NATURAL AND ARTIFICIAL VISION MODULE

387AA – ROBOTICS [WIF-LM]

Marcello Calisti
marcello.calisti@santannapisa.it

The BioRobotics Institute
Scuola Superiore Sant'Anna

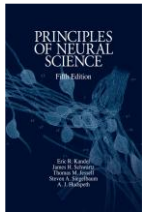
Content of the lectures

- Natural part – theorectical (3h)
- Artificial part – theorectical (3h)
- Hands-on in Matlab (6h)

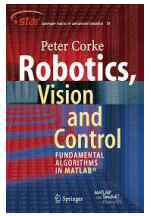
- VISUAL PATH
- RET → CGL → V1 → DORSAL / VENTR
- IMAGE EN
- 2D FILTER
- FEATURES
- RVC TOOL BOX
- IMAGE PROC TOOL BOX
- OPENCV

The material on natural vision can be found on:

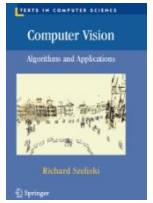
Kandel et al., “Principle of Neuroscience”, McGraw Hill, 2013



Most of the material presented in these lessons can be found on the brilliant, seminal books on robotics and image analysis reported hereafter:

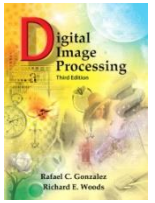


1. P.I. Corke, “Robotics, Vision & Control”, Springer 2011, ISBN 978-3-642-20143-1



2. R. Szeliski, “Computer Vision: Algorithms and Applications”, Springer-Verlag New York, 2010

PREAPRINT

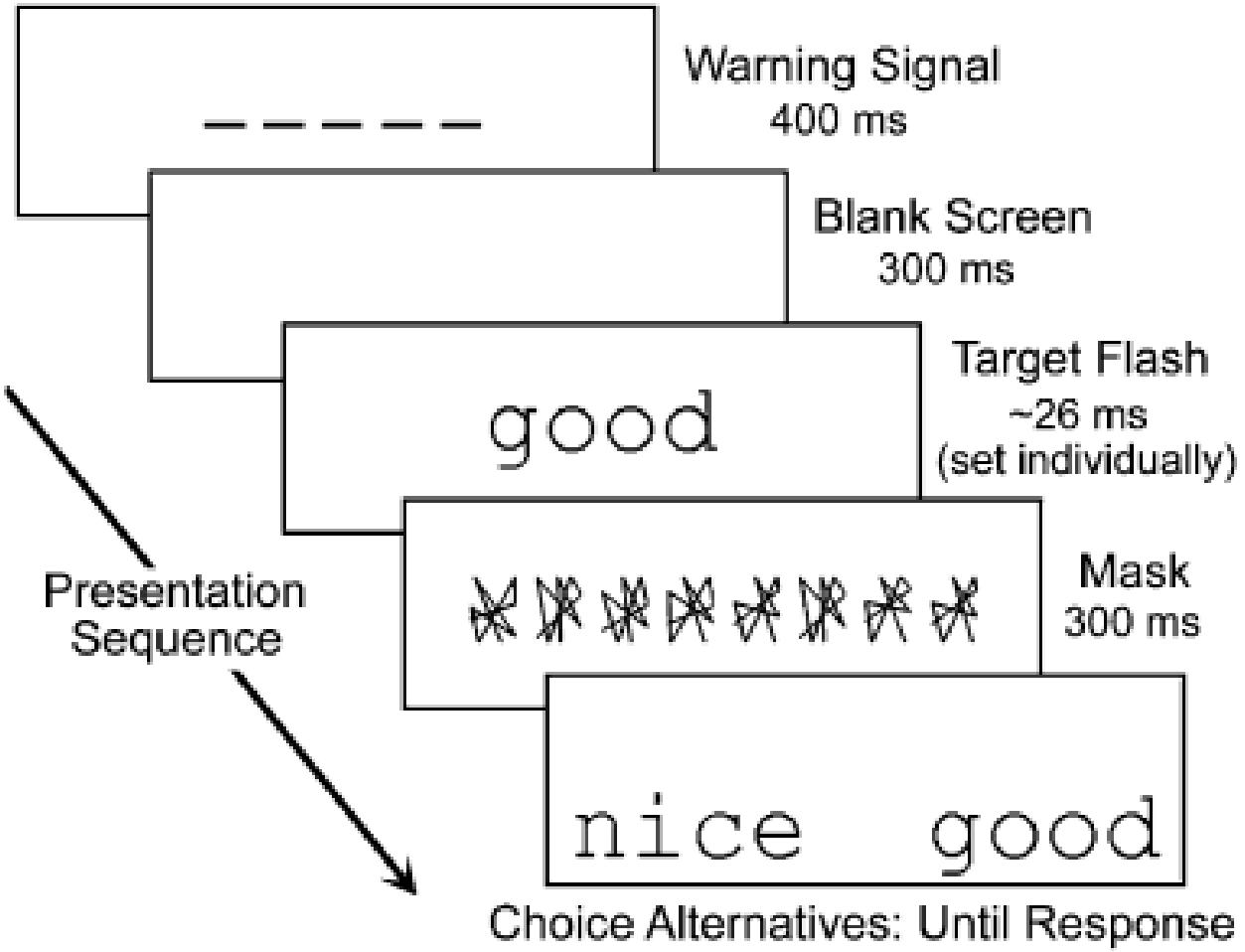


3. R.C. Gonzalez & R.E. Woods, “Digital Image Processing (3rd edition)”, Prentice-Hall, 2006

Most of the images of these lessons are downloaded from RVC website http://www.petercorke.com/RVC/index.php and, despite they are free to use, they belong to the author of the book.

← TAG ?

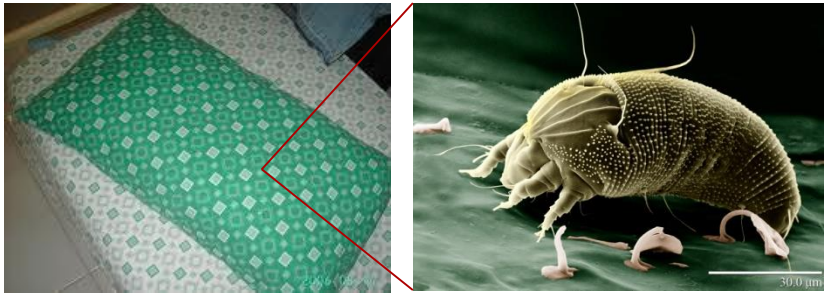
Human vision is a **creative process**, far beyond simple trasduction



Psychol Sci. 2006 Apr;17(4):287-91. **The impact of emotion on perception: bias or enhanced processing?** Zeelenberg R[1], Wagenmakers EJ, Rotteveel M.

Psychol Sci. 2003 Jan;14(1):7-13. **Emotional facilitation of sensory processing in the visual cortex.** Schupp HT[1], Junghöfer M, Weike AI, Hamm AO.

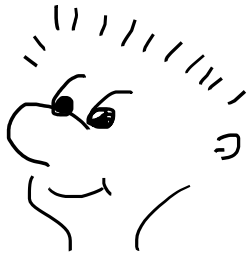
what we can



RESOLUTION

SPAT / TEMP

What we learnt



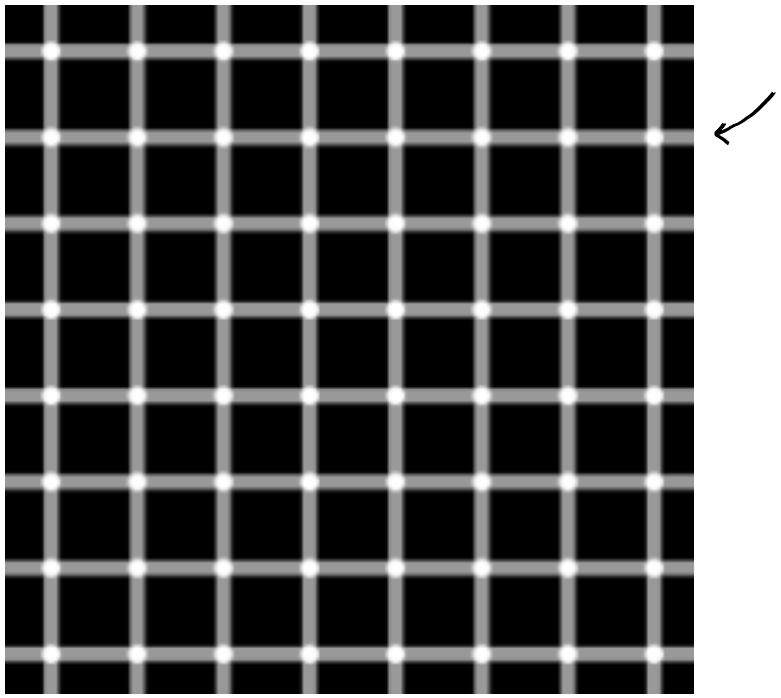
What we want

Light is «different» depending on the animal

we

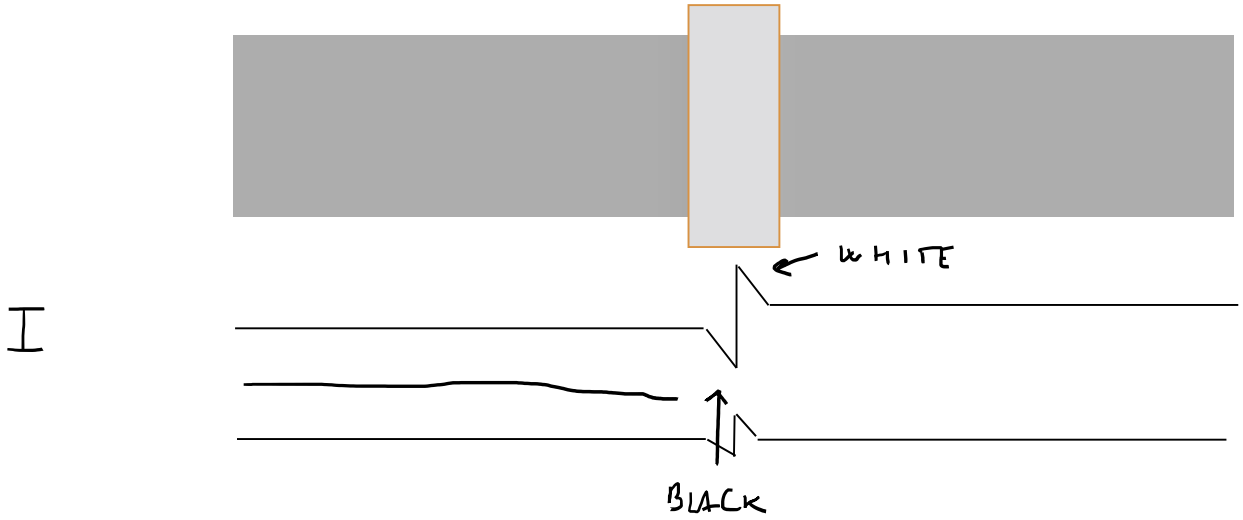
insect

Ludimar Hermann

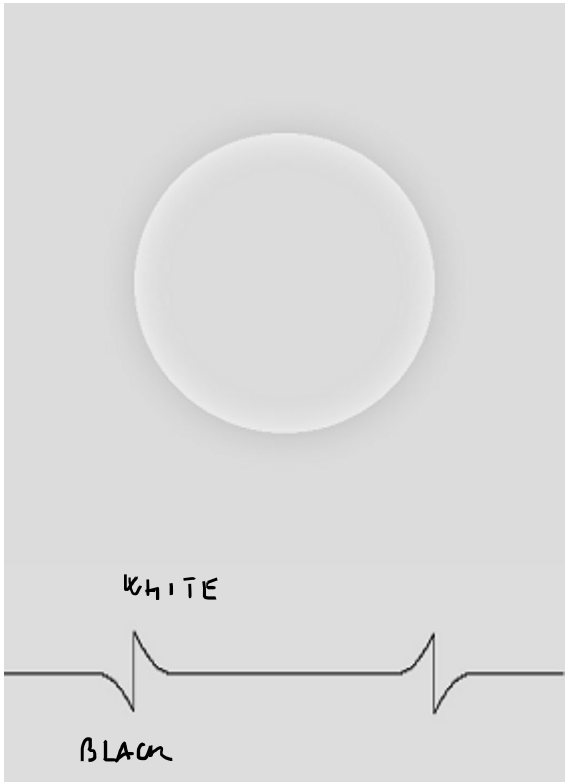


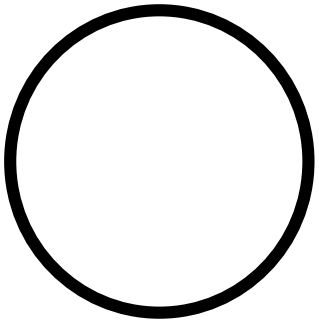
Shiny Herman grid

We saw contrast, not absolute value of illumination

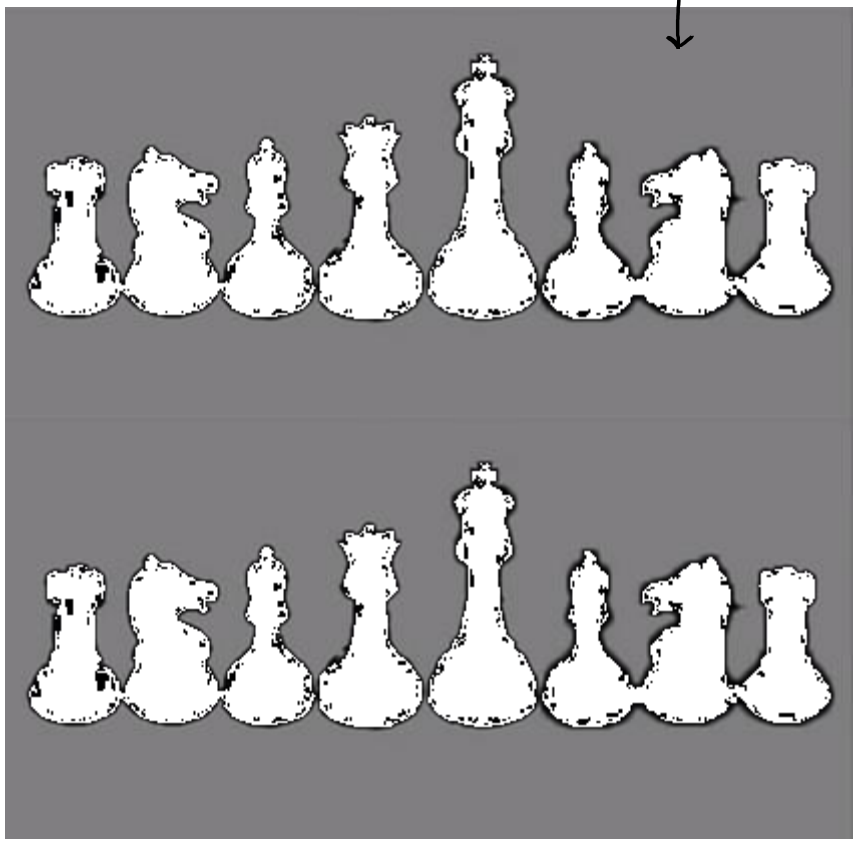


The Craik-Cornsweet illusion

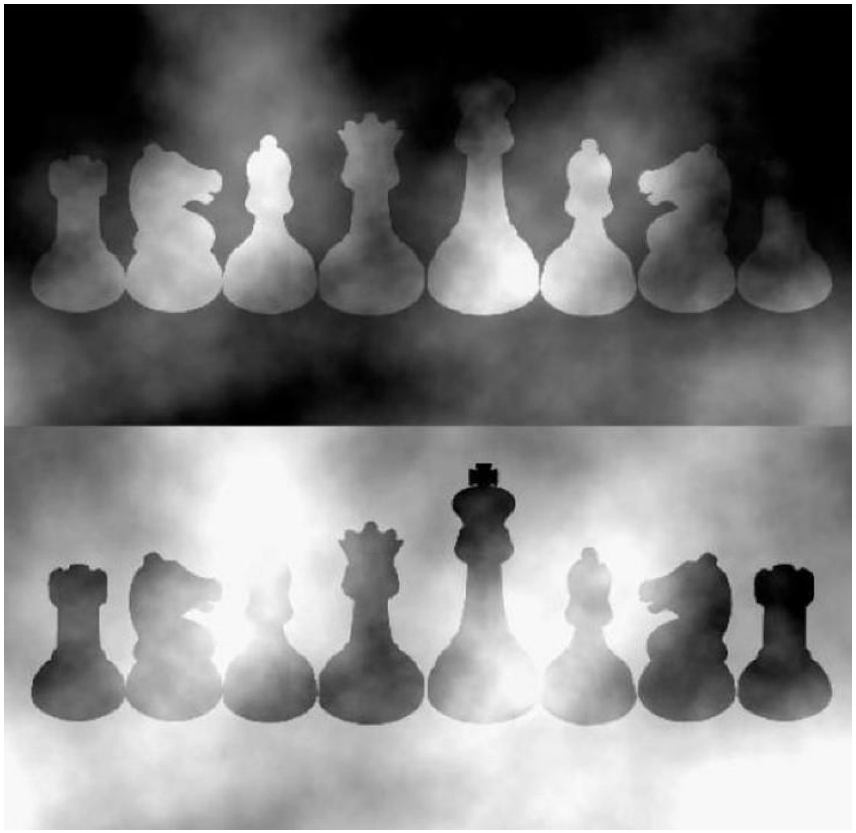




mask

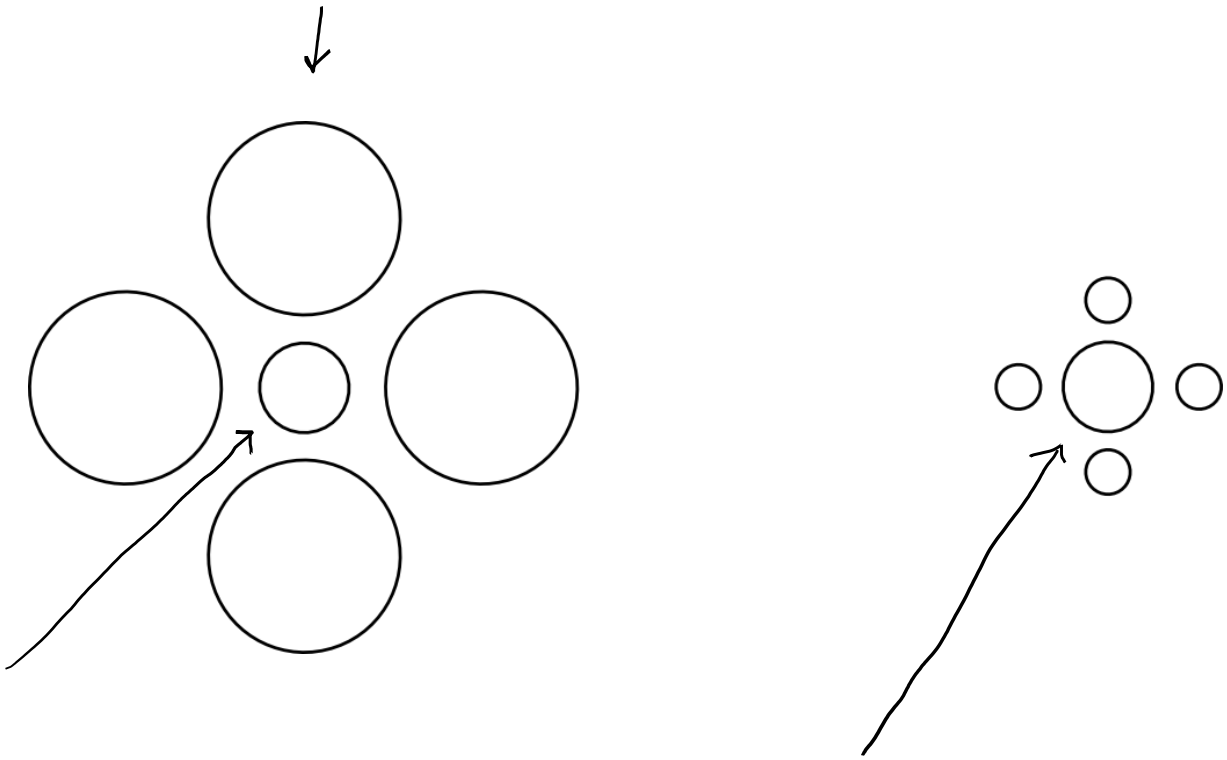


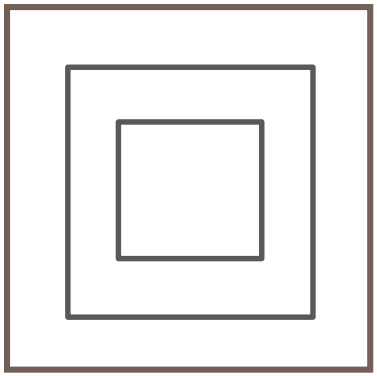
white

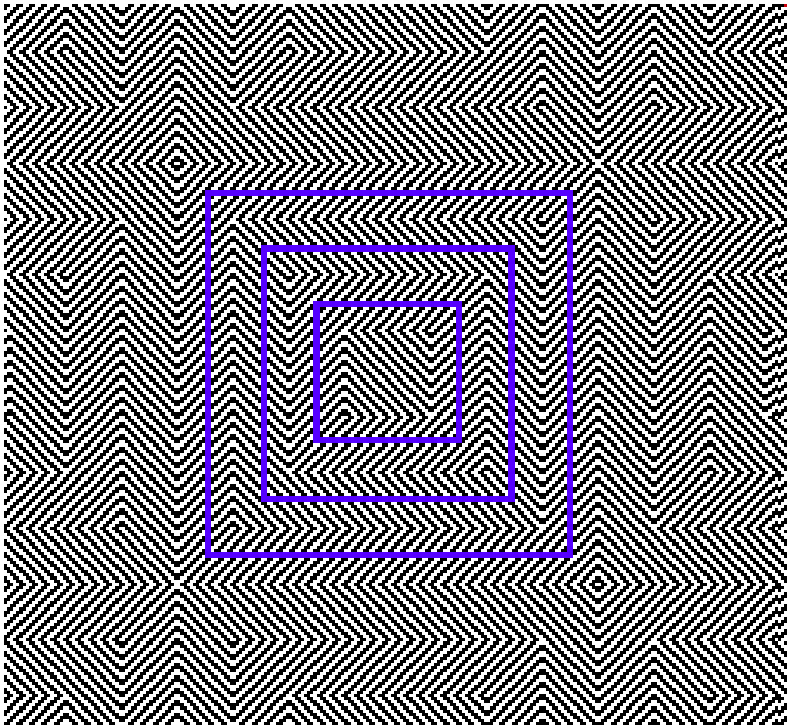


black

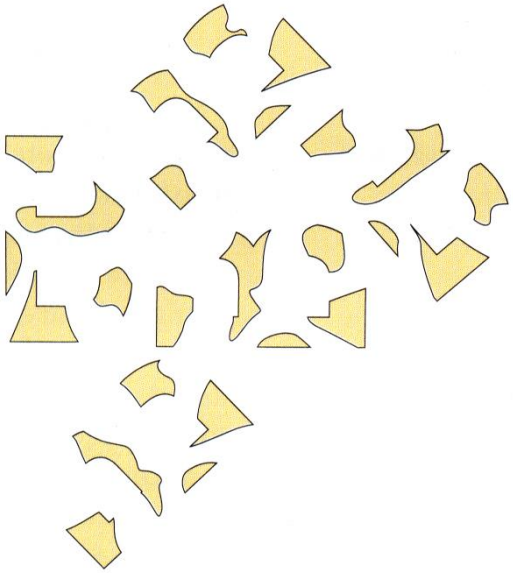
Context effects

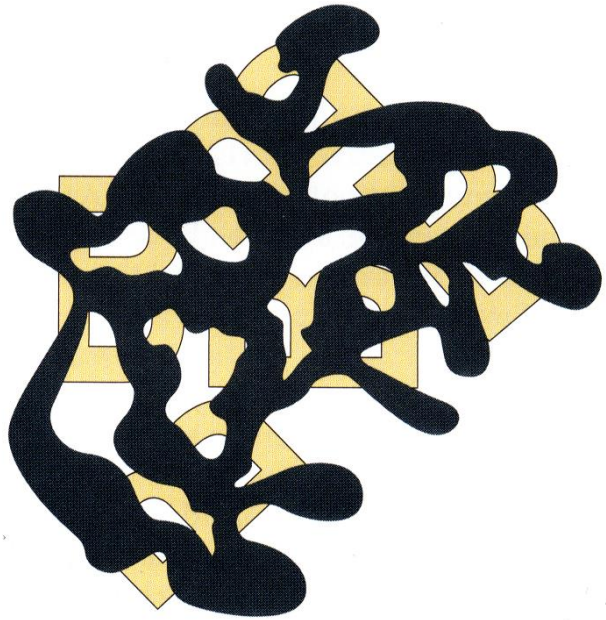






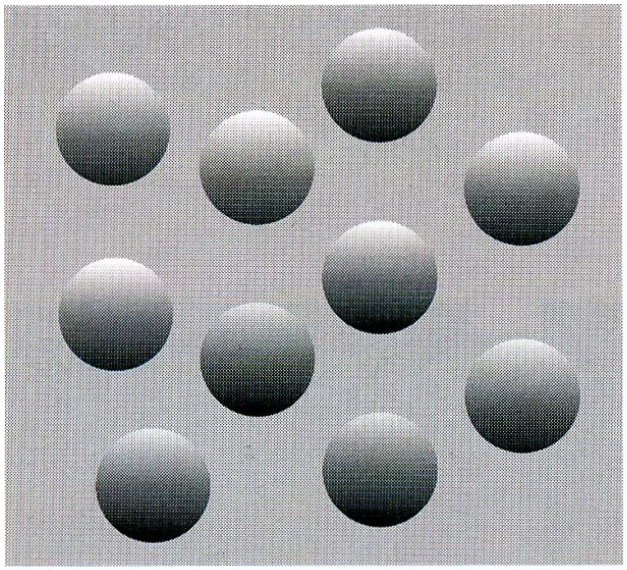
We perceive what we know



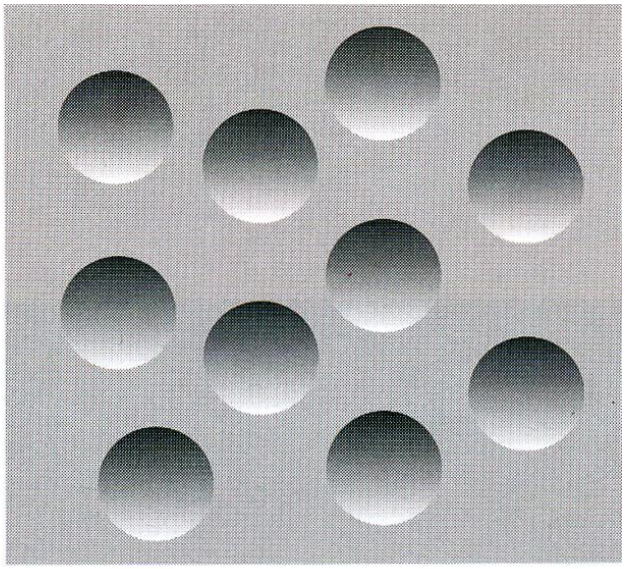


We «abstract» properties: tridimensional shapes

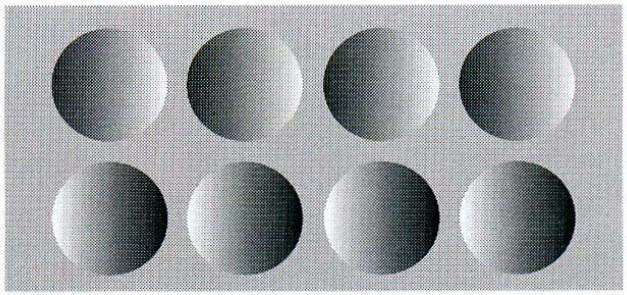
Convex



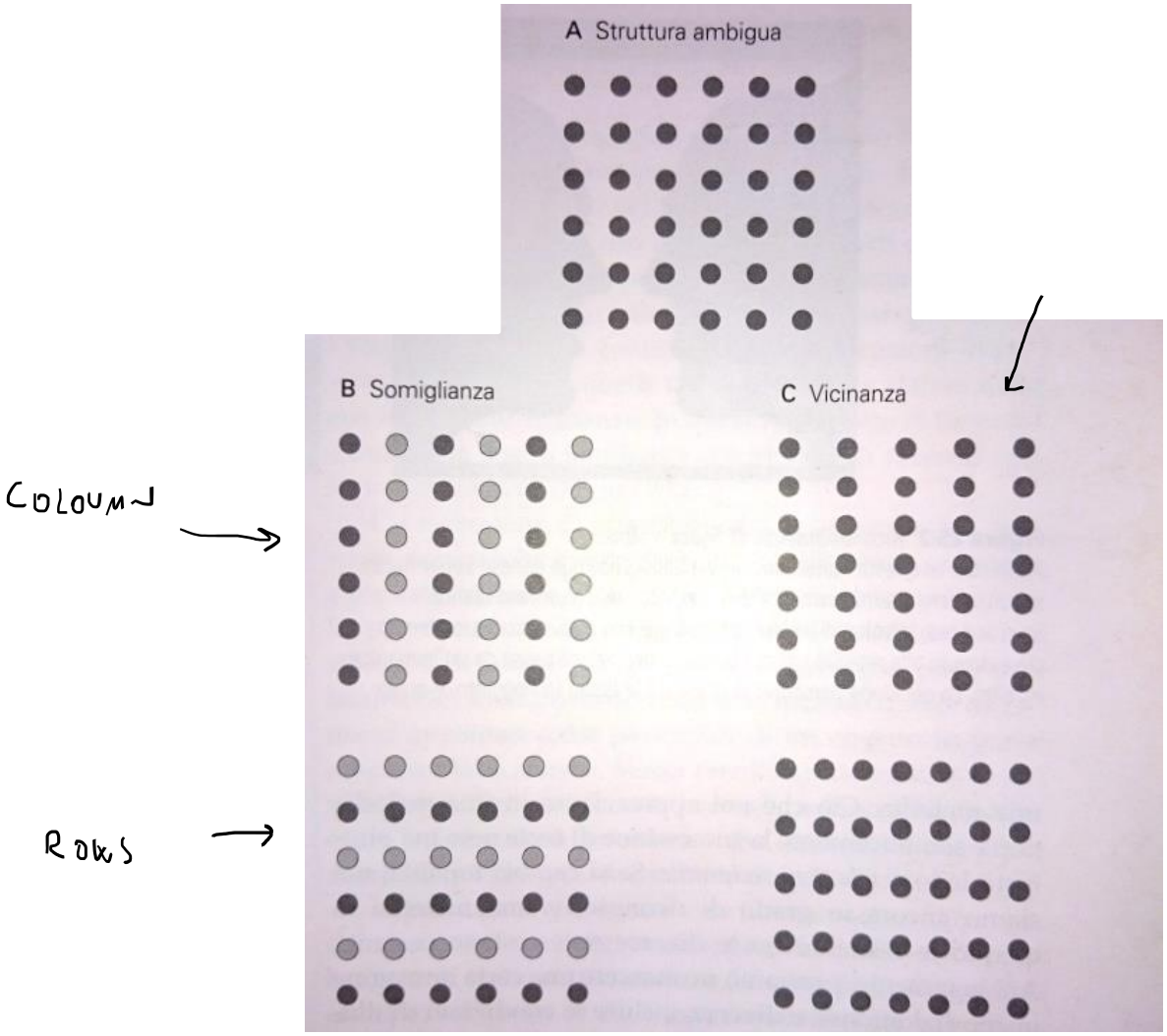
Concave

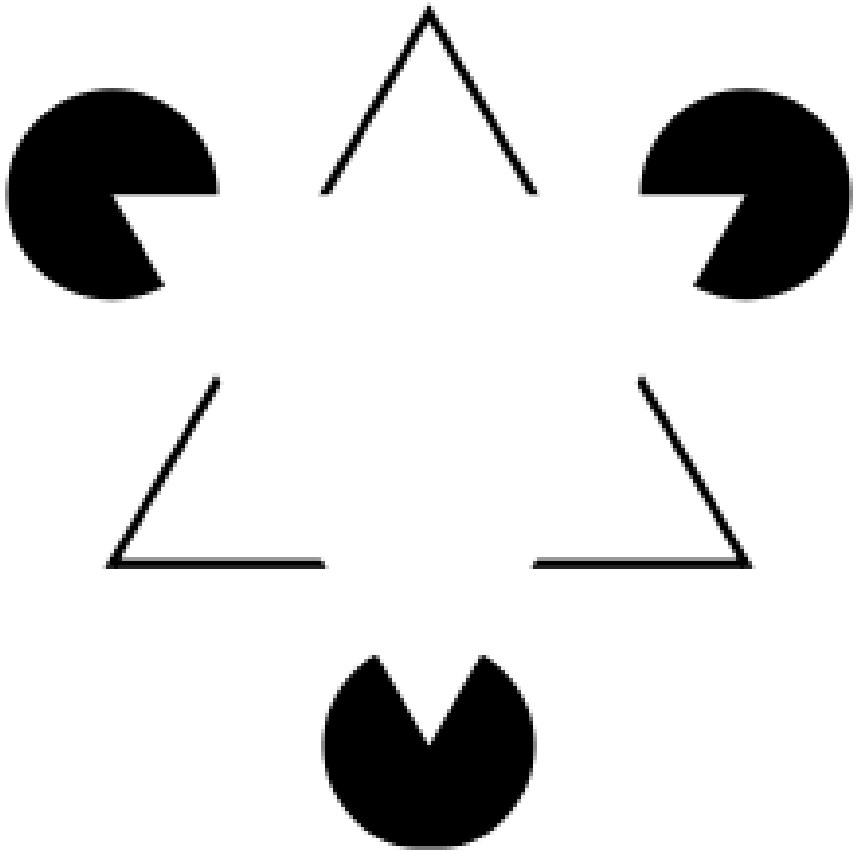


Not defined



Structure by association

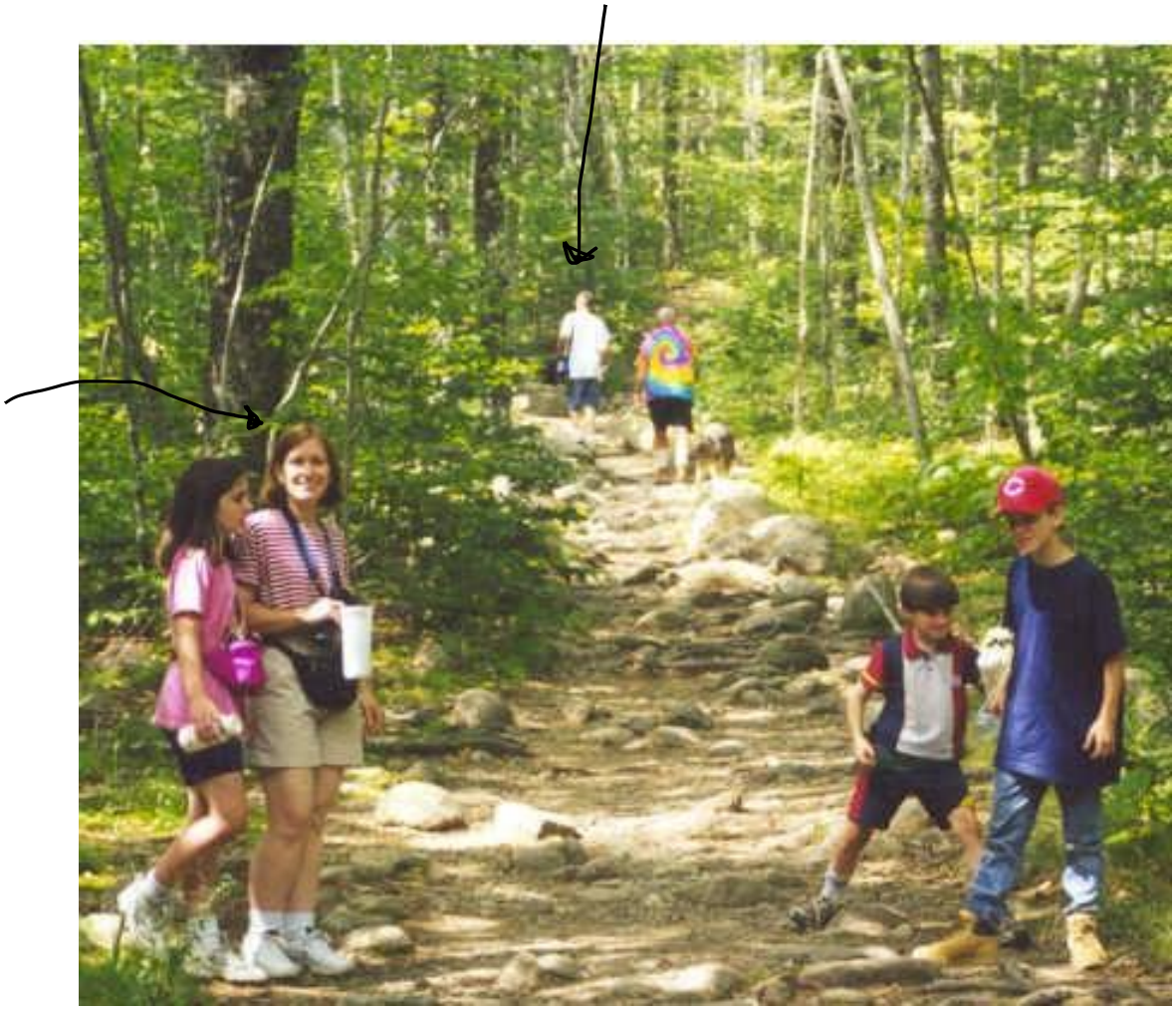


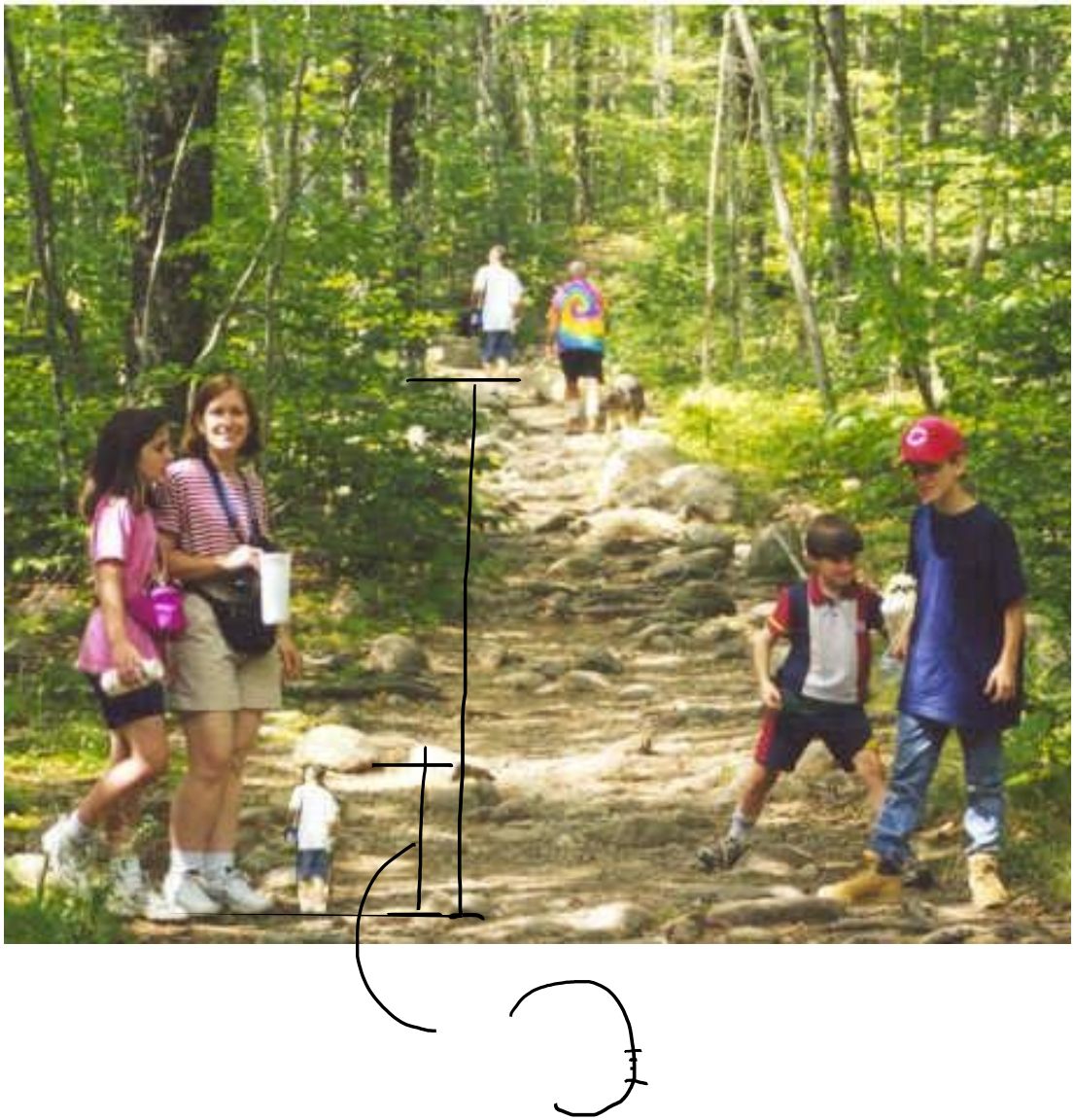


Kanizsa triangle

Illusory edges

Same dimensions, different context



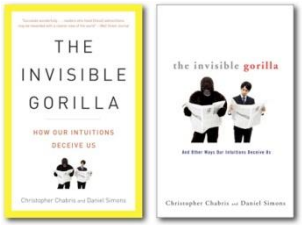


Attention perception

- **Count** ball passages from the **player with white t-shirt**

Video Simons

https://www.youtube.com/watch?v=IGQmdoK_ZfY

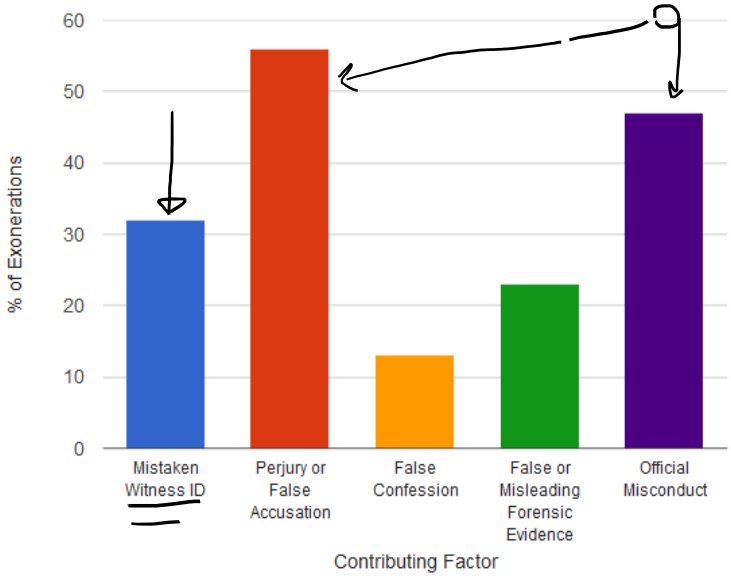


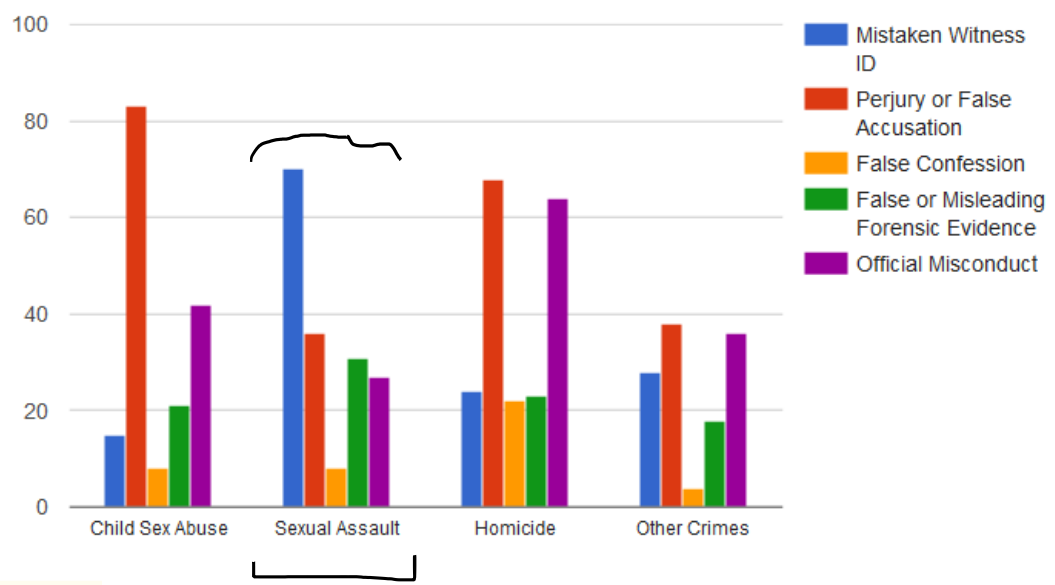
Christopher Chabris and Daniel Simons

Attention perception

Video Simons 2

https://www.youtube.com/watch?v=FWSxSQsspiQ



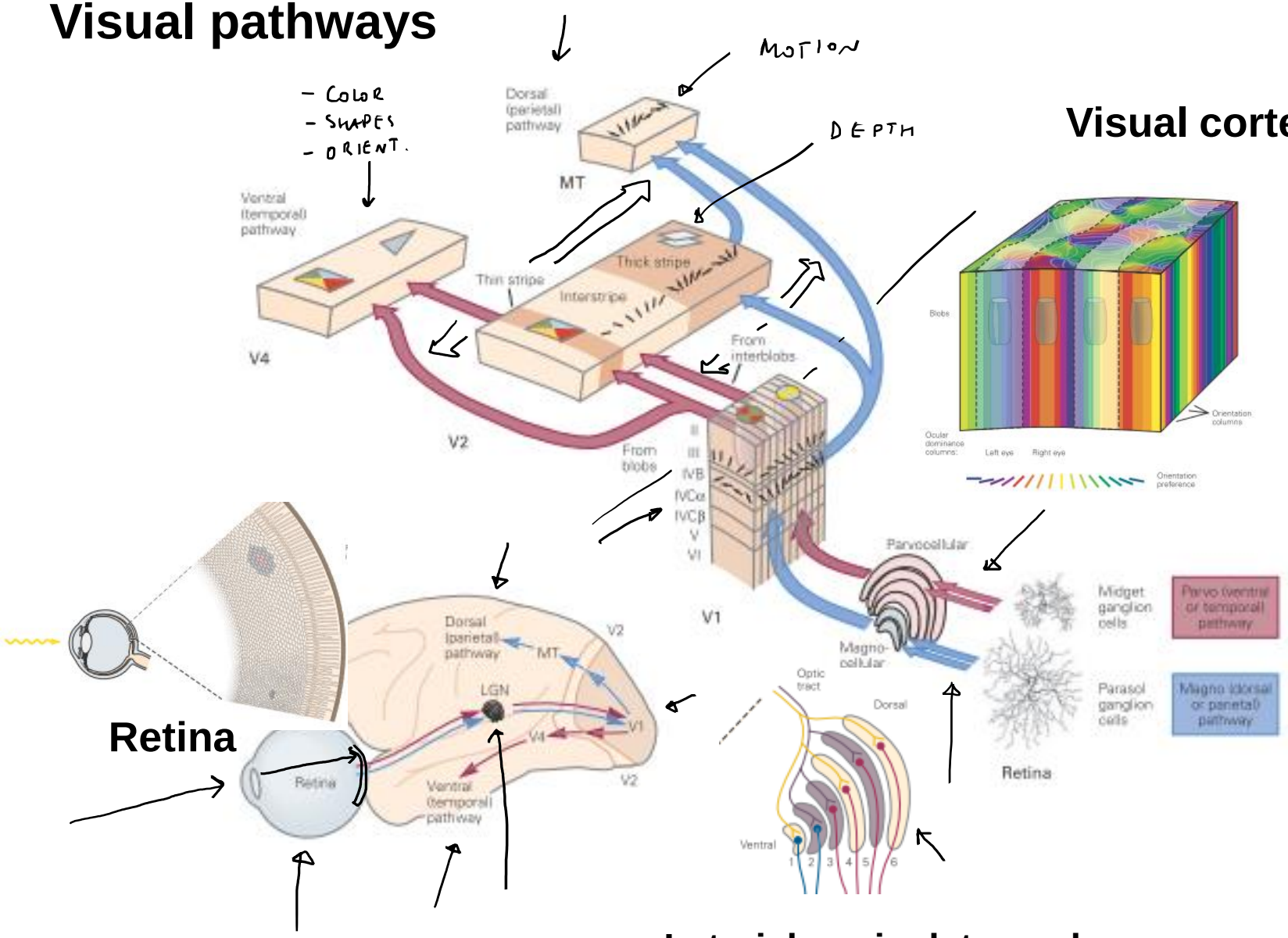


*In general, an **exoneration** occurs when a person who has been convicted of a crime is officially cleared based on new evidence of innocence.*

A more precise definition follows.
Exoneration—A person has been exonerated if he or she was convicted of a crime and later was either: (1) declared to be factually innocent by a government official or agency with the authority to make that declaration; or (2) relieved of all the consequences of the criminal conviction by a government official or body with the authority to take that action. The official action may be: (i) a complete pardon by a governor or other competent authority, whether or not the pardon is designated as based on innocence; (ii) an acquittal of all charges factually related to the crime for which the person was originally convicted; or (iii) a dismissal of all charges related to the crime for which the person was originally convicted, by a court or by a prosecutor with the authority to enter that dismissal. The pardon, acquittal, or dismissal must have been the result, at least in part, of evidence of innocence that either (i) was not presented at the trial at which the person was convicted; or (ii) if the person pled guilty, was not known to the defendant, the defense attorney and the court at the time the plea was entered. The evidence of innocence need not be an explicit basis for the official action that exonerated the person.
Exoneree—A person who was convicted of a crime and later officially declared innocent of that crime, or relieved of all legal consequences of the conviction because evidence of innocence that was not presented at trial required reconsideration of the case.

The National Registry of Exonarations, https://www.law.umich.edu/special/exoneration/Pages/about.aspx

Visual pathways

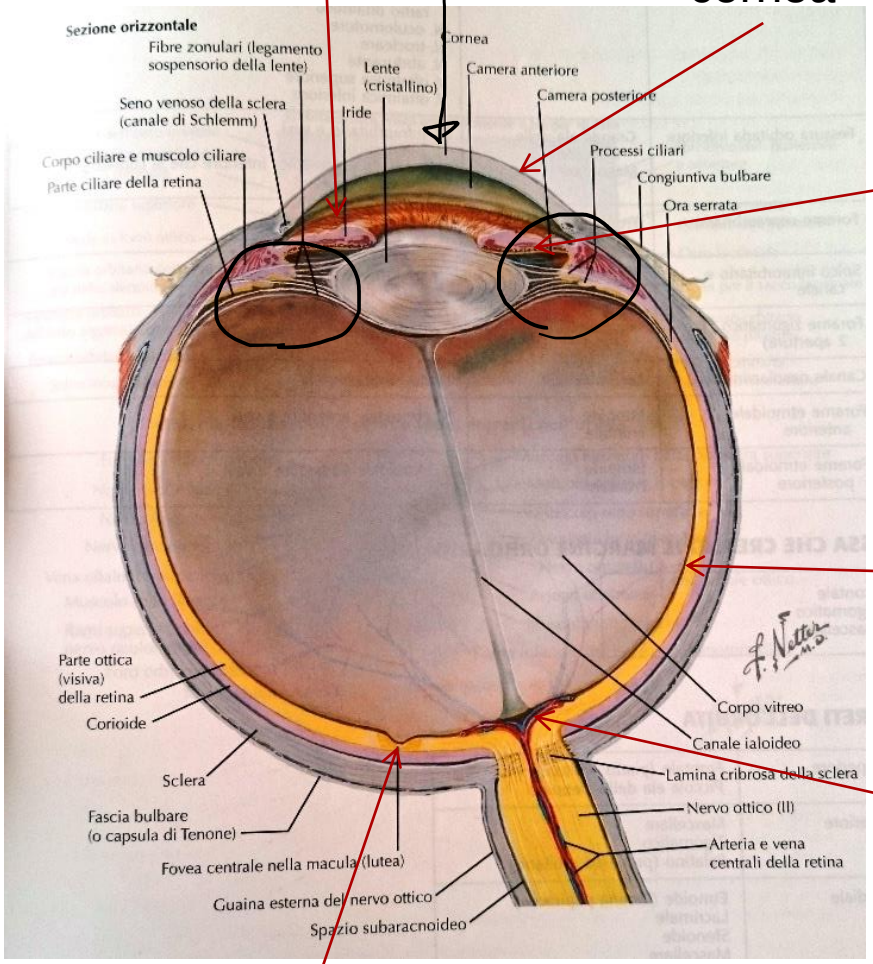


Visual cortex V1

Retina

Laterial geniculate nucleus

Anatomy of the eye



iris

LIGHT

cornea

Focalization function

- CORNEA
- IRIS → PUPIL
- LENS

Lens: transparent biconvex

retina

Transduction fuction

Optic disc

Fovea

Some images from:
"Anatomia della testa e del collo di Netter", Neil S. Norton

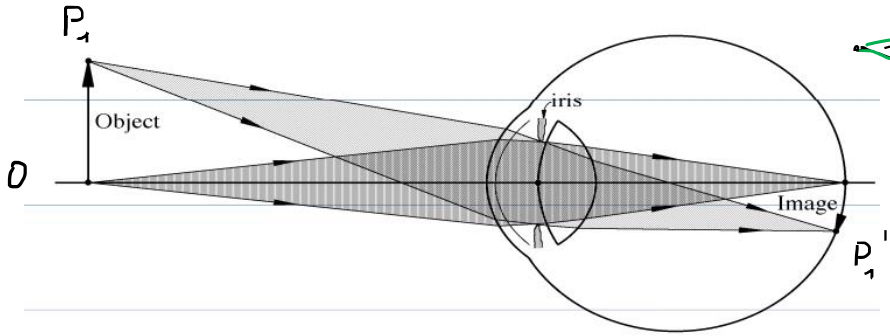
The **focalization function**

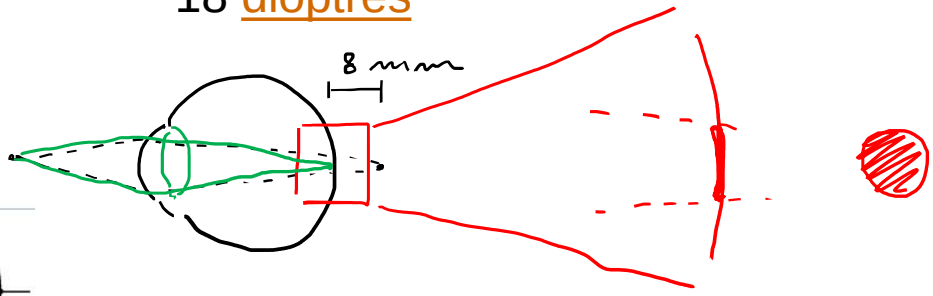
→ **Cornea + lens cortex**

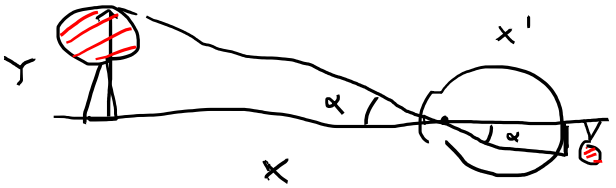
→ 43 dioptres

8 mm behind the retina

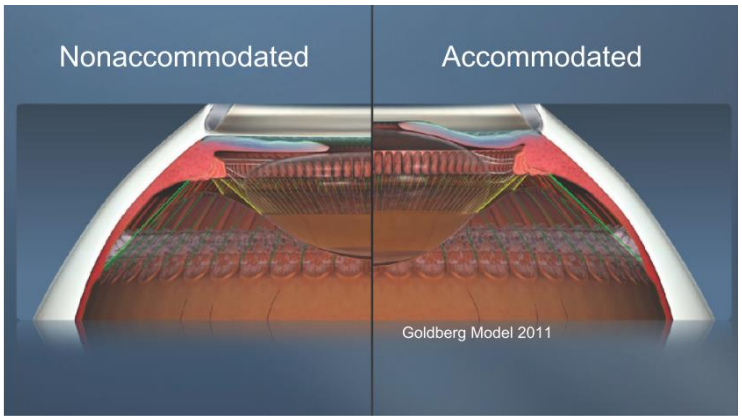
18 dioptres







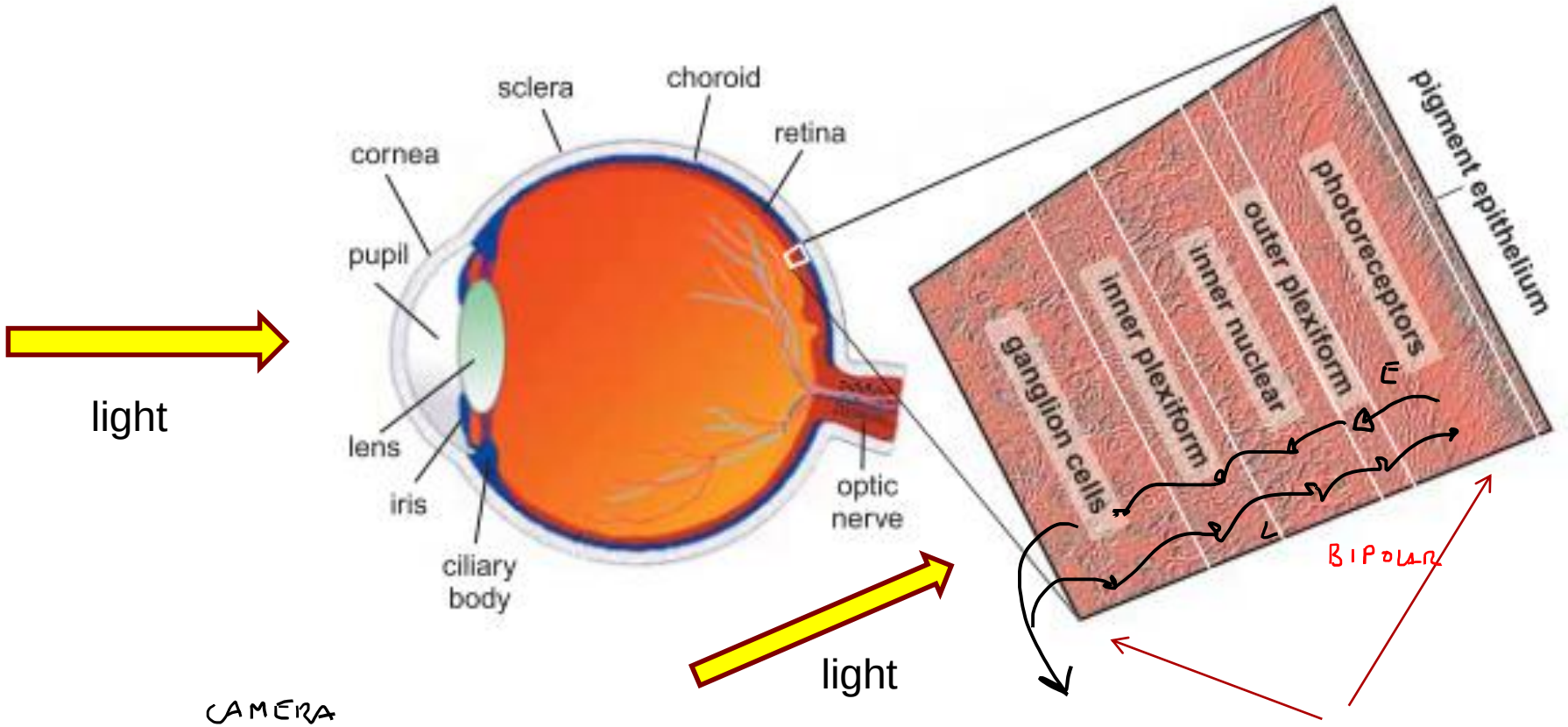
$$\operatorname{tg}\alpha = \frac{y}{x} = \frac{y'}{x'} \rightarrow y' = \frac{x'}{x}\,y$$

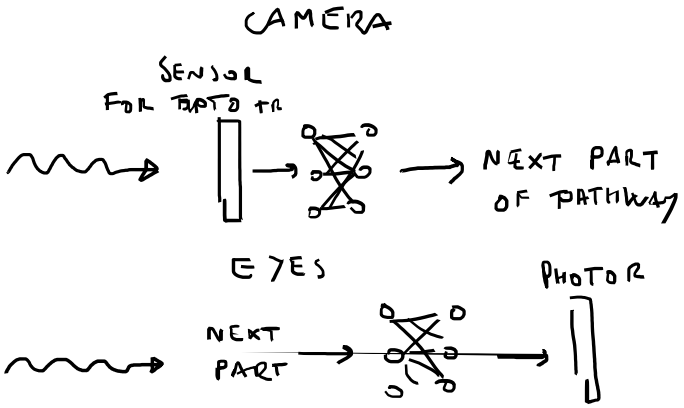


Clin Ophthalmol. 2011;5:1559-66. doi: 10.2147/OPTH.S25983. Epub 2011 Oct 28. **Computer-animated model of accommodation and theory of reciprocal zonular action.** Goldberg DB[1].

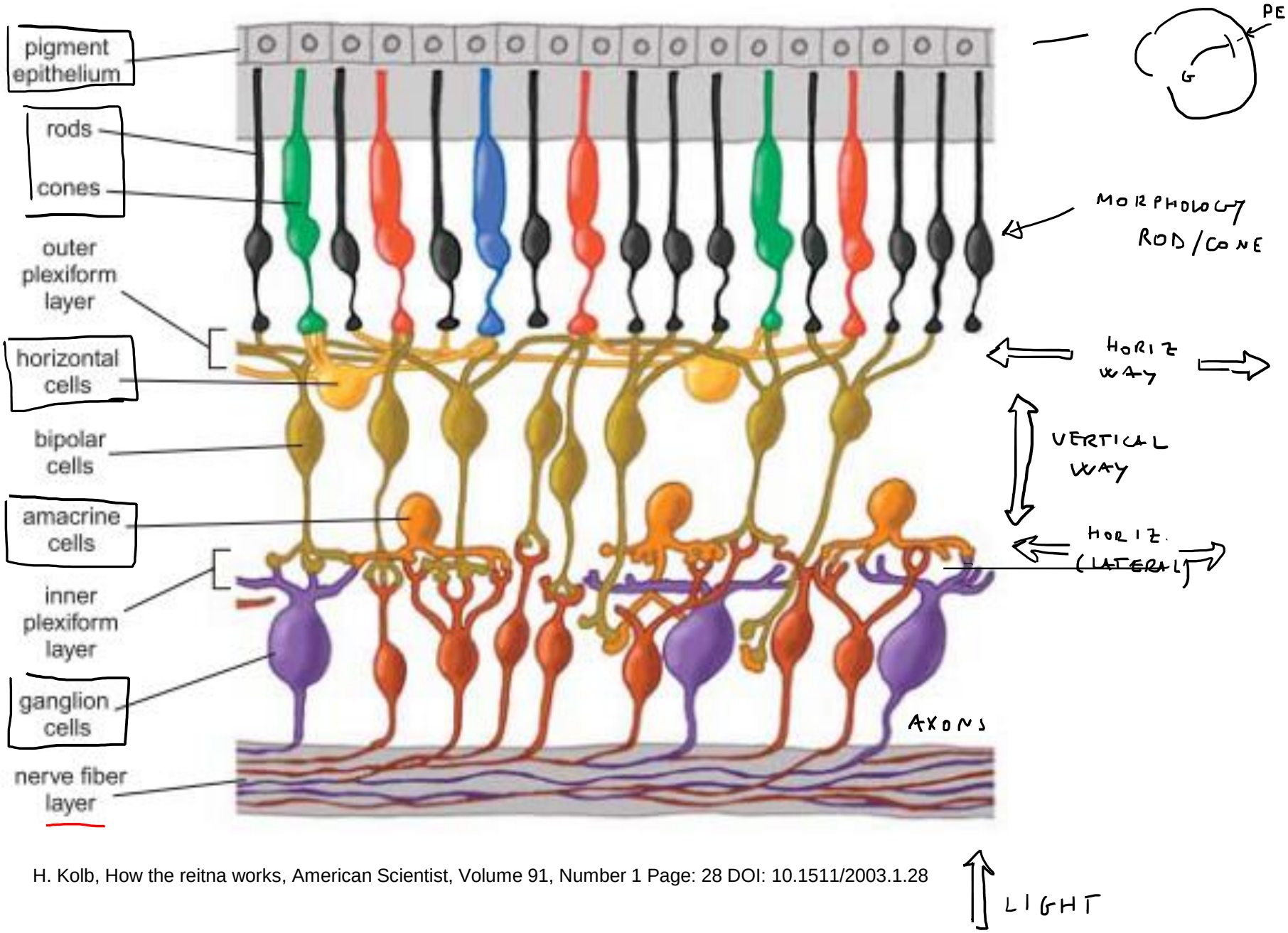
The **transduction function**

H. Kolb, How the retina works, American Scientist,
Volume 91, Number 1 Page: 28 DOI: 10.1511/2003.1.28

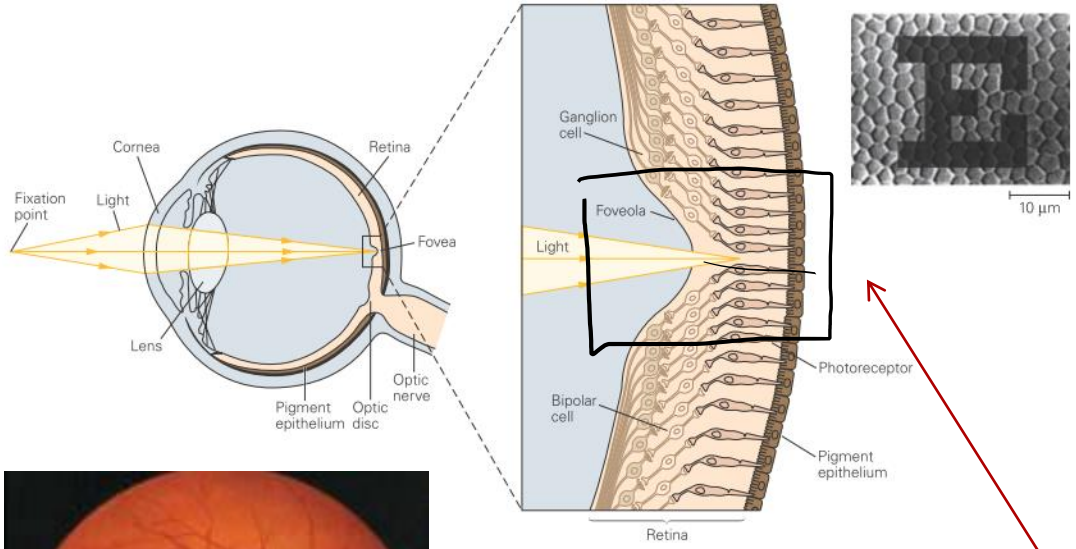


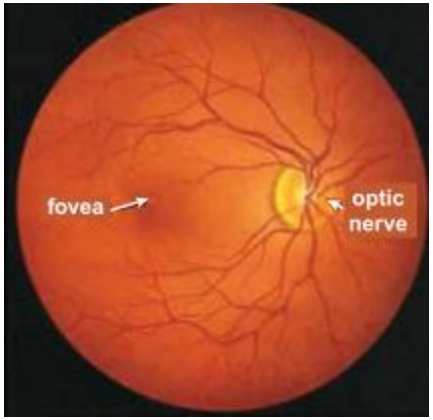


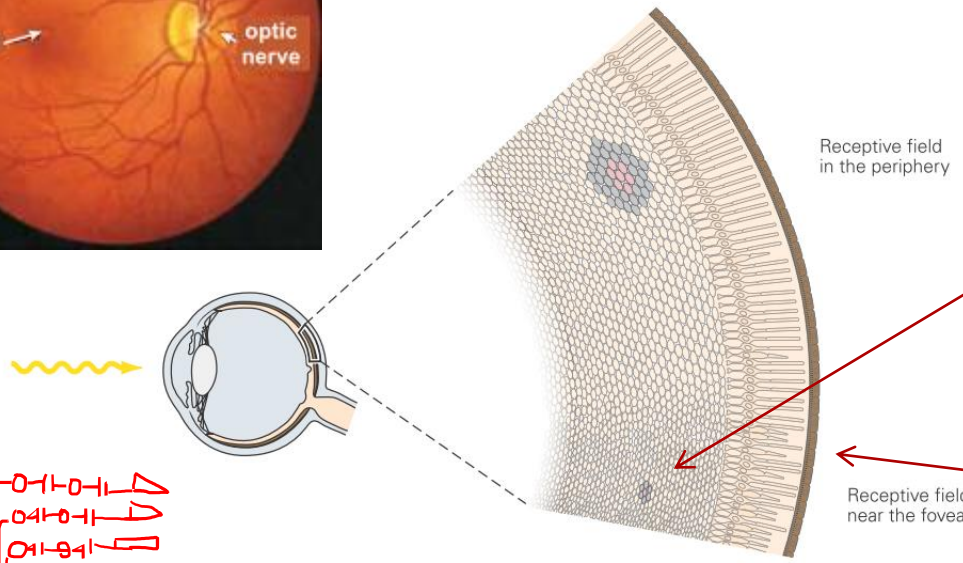
Light passes from glanglion to photoreceptors



H. Kolb, How the reitna works, American Scientist, Volume 91, Number 1 Page: 28 DOI: 10.1511/2003.1.28







Detailed vision in the fovea:

- Bipolar and gangliar cell open
- High density of photoreceptors

Pigment epitelium

Topography of the retina

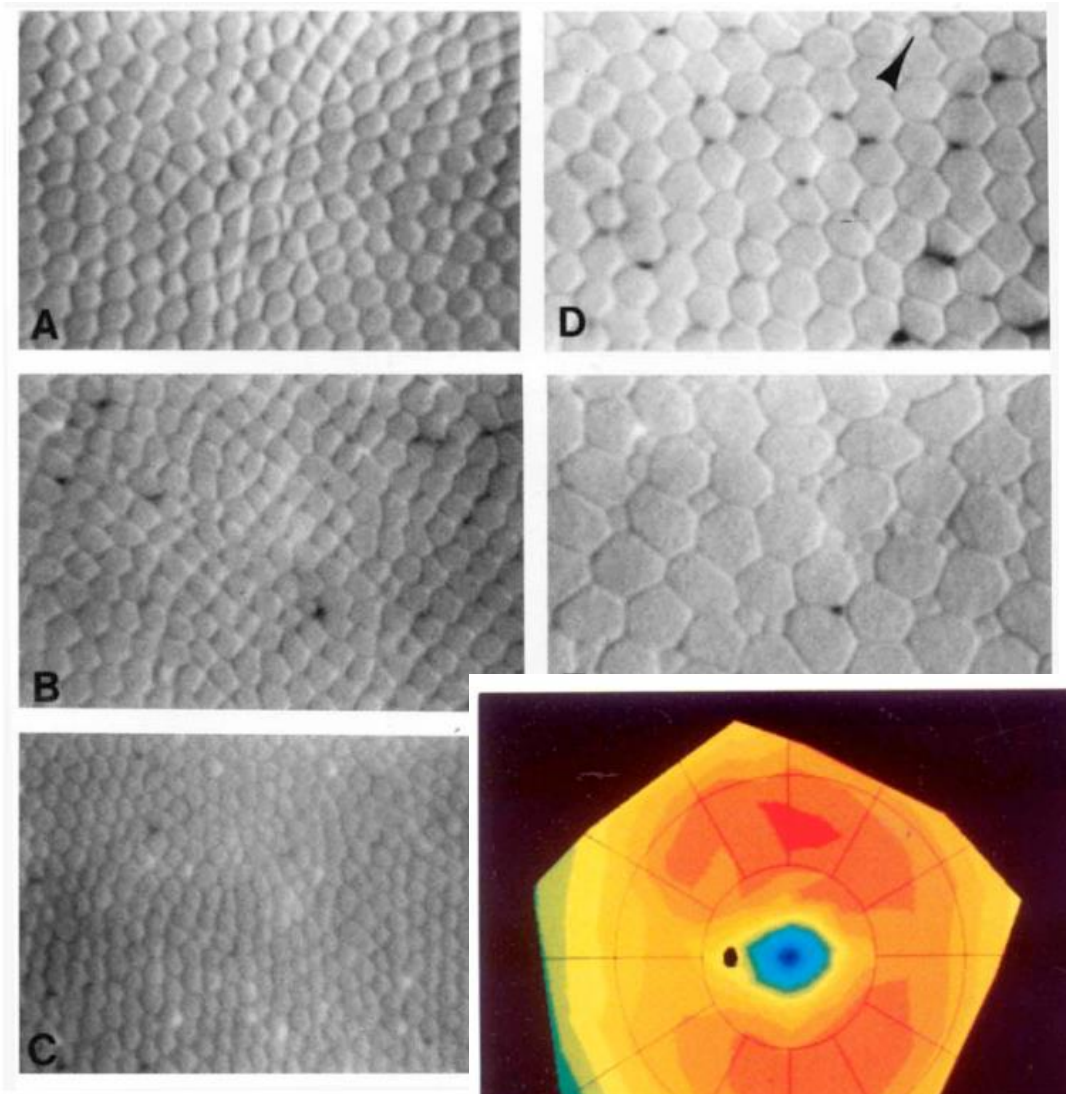
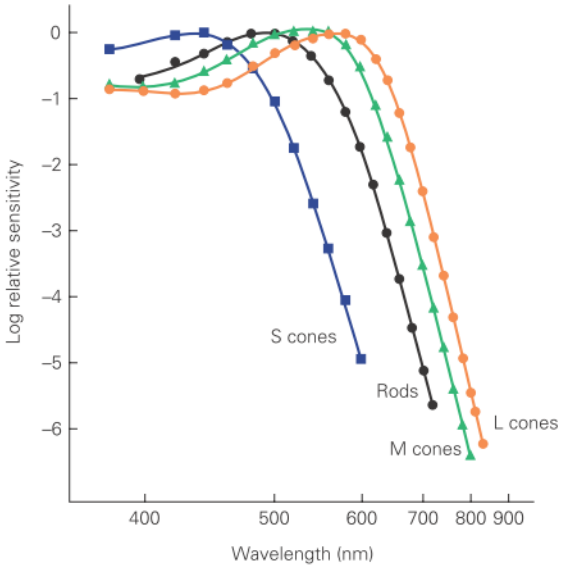
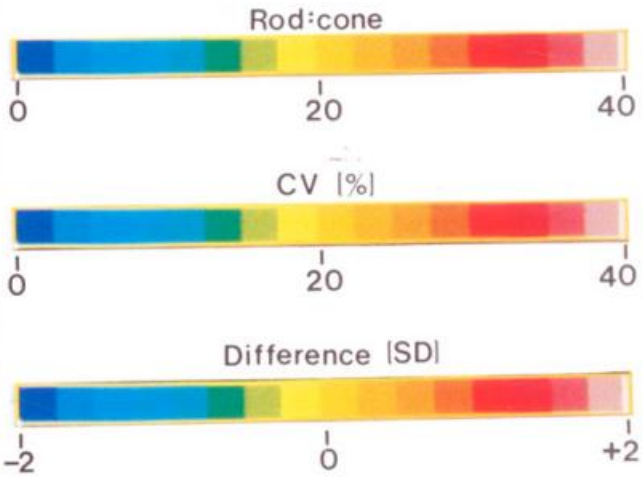


Fig. 2. Optical sections of the foveal cone mosaic, showing betw
individual variability at the foveal center and variation in cell den
and size with eccentricity. **A–C:** Foveal centers, containing only co
of H5L (A), H4 (B), and H6 (C). Note much higher density of cone
H6. **D:** Edge of rod-free zone in H4, 0.125 mm temporal to the fo
center. Arrowhead points to one rod. Note that cone inner segments

Photoreceptors sensitivity

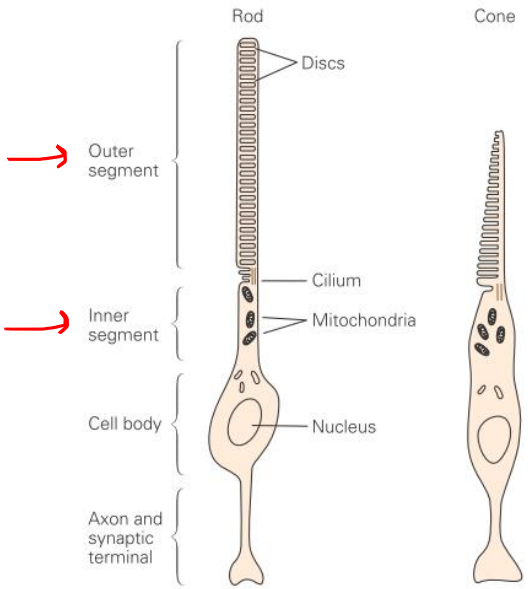


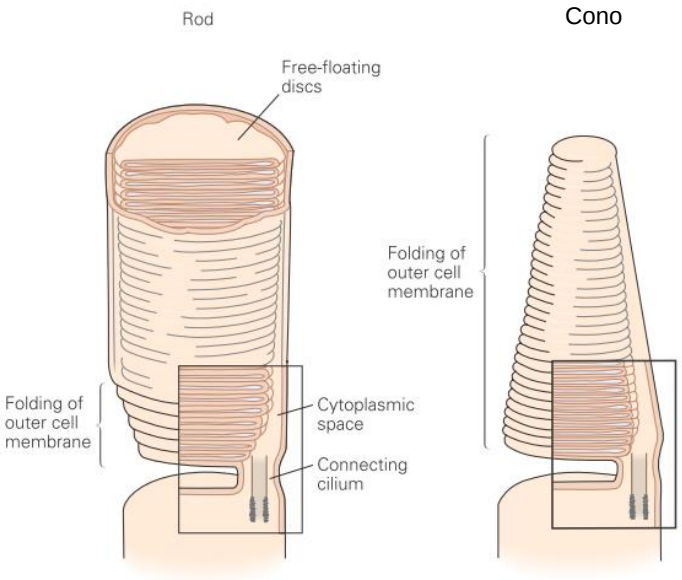


Optic disc and blind spot

- How to redesign the eye to avoid?
- Why the eys is designed in this way?

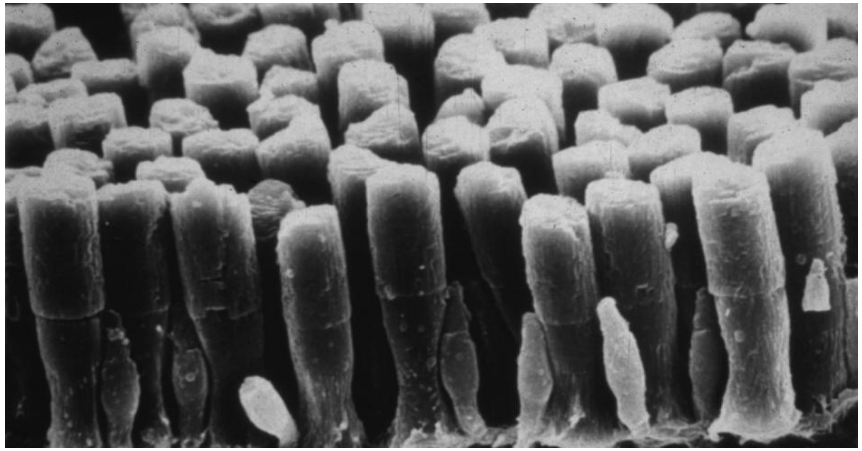
Cones and rods

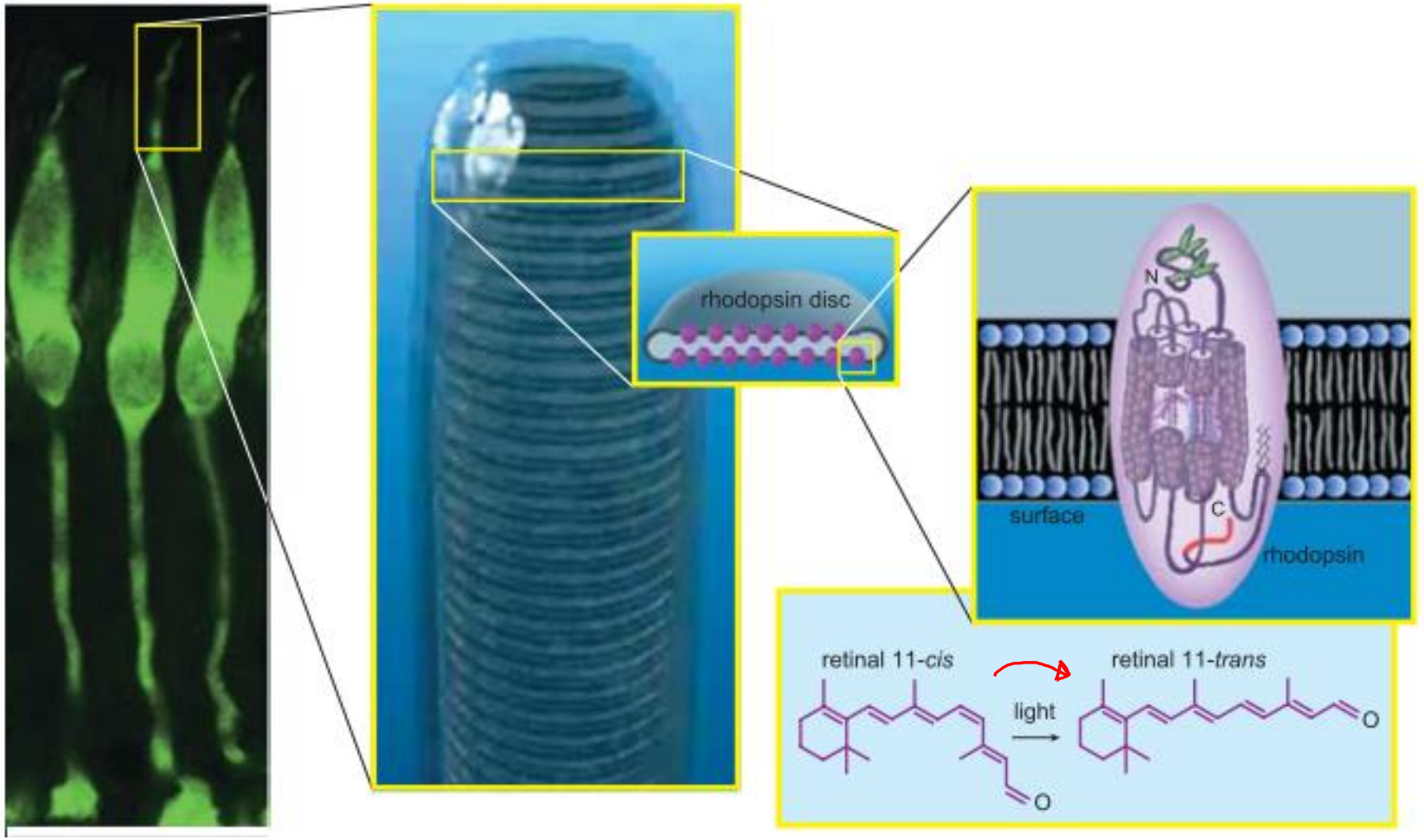




Anatomical and functional differences

About:
100 milion rods
6 miilioni di cones

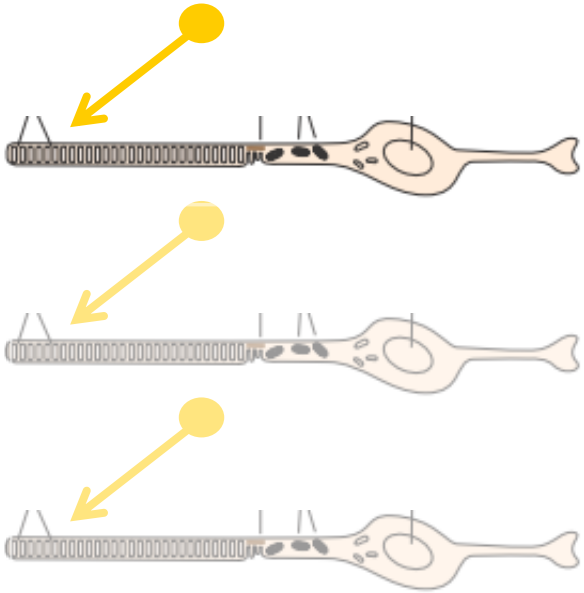




H. Kolb, How the reina works, American Scientist, Volume 91, Number 1 Page: 28 DOI: 10.1511/2003.1.28

Different sensitivity

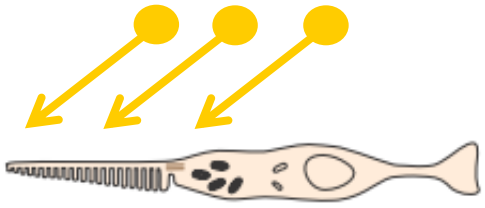
One photon activates the transduction



More rods to the same bipolar cell

Nocturnal

More photons are required to activete the transduction



One cone one bipolar cell

Diurnal

Rod	Cone
High sensitivity	Low sensitivity
Photopigment	Low amount of photopigment
High amplification	Low amplification
Low temporal resolution	High temporal resolution
Diffuse light	Direct ligth

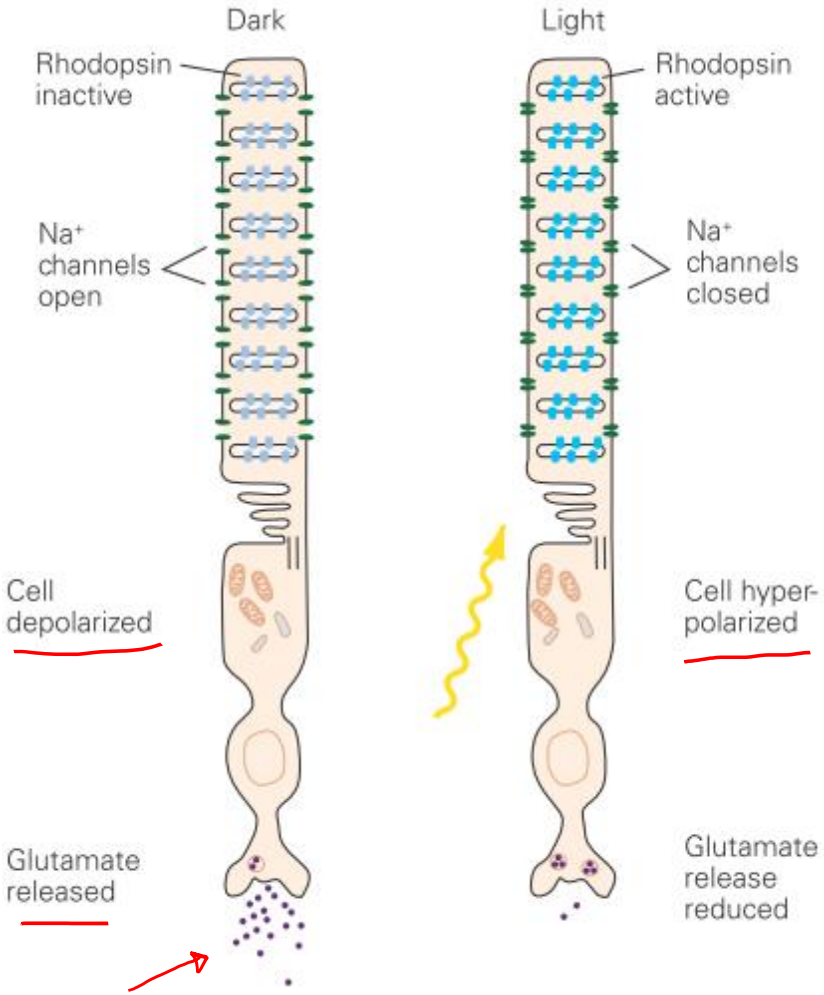
Rods system	Cones system
Low spatial resolution	High spatial resolution
Monocromatic: one kind of photopigment	Cromatic: different kinds photopigments

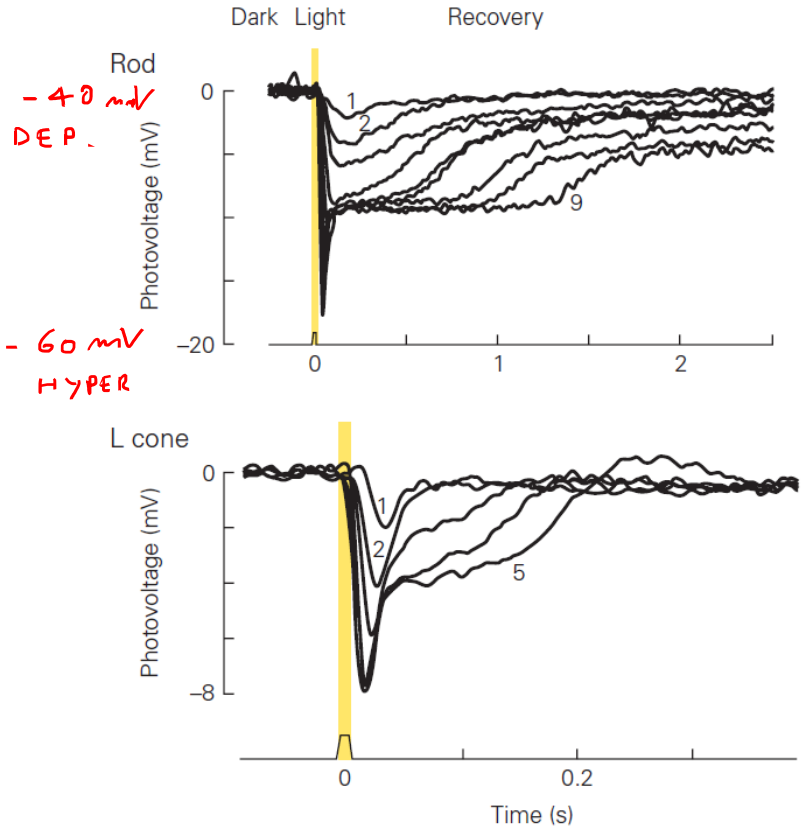
cyclic guanosine monophosphate (cGMP)

Phototransduction

Three steps:

1. Light activates the photo-pigments
2. Photo-pigments reduce the number of GMPc
3. With less GMPc, Na+ channles close, photoreceptor hyperpolarizes





Phototransduction: phase 1 (photopigment activation)

Visual pigment rods: Rodopsina

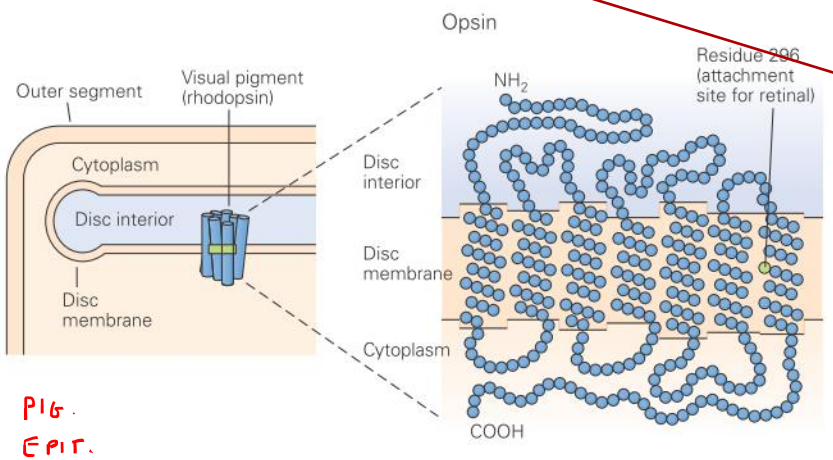
Opsin: bounded to the disc membrane and it does not react to light

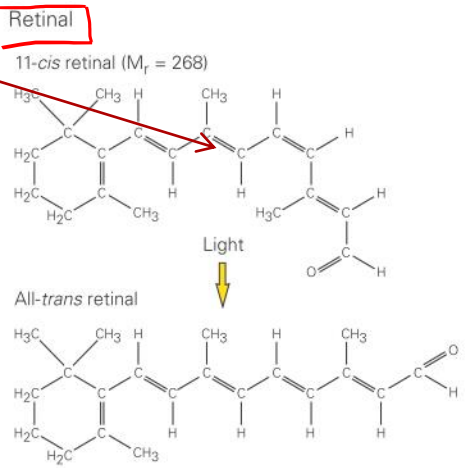
Retinal: it absorbse light and it derives from vitamin A

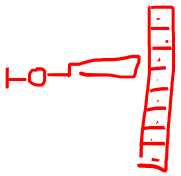
This is the only reaction to light!

Rotation along the 11-cis double link

- Light transforms retinal from 11-cis to all-trans
- Opsin (activated) become metarhodopsin II



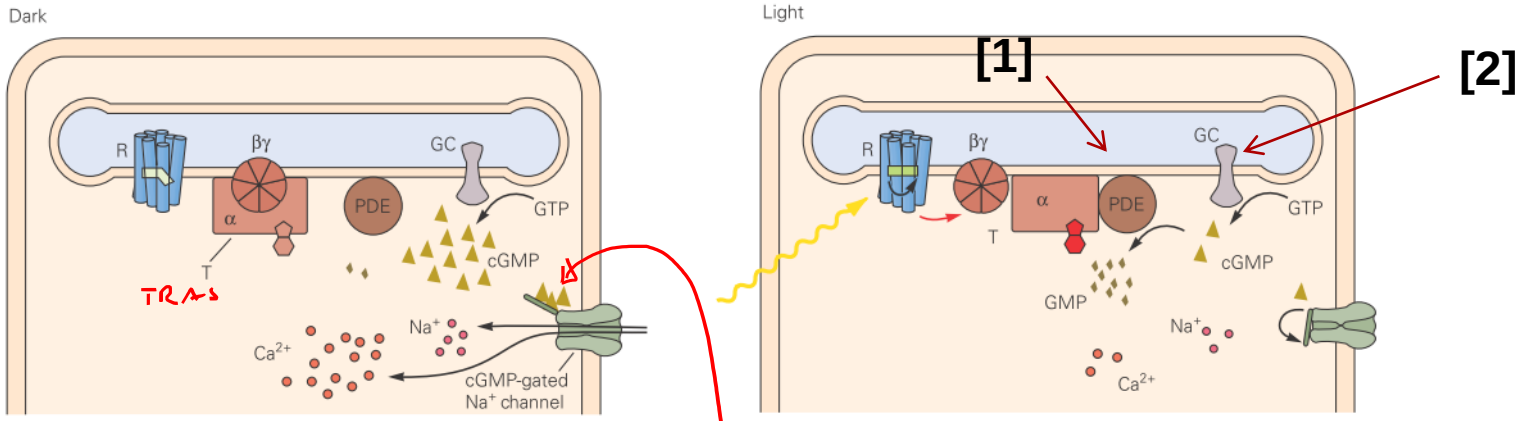




Metarhodopsina II splits in a few minutes into all-trans retinal and opsin

All-trans retinal arrives to the pigmented epitelium to be transformed back to 11-cis retinal

Phototransduction: phase 2 (GMPc concentration reduction)



- Metarhodopsina II interact with trasducin, which splits into T$\beta\gamma$ e GTP-Tα
- GTP-Tα , through other interaction, eventually increases the activity of GMPc-phosphodiesterase

GMPc-phosphodiesterase[1] reduces the amount of GMPc

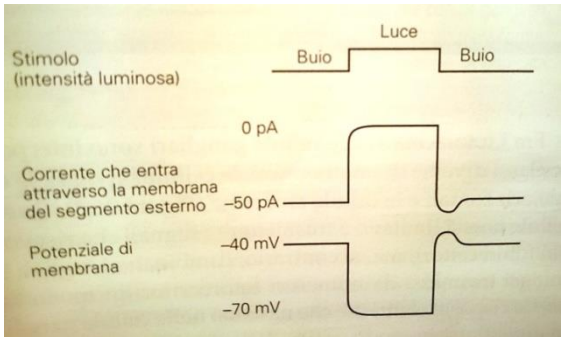
Phototransduction: phase 2 (closure of Na+ channels)

Since the channel at least 3 GMPc molecules are need to open the channel, as their concentration diminishes the channel close

The reduction of GMPc hyperpolarize the photoreceptor

Summary of phototrasduction:

light



Rodopsin → Metarhodopsin II

Trasducin → Trasducin activated

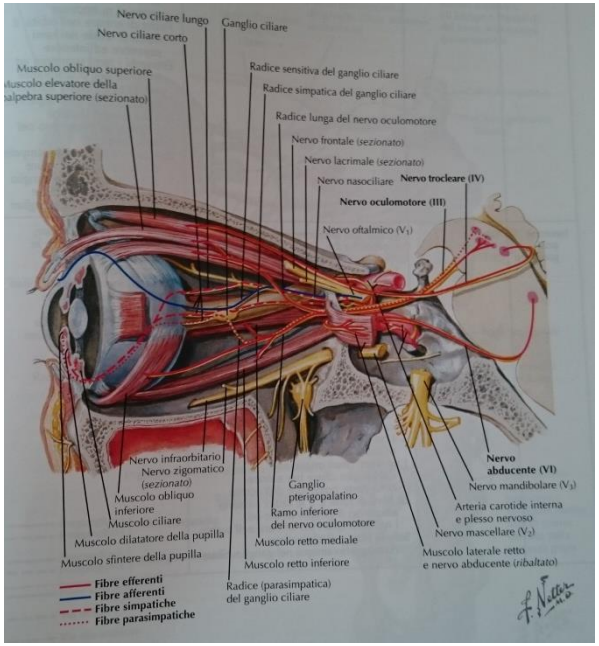
PDE not active → PDE activated

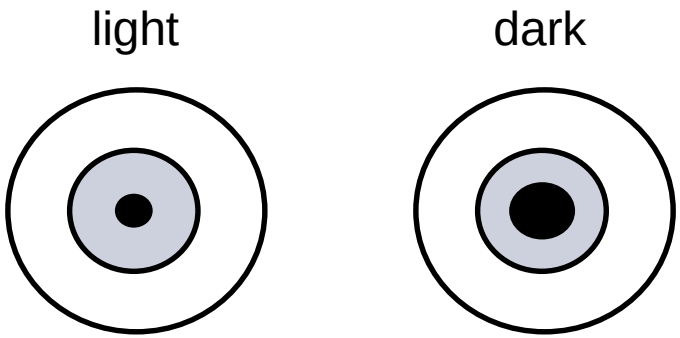
GMPc → ⇩ GMPc → Na+ channels close

Adaptation to light intensity

Two complementary activities:

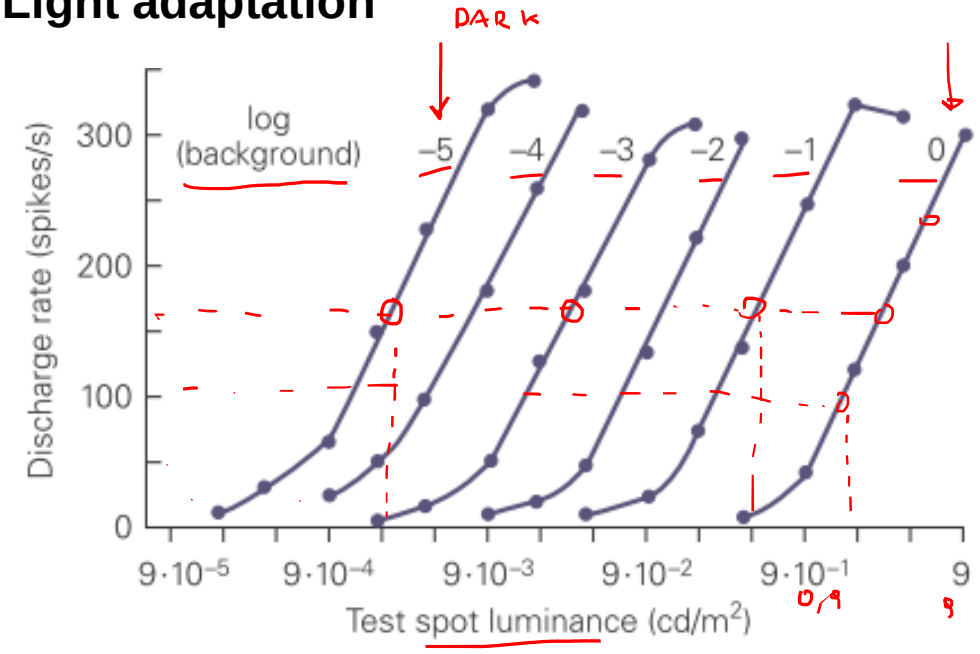
- Reduction on the pupil's size





- Secondary mechanism to depolarize the phtoreceptors

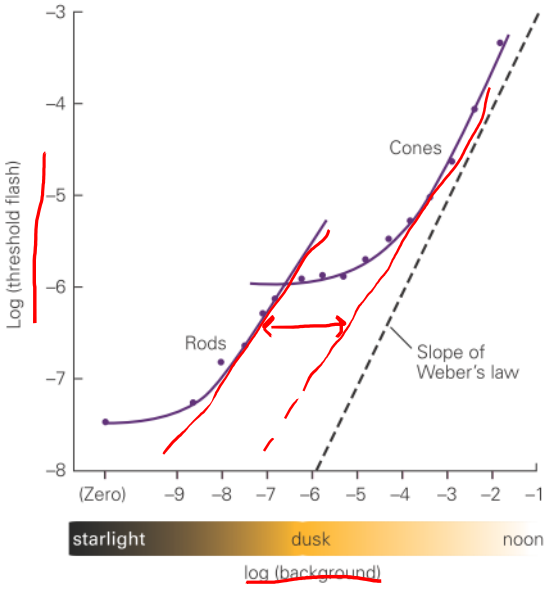
Light adaptation



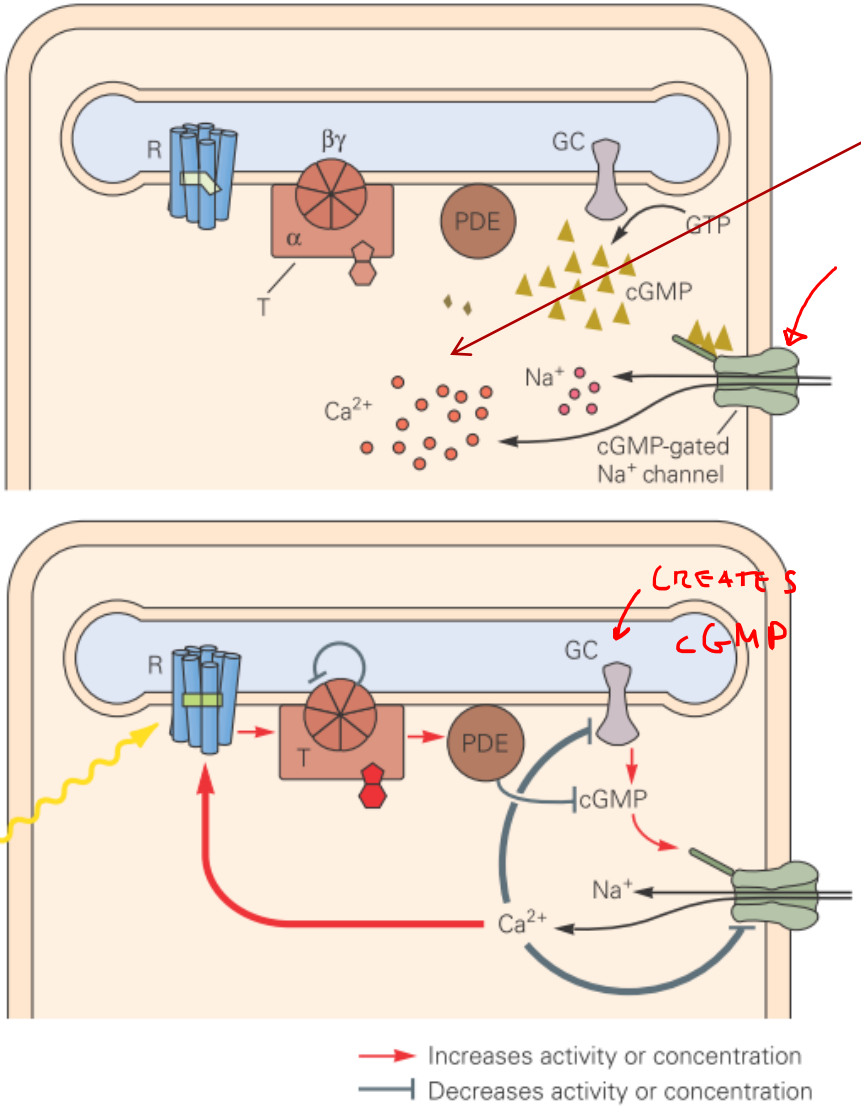
Weber law

$$\Delta S = K \cdot S$$

where ΔS is the minumum difference between a reference S and a second input $\hat{S}$ with enough intensity which can be distinguished as different



How can we perceive a new stimulation?



Sodium (Na+) channles allows the entrance of Ca2+ iones as well

Ca2+ has two roles: **ecitatory** and **inibitory**, respectively increasing activity of visual pigments and lowering the activity of guanylyn-cyclase

↓

- Ca^{2+} → reduce the amount of light for a new stimulus (retinal changes)

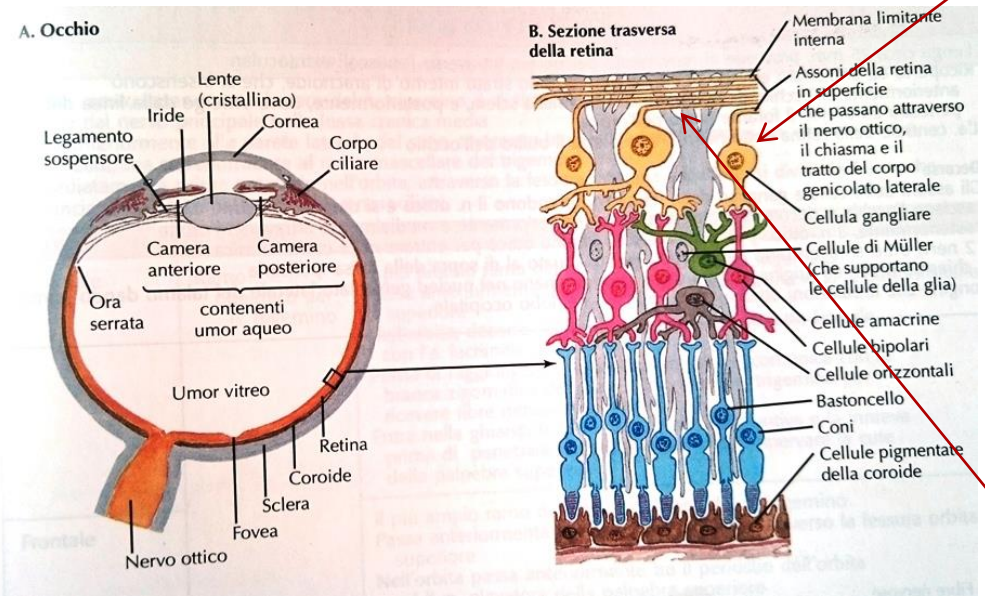
↓

- Ca^{2+} → cGMP ↑ → OPEN

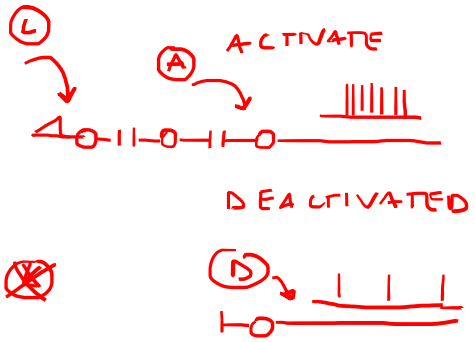
All these effects promote the return of the photoreceptor to the dark state.

Early signal processing

Retina function is to transform light into electric signals, and to provide an **early pre-processing** of the information



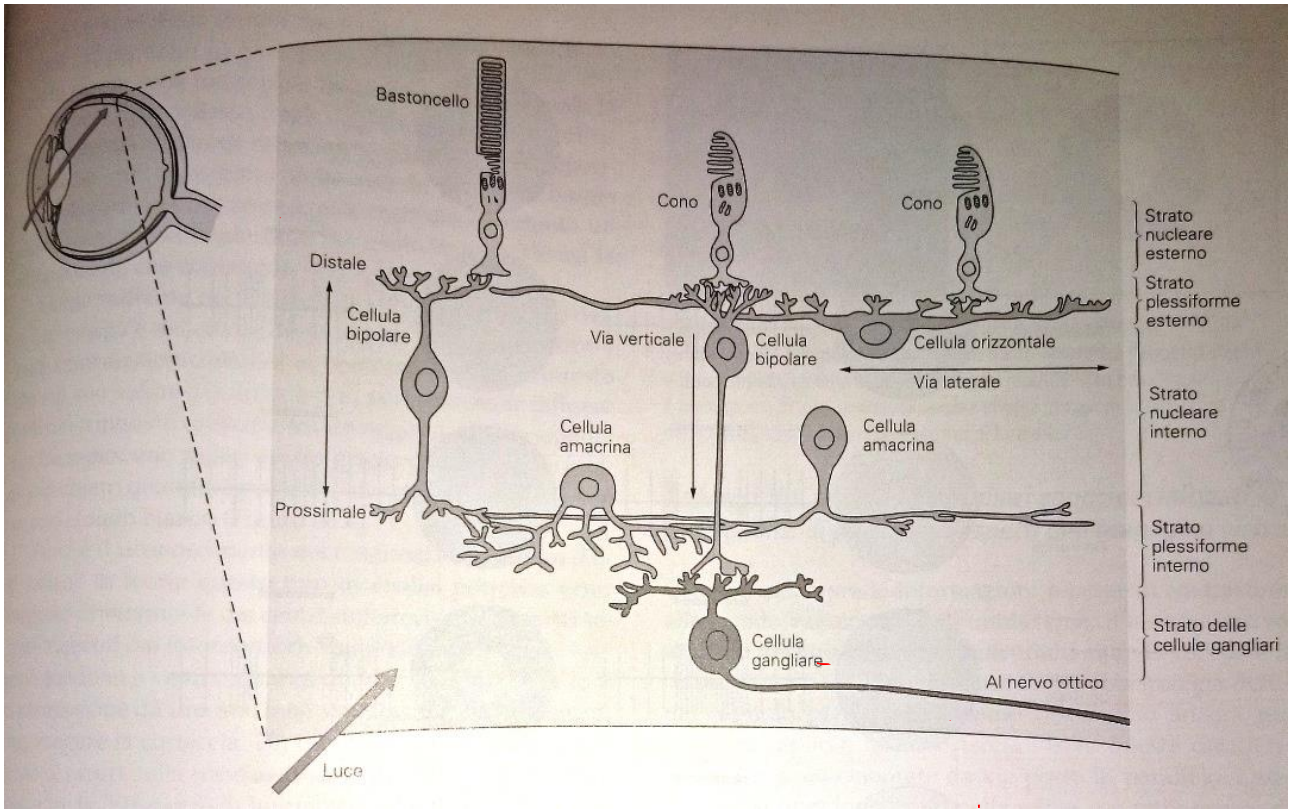
Ganglion cells are the **output neurons** of the retina: **they produce a train of spikes**



Axons of the ganglion cell are collected into the **optic nerve**, which reaches the **lateral geniculate nucleus**, the **superior colliculus**, al **pretectum** and other targets.

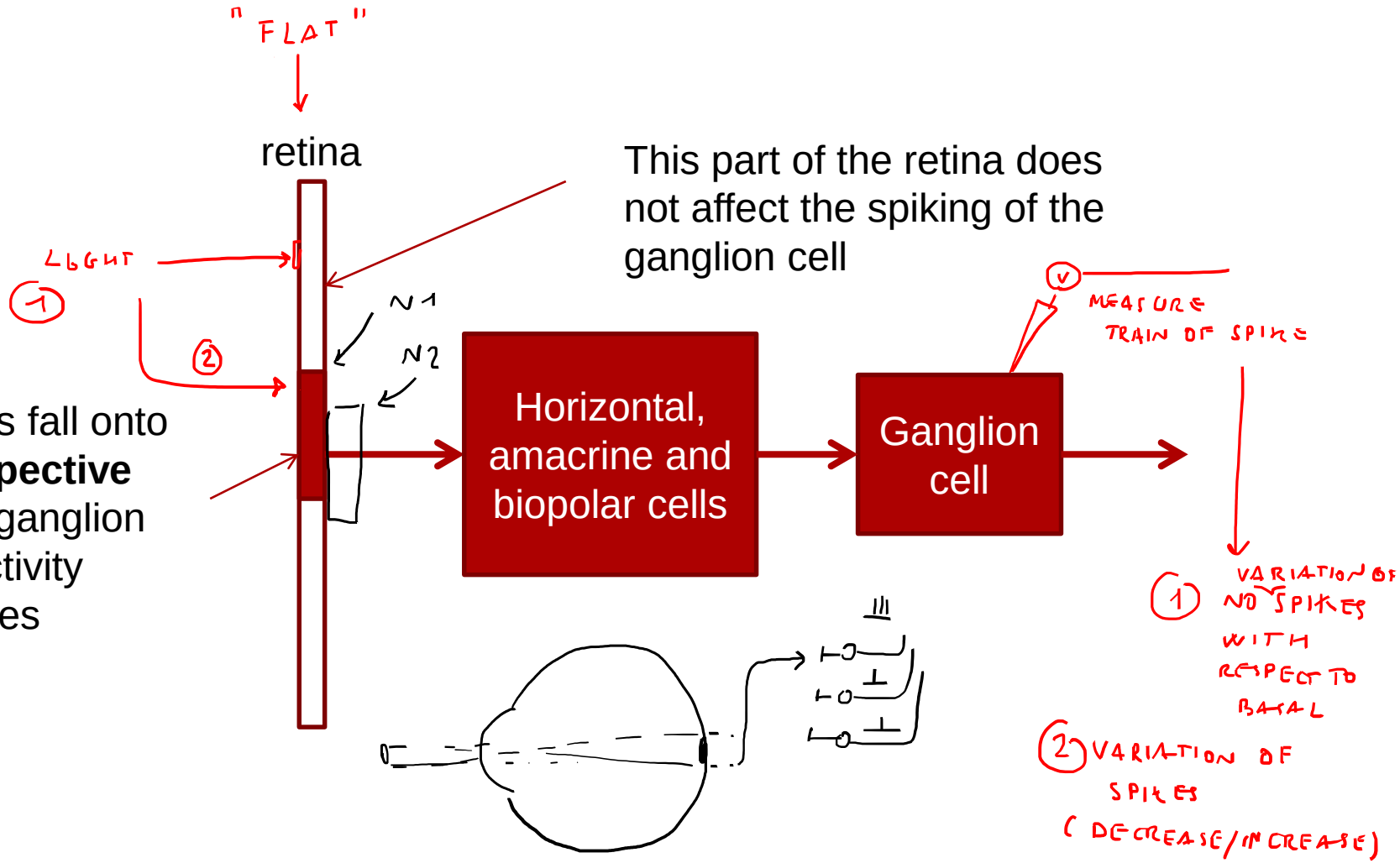
Within the optic nerve, we have **1 axon** each **100 photoreceptors**!

Between photoreceptors and ganglion cells we have: bipolar, horizontal and amacrine cells.

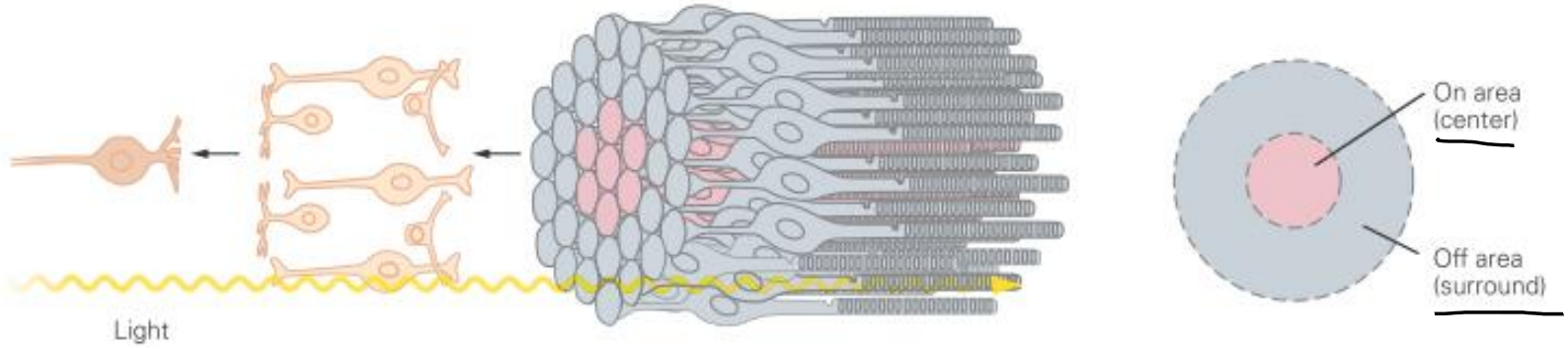


Those neurons **process the signal** which is further projected to the ganglion cells

Each ganlglion cell has a **receptive field**. → CIRCULAR RECEPTIVE FIELD

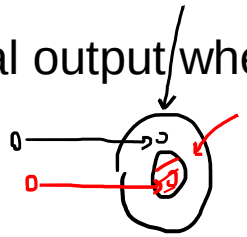


Receptive field are circular in the retina, and we have an **antagonistic** behaviour between the centre and the surround (peripheral area)



REC FIELD

Ganglion cells have optimal output when the illumination is different between the centre and the surround.

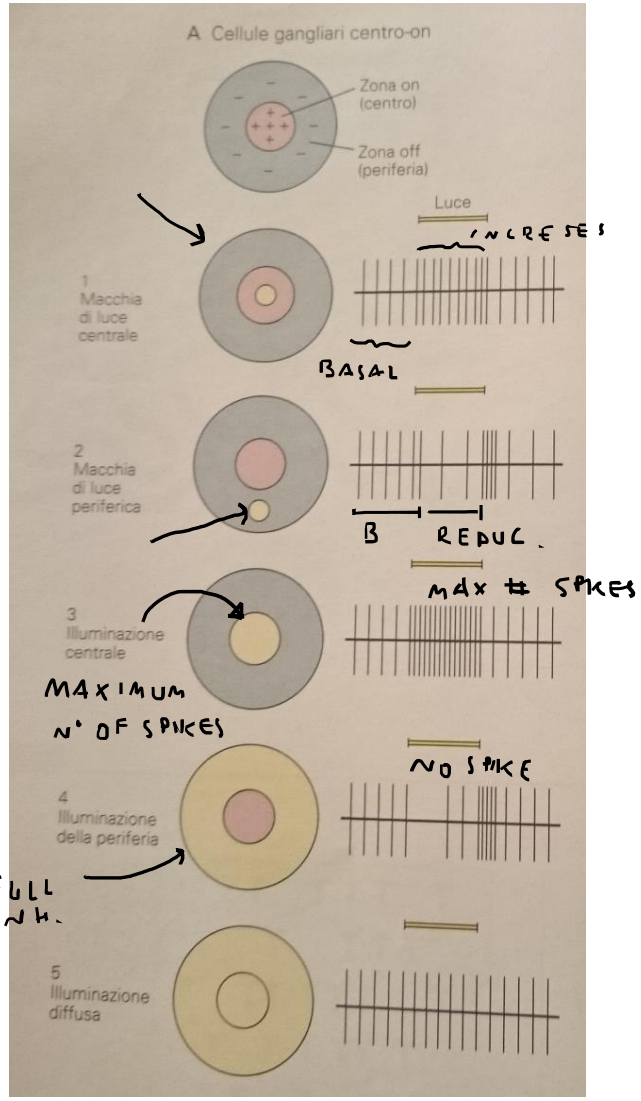


With respect to the behaviour of centre-surround, we can classify two different kinds of ganglion cells:

- **centre-on**
- **centre-off**

Centre-ON

Position of the light

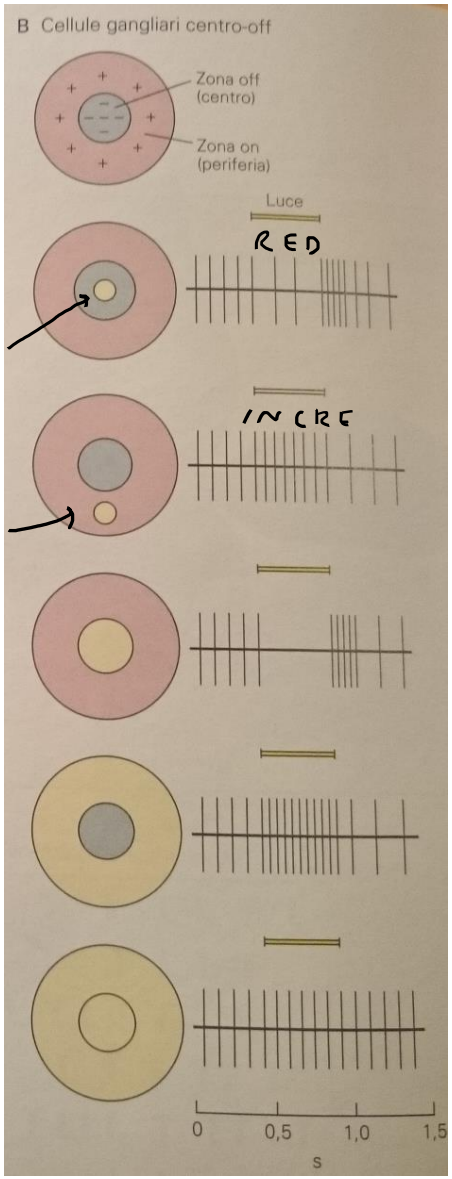


Duration of the light

Light on the centre, increases the number of spikes

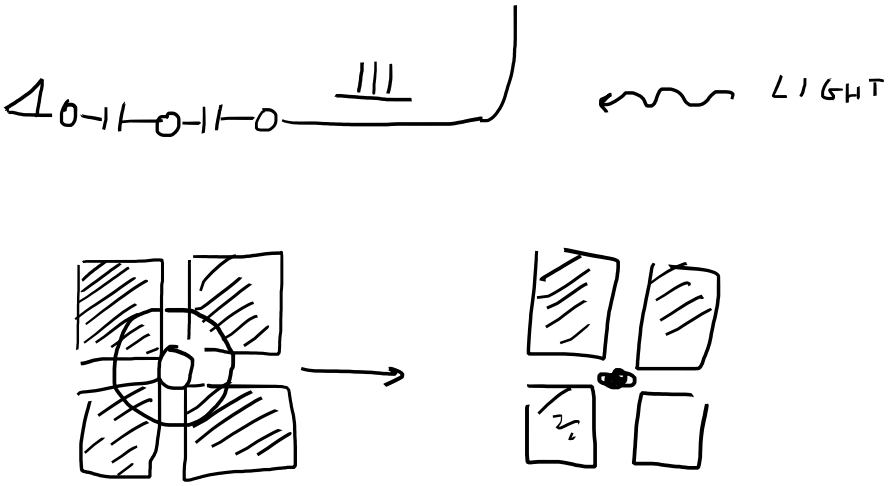
Light on the surround, decreases the number of spikes

Centre-OFF



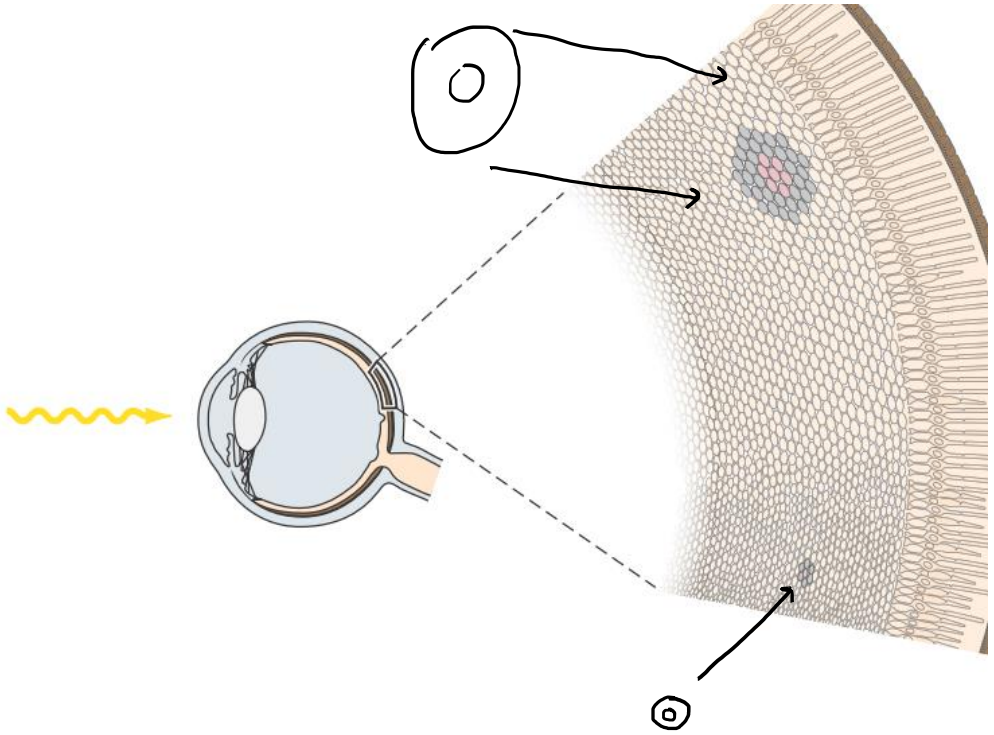
Centre-OFF cells have opposite behaviour: when light hits the centre, spike frequency decreases.

On the other end, when lights hit the surround, spike frequency increases.

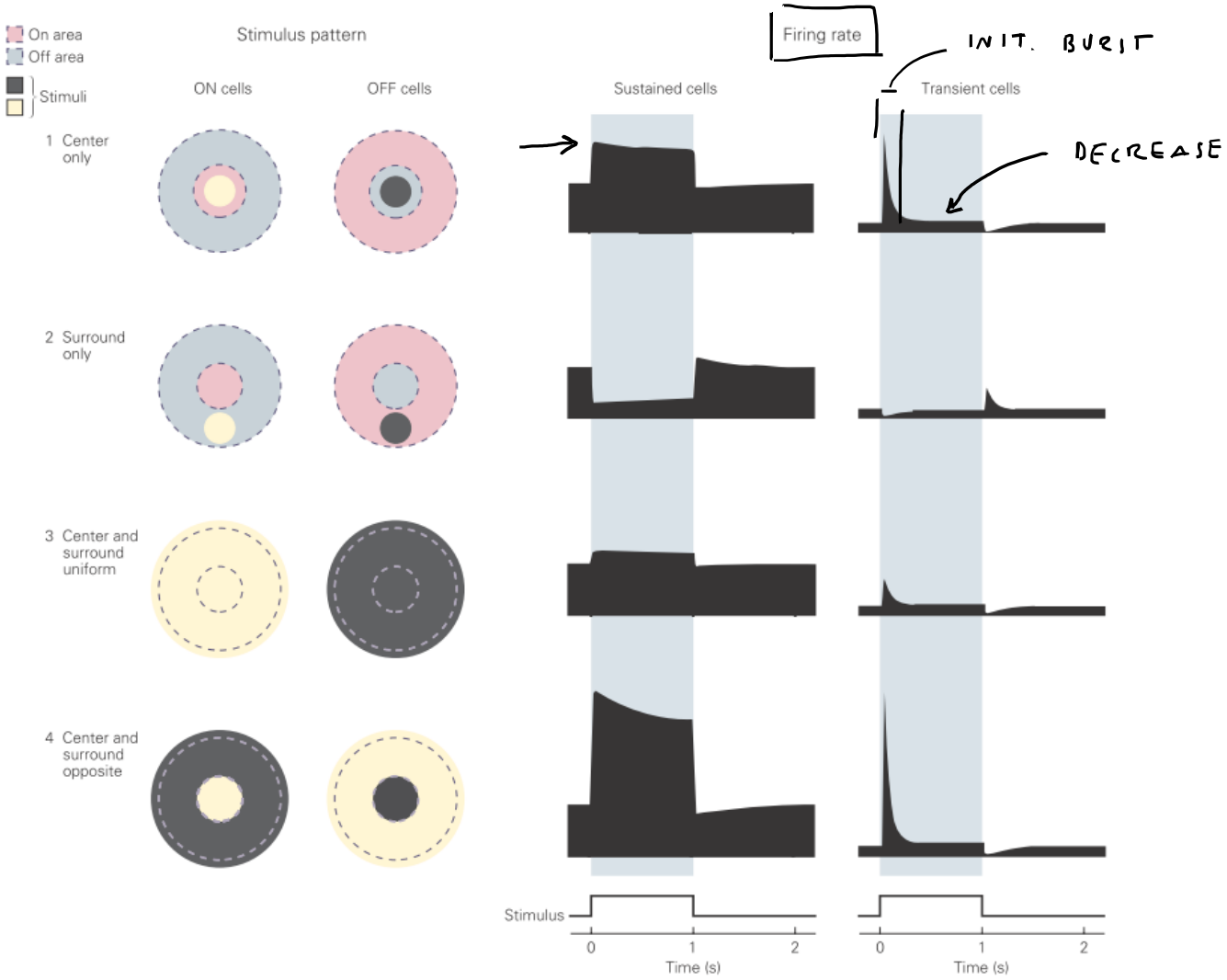


We have a similar number of center-on e center-off cells, thus they elaborate the information together

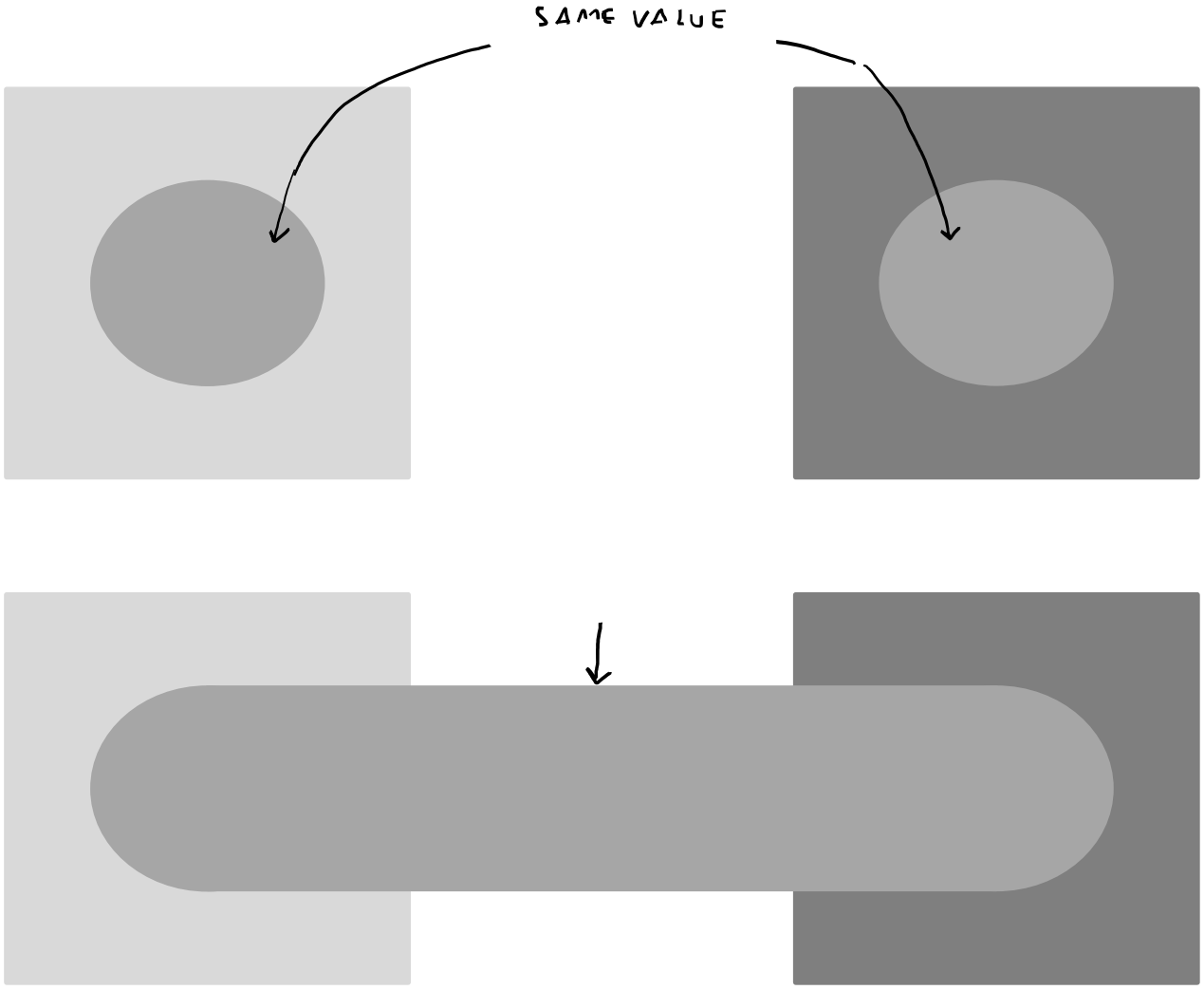
Receptive fields are **smaller in the fovea rather than in the periphery of the retina**



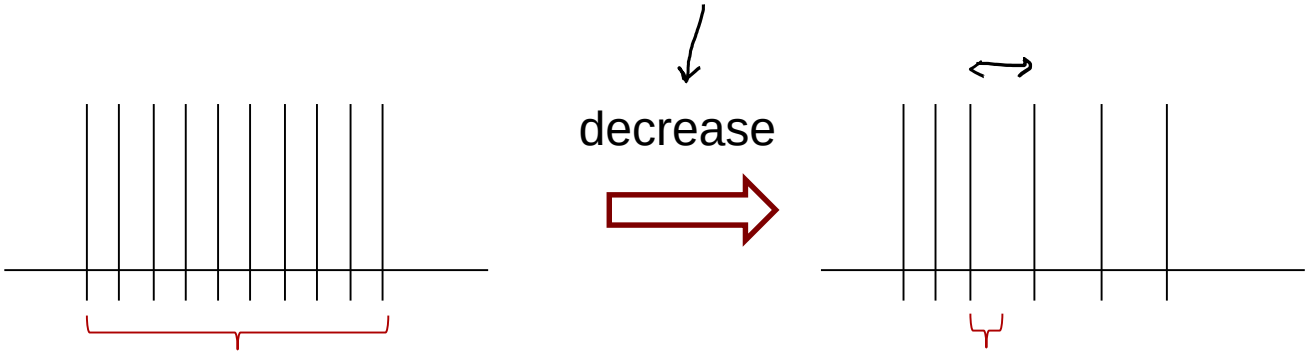
We have another classification if excitation persists throughout stimulation. We call **sustained cells** if the excitation persists, **transient cells** otherwise.



We do not perceive the absolute value of illumination. Our eye detects **intensity contrast** within the scene.

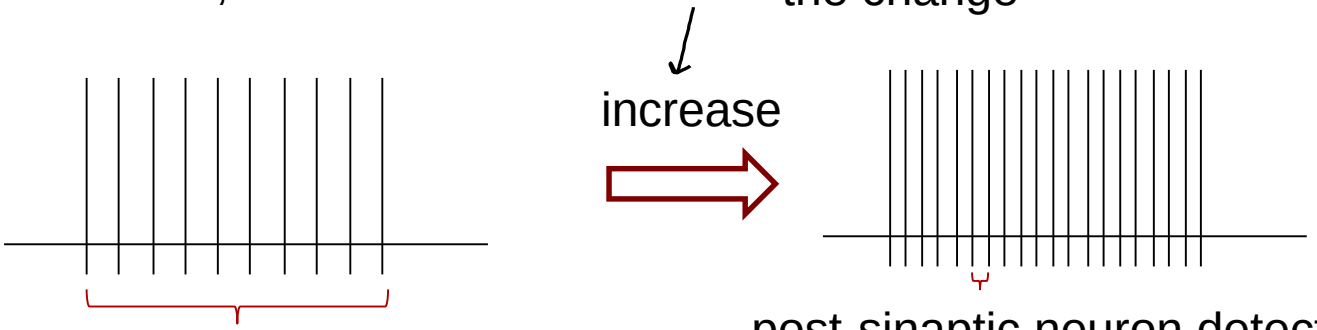


❖ Why two complementary systems?



10 spike /s = 1 spike each 0,1s

→ post-sinaptic neuron needs 0,1s to identify the change

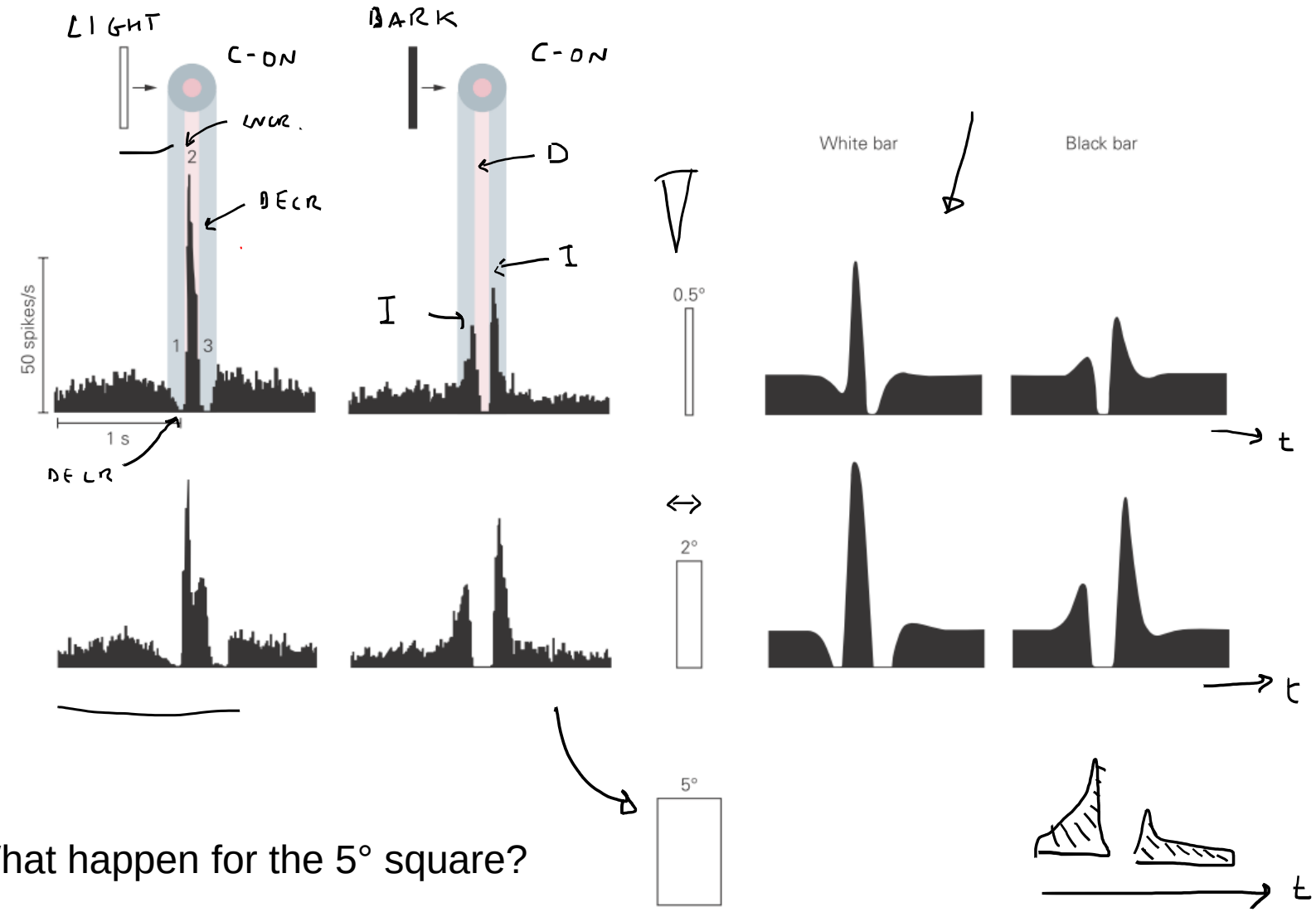


10 spike / s = 1 spike ogni 0,1s

post-sinaptic neuron detects in less than 0,1s the change

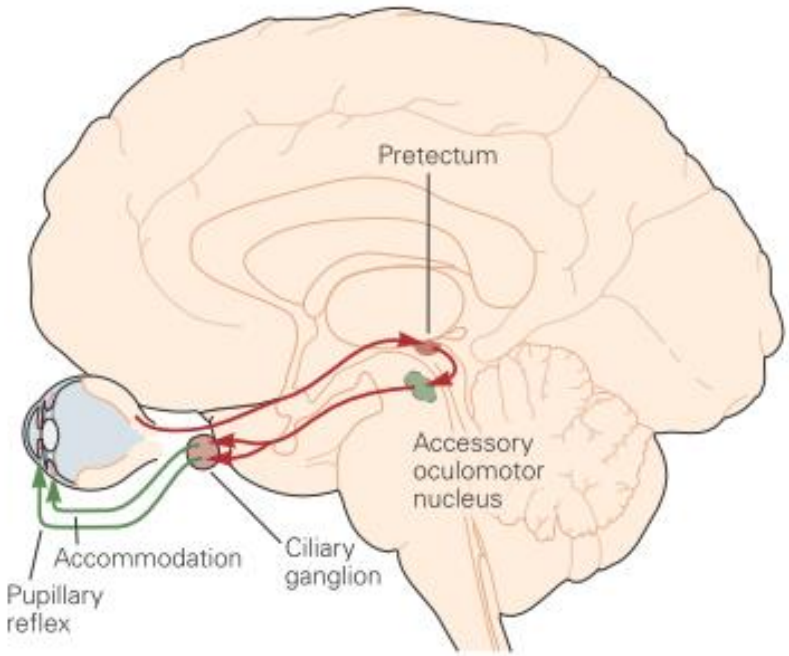
This is a hypothetic explanation of the existence of two complementary systems

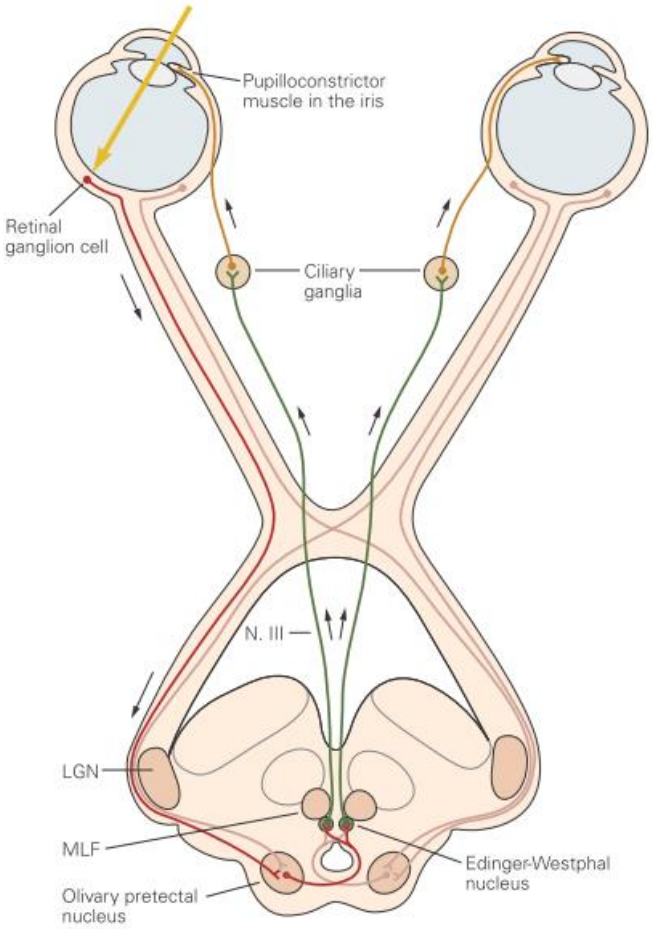
This system allows also to detect **movements** and **gross dimension**:



- What happen for the 5° square?
- What happen if the objects appear from the right?

- There exist ganglion cells without the center/periphery behavior. Which could be their function?

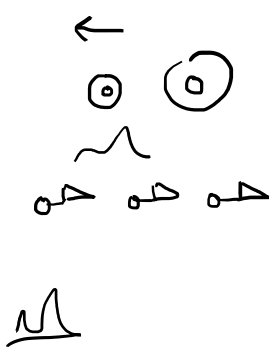




There are about 20 different types of ganglion cells

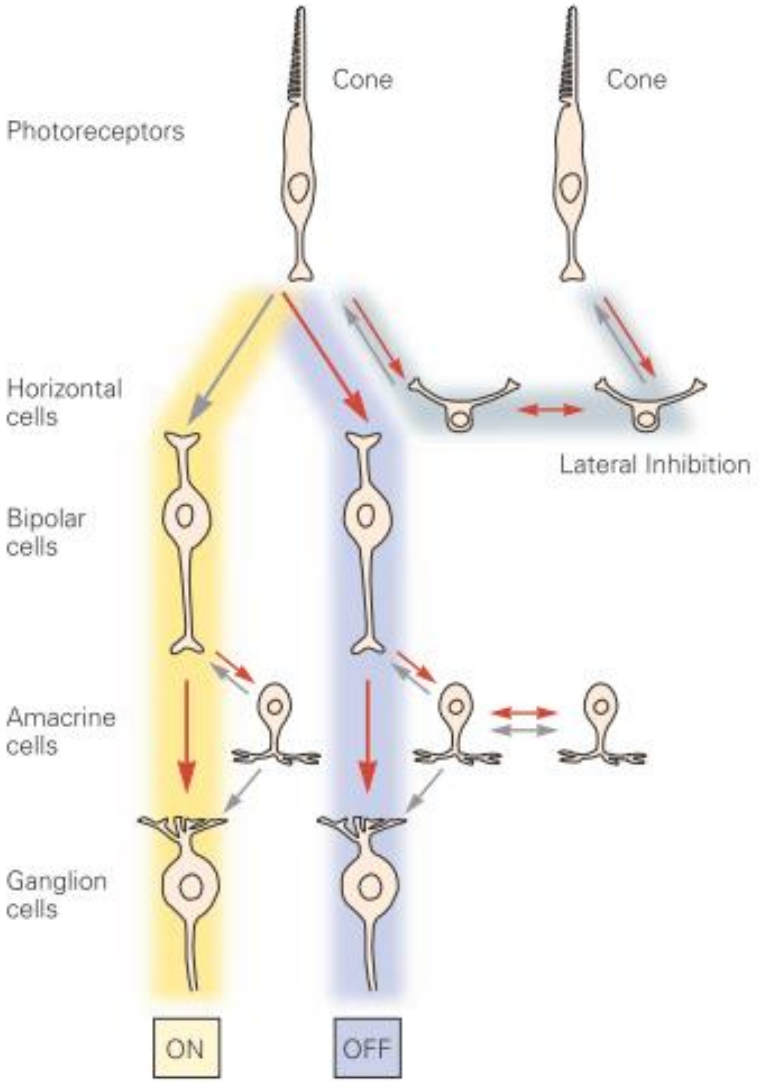
If we consider different features of the image, we can state the optic nerve transmit 20 representation of the worlds, different for:

- On / Off behaviour
- Spatial resolution
- Temporal resolution
- Spectral filtering
- Diffuse light
- Movements



How is possible to transform the relatively simple photoreceptors behaviour, in such a rich features extraction system?

The interneuron layers provide this «features extraction» property



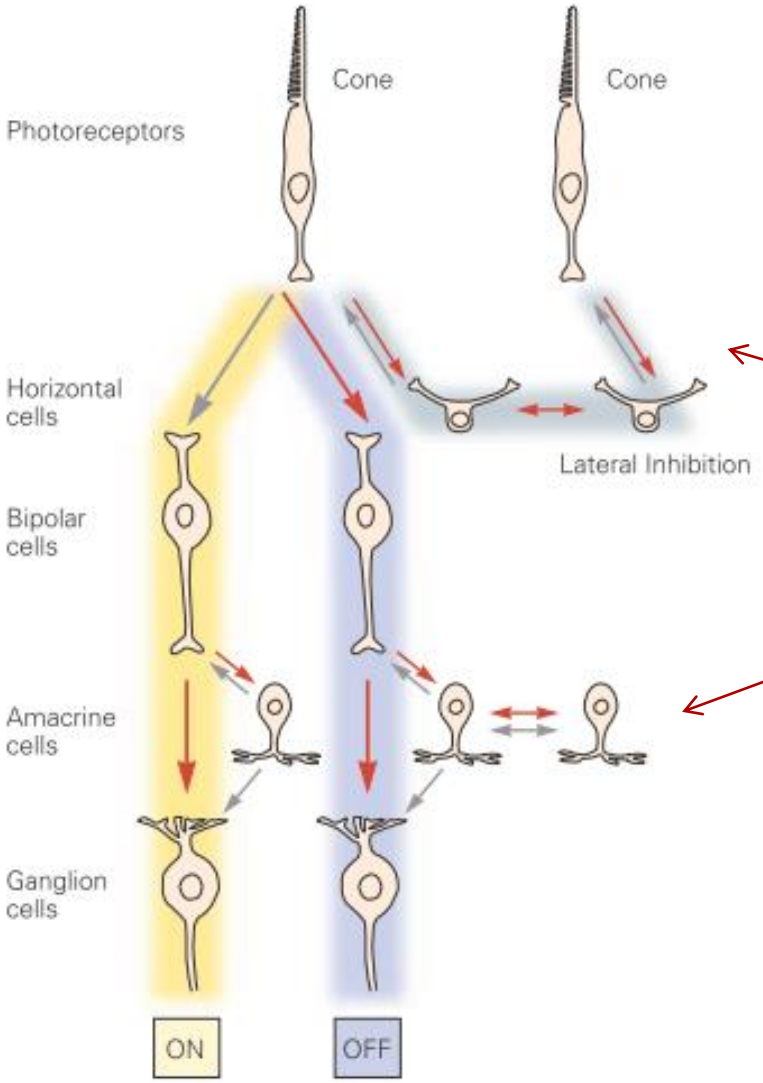
We have two pathways in the retina: one vertical and one horizontal

Cone -> Bipolar -> Ganglion cells is called **direct pathway** (vertical)

Cone -> Horizontal -> Bipolar -> Ganglion cells,

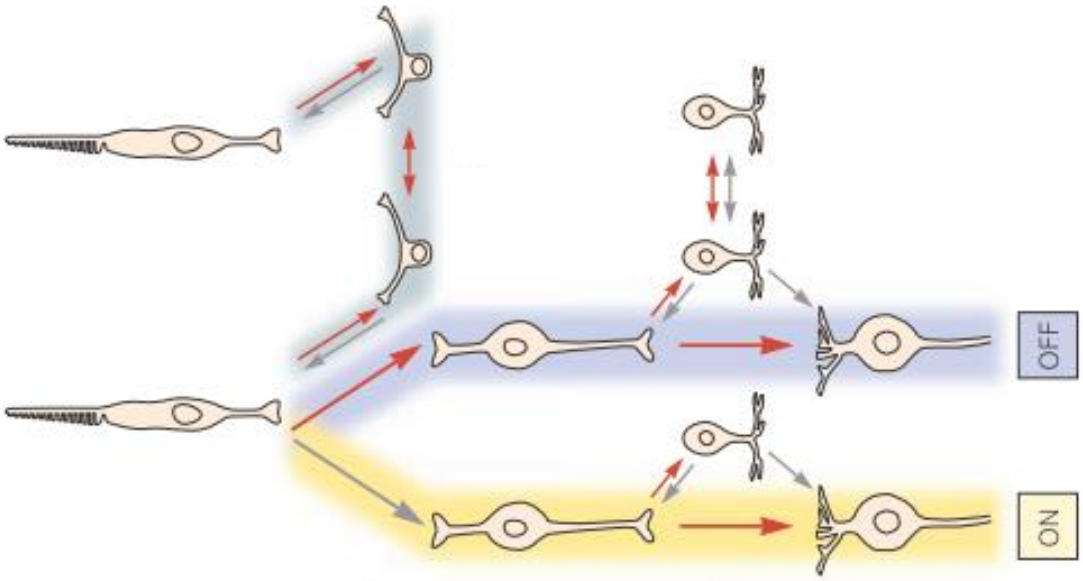
Bipolar <-> Amacrine -> Ganglion ->

They are indirect ways, called **lateral pathways** (horizontal)



We have horizontal pathways both at the level of horizontal cells and at amacrine cells

Amacrine cells allow to create direction-selective behaviour for Magno ganglion cells

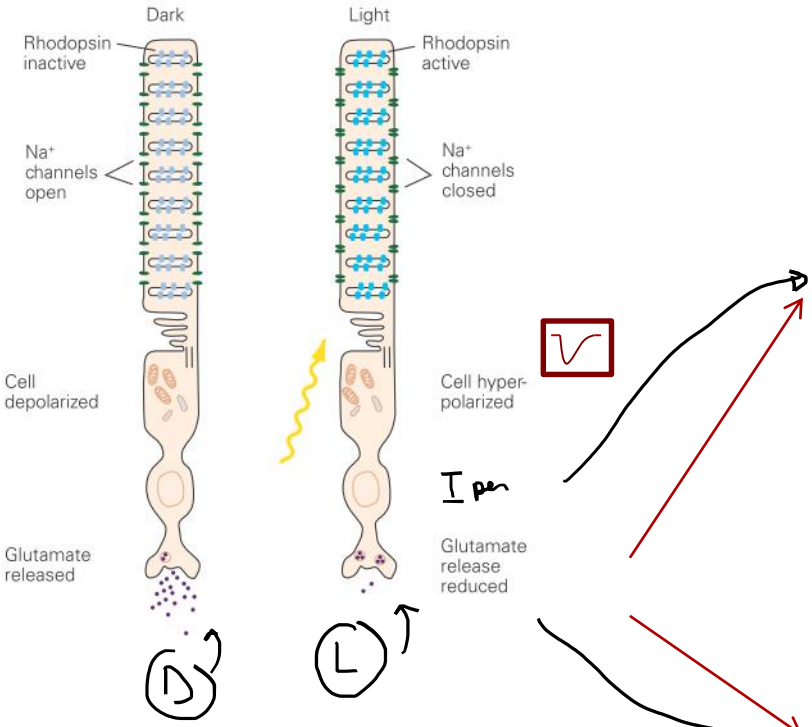


Also bipolar cells have a different behavior on center-on and center-off receptive field

Bipolar center-on cells depolarize when subject to illumination, and further **depolarize center-on ganglion cells**.

Bipolar center-off cells hyper-polarize when subject to illumination, and **hyperpolarize center-off ganglion cells**

Bipolar (centre-on and centre-off) have a different response to glutamate, released by photoreceptors



I)

Centre-on

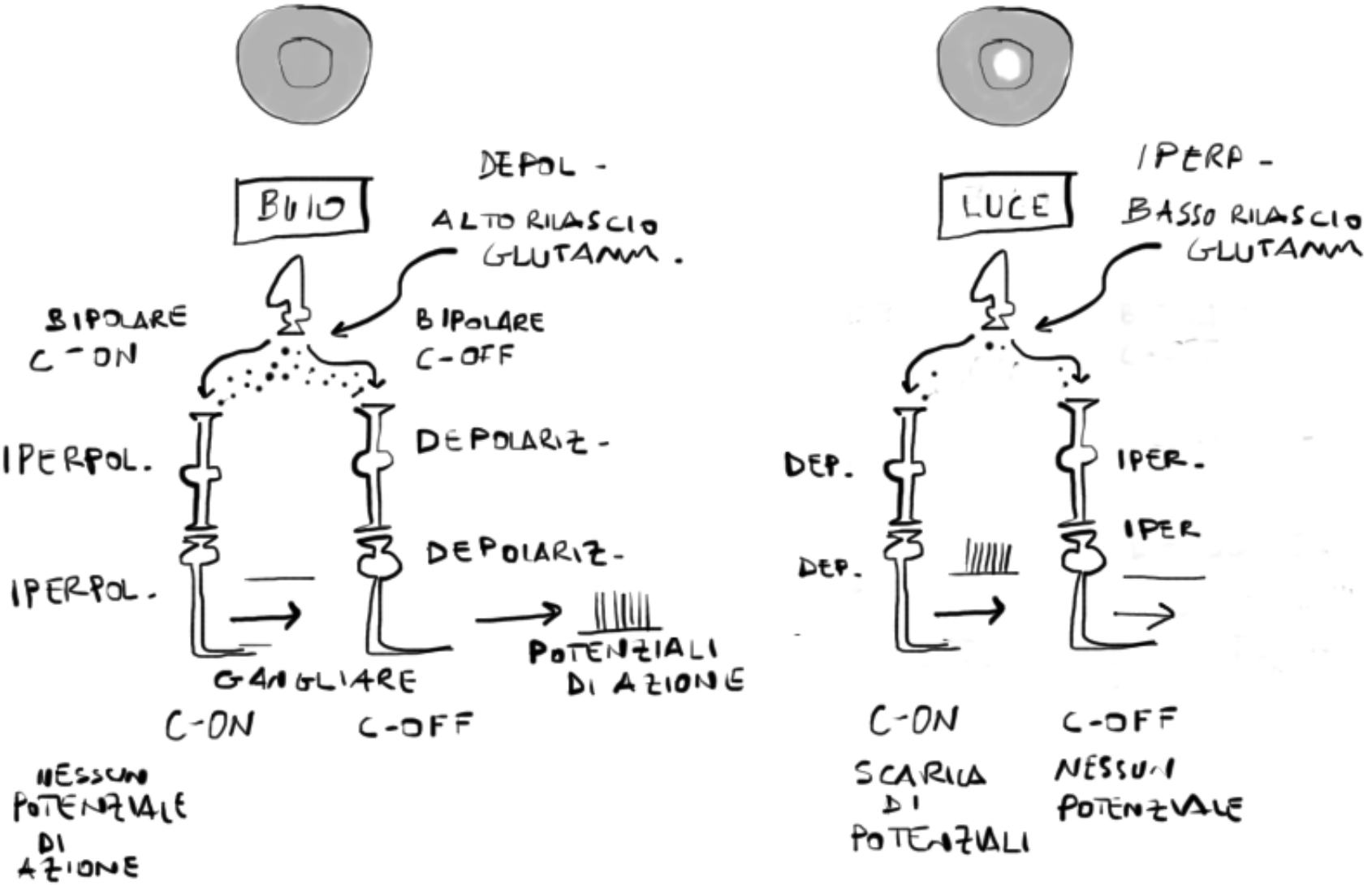
Ganglion cells

depolarized

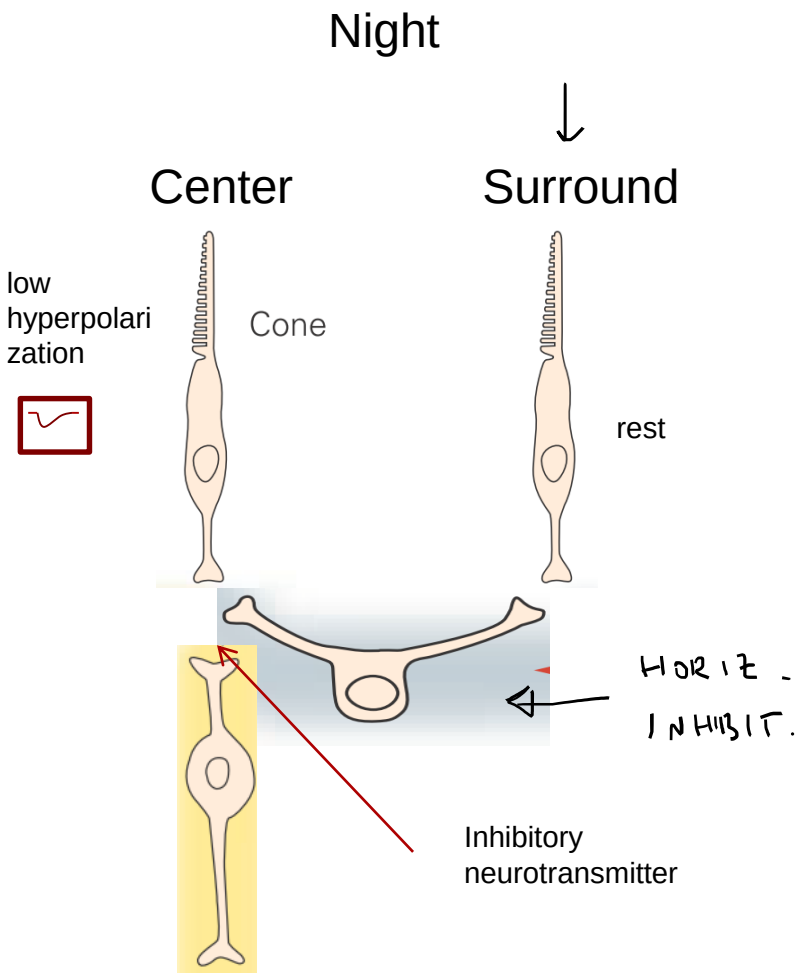
Centre-off

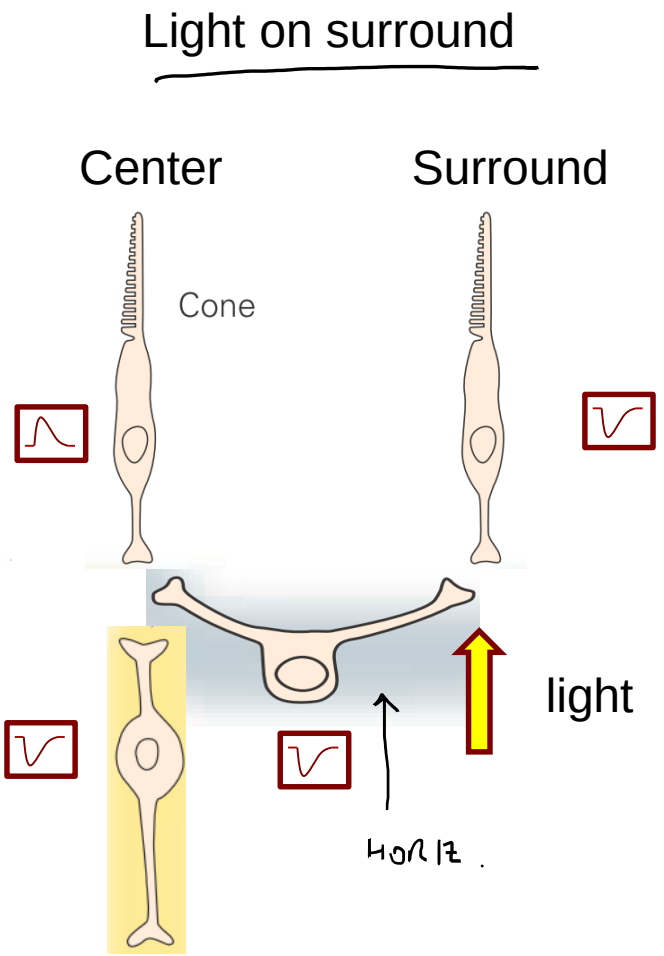
Hyper-polarized

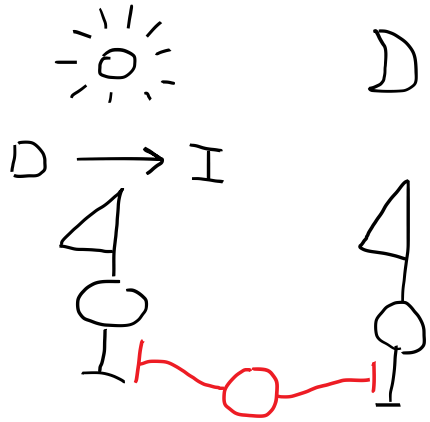
VIA VERTICALE

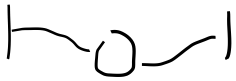


Lateral **inhibitory** connections create the antagonistic mechanism of receptive field

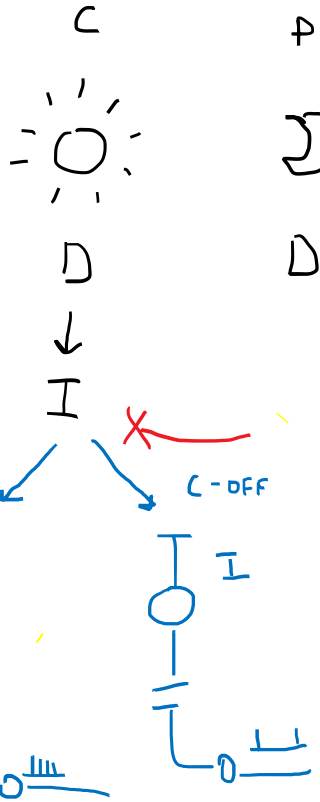


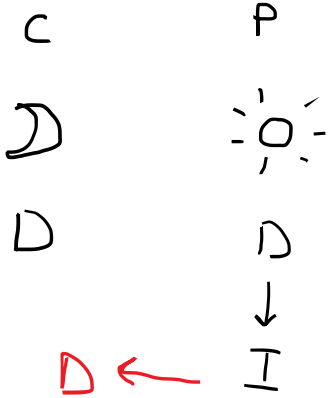


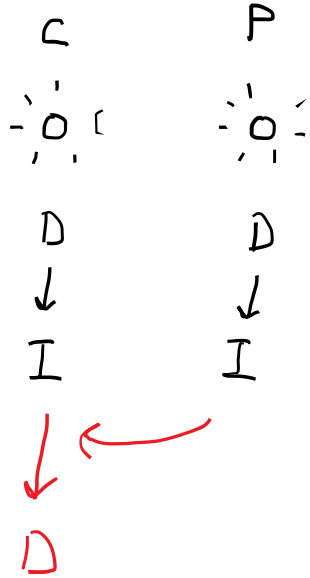


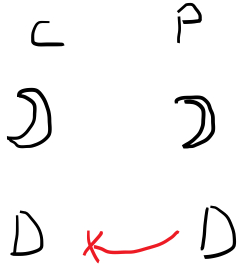


WHEN "ACTIVATED",
TENDS TO BRING
PHOTORECEPTOR TO
DEPOLARIZED STATE

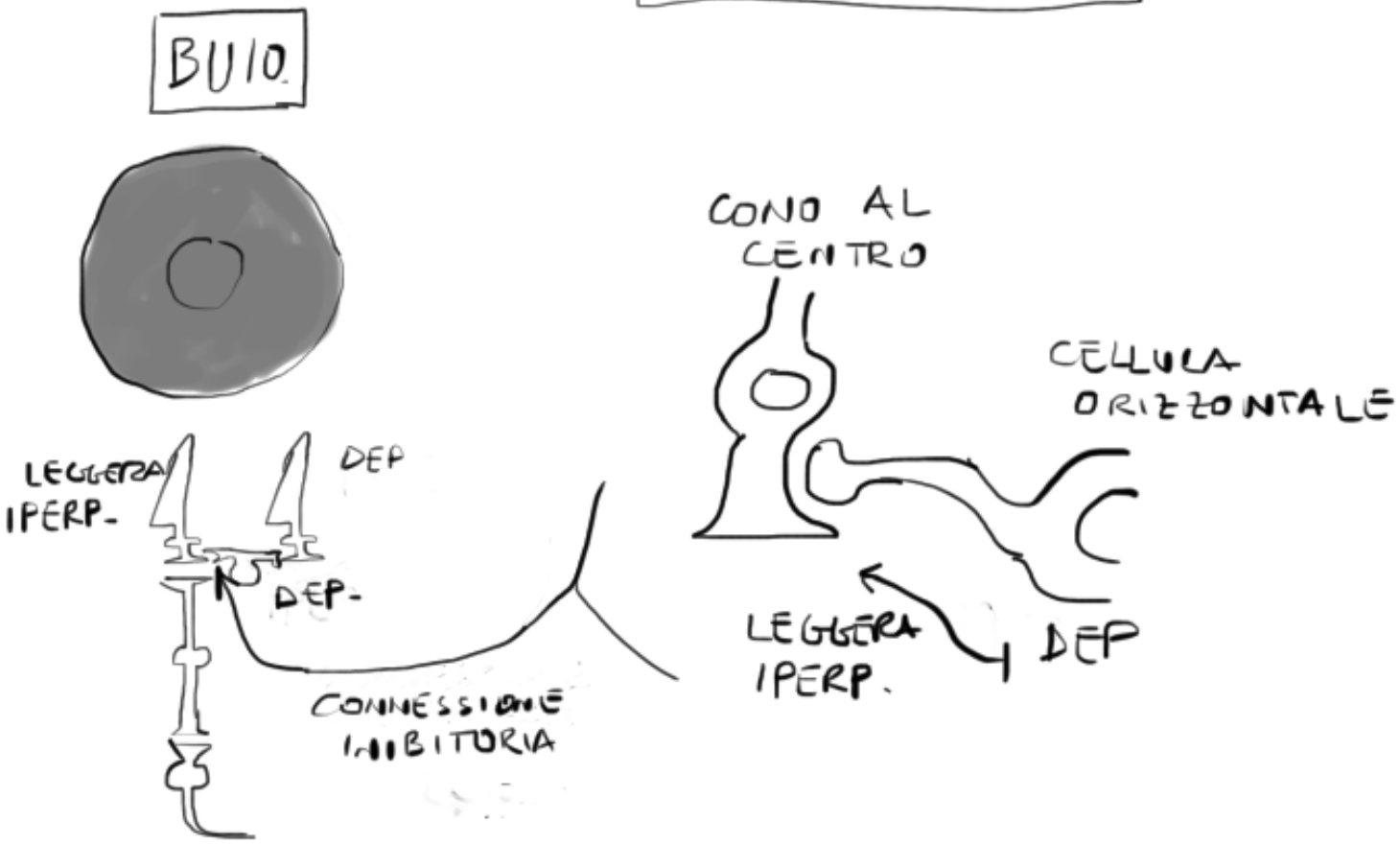


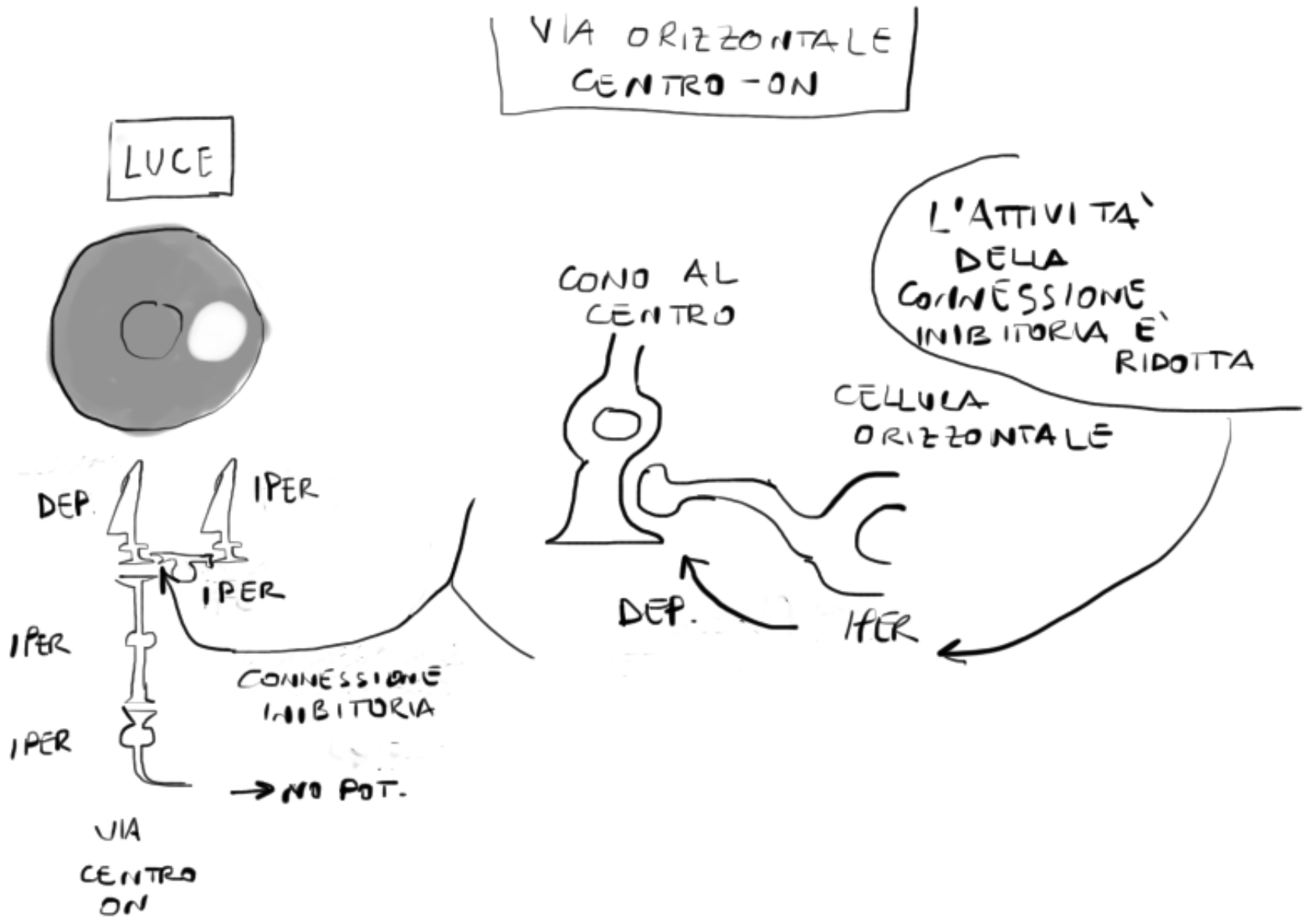


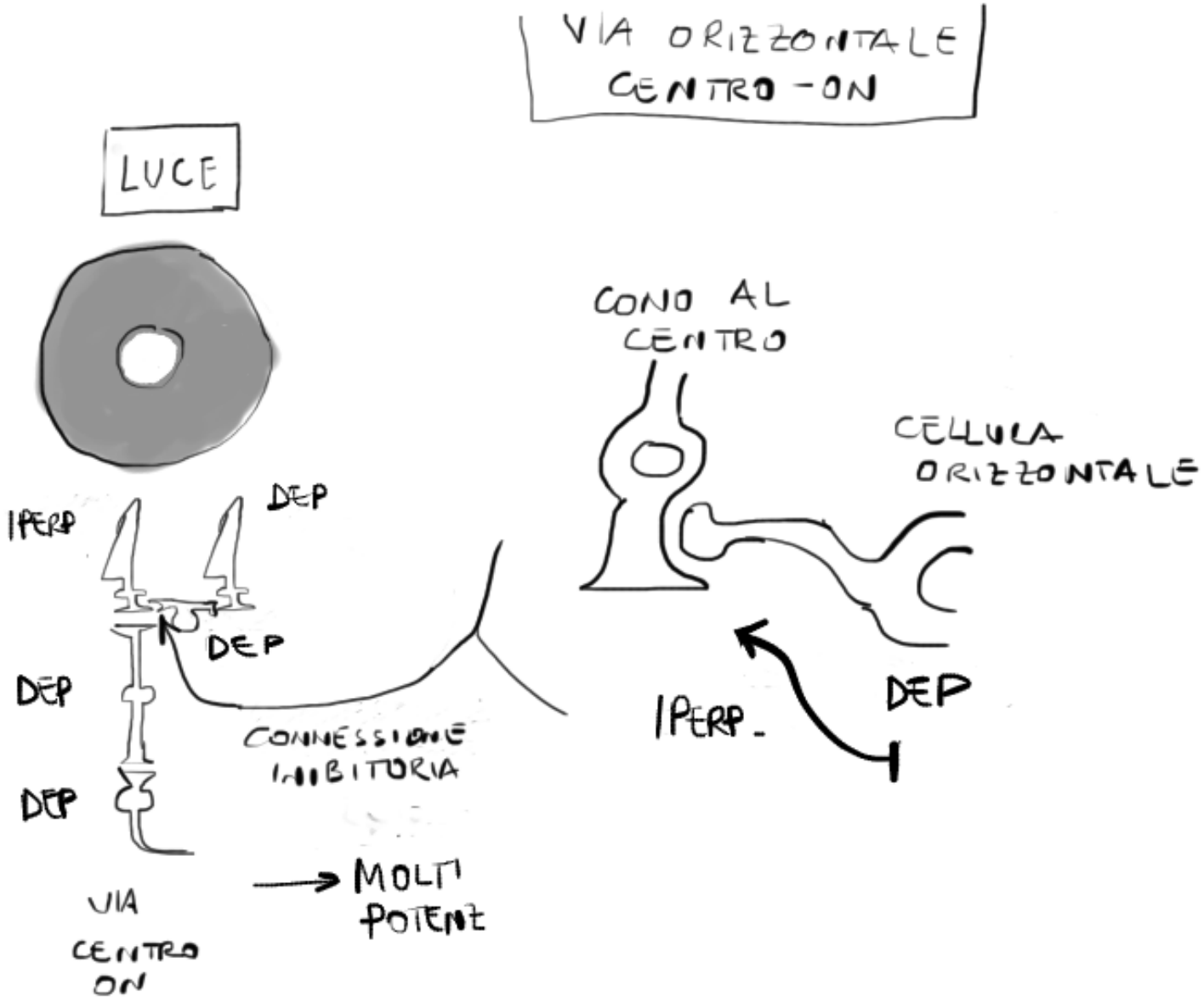




VIA ORIZZONTALE CENTRO-ON

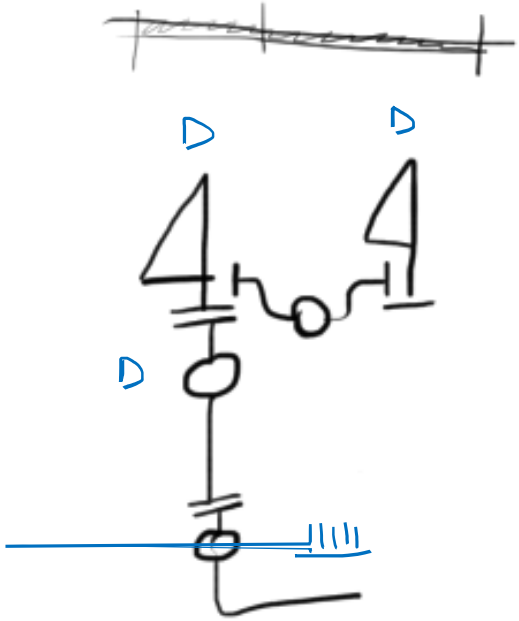



VIA ORIZZONTALE
CENTRO -ON
LUCE
L'ATTIVITA' DELLA CONNESSIONE INIBITORIA È RIDOTTA
CONO AL CENTRO
CELLULA ORIZZONTALE
DEP.
IPER
IPER
IPER
IPER
DEP.
IPER
CONNESSIONE INIBITORIA
→NO POT.
VIA CENTRO ON


VIA ORIZZONTALE
CENTRO -ON
LUCE
CONO AL CENTRO
CELLULA ORIZZONTALE
IPERP
DEP
DEP
DEP
DEP
DEP
IPERP
CONNESSIONE INIBITORIA
VIA CENTRO ON
MOLTI POTENT

VIA CENTRO OFF

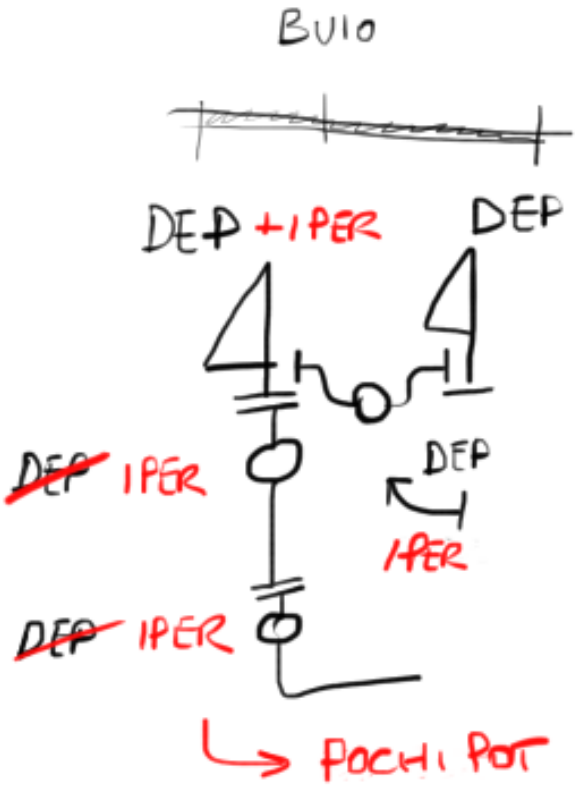
BUIO

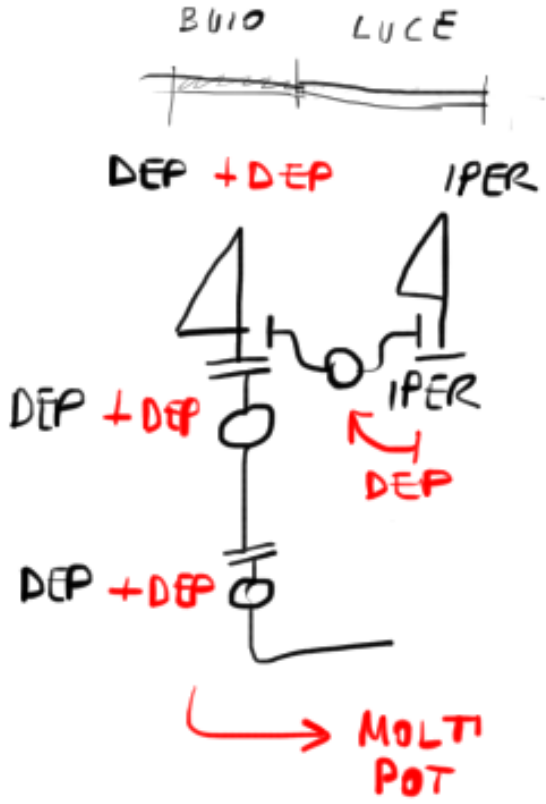


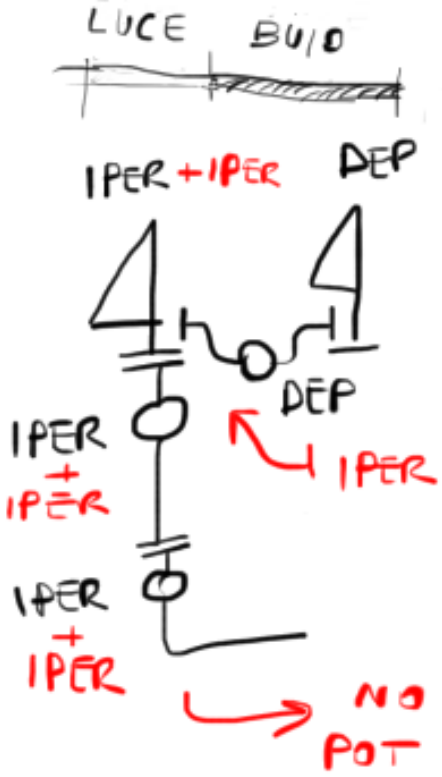
BUIO LUCE

LUCE BUIO

VIA CENTRO OFF



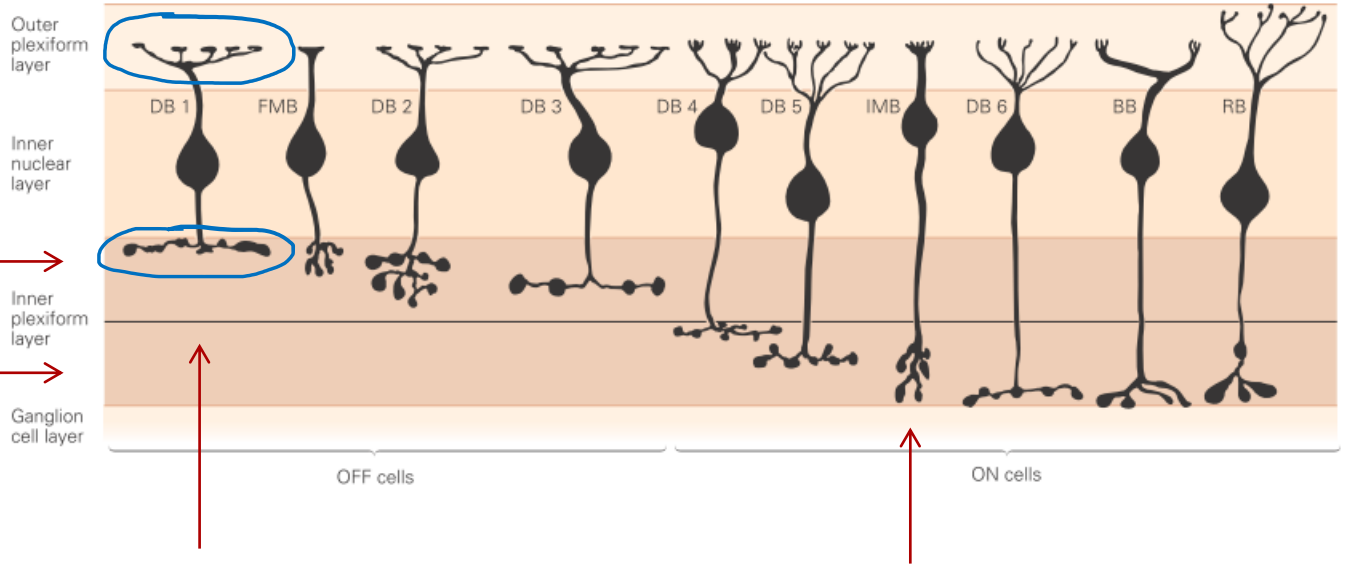




Morphology of dendritic tree and termination make a further distinction among bipolar cells

distal

proximal

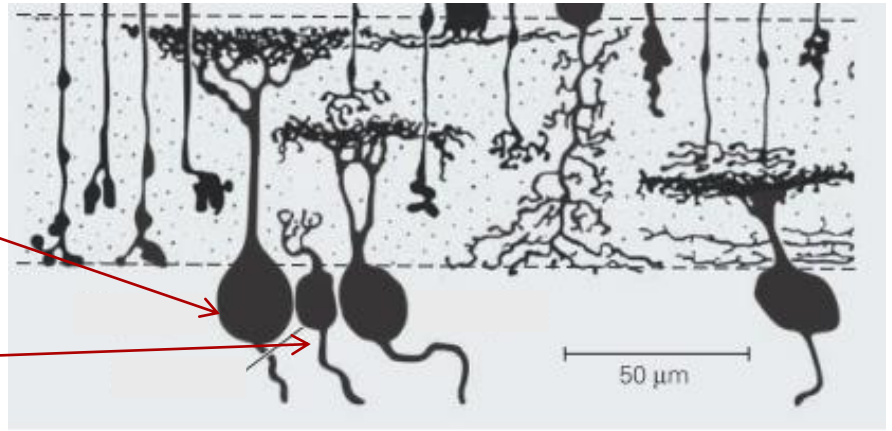


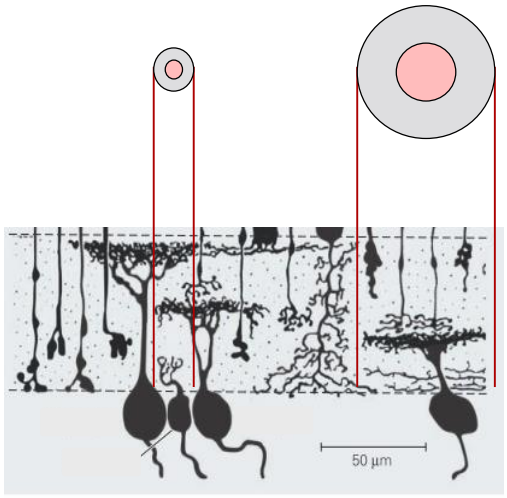
Diffuse bipolar cell: more than one cone links to the bipolar cell, which eventually connects to M ganglion cell

Midget (nane) bipolar: one cone with a P cell

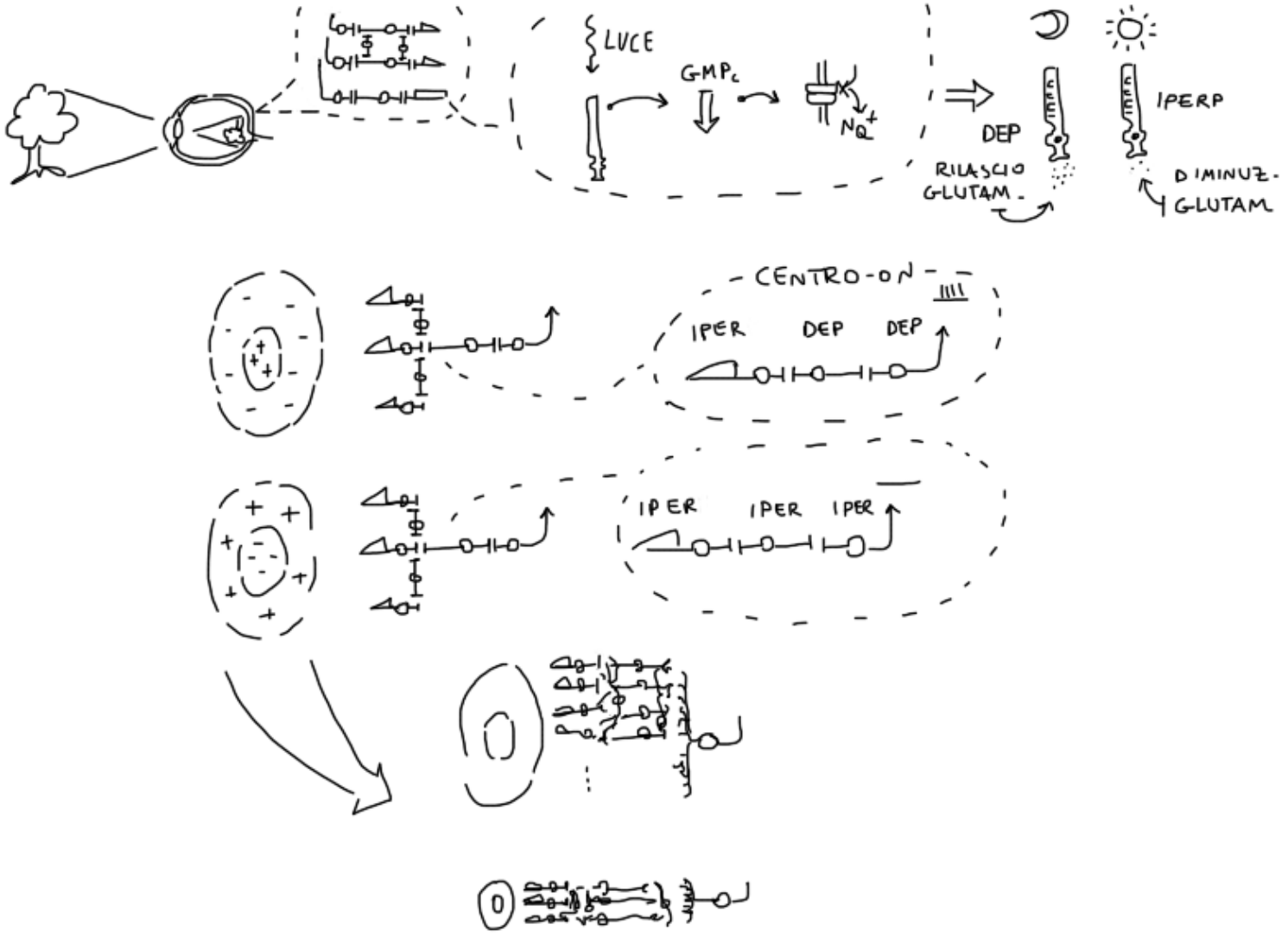
Another difference is related to the size of the ganglion cells, **M cells** (magnae, big) and **P cells** (parvae, small).

With have either centre-on and centre-off cells

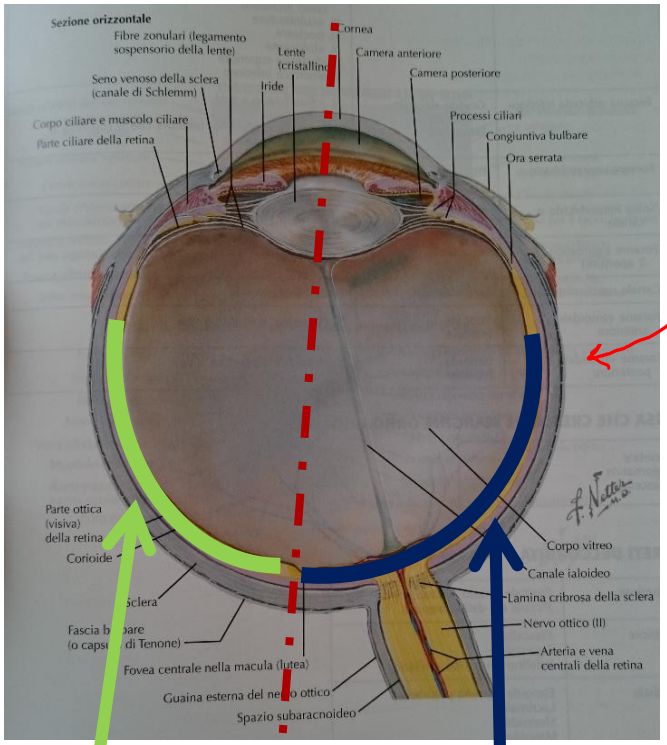




Receptive fields are different for M and P cell. They are big for M cell, and small for P.


LUCE
GMPc
Na+
DEP
RILASCIO GLUTAM.
IPERP
DIMINUZ. GLUTAM
CENTRO-ON
IPER DEP DEP
IPER IPER IPER
0

Meridian plane

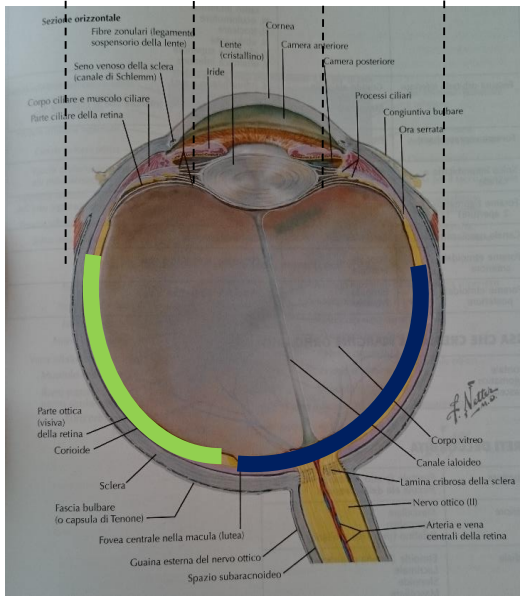


Temporal side

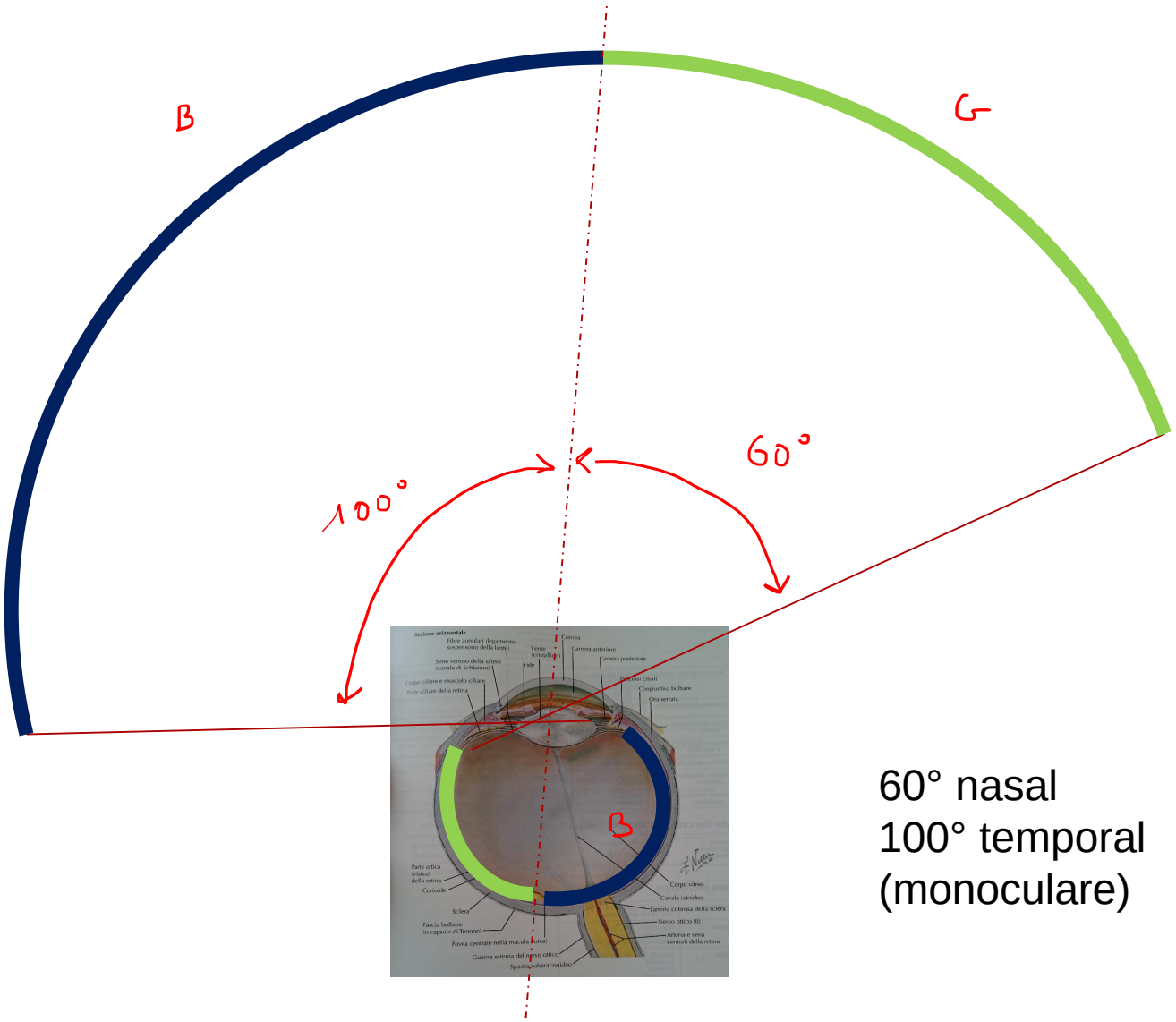
Nasal side

Dorsal side (top)

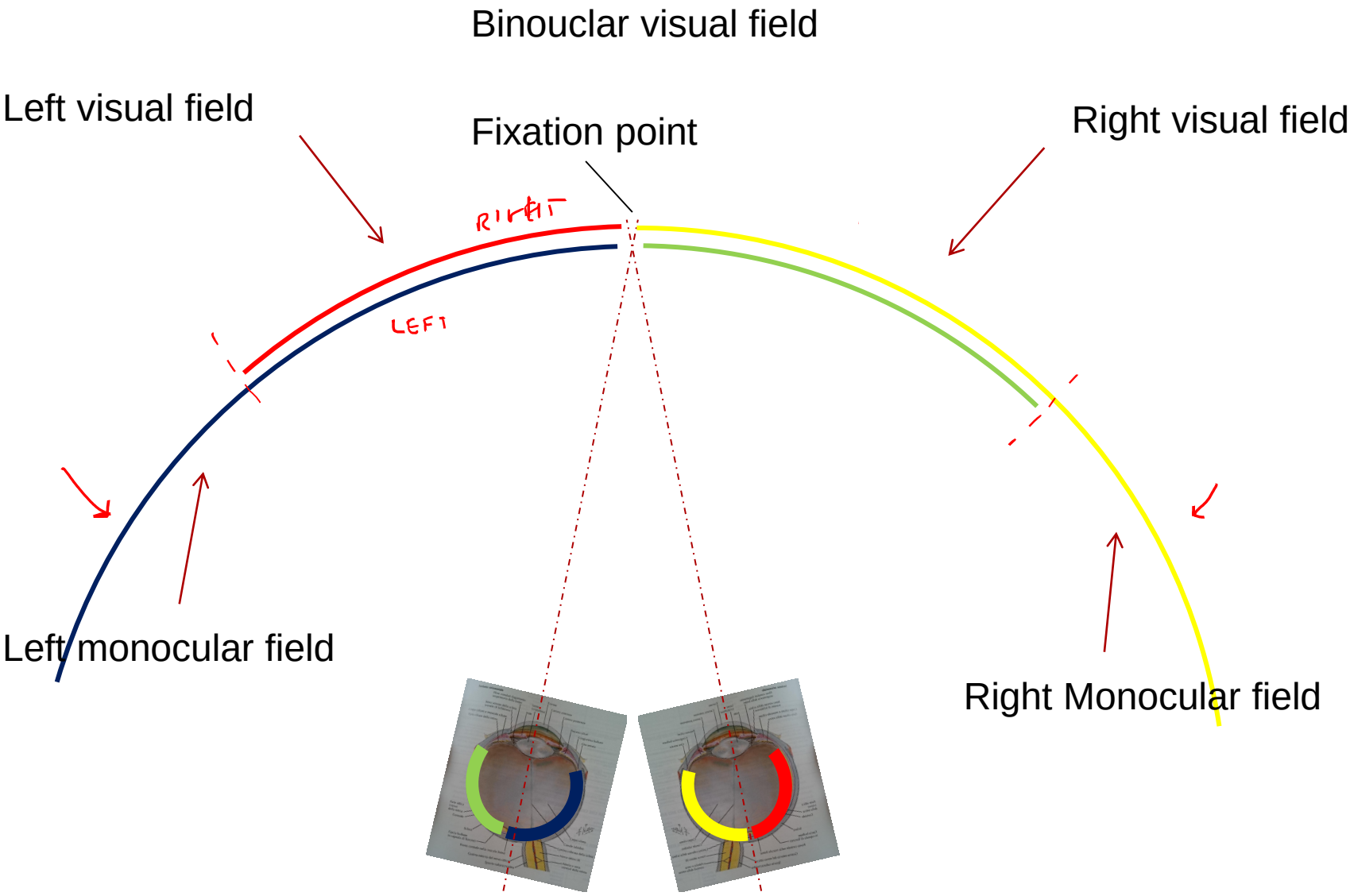
Ventral side (bottom)



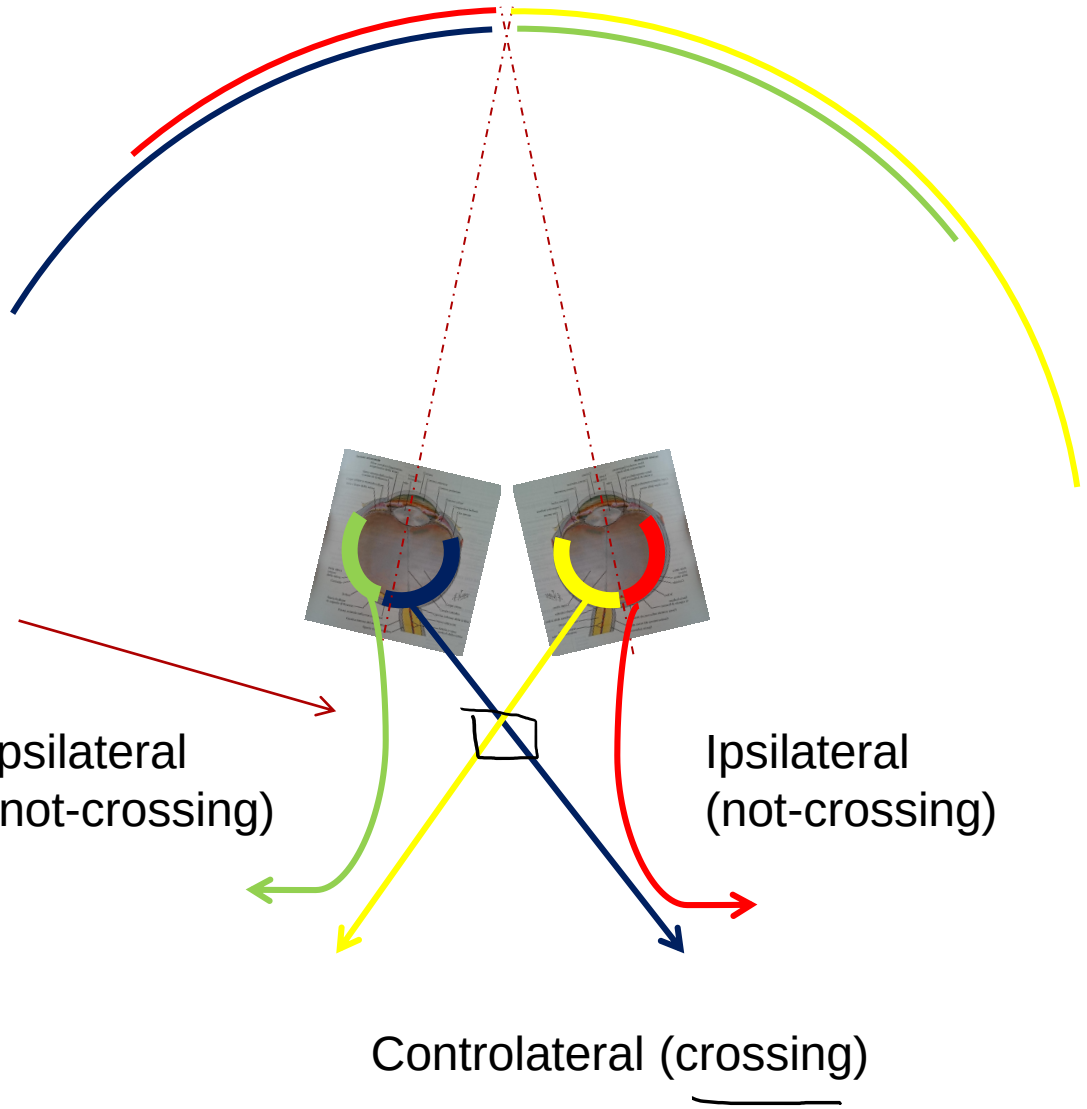
Visual field

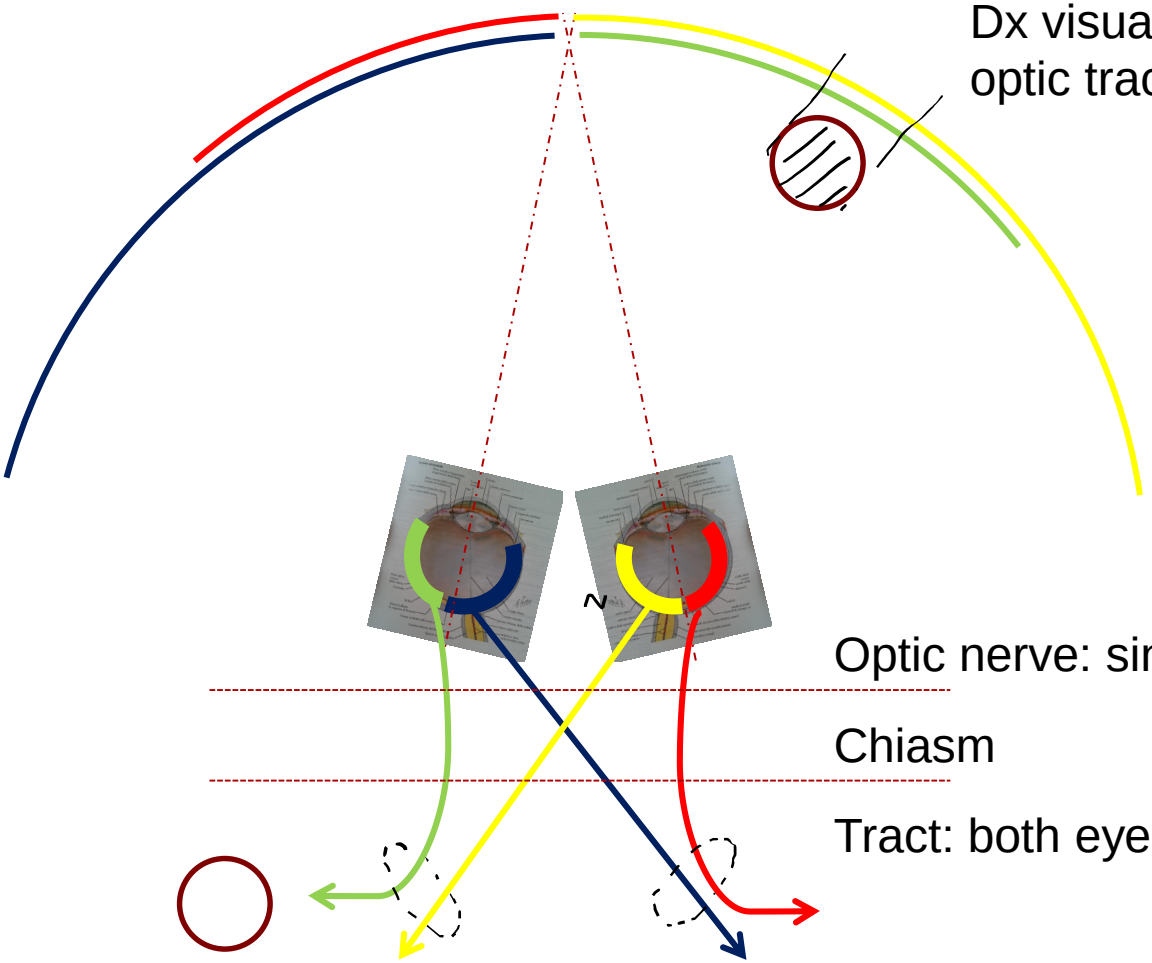


60° nasal
100° temporal
(monoculare)


Binouclar visual field
Left visual field
Fixation point
Right visual field
RIGHT
LEFT
Left monocular field
Right Monocular field

Crossing of visual pathways





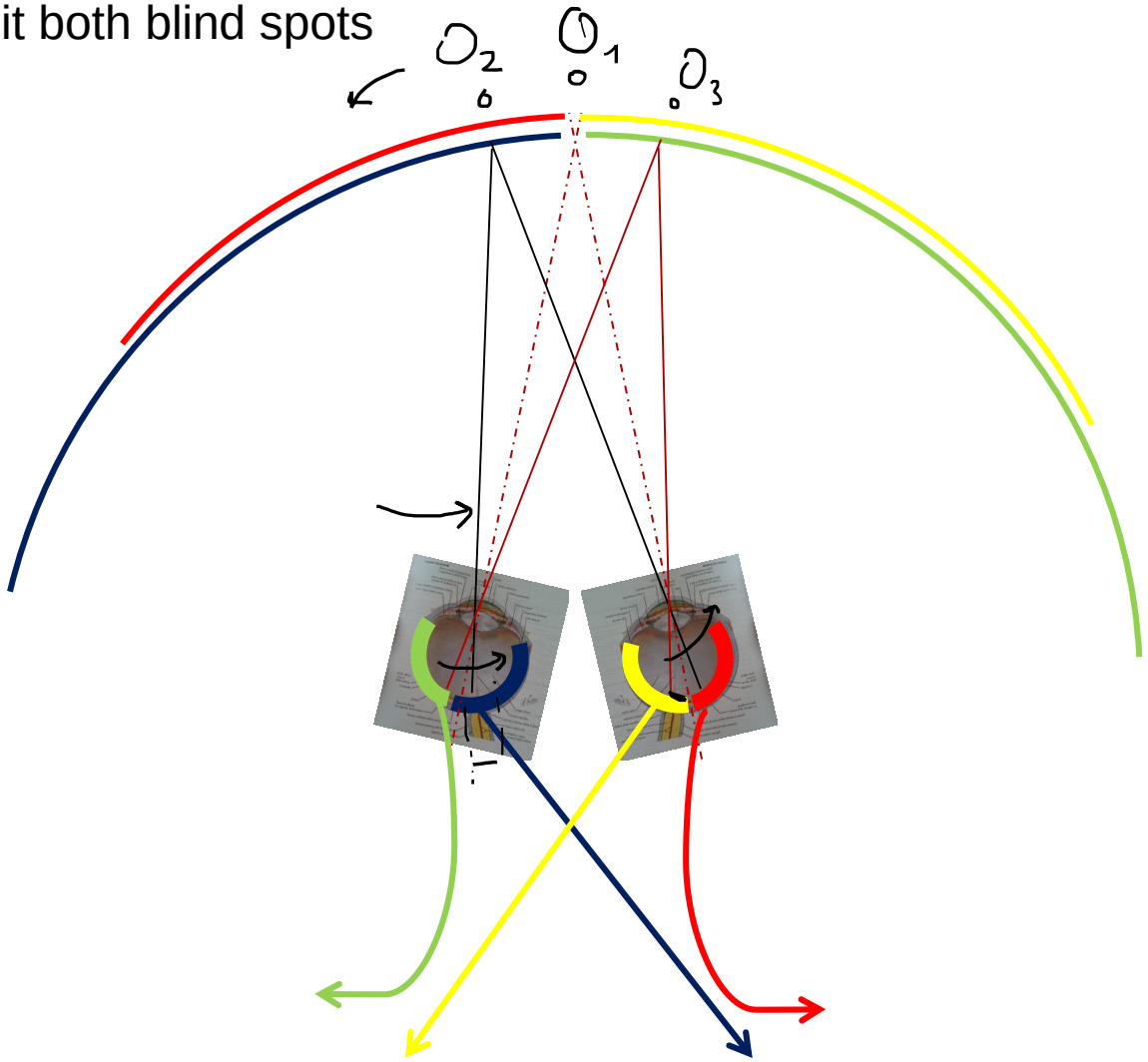
Dx visual field is projected to left optic tract

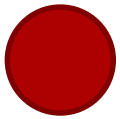
Optic nerve: single eye information

Chiasm

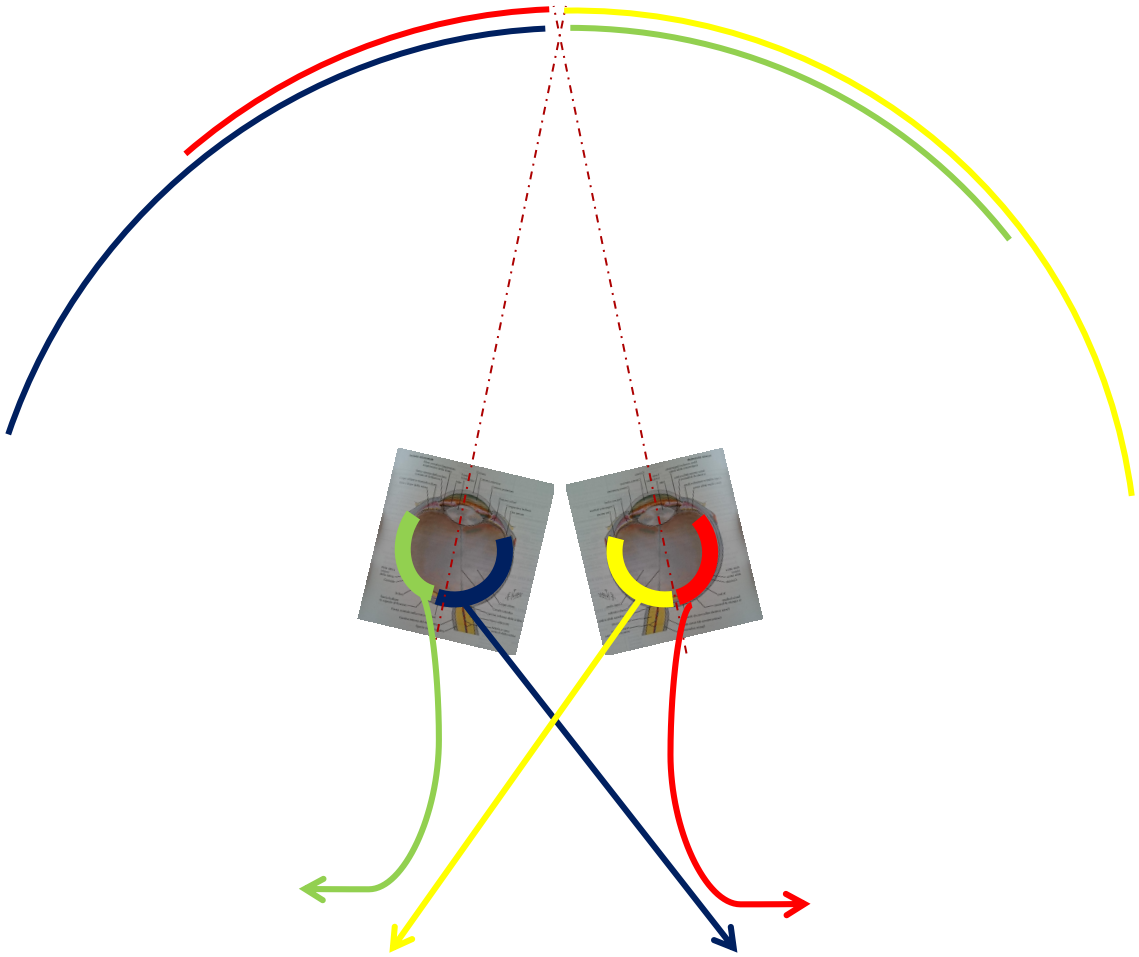
Tract: both eyes information

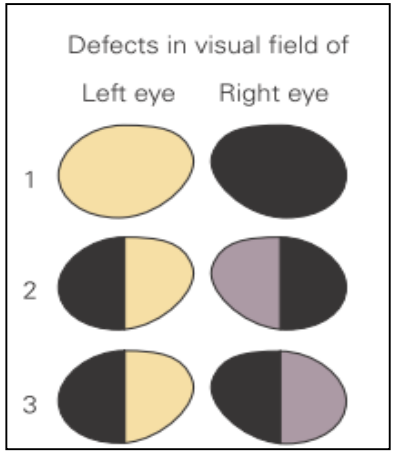
The optical disk is in the nasal part with respect to the fovea: within the binocular view, it will never hit both blind spots

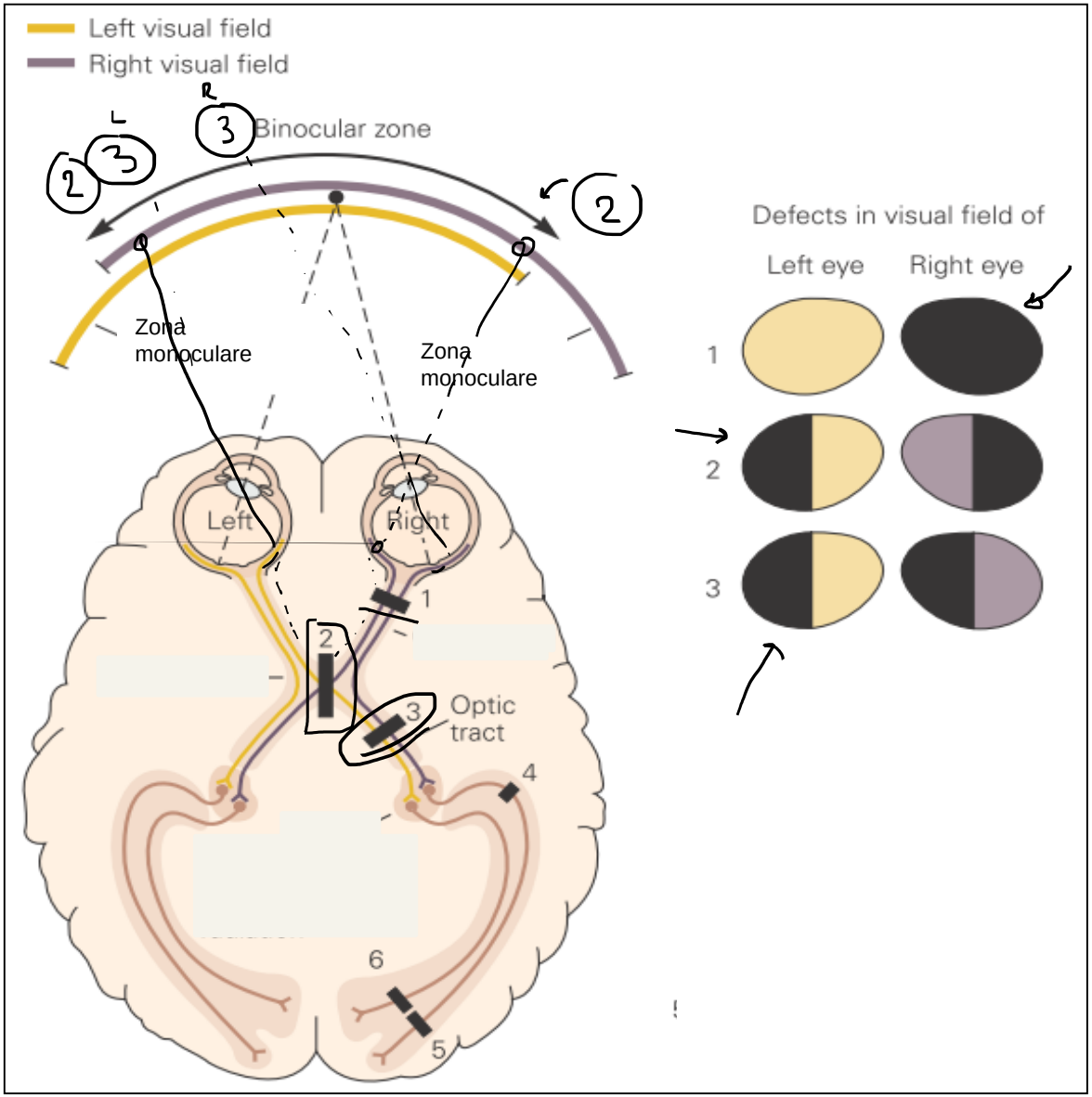




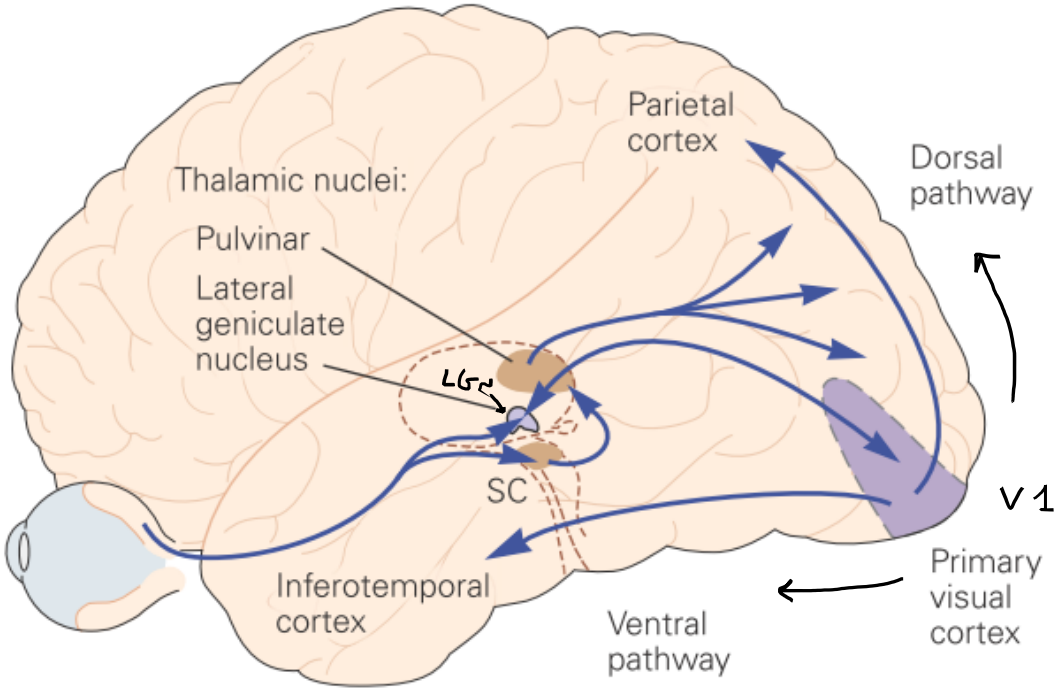
Where are the damages to the visual pathway?






Left visual field
Right visual field
Binocular zone
Zona monoculare
Zona monoculare
Left
Right
1
2
3
Optic tract
4
5
6
Defects in visual field of
Left eye
Right eye
1
2
3

About 90% of axons from the retina reach the lateral geniculate nucleus (LGN)



→ Withouth LGN, visual perception is almost lost. The residual perception is called: blind vision

LGN is preserving the topology of the retin, but with different relative spatial distribution

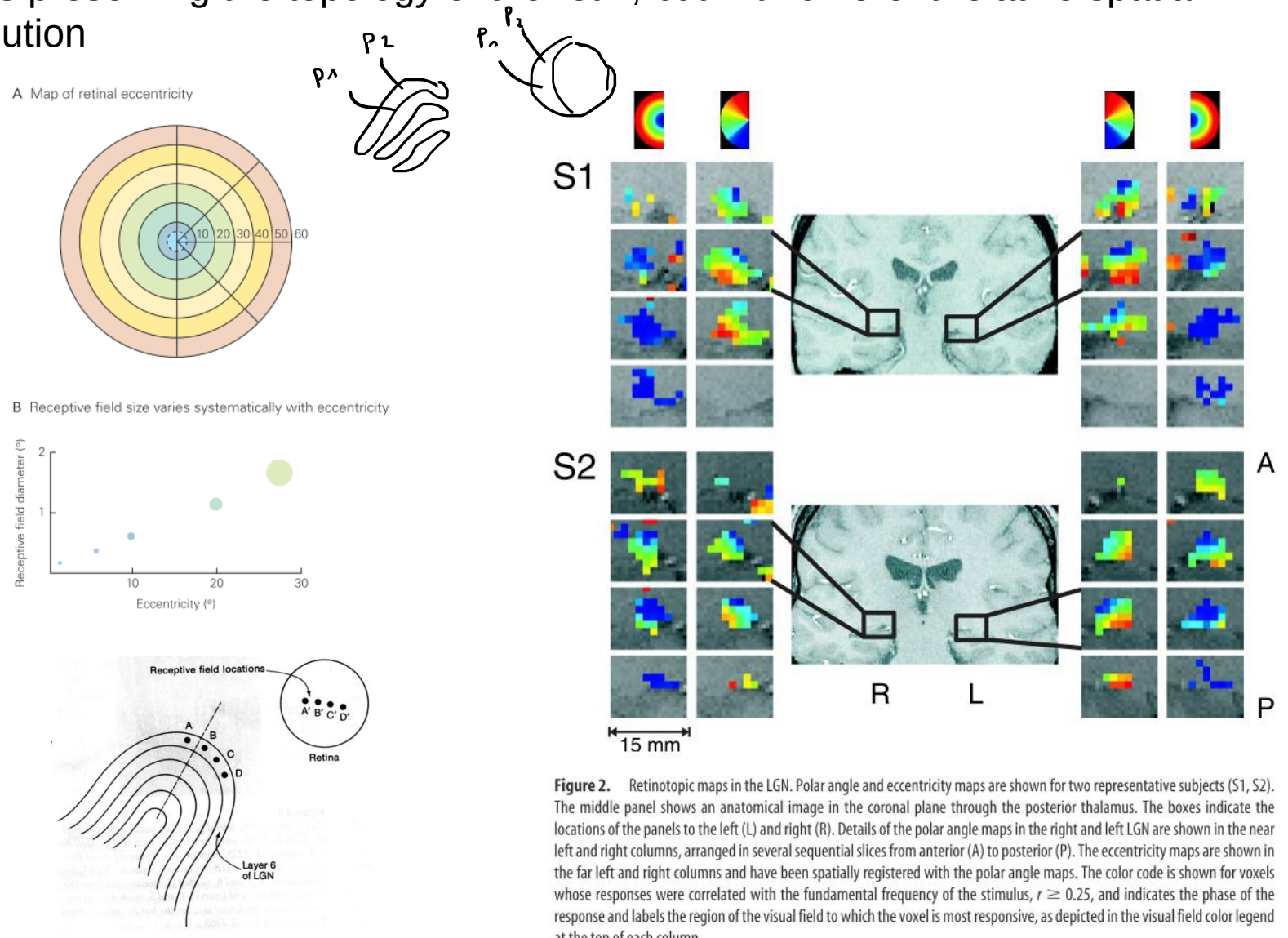
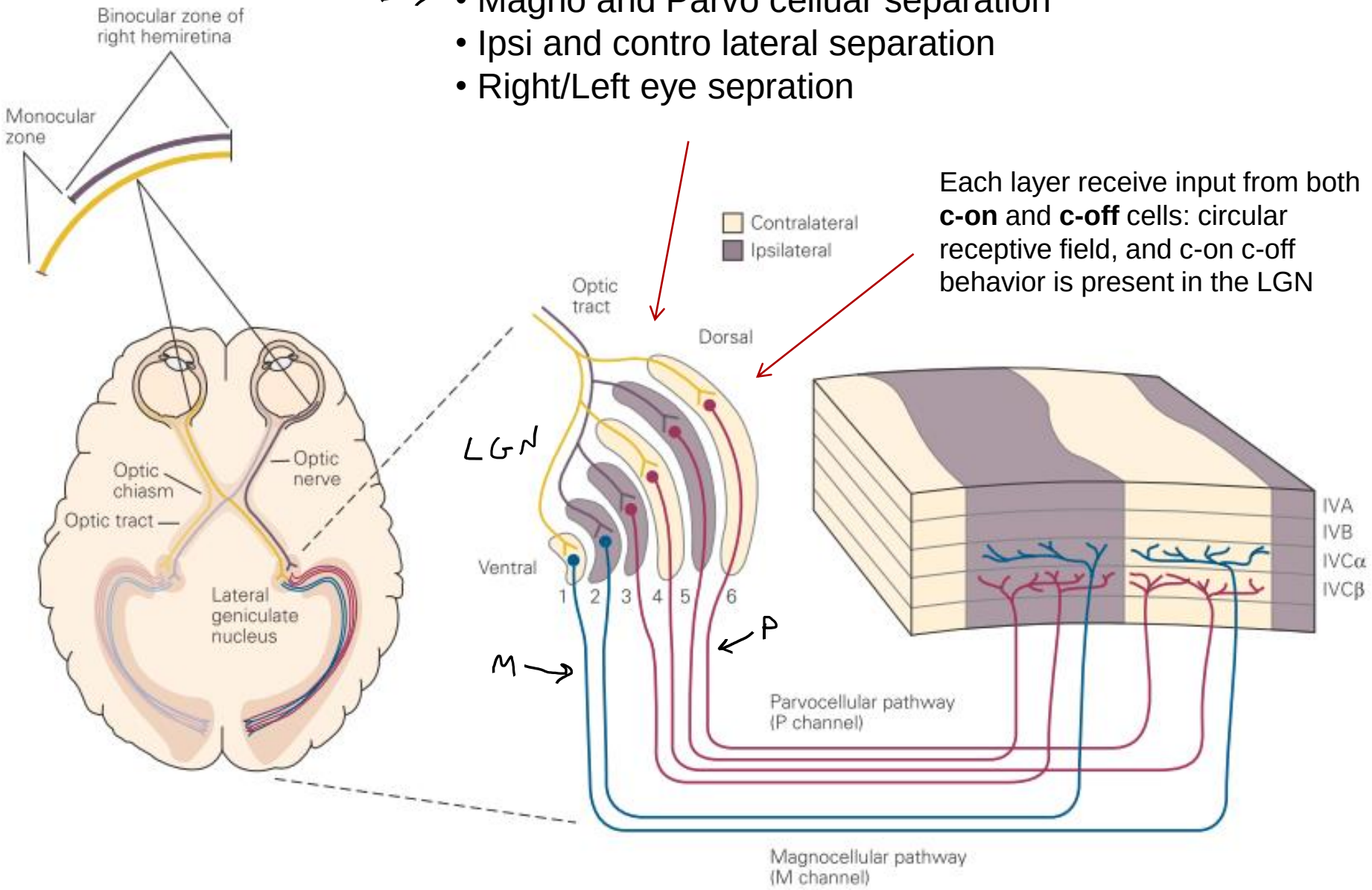


Figure 2. Retinotopic maps in the LGN. Polar angle and eccentricity maps are shown for two representative subjects (S1, S2). The middle panel shows an anatomical image in the coronal plane through the posterior thalamus. The boxes indicate the locations of the panels to the left (L) and right (R). Details of the polar angle maps in the right and left LGN are shown in the near left and right columns, arranged in several sequential slices from anterior (A) to posterior (P). The eccentricity maps are shown in the far left and right columns and have been spatially registered with the polar angle maps. The color code is shown for voxels whose responses were correlated with the fundamental frequency of the stimulus, $r \geq 0.25$, and indicates the phase of the response and labels the region of the visual field to which the voxel is most responsive, as depicted in the visual field color legend at the top of each column.

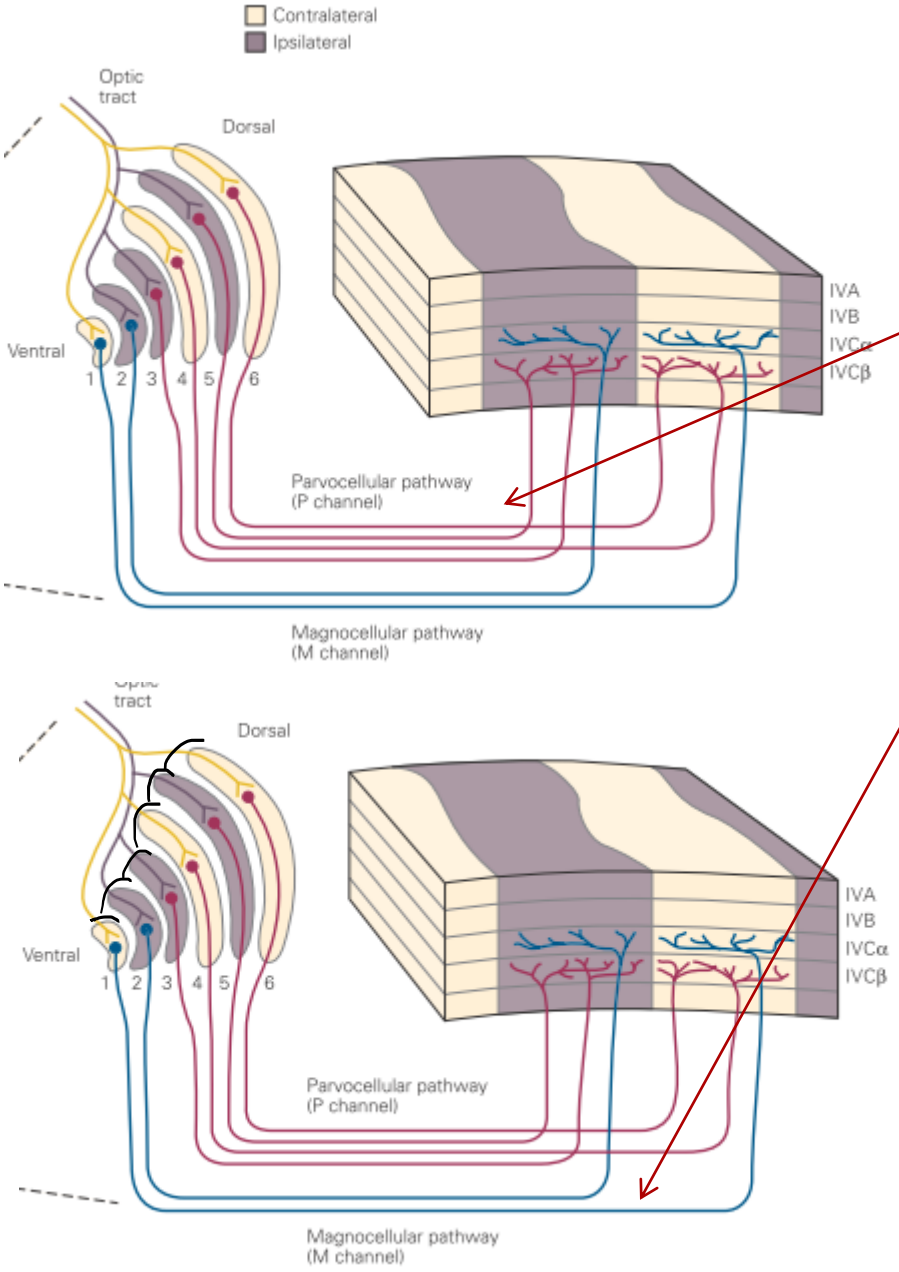
LGN in humans

→ Six layers, which preserve:

→ • Magno and Parvo celluar separation
• Ipsi and contro lateral separation
• Right/Left eye sepration



Each layer receive input from both **c-on** and **c-off** cells: circular receptive field, and c-on c-off behavior is present in the LGN

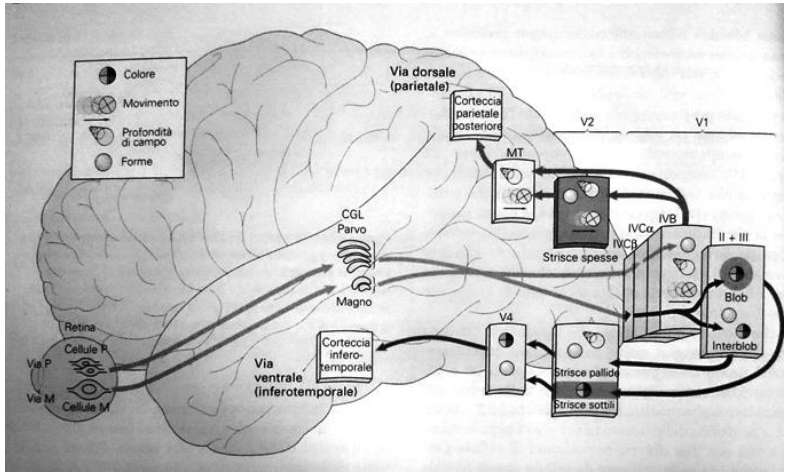


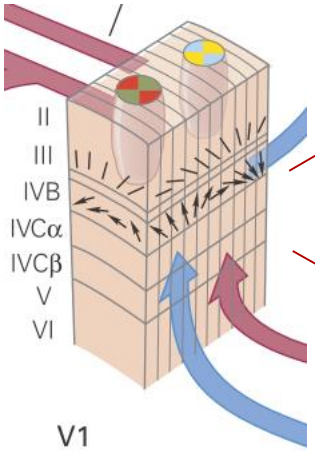
P and M channels are different from anatomical and functional point of views:

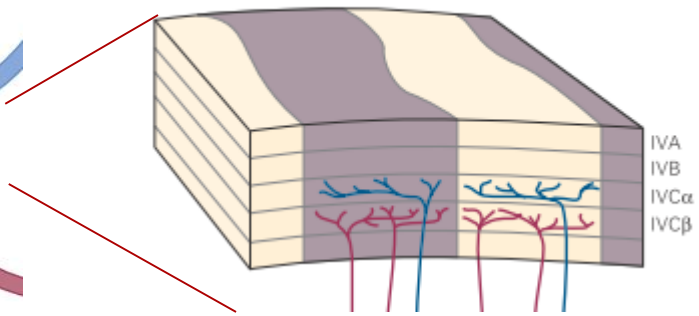
- Color contrast
- Spatial frequency

- Brightness contrast
- Temporal frequency

- ❑ P-cells (parvocellular)
  - Small medium sized cell body
  - Reaches layers 3,4,5,6
  - Responsible for color, fine textures, patterns and details vision
- ❑ M-cells (magnocellular)
  - Larger cell bodies
  - Reaches layers 1,2
  - Responsible for motion detection
- ❑ K-cells (koniocellular)
  - Largest cell bodies
  - Reaches all the six layers

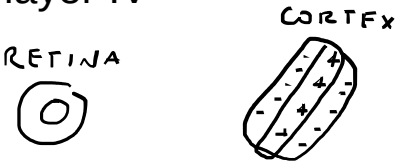






From LGN to visual cortex (V1)

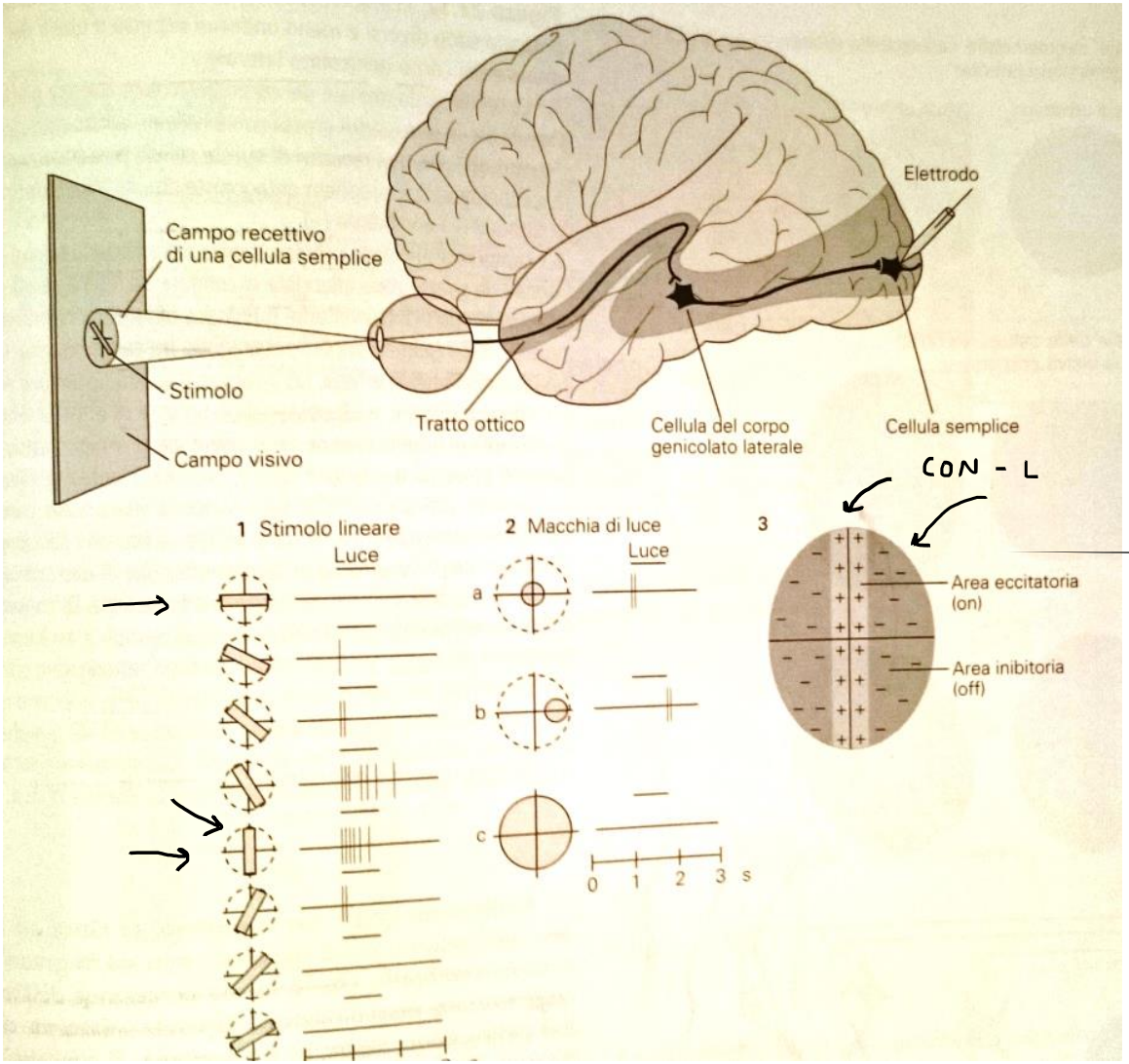
V1 is made of six layers (about 2 mm thick): fibers reach mostly layer IV



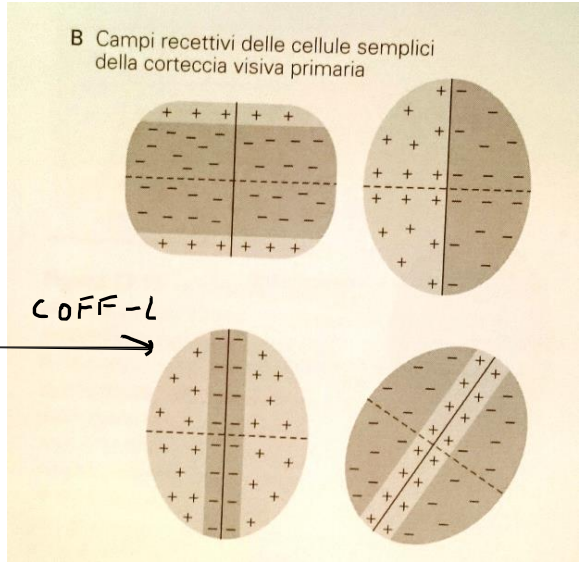
Change in receptive field shapes and function: from circular receptive fields to **linear receptive fields**.

Two different kind of cells increase their activity subject to linear stimulations: **simple** cells and **complex** cells

Simple cells: orientation selective

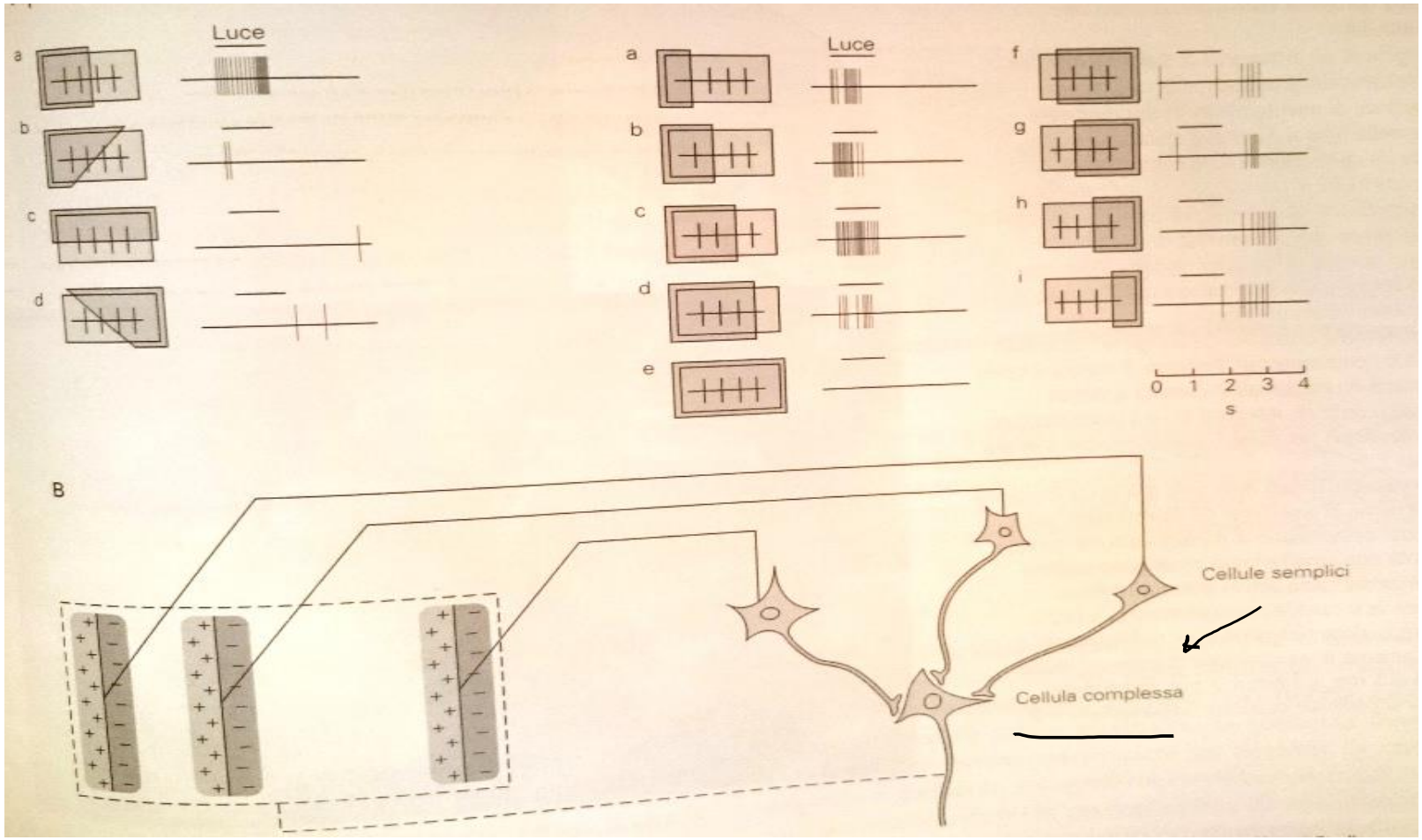


Simple cells when a **specific orientation is presented**

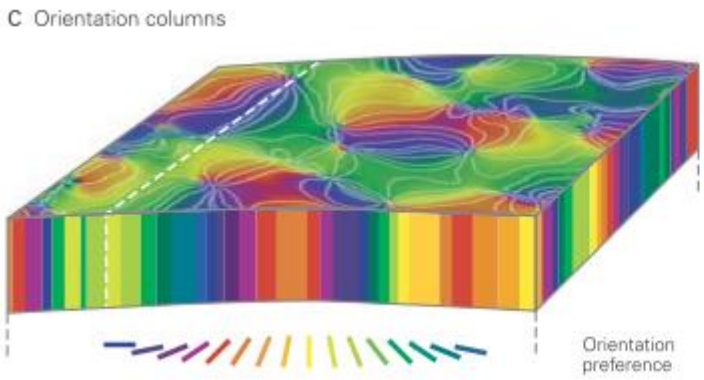


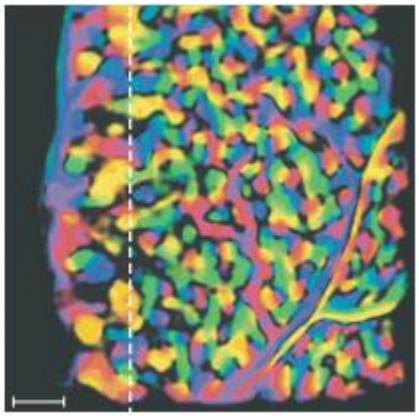
1. Right spot
2. Right shape
3. Right orientation

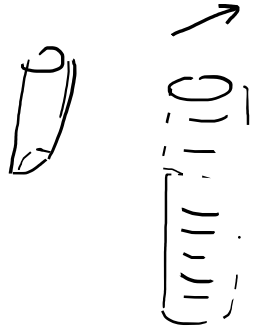
Receptive fields of **complex cells** are **orientation dependent** but **do not have antagonistic behavior (excitatory or inhibitory) neither position dependence**



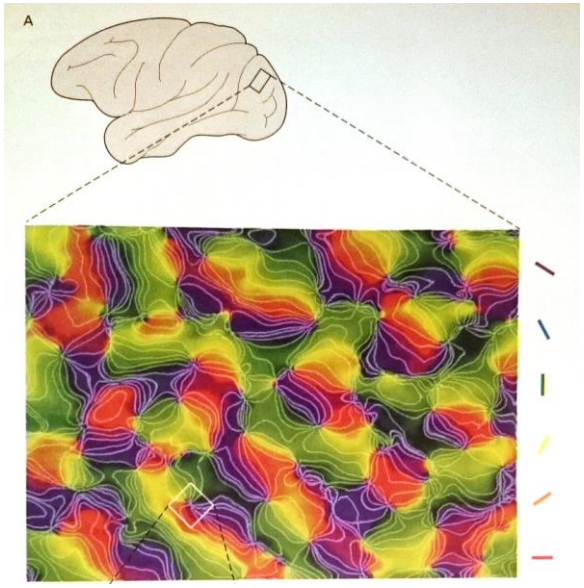
Cells in cortex (simple or complex) are aggregated in **orientation columns**

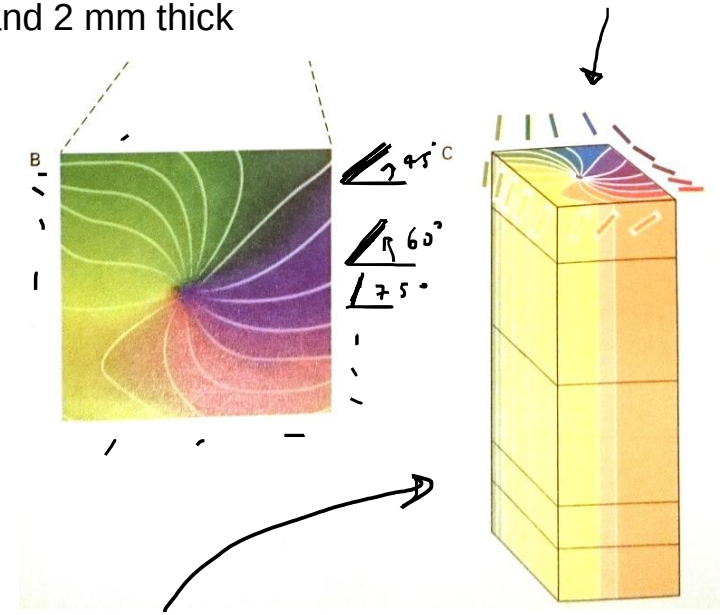






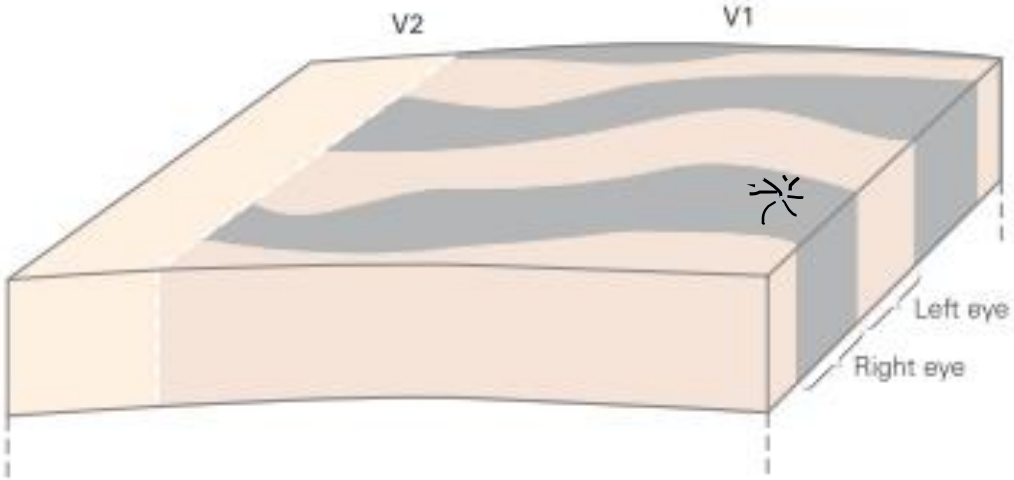
30-100 micrometres width, and 2 mm thick

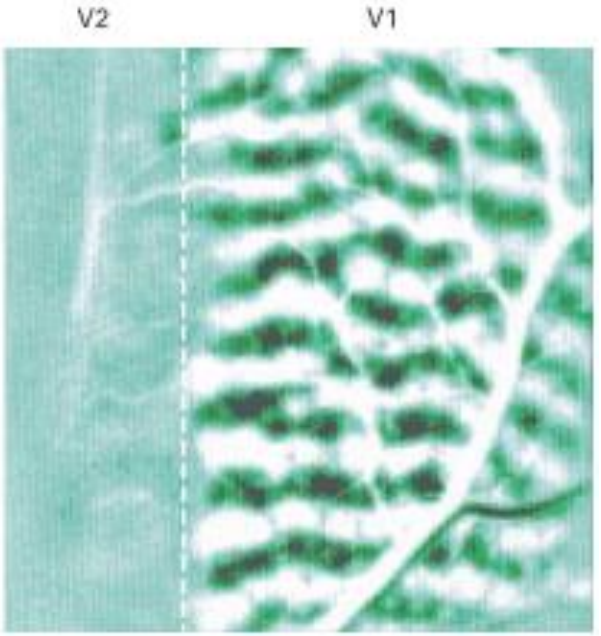




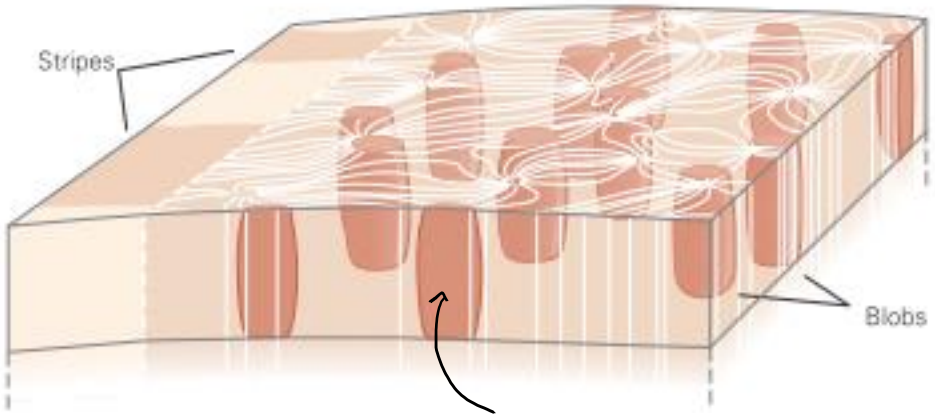
Hyper coloumn: all orientations are represented

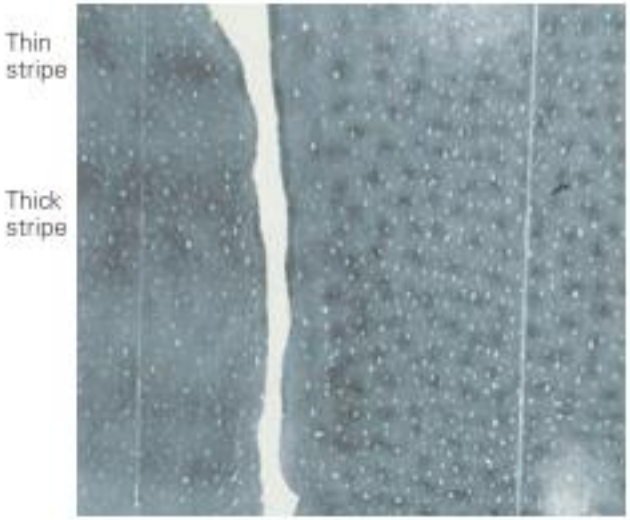
B Ocular dominance columns

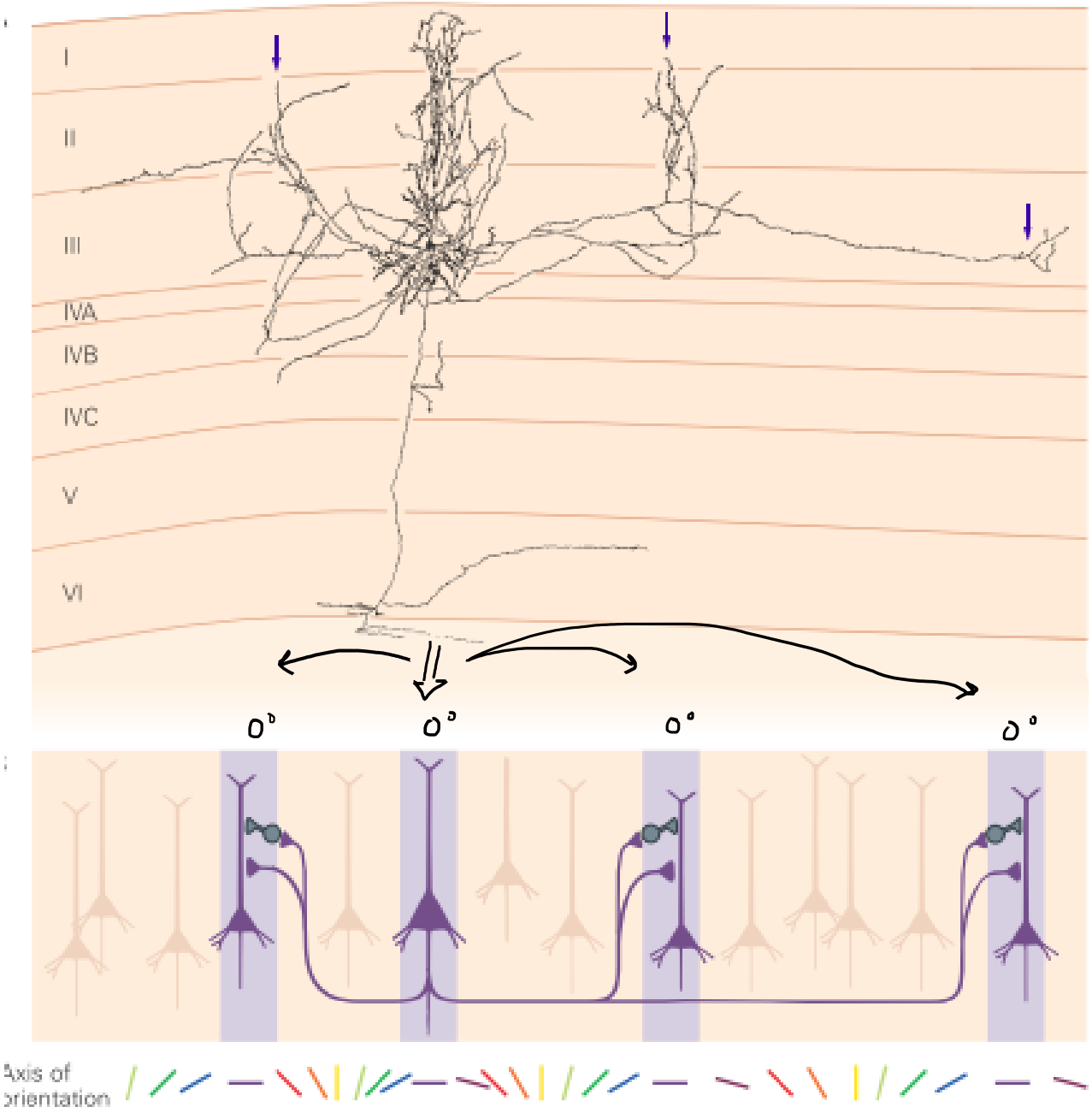


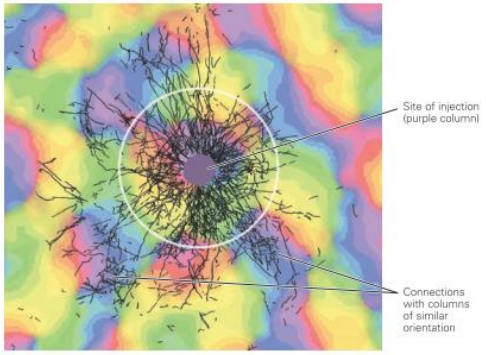


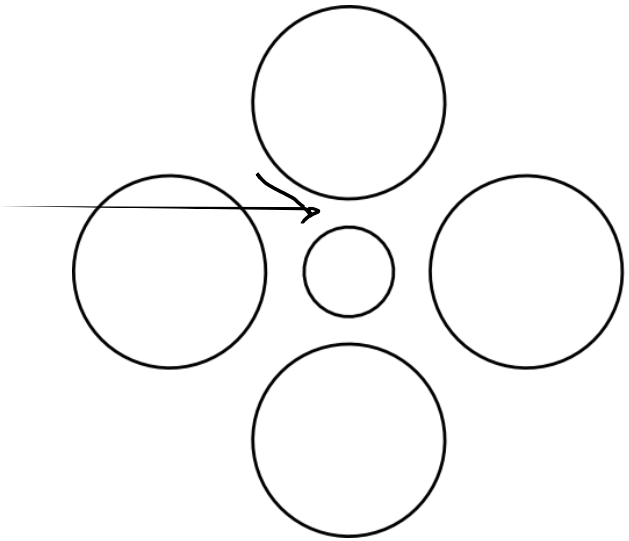
D Blobs, interblobs (V1), and stripes (V2)

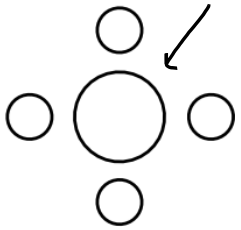


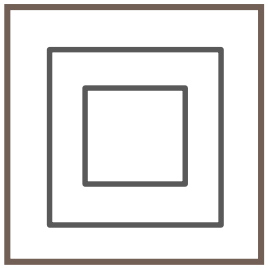


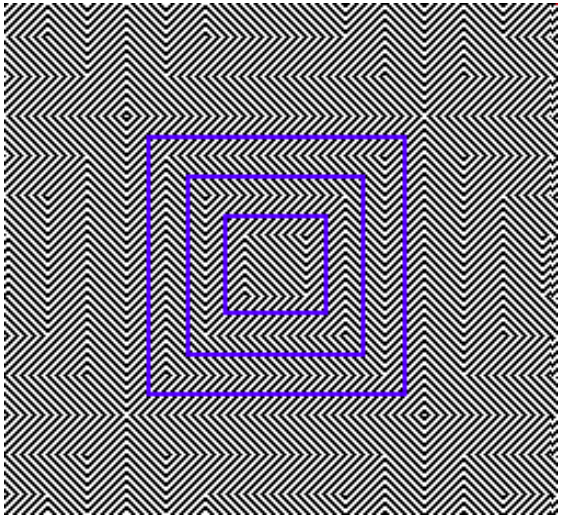

I
II
III
IVA
IVB
IVC
V
VI
0°
0°
0°
0°
Axis of orientation


Site of injection (purple column)
Connections with columns of similar orientation

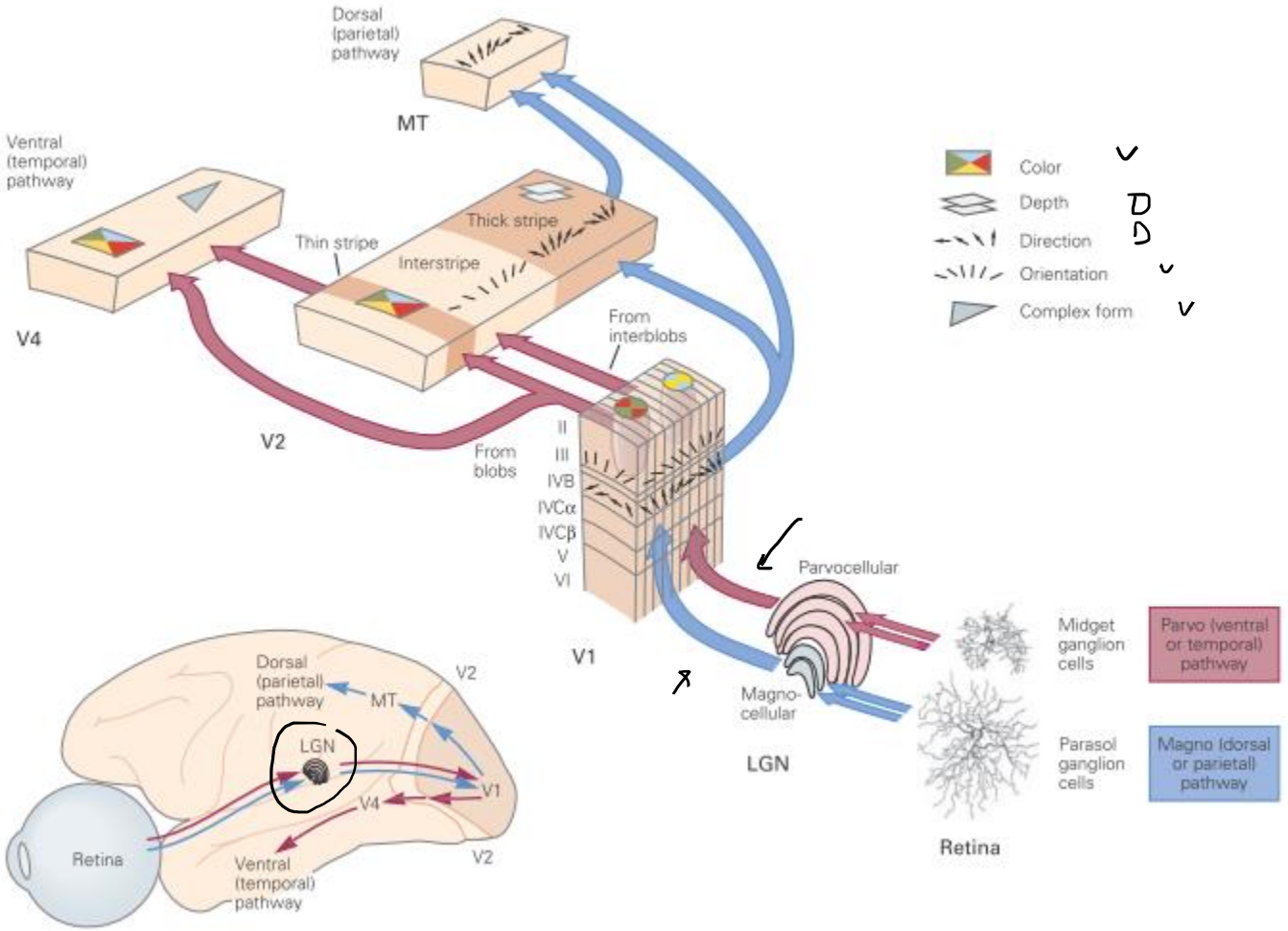




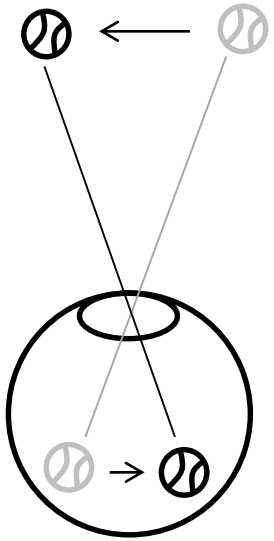




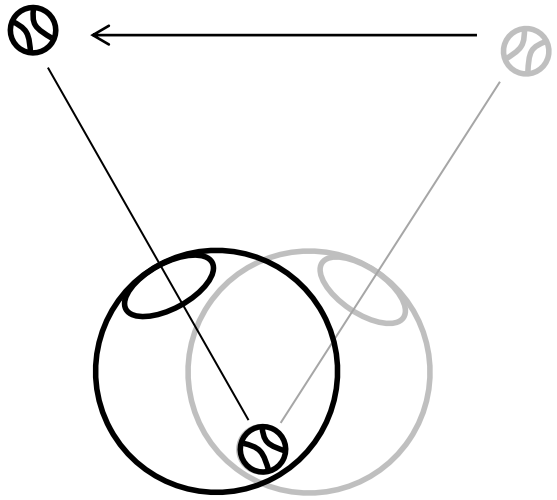
Dorsal pathway: movement and depth



What is movement?

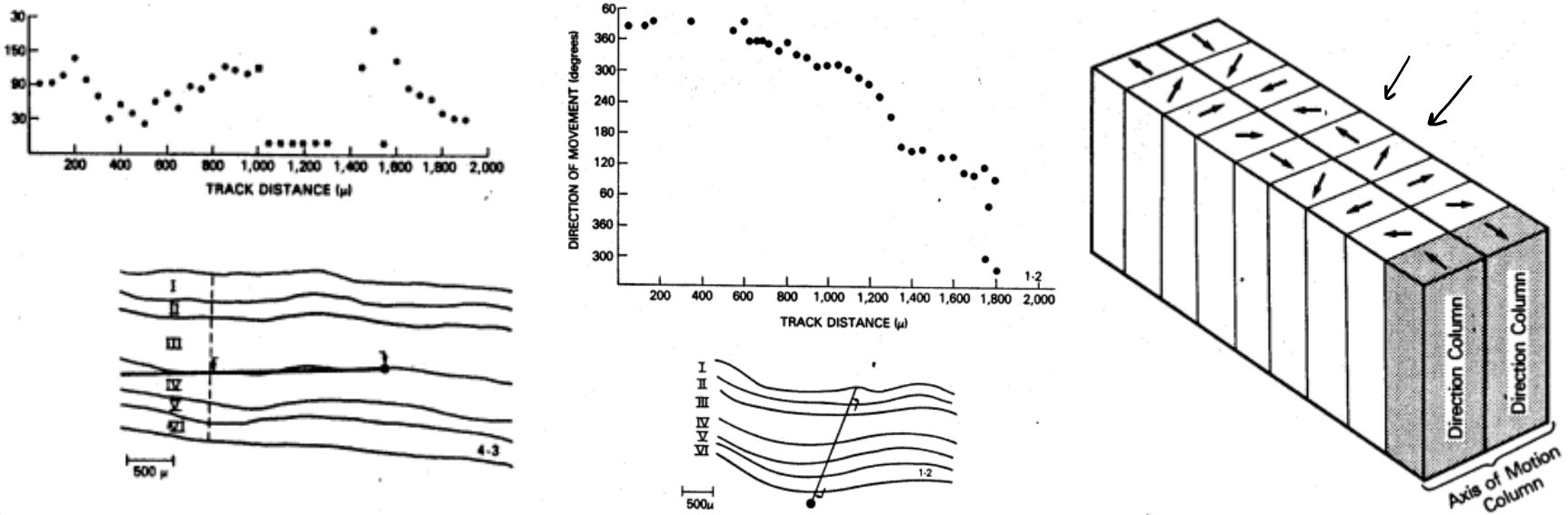


Effective translation of the image on the retina



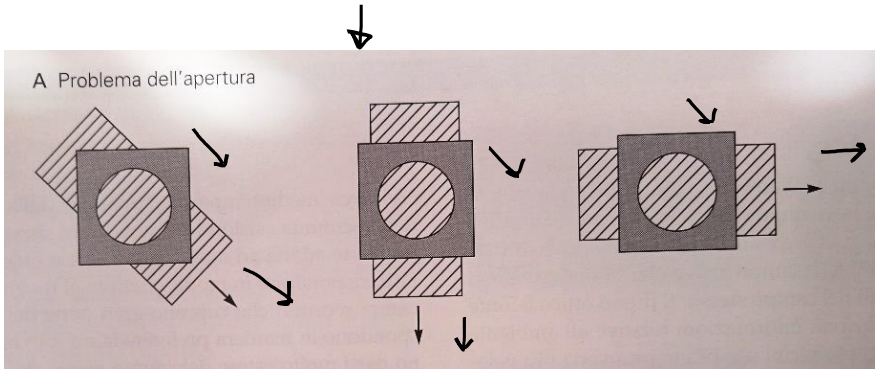
Eye movement while the image is still

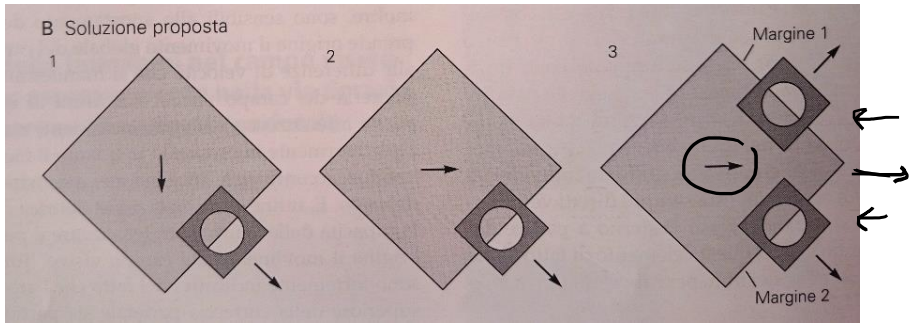
Within the Medio Temporal area we have **coloumns** to detect direction of the movements



J Neurophysiol. 1984 Jan;51(1):16-31. **Columnar organization of directionally selective cells in visual area MT of the macaque.** Albright TD, Desimone R, Gross CG.

Only oriented movements is not enough to detect motions of complex objects: **aperture problem**

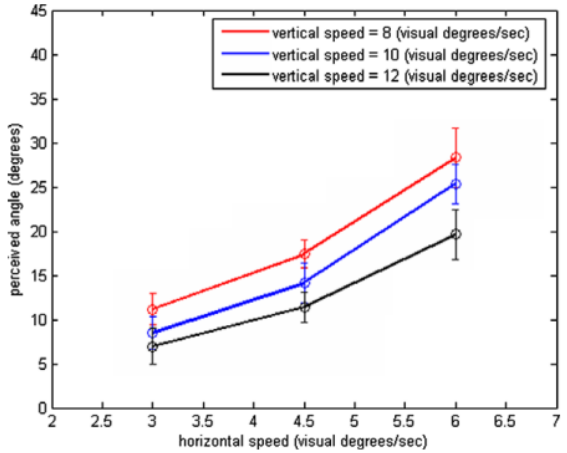




How to trick your movement systems

Motion Integration Unleashed New Tricks for an Old Dog

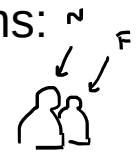
https://www.youtube.com/watch?v=Jri0del_6t4



Tse and Hsieh, *The infinite regress illusion reveals faulty integration of local and global motion signals* (2006), Vision Research (46) 3881–3885

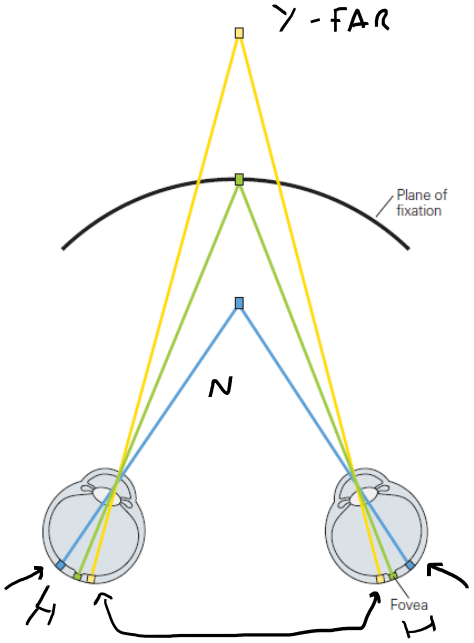
Several clues are used to obtain depth information:

- Monocular mechanisms:
  - Familiarity
  - Occlusion
  - Perspective
  - Object size
  - Lights/shadows distribution
  - movements
  - accommodation

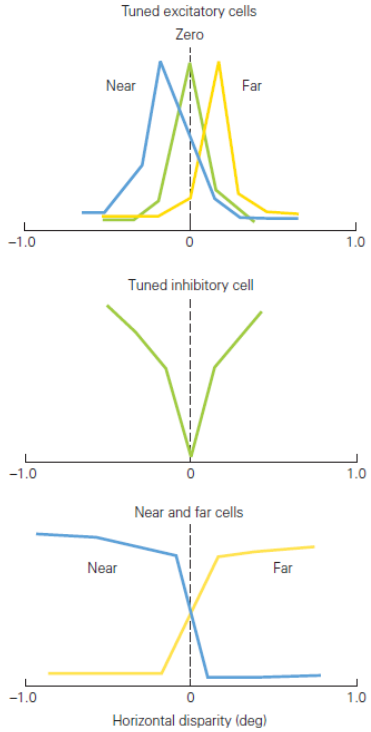


- Binocular mechanisms
  - disparity

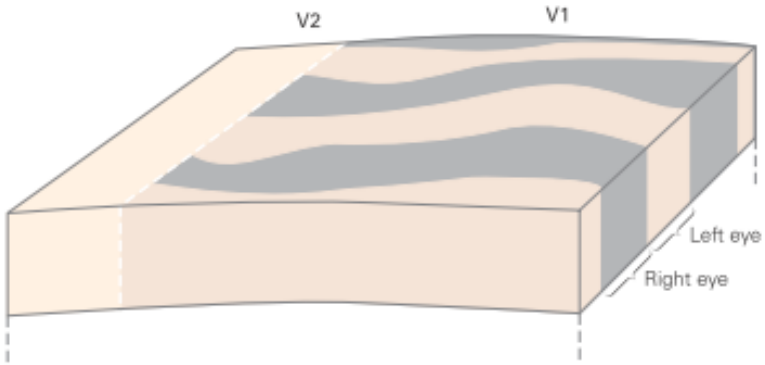
A Binocular disparity of retinal images



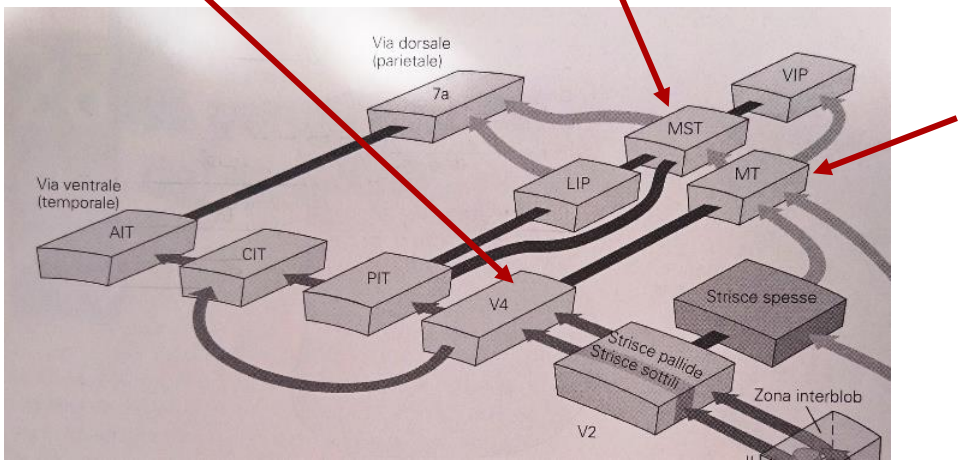
B Disparity-selective neurons



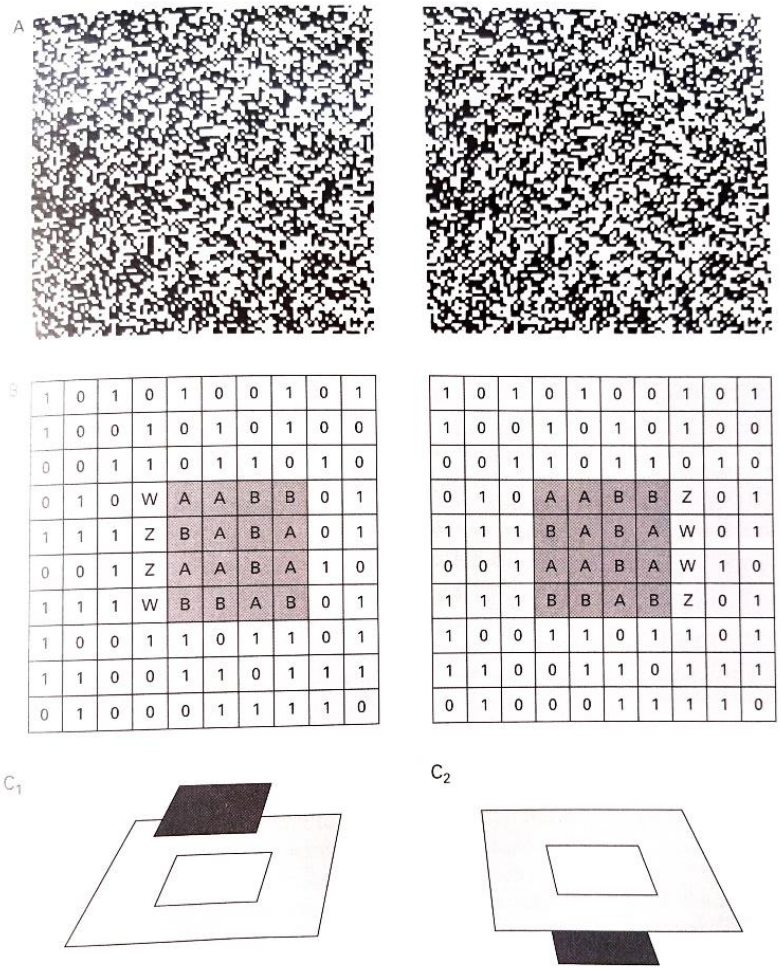
B Ocular dominance columns



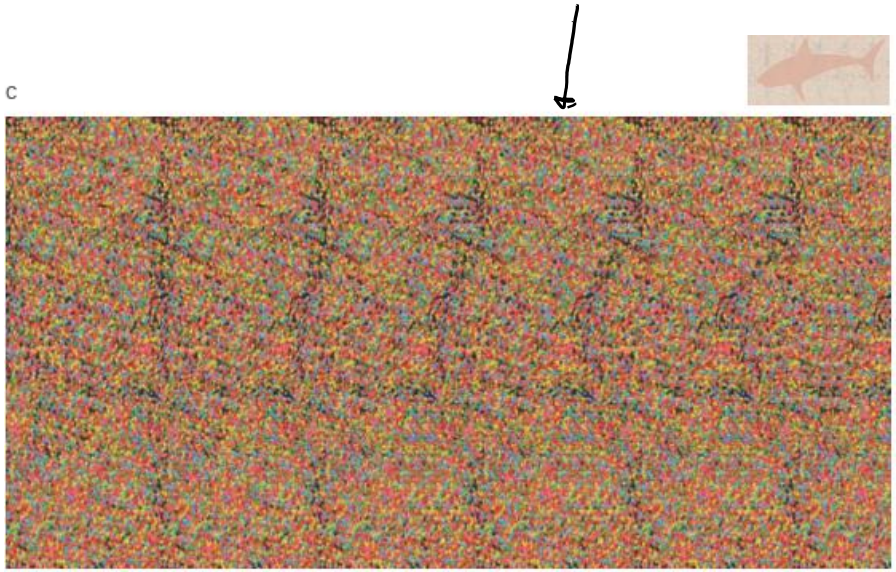
V4



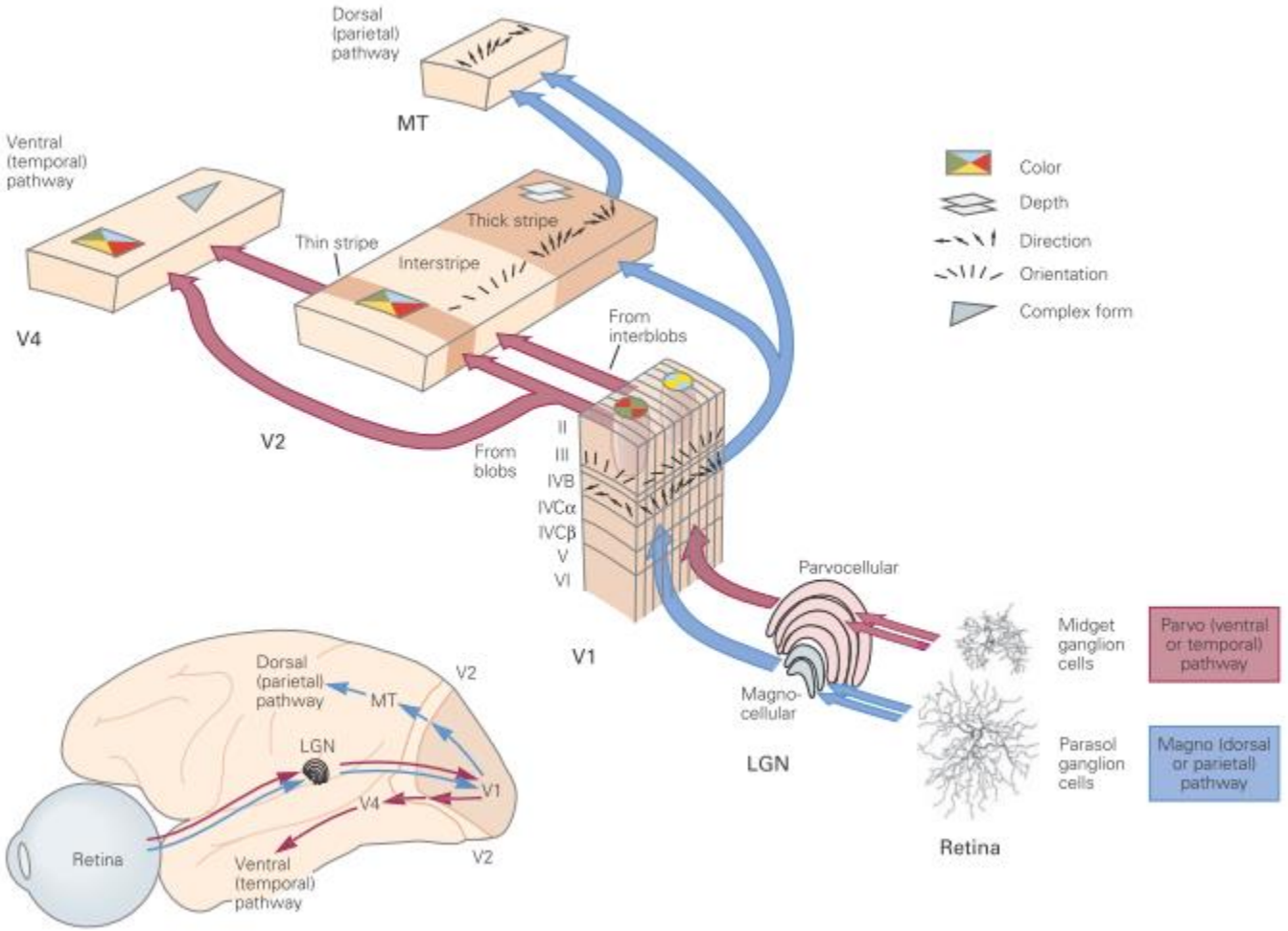
Stereoscopic vision is not dependent on the shape recognition!



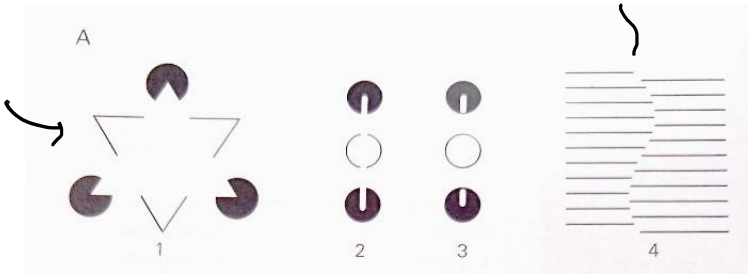
Retinic disparity is the only needed element (Julesz, 1960)

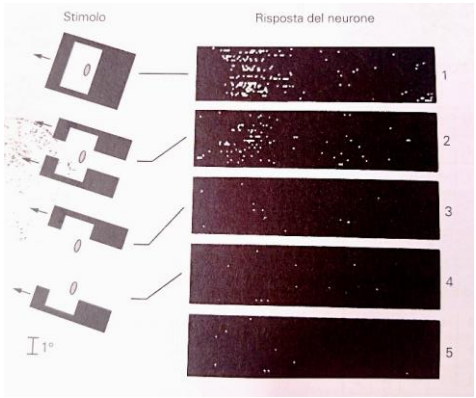


Ventral pathway: shape and color

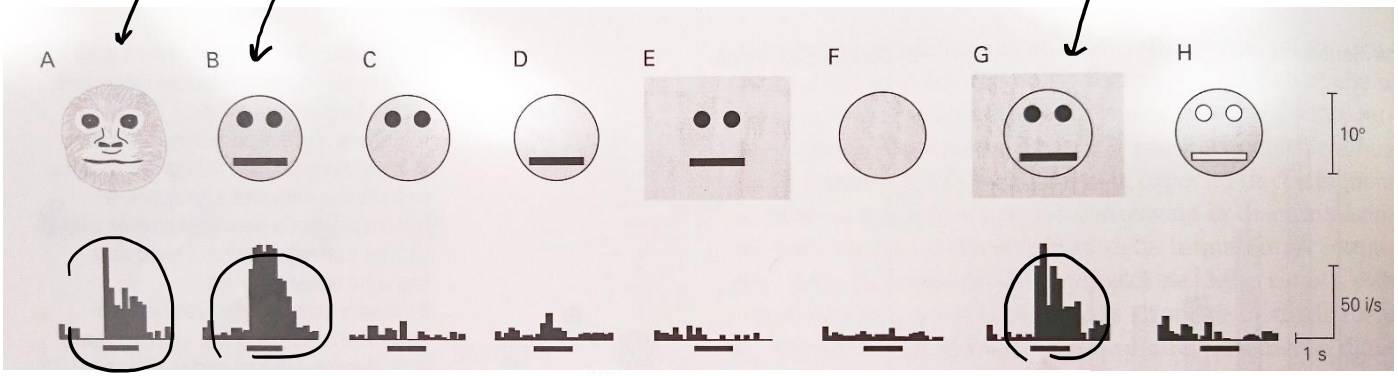


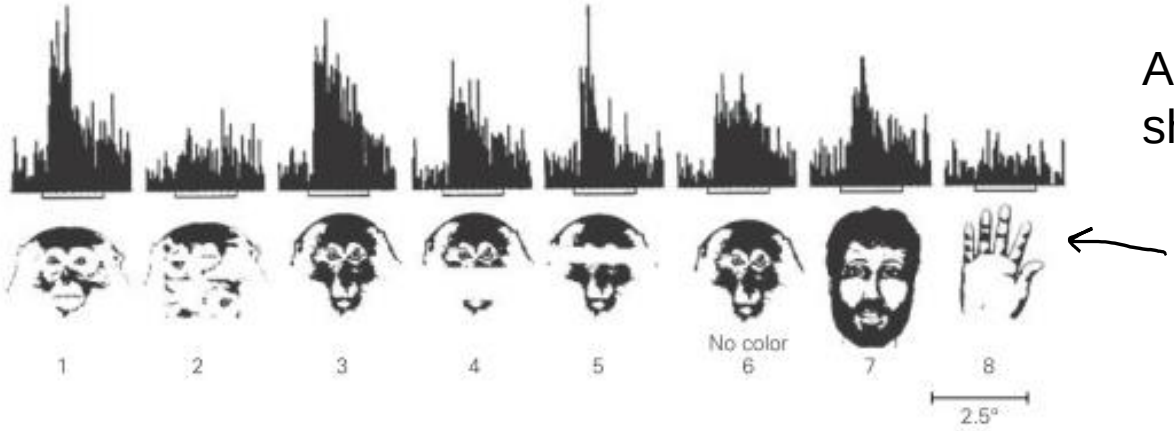
Cells respond to both virtual and real edges





Von der Heydt, 1984





And to very specific complex shapes and colour

The construction of the perception is a complex porcess which starts from low level features extraction, and it is linked to high level processing by intermediate connection of the low level features.