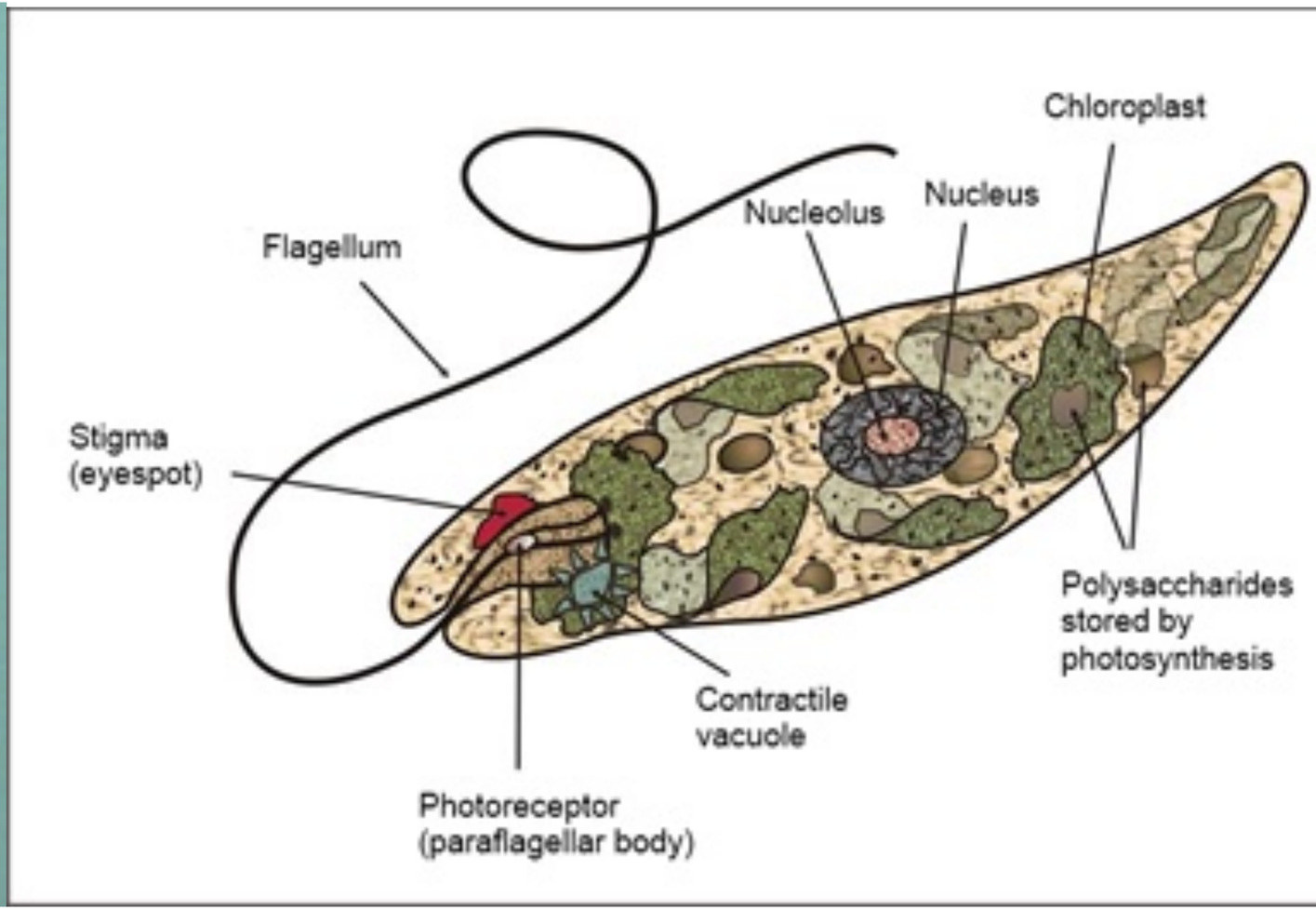


Discovered by George Wald in 1958 → Nobel prize for Chemistry 1967



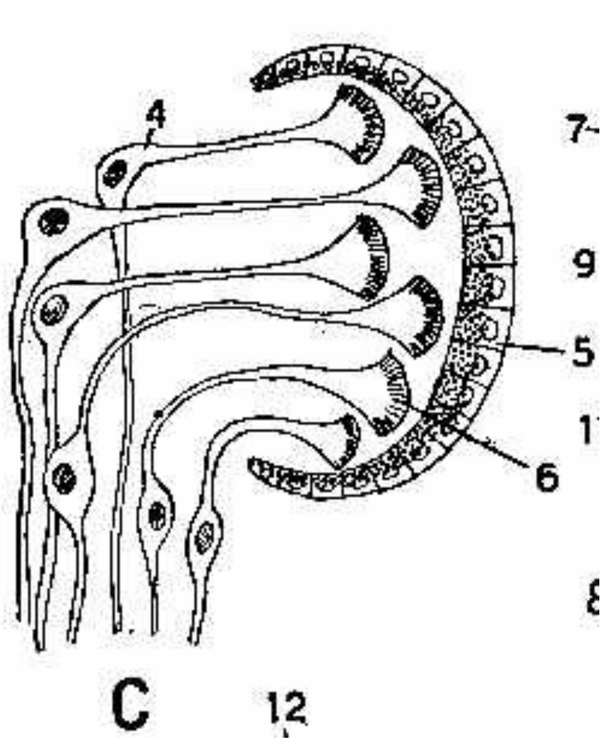
It was a really good idea





Evolution *invents* the eyespot independently 40-65 times

Multicellular organisms with eyes rapidly followed

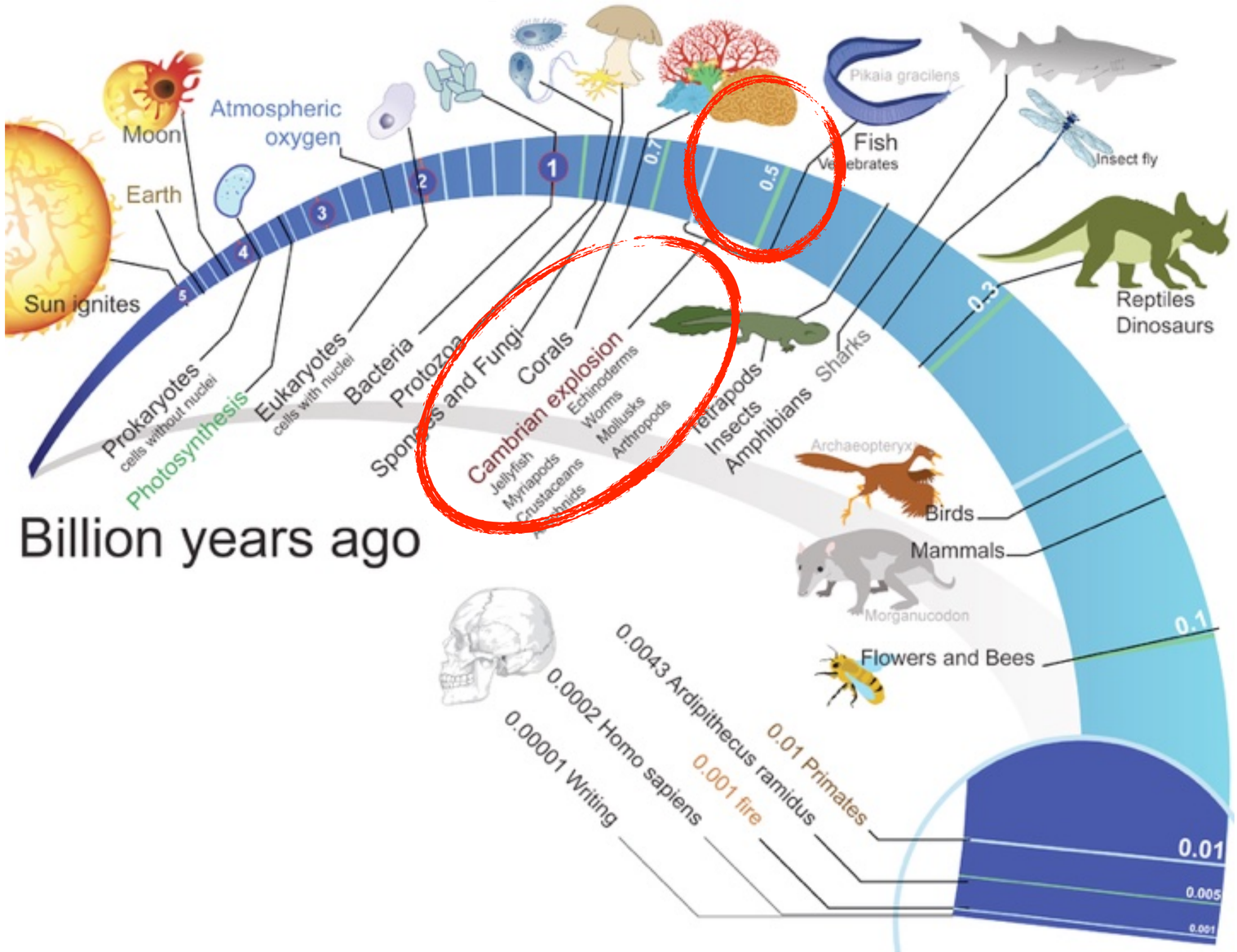


Planarium flat worm has simple eye pits which give some directional light sensing capability

Compound eyes developed early on

The trilobites were among the most successful of all early animals, appearing 521 million years ago and roamed the oceans for over 270 million years.

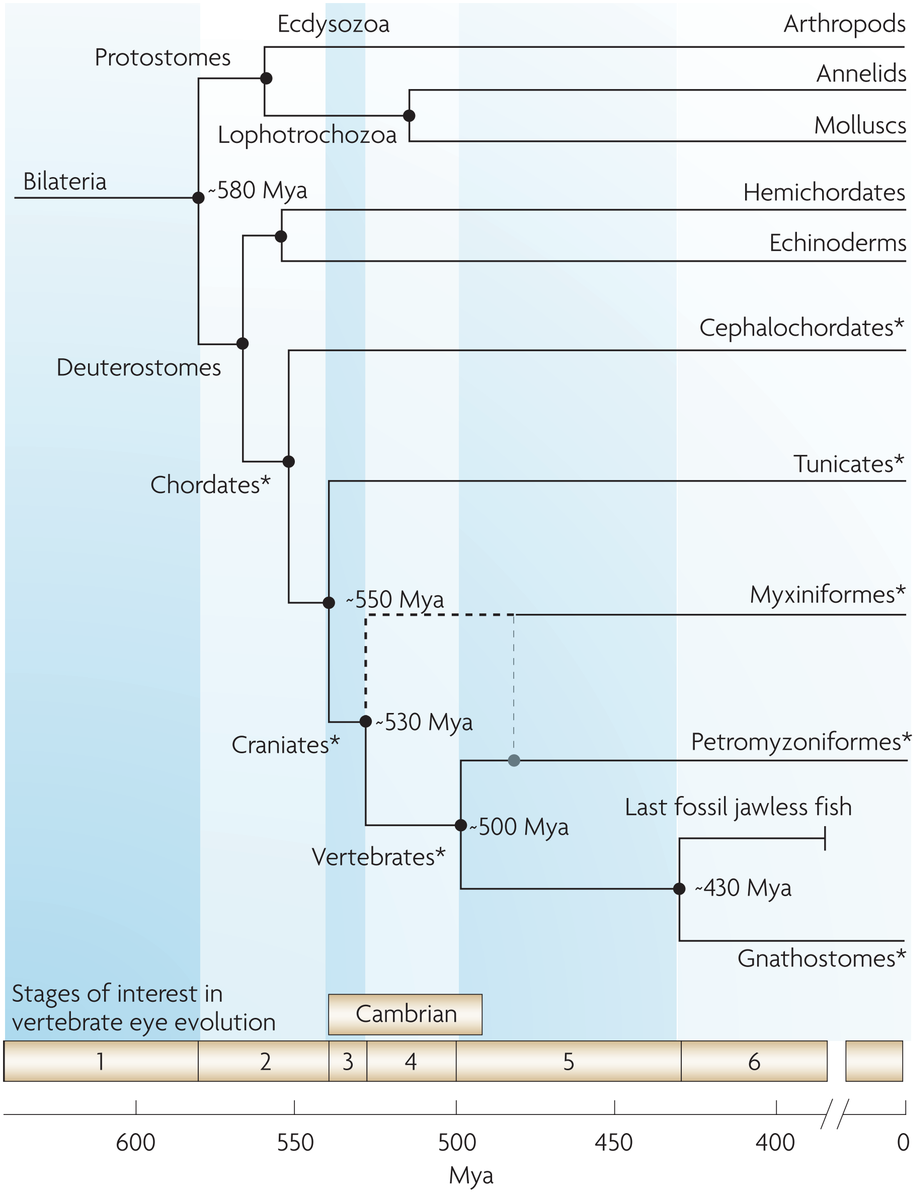
Which in turn led to an explosion of life forms on Earth. We call this the Cambrian explosion.

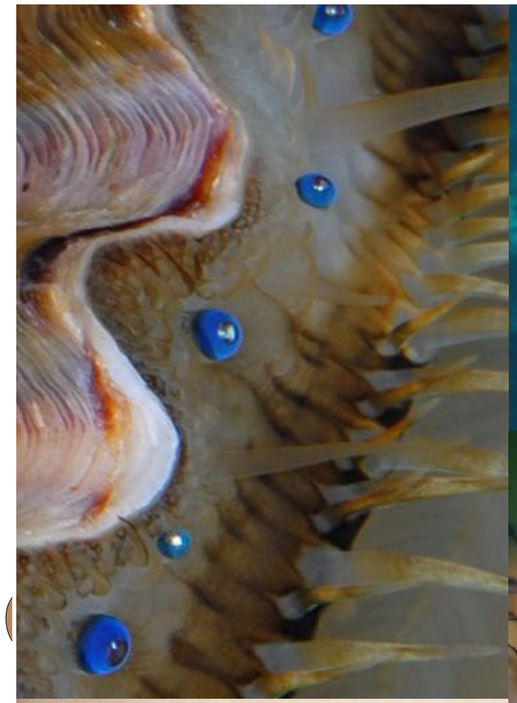


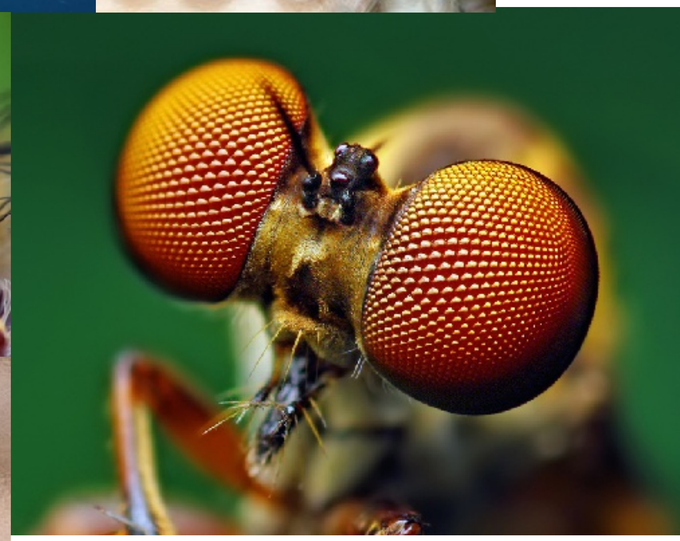
The lightswitch theory of evolution holds that the sense of vision led to an evolutionary arms race and a rich diversity of life on Earth.

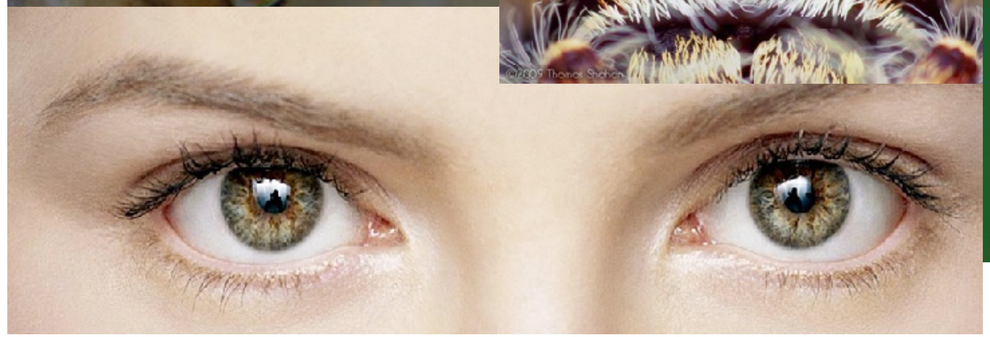
Parker

Today, eyes are everywhere and they are diverse









- 10 different eye designs
- 96% of animal species have eyes

As technologists we want to try and replicate this capability

It was once thought to be easy…

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966

Vision Memo. No. 100.

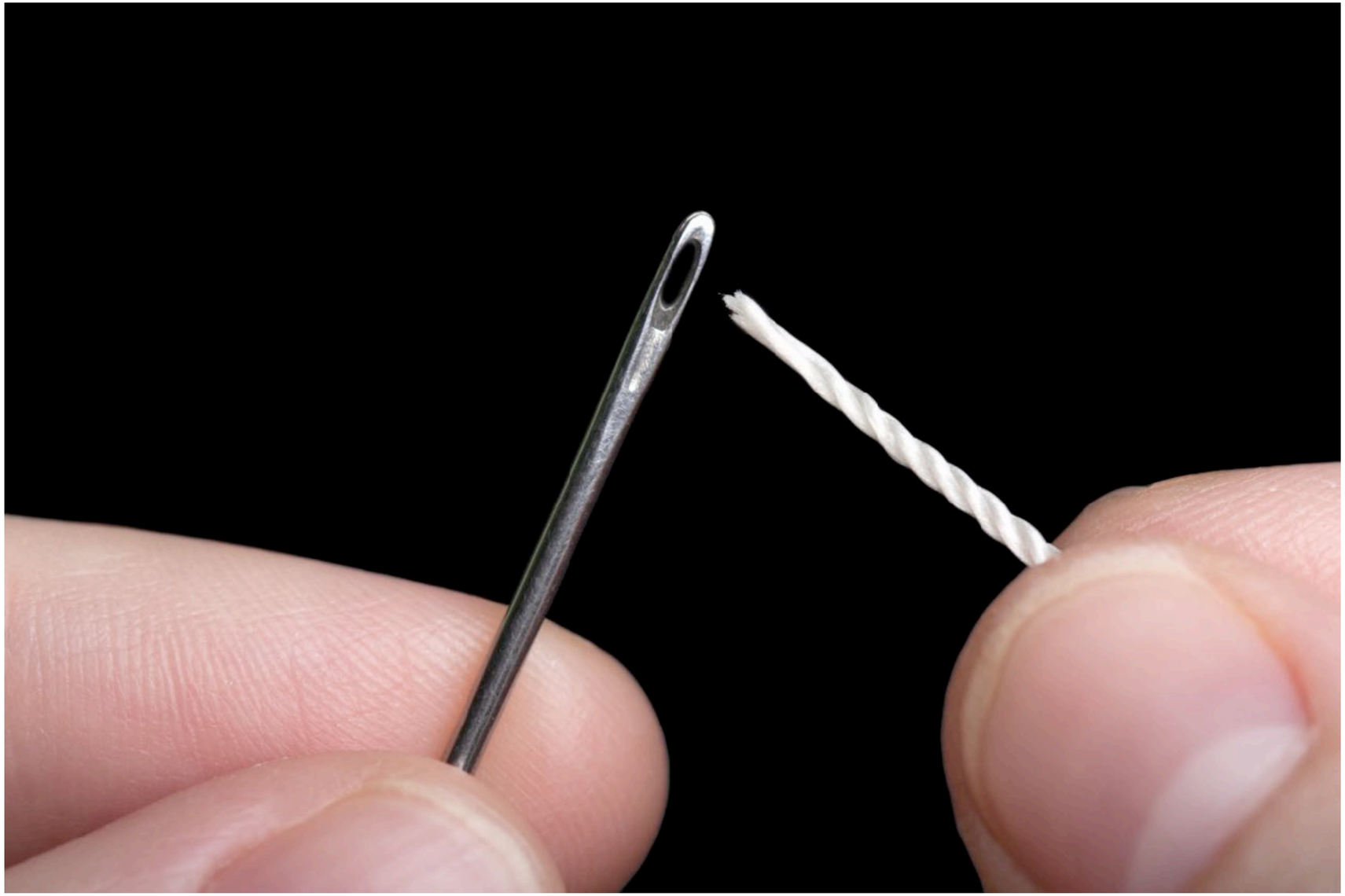
THE SUMMER VISION PROJECT

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".

1966

We can solve many 3D problems without knowing 3D pose



We’re here to learn about robotic vision

Vision is the process of discovering from images what is present in the world and where it is.

David Marr

Robotic vision is concerned with using information from cameras to understand and predict the world so that a robot can intelligently perform useful actions in the real world.