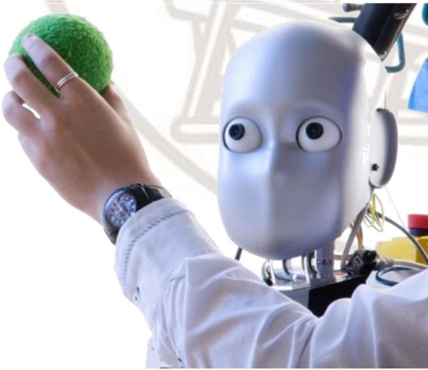
Master in Bionics Engineering
University of Pisa and Scuola Superiore Sant'Anna
Human and Animal Models for BioRobotics

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

Human Vision and Eye Movements

Cecilia Laschi
The BioRobotics Institute
cecilia.laschi@santannapisa.it



Summary of bioinspired approaches to robotics (in this course…)

Robot vision

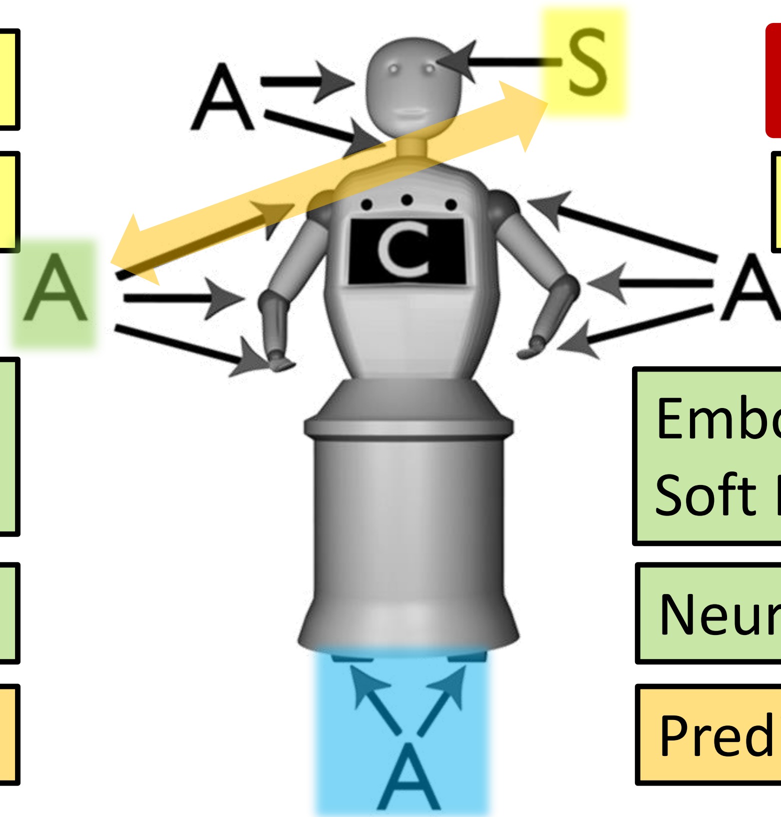
Robot sensors

Robot mechanics and kinematics

Robot control

Robot behaviour

Robot navigation



Bioinspired vision

Vestibular system

Embodied Intelligence, Soft Robotics

Neurocontrollers

Predictive behaviour

Bioinspired navigation, Soft locomotion

Image generation in the eye

THE BIOROBOTICS INSTITUTE
Scuola Superiore Sant'Anna

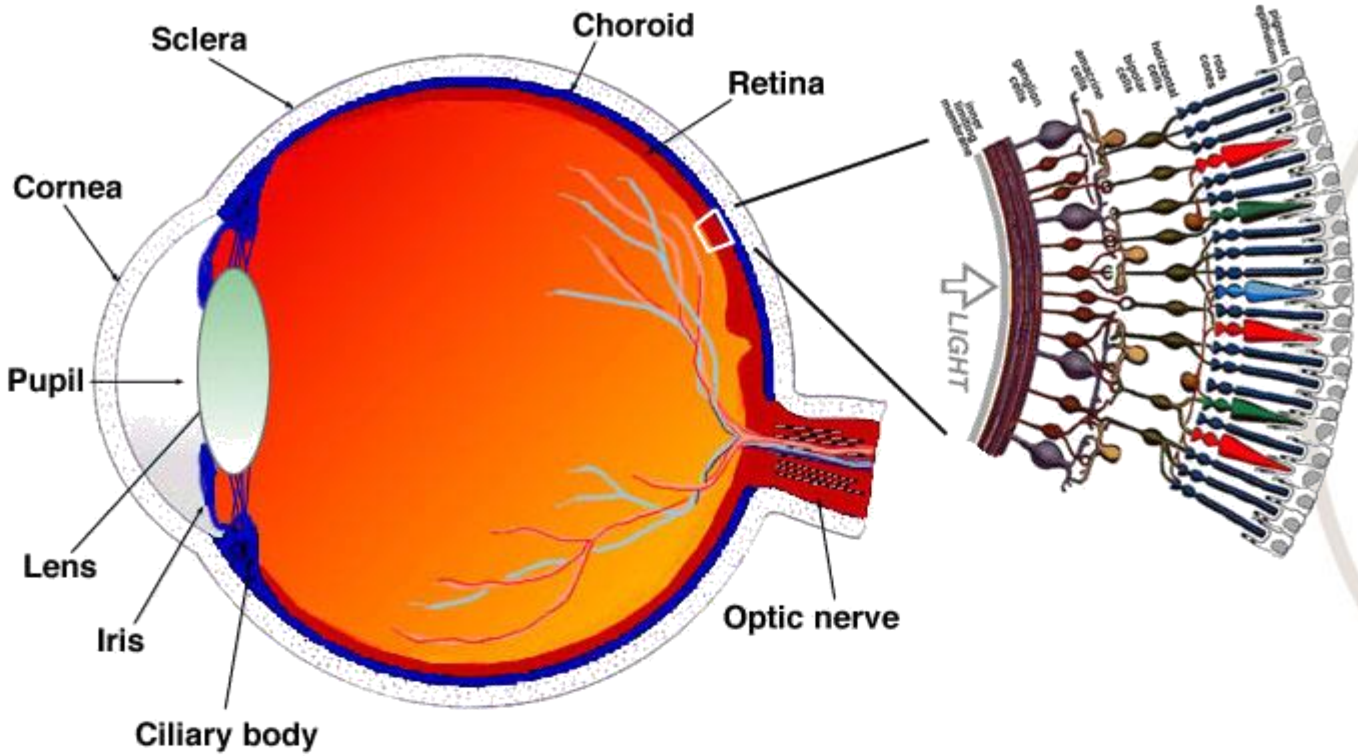


Fig. 1.1. A drawing of a section through the human eye with a schematic enlargement of the retina.

Image generation in the eye

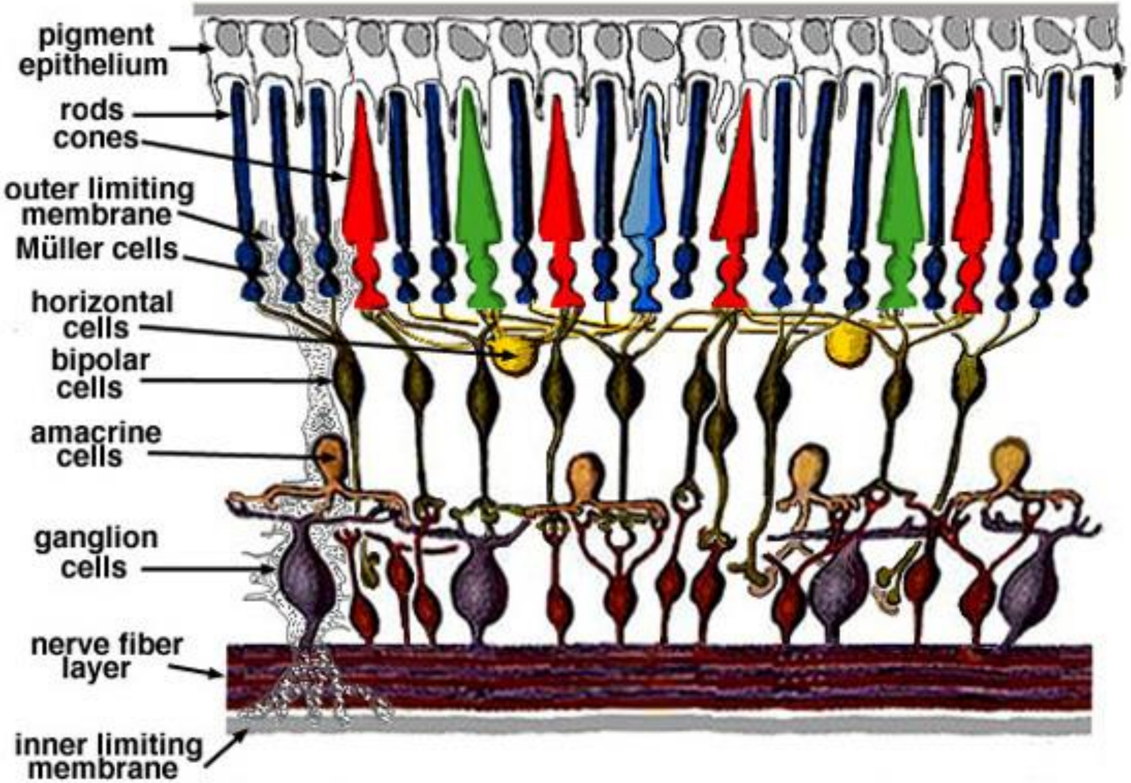
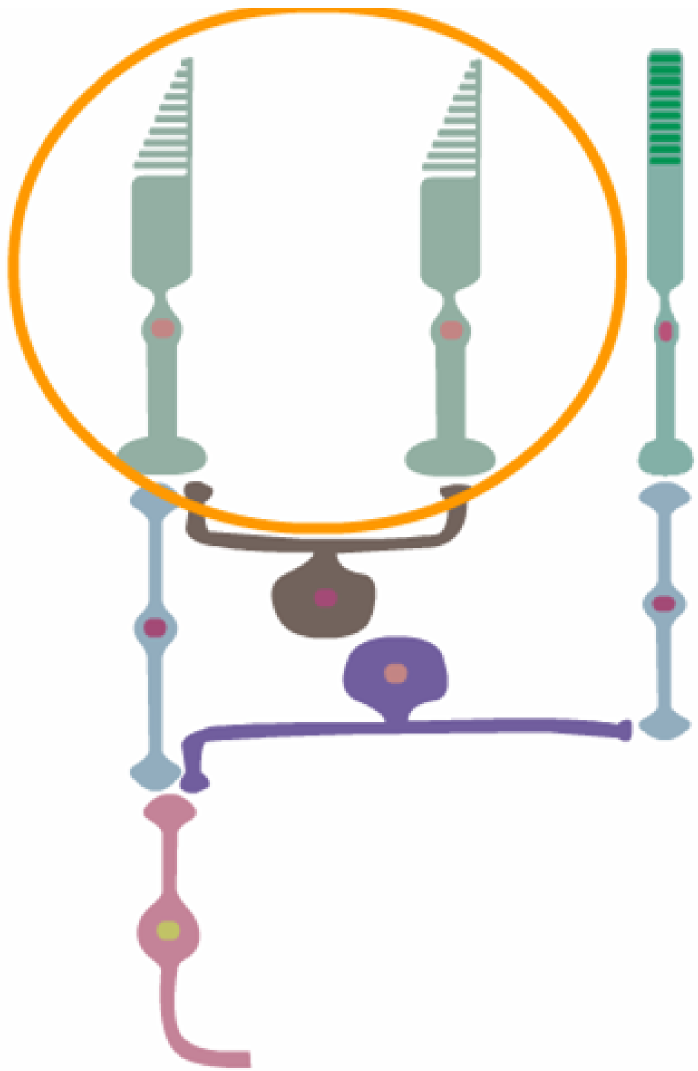


Fig. 2. Simple diagram of the organization of the retina.

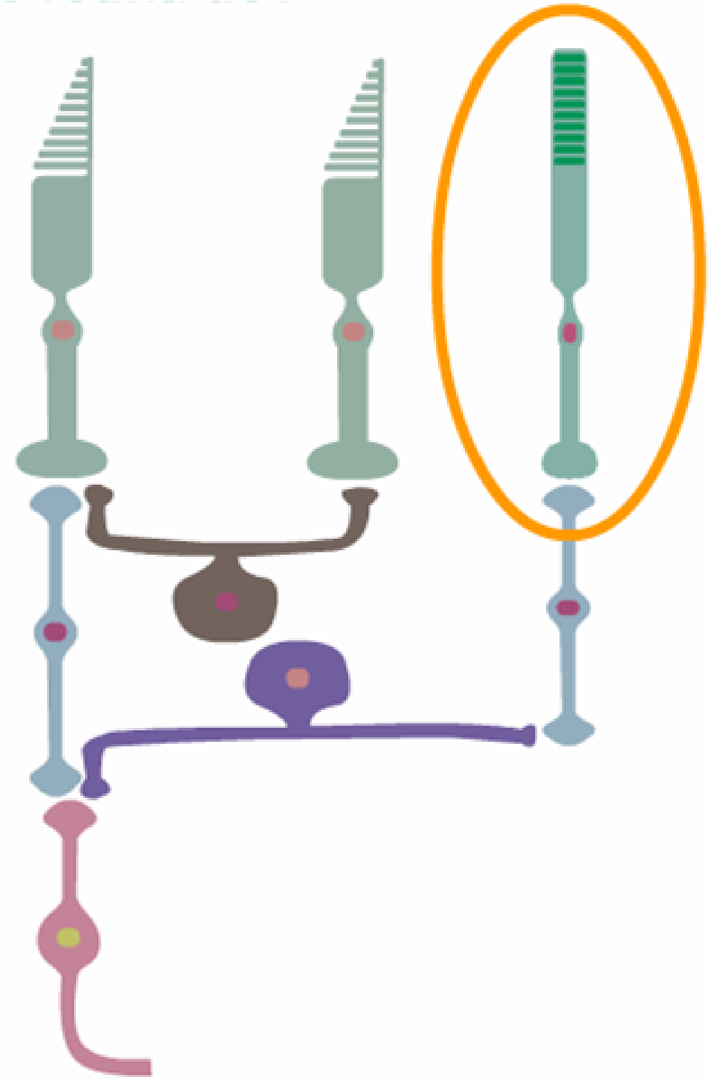
Photoreceptors: cones and rods

1. Light activates sensitive receptors
   Cones by different colors



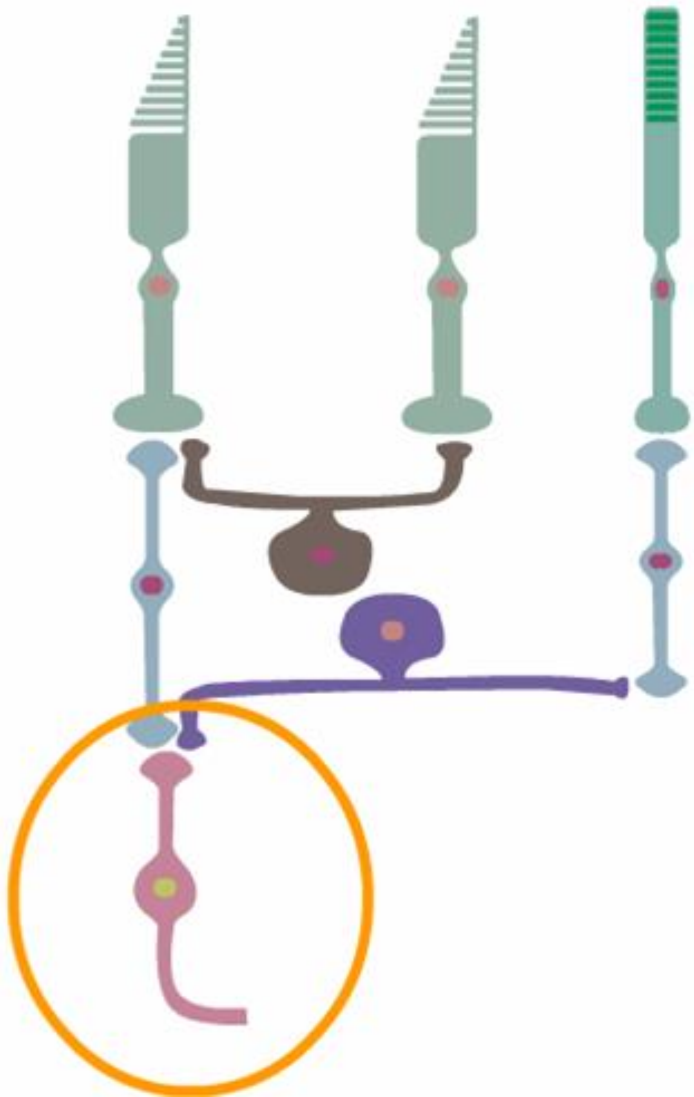
Photoreceptors: cones and rods

1. Light activates sensitive receptors
   Cones by different colors
   Rods by black and white



Photoreceptors: cones and rods

1. Light activates sensitive receptors
   Cones by different colors
   Rods by black and white
2. Ganglion cells are
   the only output from the eye.

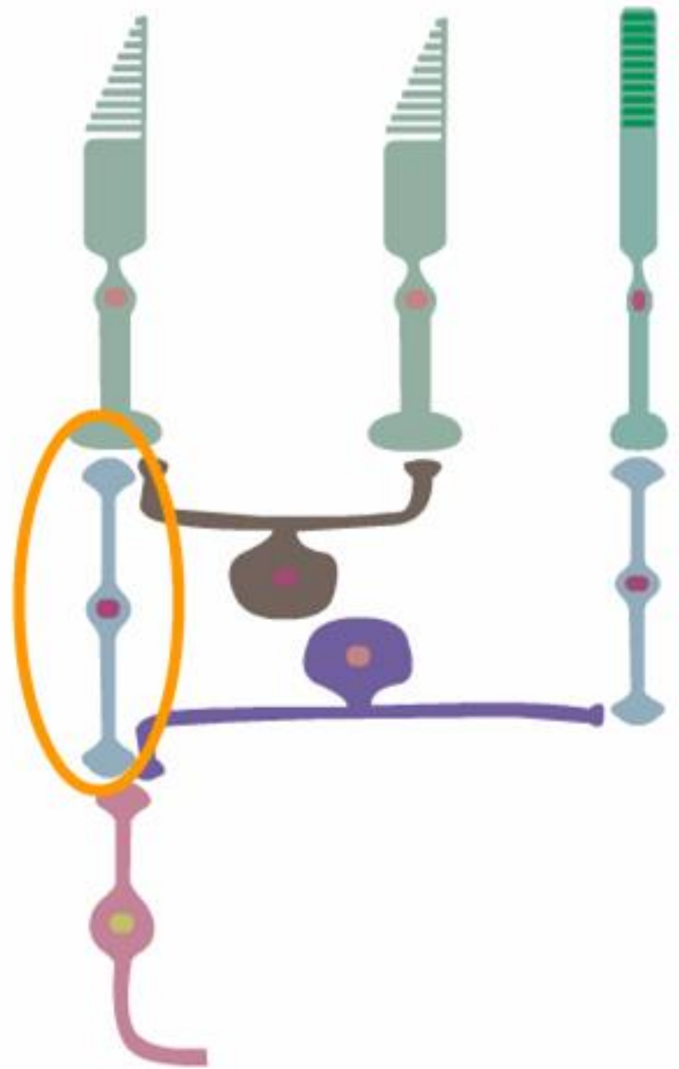


Photoreceptors: cones and rods

THE BIOROBOTICS INSTITUTE

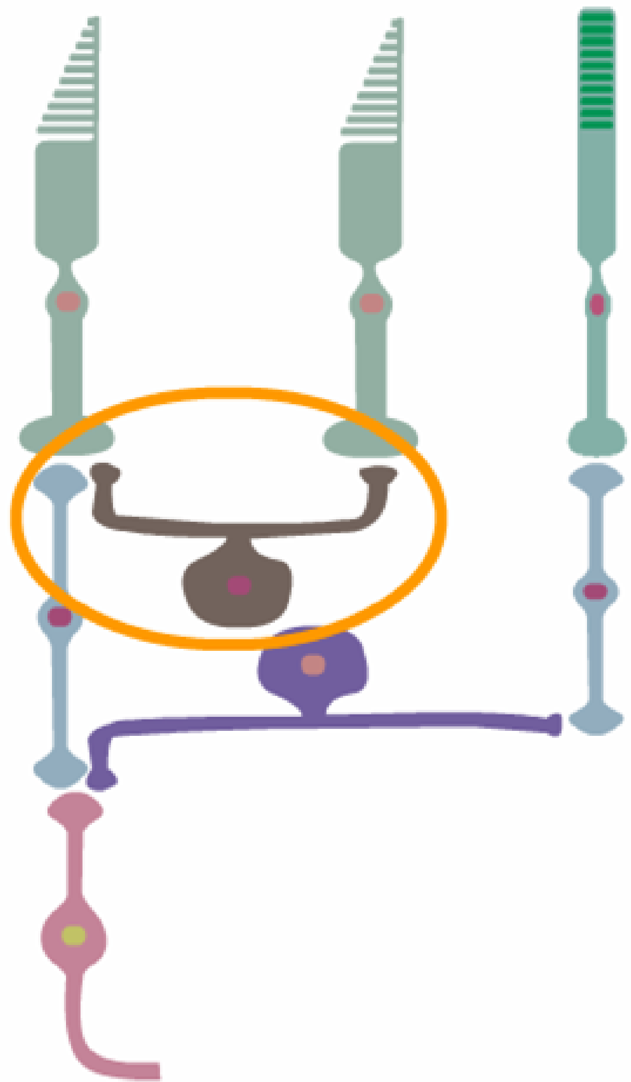
Scuola Superiore Sant'Anna

1. Light activates sensitive receptors
   Cones by different colors
   Rods by black and white
2. Ganglion cells are
   the only output from the eye.
3. Bipolar cells connect
   the receptors to the ganglion cells.



Photoreceptors: cones and rods

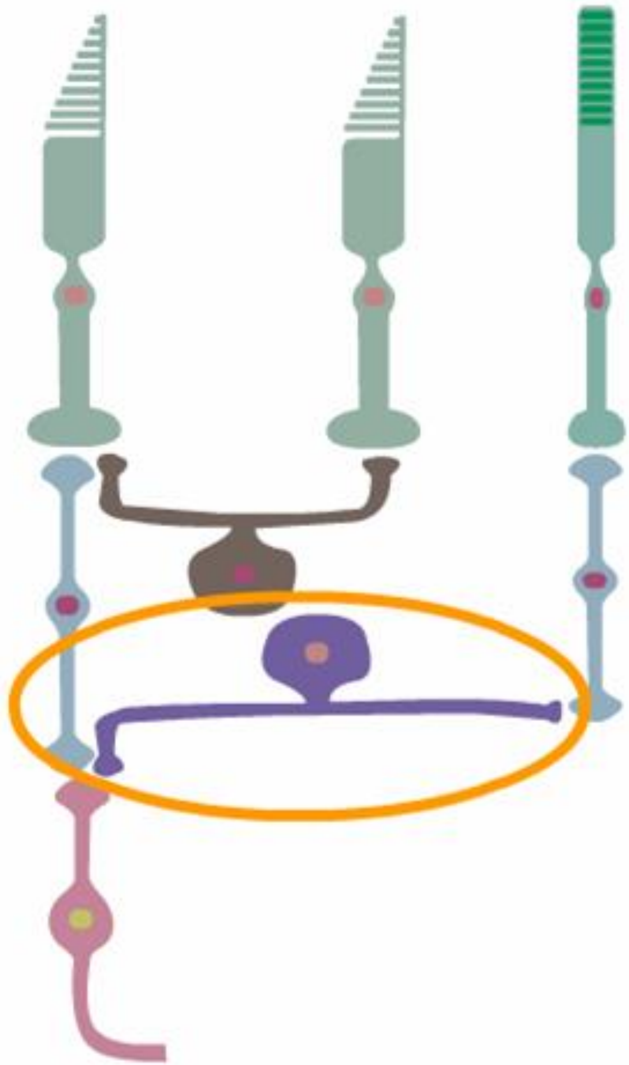
1. Light activates sensitive receptors
   Cones by different colors
   Rods by black and white
2. Ganglion cells are
   the only output from the eye.
3. Bipolar cells connect
   the receptors to the ganglion cells.
4. Horizontal cells
   converge signals from several cones.
   They determine how many receptors
   each ganglion cell sees.



Photoreceptors: cones and rods

THE BIOROBOTICS INSTITUTE
Scuola Superiore Sant'Anna

1. Light activates sensitive receptors
   Cones by different colors
   Rods by black and white
2. Ganglion cells are
   the only output from the eye.
3. Bipolar connect
   the receptors to the ganglion cells.
4. Horizontal cells
   converge signals from several cones.
   They determine how many receptors
   each ganglion cell sees.
5. Amacrine cells do the same
   from peripheral rods.

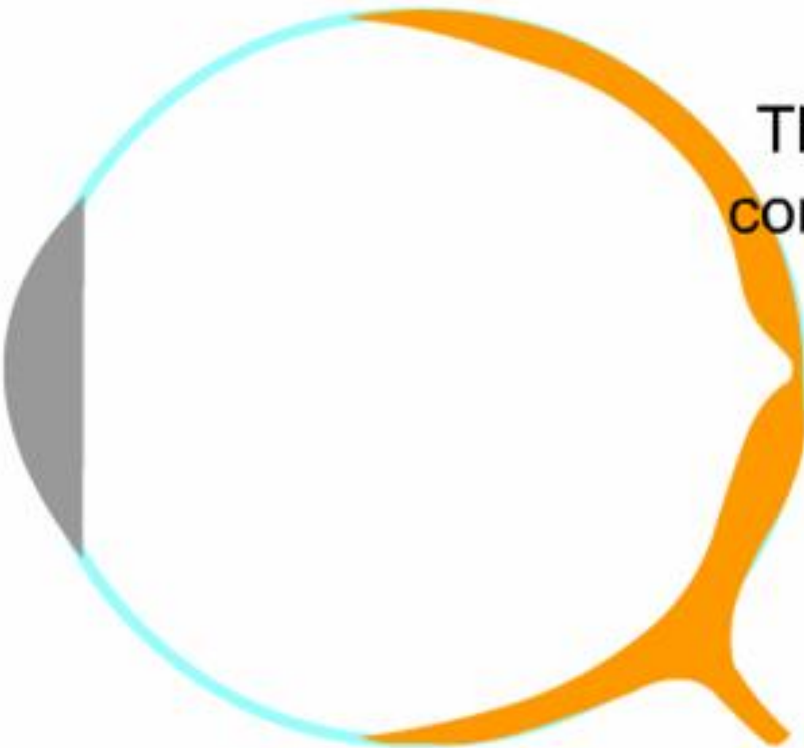


THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

Distribution of photoreceptors in the retina

the retina is not uniform.



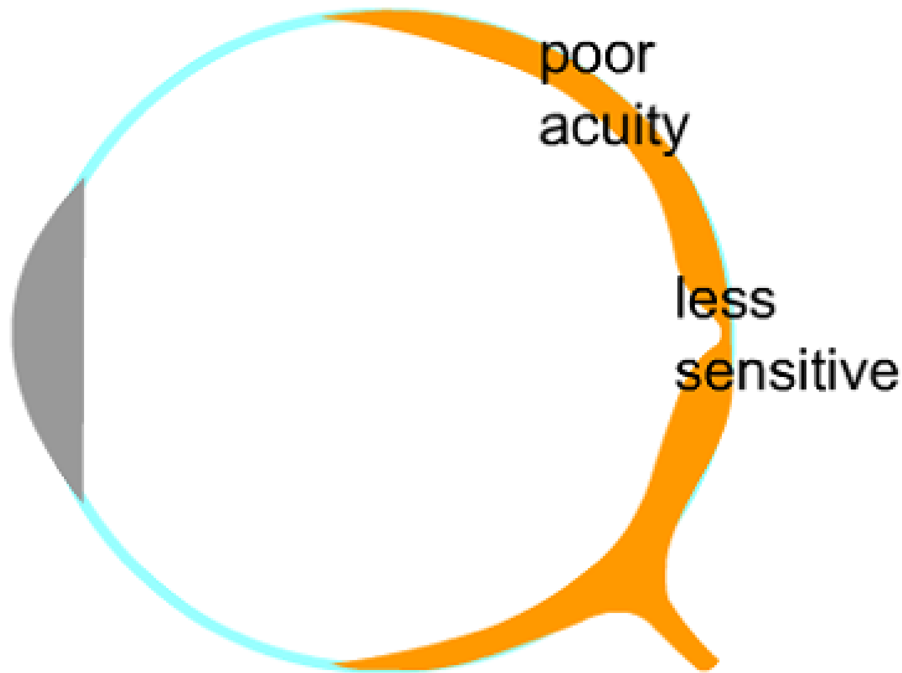
The **peripheral** retina contains primarily **rods**.

The **fovea**, in the center of the eye, contains only **cones**.

Sensitivity of photoreceptors in the retina

the rods and cones are
not equally sensitive to low light levels.

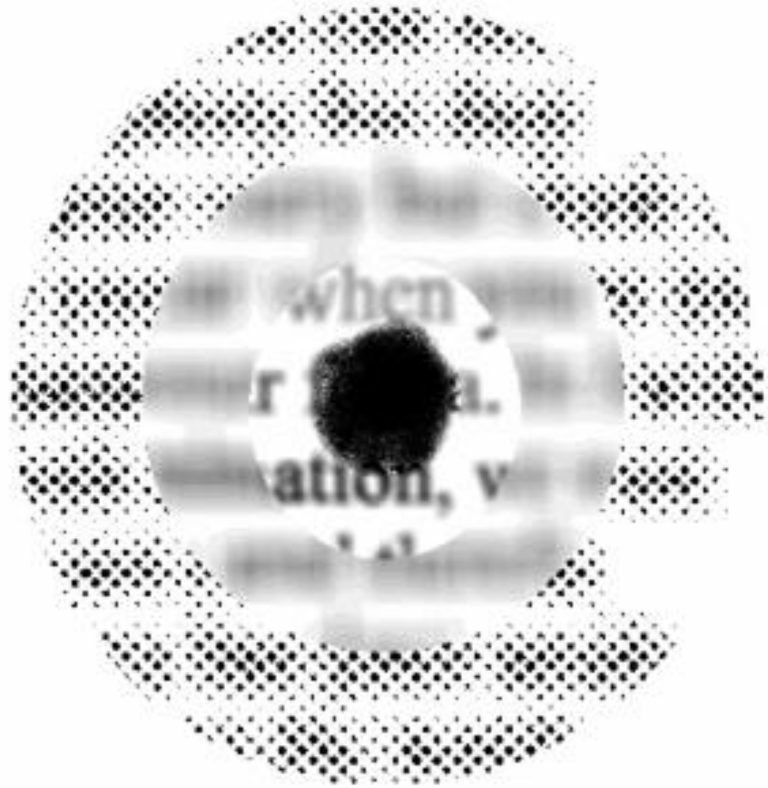
Cones are less sensitive to light.
e.g. Looking at dim stars,
one can see stars in the periphery
but they disappear when you
look at them with your fovea.
In very low levels of illumination,
we see only with our rods and
therefore see greys not colours.



the periphery has poor acuity.

What the eye sees

By daylight,
only the central fovea sees
clearly and in color

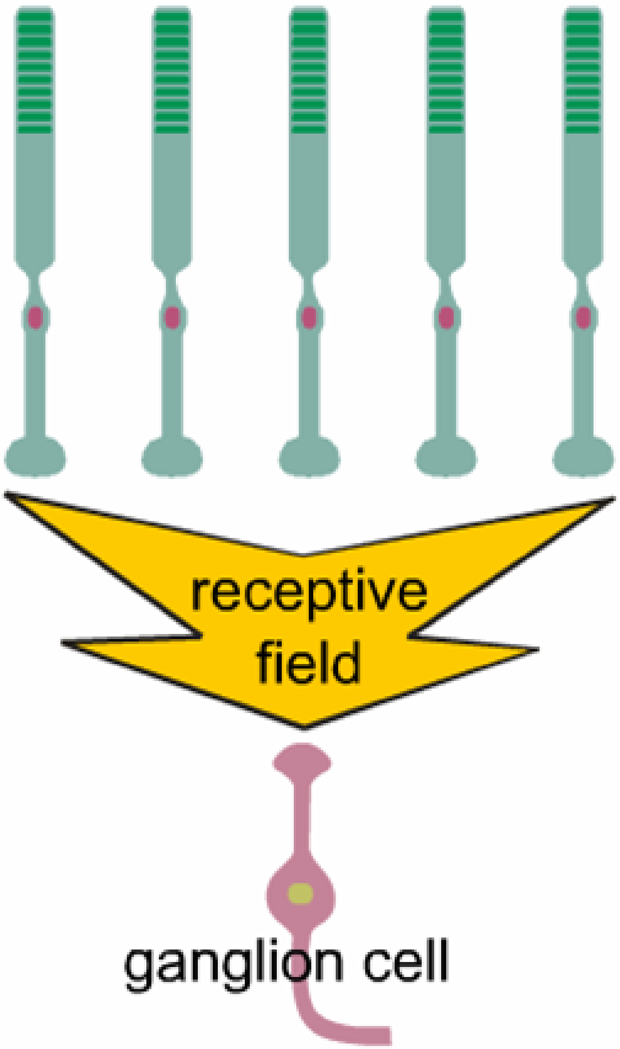


On a dark night,
only the periphery sees
only in
black & white,
and with poor resolution.
The fovea is blind

Receptive fields

Definition of the receptive field
of a ganglion cell:
*"That area of retina
over which light stimuli
change the activity
of a particular ganglion cell."*

The receptive field shows
which rods & cones
are connected to the ganglion cell.



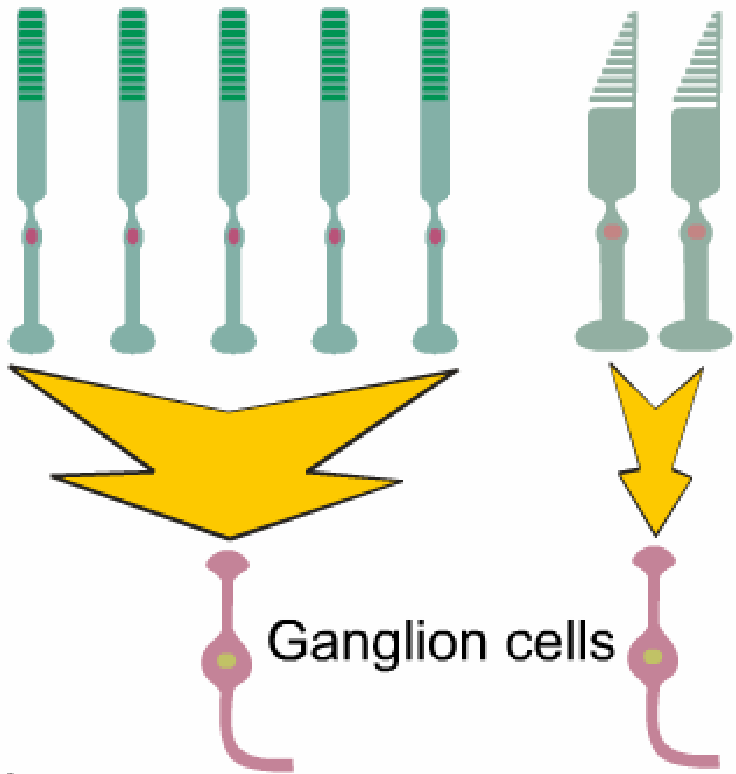
Distribution of photoreceptors in the retina

1) peripheral rods
large spacing
(lower density)

foveal cones
high density

2) large
convergence

small
convergence



integrate information
from a large area of retina (3 deg)

integrate information
from a small area of retina (.03 deg)

large spacing and large convergence
result in low acuity

small spacing and low convergence
result in high acuity.

Why is compression important?

There are 100 million rods and cones. Why not have the same number of ganglion cells instead of the actual 1 million?

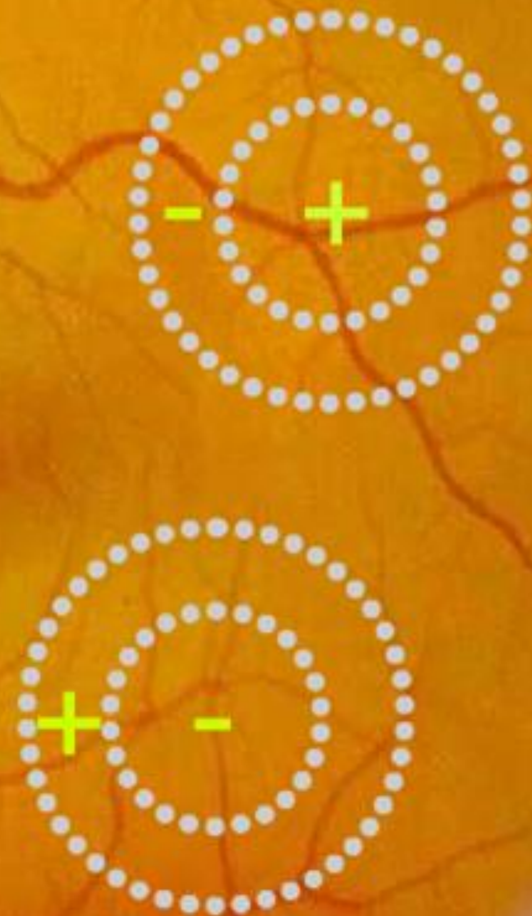
The answer is efficiency. Much of the information that the eyes sees is redundant. Through evolution the eye has been designed to remove this redundant information before sending it on to the brain. Why build and maintain a huge number of fibers when a much smaller number can convey the same information? When the eye sees a round image, it does not transmit the color and brightness of every point inside the image; only that at the edges.

This is similar to computer compression software used to transmit large images along the Internet. The ganglion cell sends compressed information down the optic nerve. The visual cortex then uncompresses this information. Suppose you were transmitting the color of a series of dots along the internet, each colour coded by a number from one to 2000000. The series uncompressed would look like: 1756333, 1756333, 1756333, 1756335, 1756335, 1756333 ...... Compressed, by coding only changes in color, it would look like: 1756333,0,0,2,0,-2.... a lot shorter.

Types of receptive fields of retinal ganglion cells

What are the two major types
of receptive fields
of retinal ganglion cells?

(a) ON center, OFF surround
which measure how much brighter
an object is than its background.
(b) OFF center, ON surround
which measure how much darker
an object is than its background.



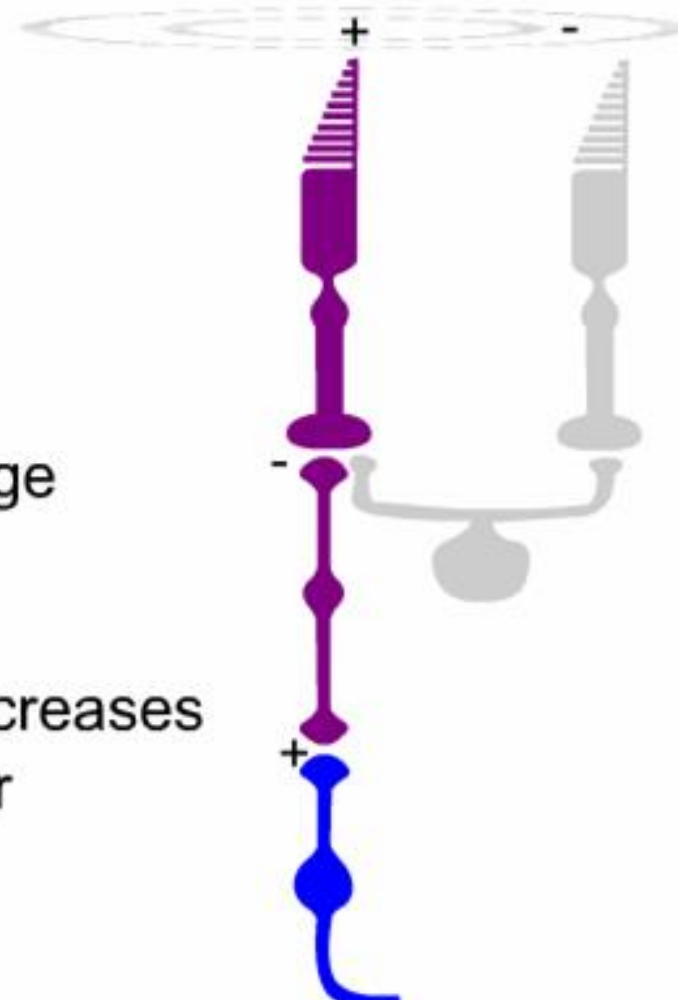
THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

ON centre, OFF surround ganglion cell

Light to a cone in the centre
produces
excitation of the ganglion cell.

This is because:
1) light decreases the cone voltage
and the cone releases
less inhibitory transmitter
2) the voltage inside the bipolar cell increases
and it releases more transmitter
3) the ganglion cell is excited
and it fires more often.



ON centre, OFF surround ganglion cell

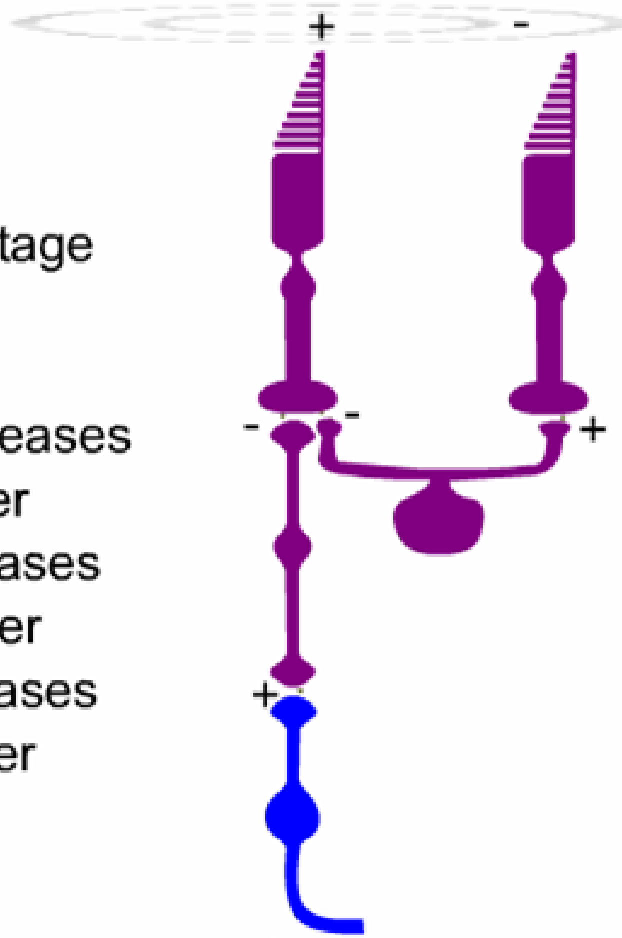
THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

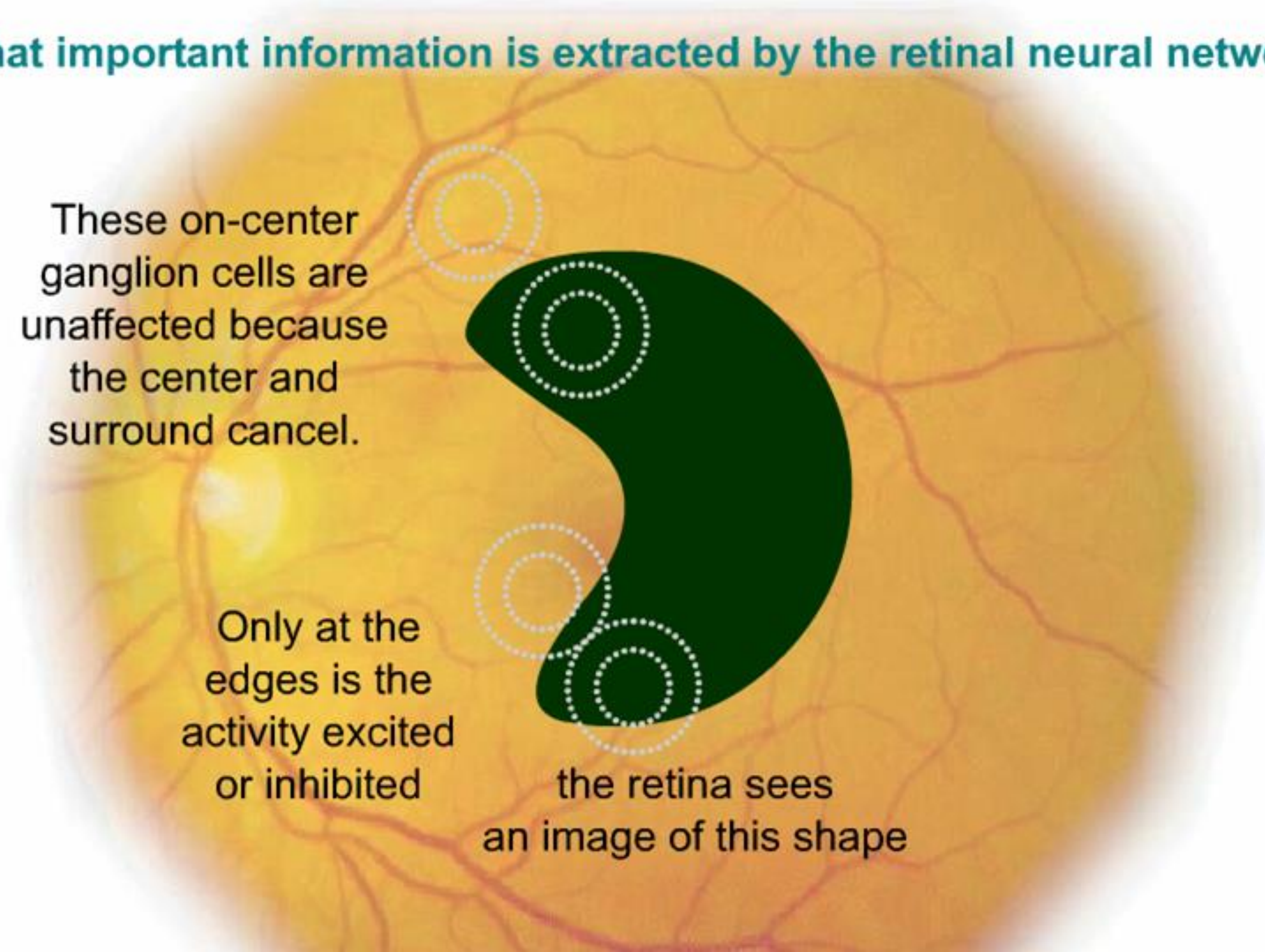
Light to a cone in the surround
produces
inhibition of the ganglion cell.

This is because:

1) light decreases the surround cone's voltage and the cone releases less excitatory transmitter
2) the voltage inside the horizontal cell decreases and it releases less inhibitory transmitter
3) the voltage inside the center cone increases and it releases more inhibitory transmitter
4) the voltage inside the bipolar cell decreases and it releases less excitatory transmitter
5) the ganglion cell is inhibited and it fires less often.



What important information is extracted by the retinal neural network?



What important information is extracted by the retinal neural network?

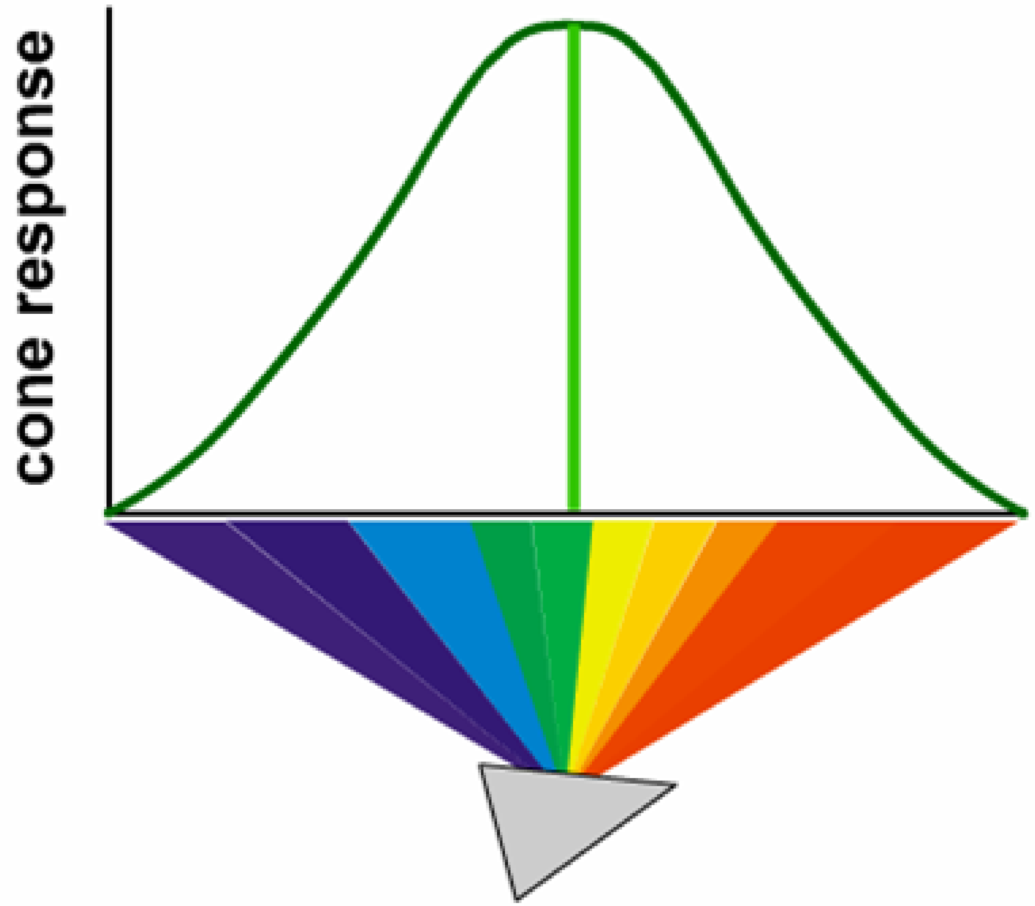
Ganglion cells exaggerate the contrast at borders (i.e. like a cartoon).

Why? By sending only the information on contours, the changes in brightness, less redundant information is sent along the small optic nerve to the CNS.

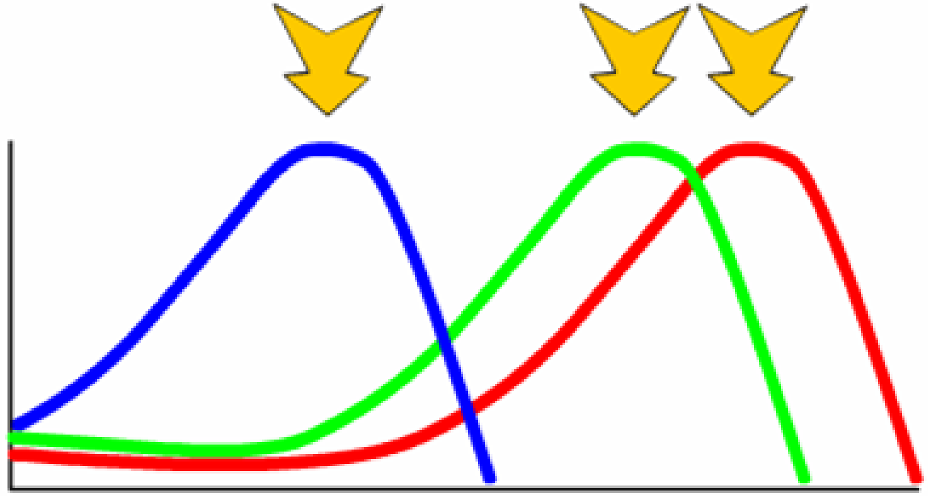
Color perception

Cones responds best to a particular wavelength of light and less to others.

Note that this one reponds best to green light.

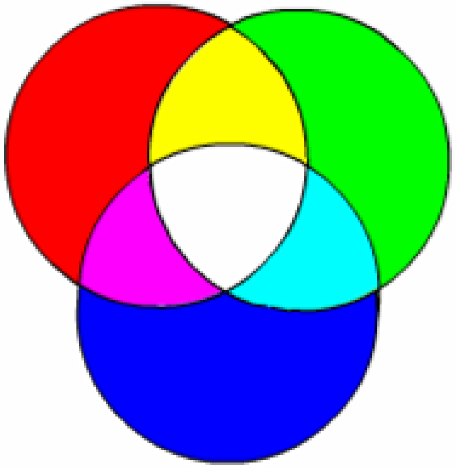


We have 3 cone types.



Mixing light is **not** like mixing paint.

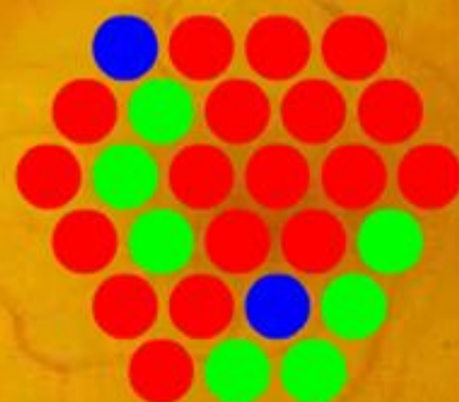
When red & green cones are stimulated one sees yellow



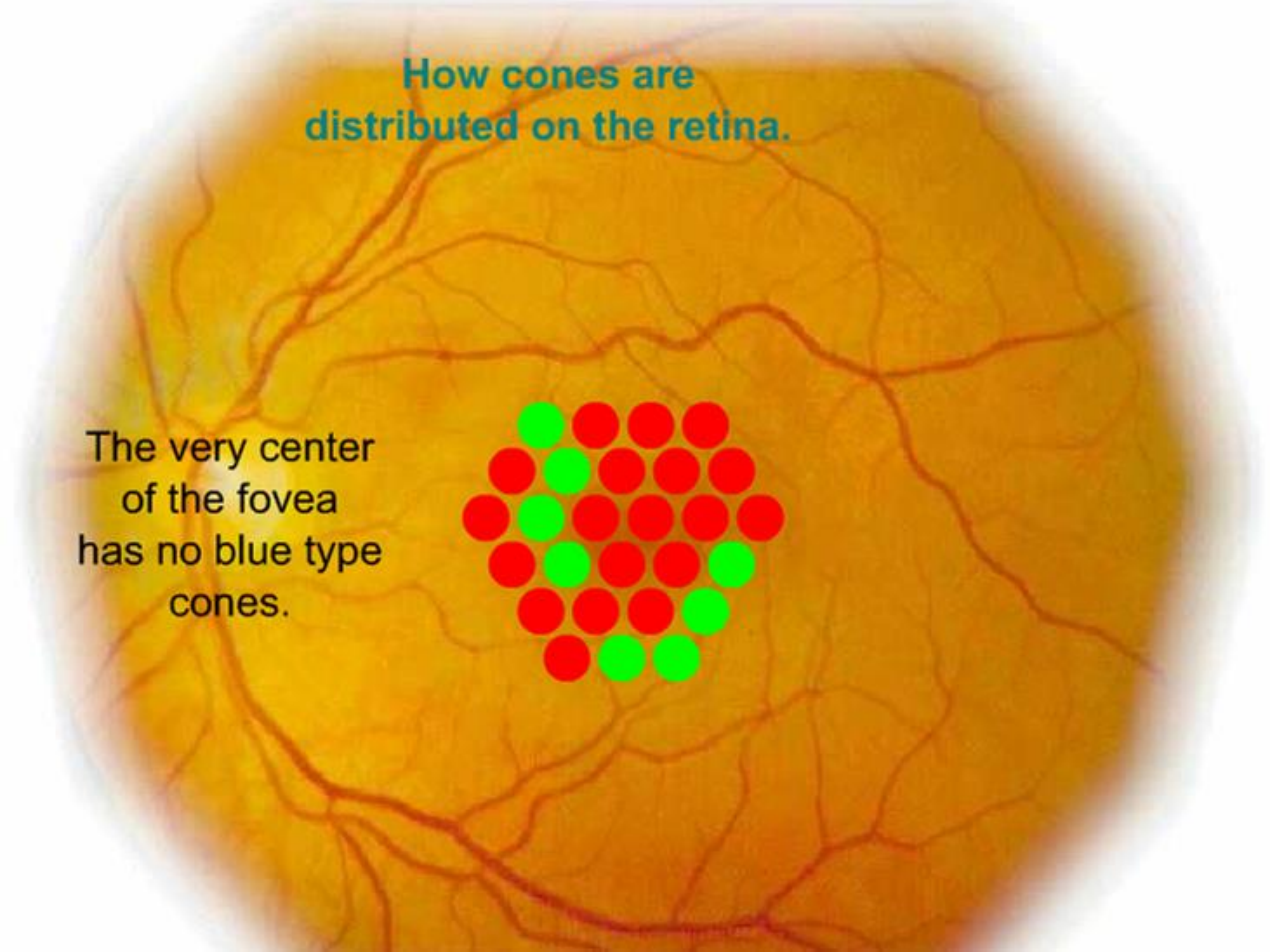
When all three cone types are stimulated one sees white.

How cones are distributed on the retina.

In the fovea
1) the # of each cone type is not equal. Usually red cones are most numerous and blue cones least numerous
2) the relative #'s vary from person to person
3) the cones of the same type form clusters

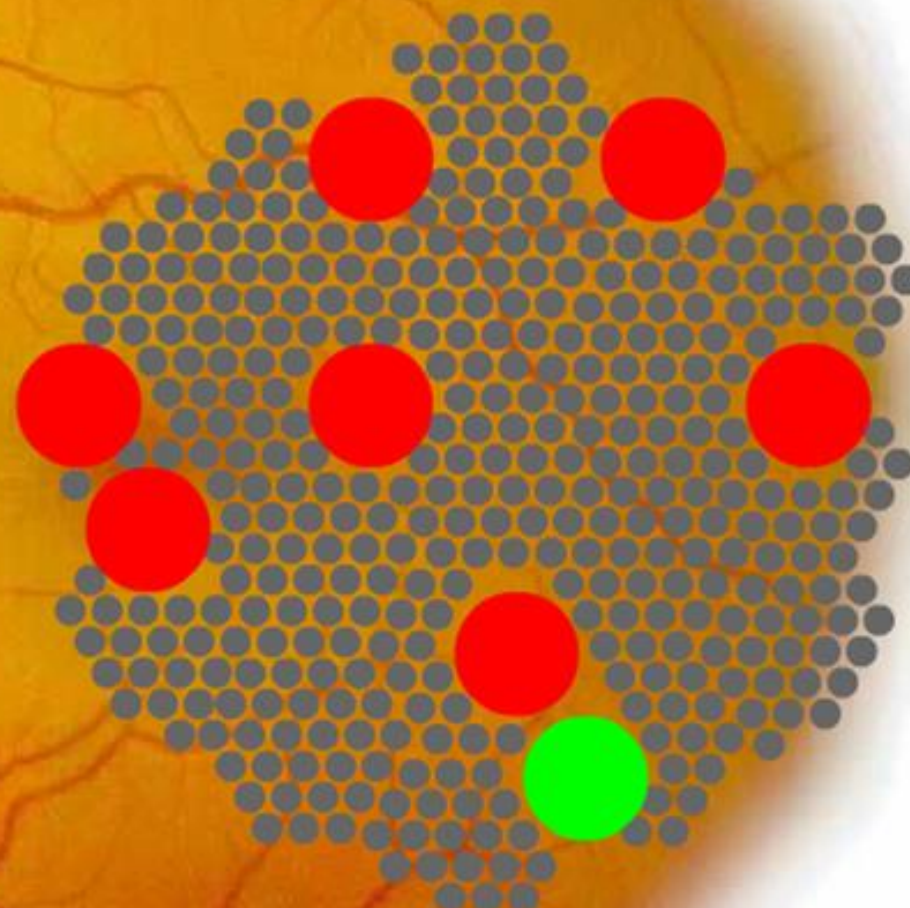


How cones are distributed on the retina.



How cones are distributed on the retina.

As one moves
away from the fovea
1) the #'s of cones drops and the #'s of rods increases
2) the size of both rods and cones increases and thus their density (# per square mm) decreases
3) cones become larger than rods.



Color perception

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

a) **200 hues**

The brain transforms the single wavelengths of light seen in rainbow into a color circle. Hues on opposite sides of the circle are complementary.

Color perception

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

a) **200 hues**

The brain transforms the single wavelengths of light seen in rainbow into a color circle. Hues on opposite sides of the circle are complementary.

b) **20 levels of saturation**

Combinations of two more wavelengths. When complementary wavelengths are combine equally one gets white.

Color perception

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

a) **200 hues**

The brain transforms the single wavelengths of light seen in rainbow into a color circle. Hues on opposite sides of the circle are complementary.

b) **20 levels of saturation**

Combinations of two more wavelengths. When complementary wavelengths are combine equally one gets white.

c) **500 brightness levels**

Any color on the circle can be made brighter or darker. But because brighter or darker colors are more difficult to distinguish, the circle becomes narrower.

500x200x20 = 2,000,000 gradations of color

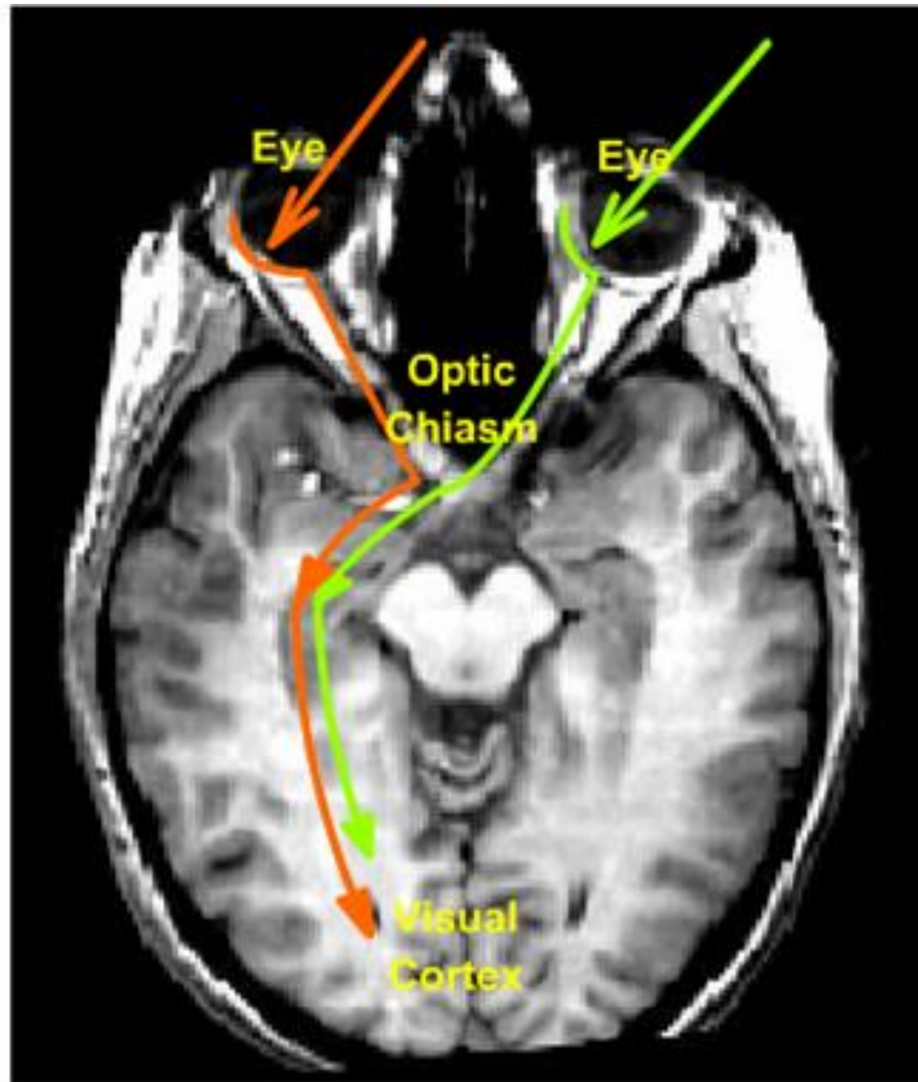
THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

Projection in the brain

Images seen on one side are processed by the opposite side of the brain.

To do this,
the ganglion cells
on the medial side of each eye,
from the middle of the fovea on,
shown in **green**,
cross at the **optic chiasm**.



THE BIOROBOTICS INSTITUTE

Scuola Superiore

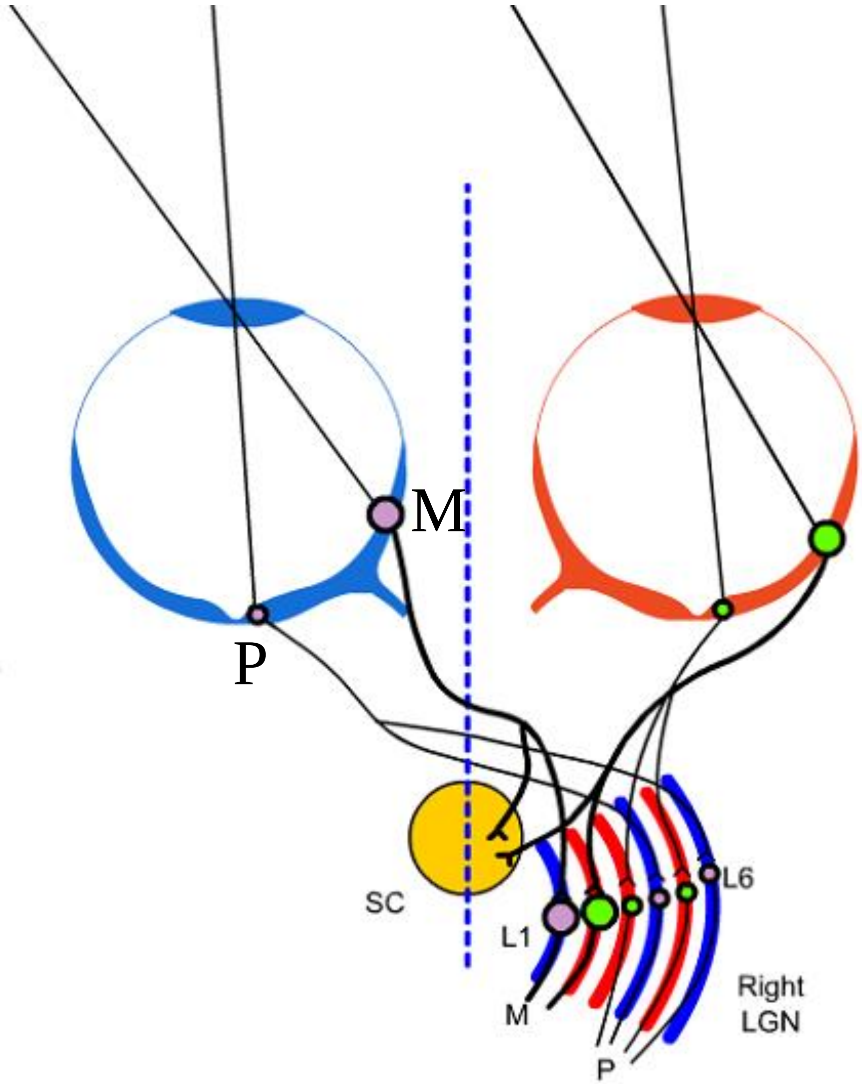
Projection in the Superior Colliculus (SC) and in the Lateral Geniculate Nucleus (LGN)

P (small) ganglion cells
primarily from the fovea
project to a part of the thalamus
called the lateral geniculate nucleus (LGN)

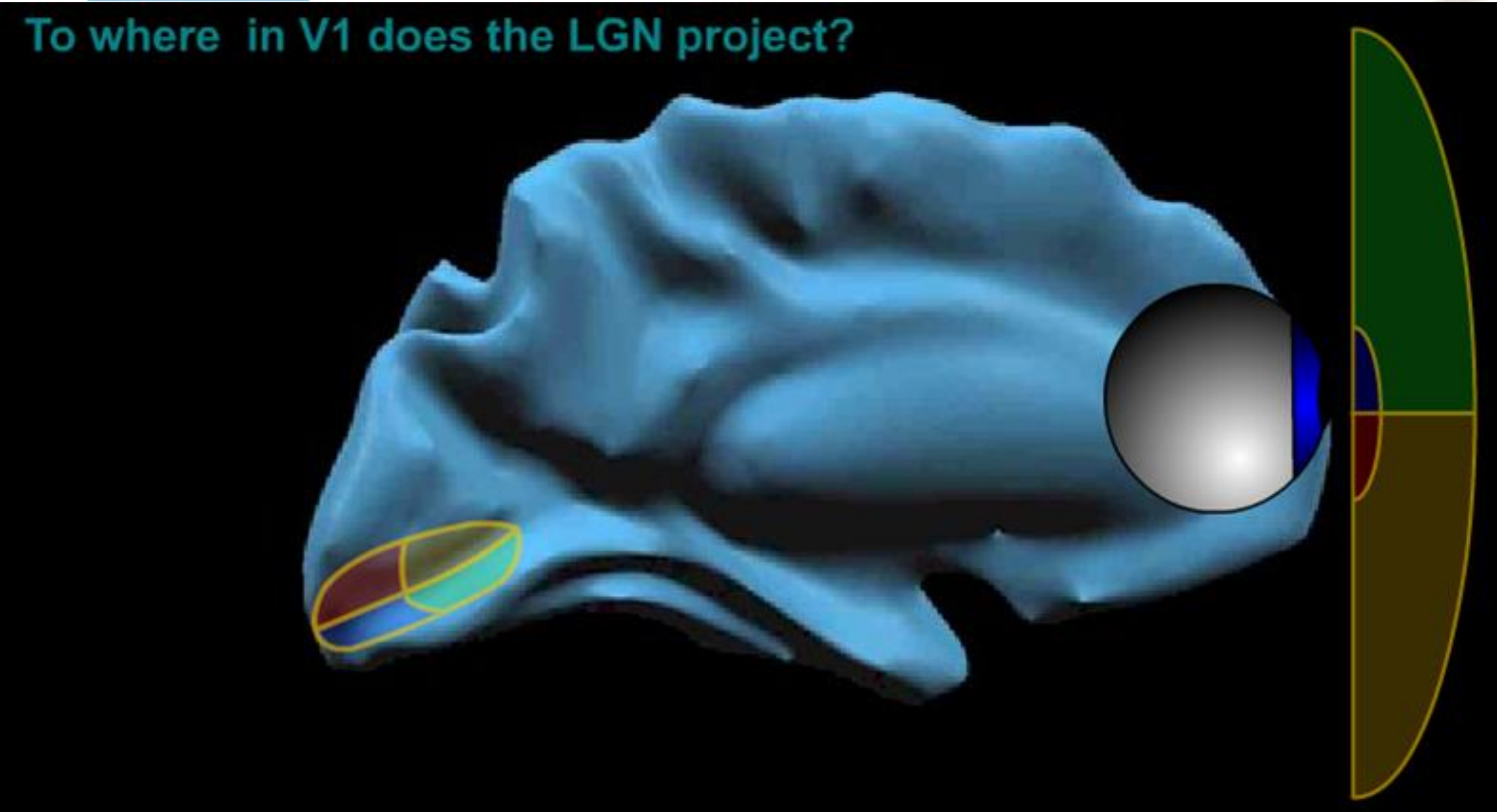
M (large) ganglion cells,
primarily from the peripheral retina,
code **where** objects are &
project both to LGN and several
structures in the brainstem, including
the superior colliculus (SC).

The SC causes the eye and head
to turn to an interesting visual object:
the “visual grasp reflex”.

The LGN sends information to visual
cortex; information as to **what** an object
is (P cells) and **where** it is (M cells).



Projection in the visual cortex

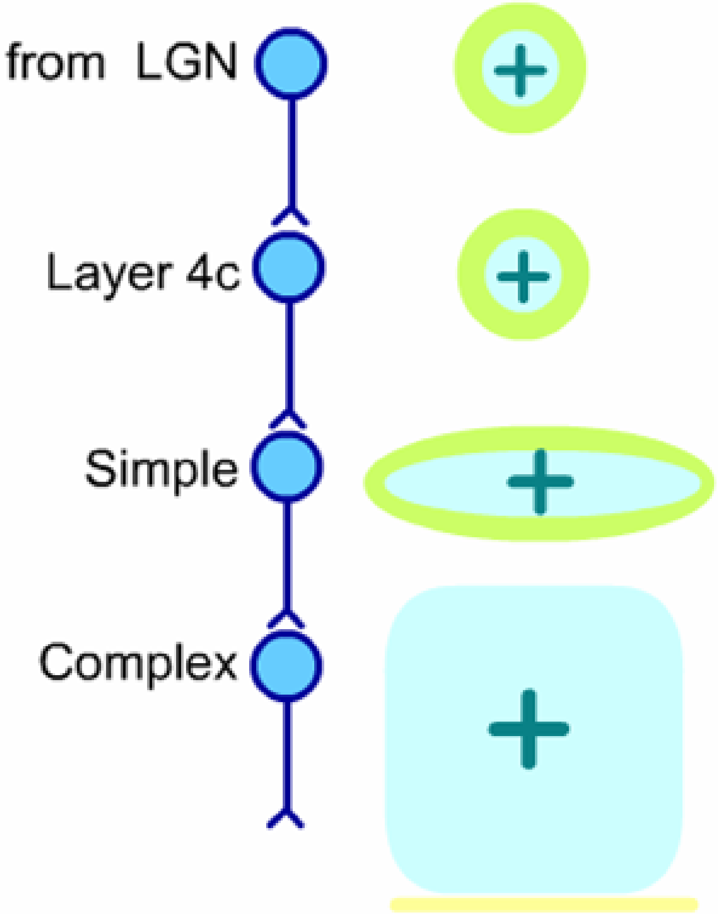


V1 contains 3 main types of cells

1. Layer 4c cells, whose receptive fields are the same as that of LGN & ganglion cell.

2. Simple cells with elongated receptive fields and thus maximally activated by a line of a particular orientation activating a particular region of the retina.

3. Complex cells whose receptive fields are similar to those of simple cells except the line can lie over a larger area of the retina (positional invariance) and they fire most to moving lines.

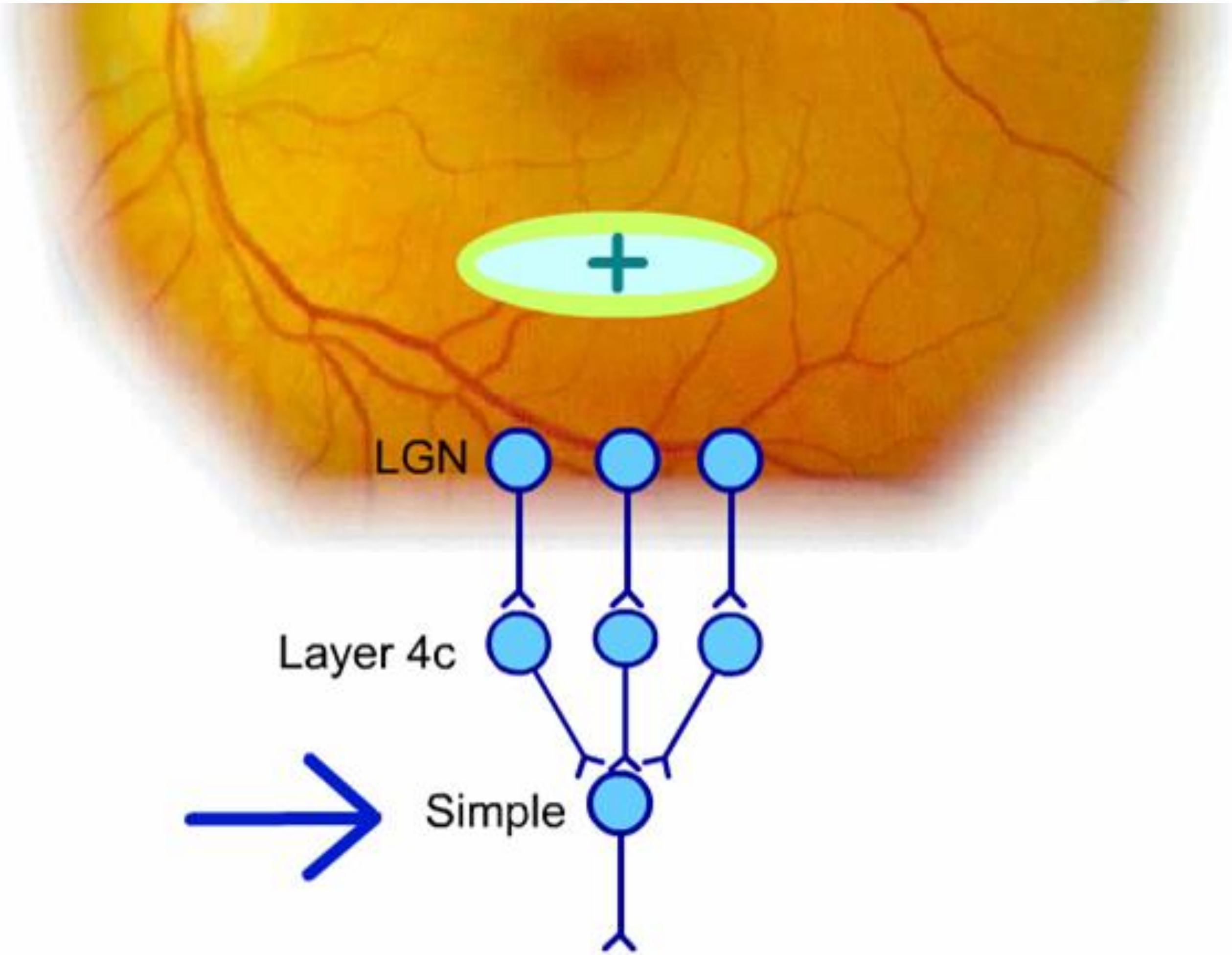


Simple cells

THE BIOROBOTICS INSTITUTE

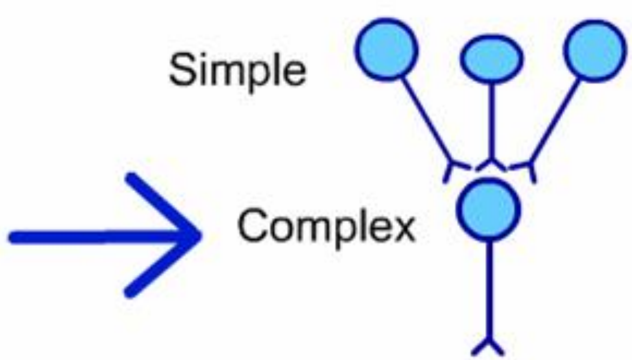
Scuola Superiore Sant'Anna

Simple cells
Several ganglion cells, whose receptive fields lie along a common line, converge by way of the LGN onto a simple cell.

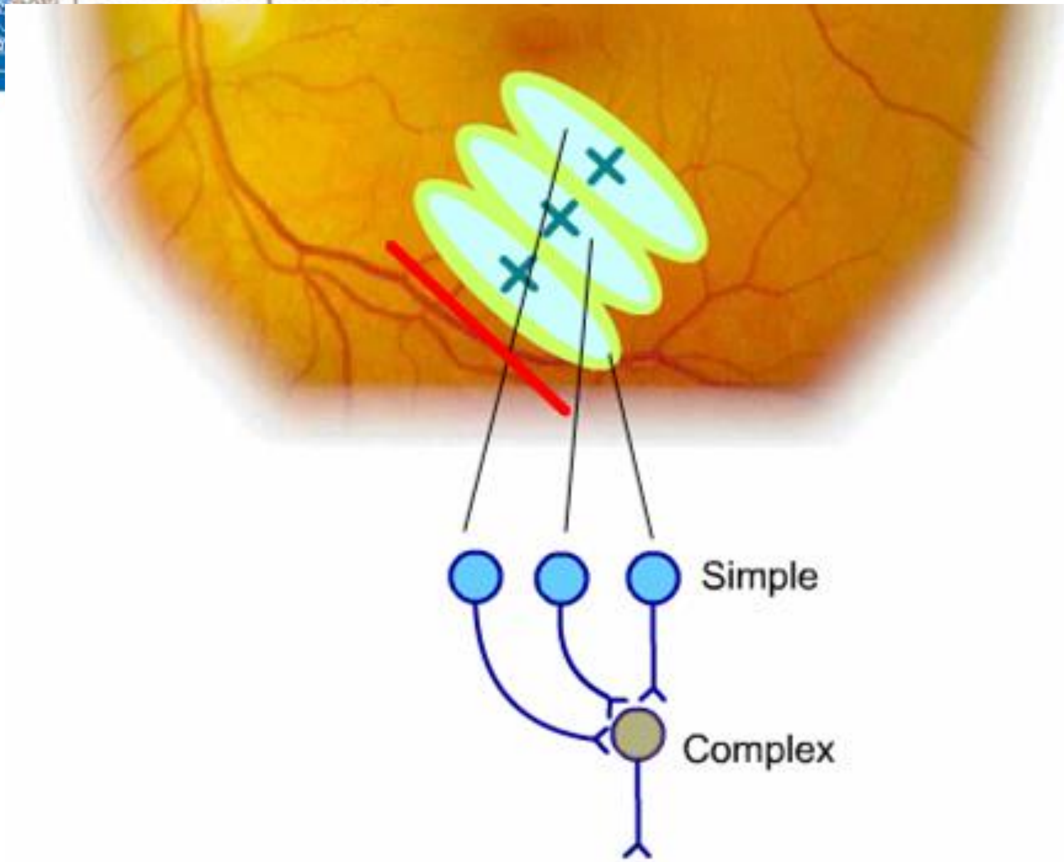


Complex cells

Complex cells: Several simple cells of the same orientation converge onto a complex cell.



Motion perception



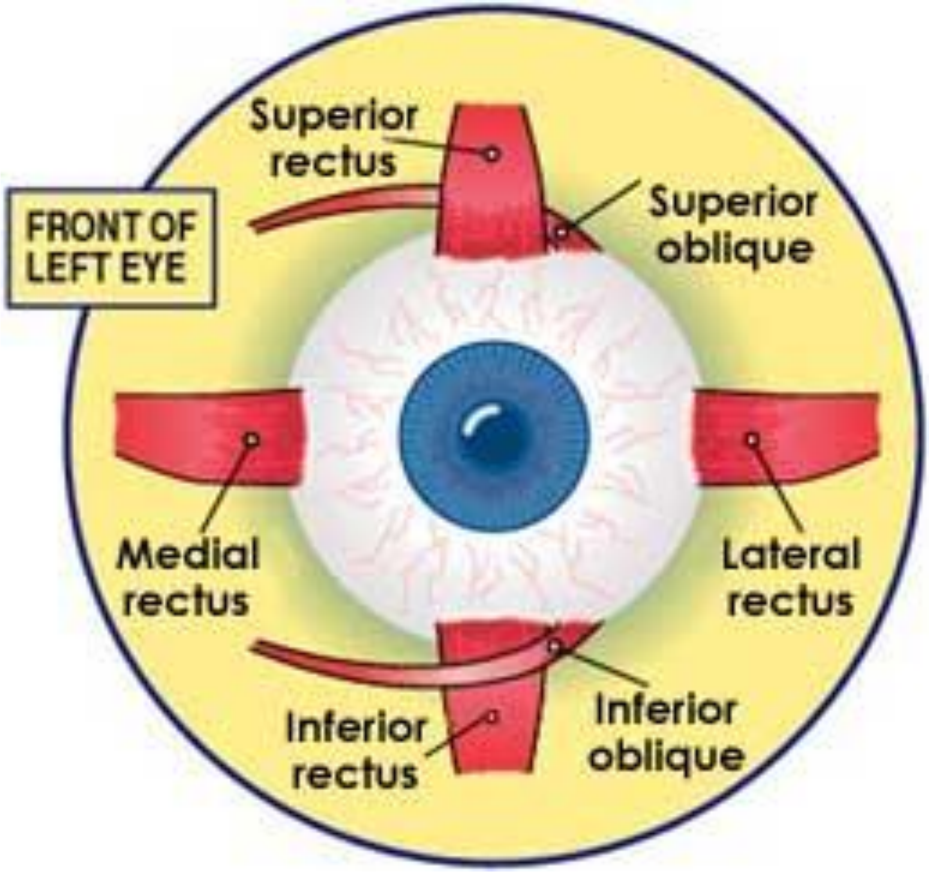
Recall that motion cells, in V1, sense motion of lines.

These cells are best tuned to lines
moving perpendicular to the line's orientation.

Eye movements

1. Saccades
2. Vergence
3. Pursuit
4. Vestibulo-Ocular Reflex (VOR)
5. Opto-Kinetic Response (OKR)

Eye movements



Eye Muscles

Saccades

If an image appears to the side,
eye movements called saccades rotate both eyes
so that the image now falls on the fovea.
Saccades are what you are using now
to point the fovea at each word in this sentence.

Because vision is poor during saccades,
saccades are very fast, up to 500 degrees per second.

Saccades

- Quick “jumps” that connect fixations
- Duration is typically between 30 and 120 ms
- Very fast (up to 700 degrees/second)
- Saccades are ballistic, i.e., the target of a saccade cannot be changed during the movement
- Vision is suppressed during saccades to allow stable perception of surroundings
- Saccades are used to move the fovea to the next object/region of interest

Example of 10° saccade (1000 samples per second)

The subject executes a saccade 150 ms after the target shift and complete the saccadic movement in 200 ms

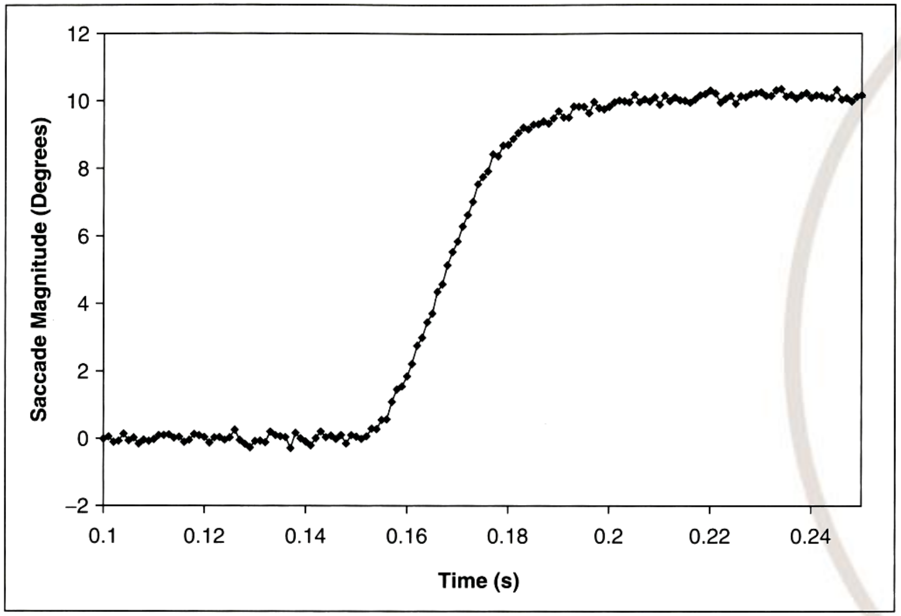


Figure 12.10 Sample saccadic eye movement of approximately 10 degrees. Data was collected with a sampling rate of 1000 samples/s.

Latency is about 150 ms and it is thought to be the time required by the CNS (Central Nervous System) to determine whether to execute the saccades, to calculate the shift and to transform the retinic error in muscle activity

Vergence

If you look (i.e. direct the foveas) from a far object to a near one, vergence eye movements are generated, convergence when looked from far to near and divergence when looking from near to far.

How do saccadic and vergence eye movements differ?

Notice that vergence movements are much slower than saccades. Also during saccades both eyes rotate in the same direction. During vergence, they rotate in opposite dirctions.

Vergence

- Slow, smooth movements changing the vergence angle (the angle between the two viewing axes)
- Used for changing gaze from a near to a far object or vice versa
- Can take up to 1 second
- Execution is often interrupted if no thorough inspection of the object is required

Smooth pursuit

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

When an object moves,
the image is kept still on the retina by means of a pursuit eye movement
(e.g. tracking a moving ball or your finger).

- Smooth movement of the eyes for visually tracking a moving object
- Cannot be performed in static scenes (fixation/saccade behavior instead)

→ *prediction*

VOR (Vestibulo-Ocular Reflex)

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

If we move our head, an eye movement very similar to pursuit is elicited whose function is also to keep the image still on the retina.
However, in spite of the fact that the movement looks similar, it is generated by a different neural circuit, the vestibular ocular reflex (VOR).

The VOR responds much faster than the pursuit system. Notice that you can read a page of text while you shake your head quickly from side to side.
To activate the pursuit system, take a page of text and try reading it while you shake the page quickly from side to side.

Also unlike the pursuit system, the VOR does not need a visual stimulus.
It works in the dark. Rotate your head with your eyes closed.
Feel your eyes move with your finger tips.

VOR (Vestibulo-Ocular Reflex)

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

- Reflex eye movement that stabilizes images on the retina during head movement by producing an eye movement in the direction opposite to head movement, thus preserving the image on the center of the visual field.
- Since slight head movements are present all the time, the VOR is very important for stabilizing vision: patients whose VOR is impaired cannot read, because they cannot stabilize the eyes during small head tremors
- The VOR reflex does not depend on visual input and works even in total darkness or when the eyes are closed
- Latency of 14 ms (time between the head and the eye movement)

OKR (OptoKinetic Response)

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

The VOR does not work well for slow prolonged movements.
In this case vision, through the optokinetic response (OKR), assists the VOR.
The OKR is activated when the image of the world slips on a large portion of the retina and produces a sense of self motion (e.g. when sitting in a car that is stopped and a car beside you starts to move, you sometimes feel like you are moving).

OKR (OptoKinetic Response)

- The optokinetic reflex allows the eye to follow objects in motion when the head remains stationary
- The OKR is activated when the image of the world slips on a large part of the retina
- This reflex is based on the visual information
- The latency is longer than in VOR

Other Eye Movements

Torsional Eye Movements:

- Rotation of the eye around the viewing axis
- Stabilization of visual scene by compensating body rotation (up to about 15 degrees)

Other Eye Movements

Fixations:

- The eye is almost motionless, for example, while reading a single, short word
- The information from the scene is almost entirely acquired during fixation
- Duration varies from 100-1000 ms, typically between 200-600 ms
- Typical fixation frequency is about 3 Hz
- Fixations are interspersed with saccades

Other Eye Movements

Tremor:

- Fast, low-amplitude (seconds of arc) eye-movement “jitter”
- Improves the perception of high spatial frequencies
- Prevents the fading of static images during fixations

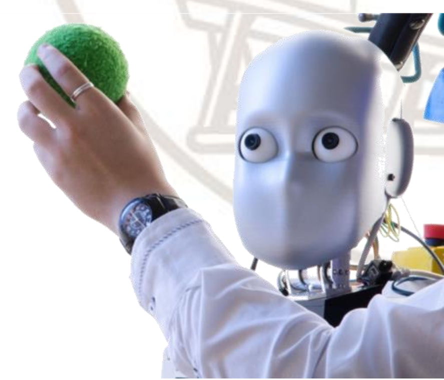
Summary

- Human vision is based on spatio-variant resolution images, formed by a non-uniform retina
- Retina-like images reduce data amount, by performing some early processing of images
- This strategy requires eye movements

Master in Bionics Engineering
University of Pisa and Scuola Superiore Sant'Anna
Human and Animal Models for BioRobotics

Retina-like vision and Eye Movements in Robots

Cecilia Laschi
The BioRobotics Institute
cecilia.laschi@santannapisa.it



Basic principles of retina-like vision

Standard image

Retina-like image

Log-polar image (magnified to 200% for display)

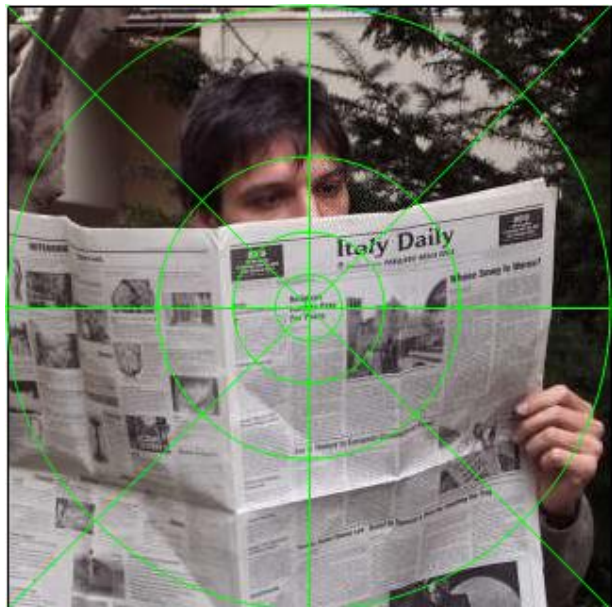
Log-polar projection

THE BIOROBOTICS INSTITUTE

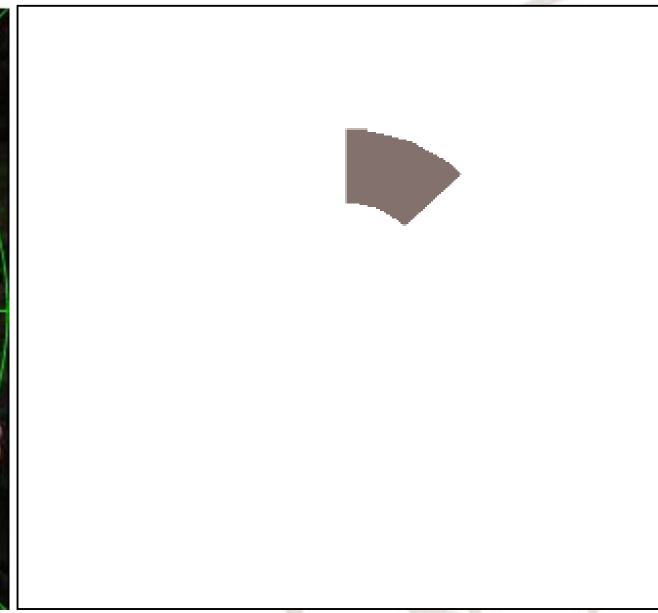
Scuola Superiore Sant'Anna

Building a retina-like image

Cartesian image

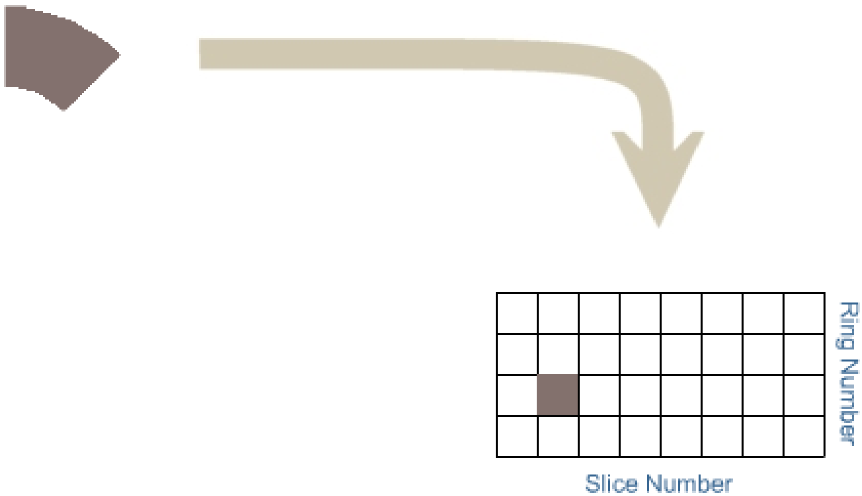


Cutting in circles and slices



Computing the average value in each sector

Building a retina-like image



Copying the average value of a sector in a polar image

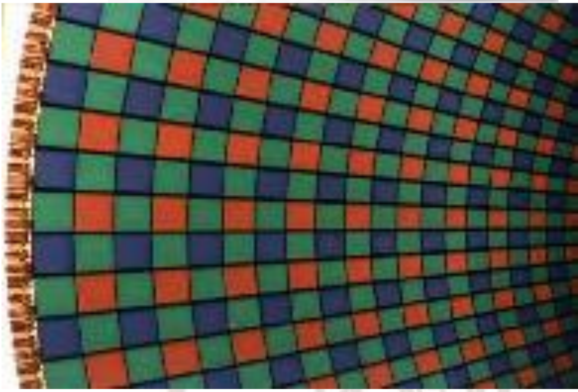
Resulting polar image

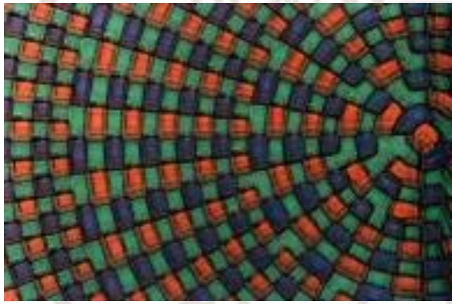
Cartesian image re-built from the polar image

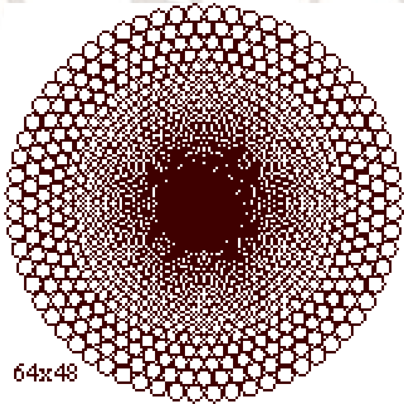
THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

The Retina-like Giotto cameras

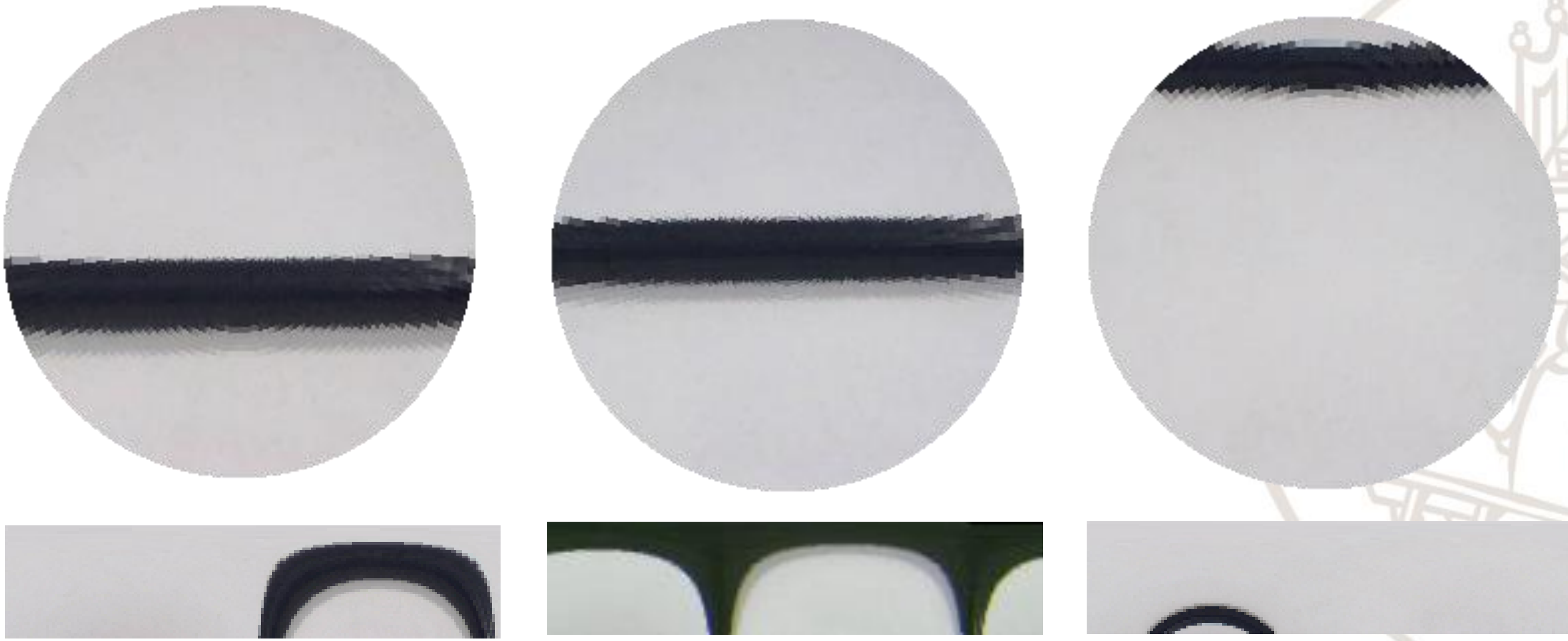




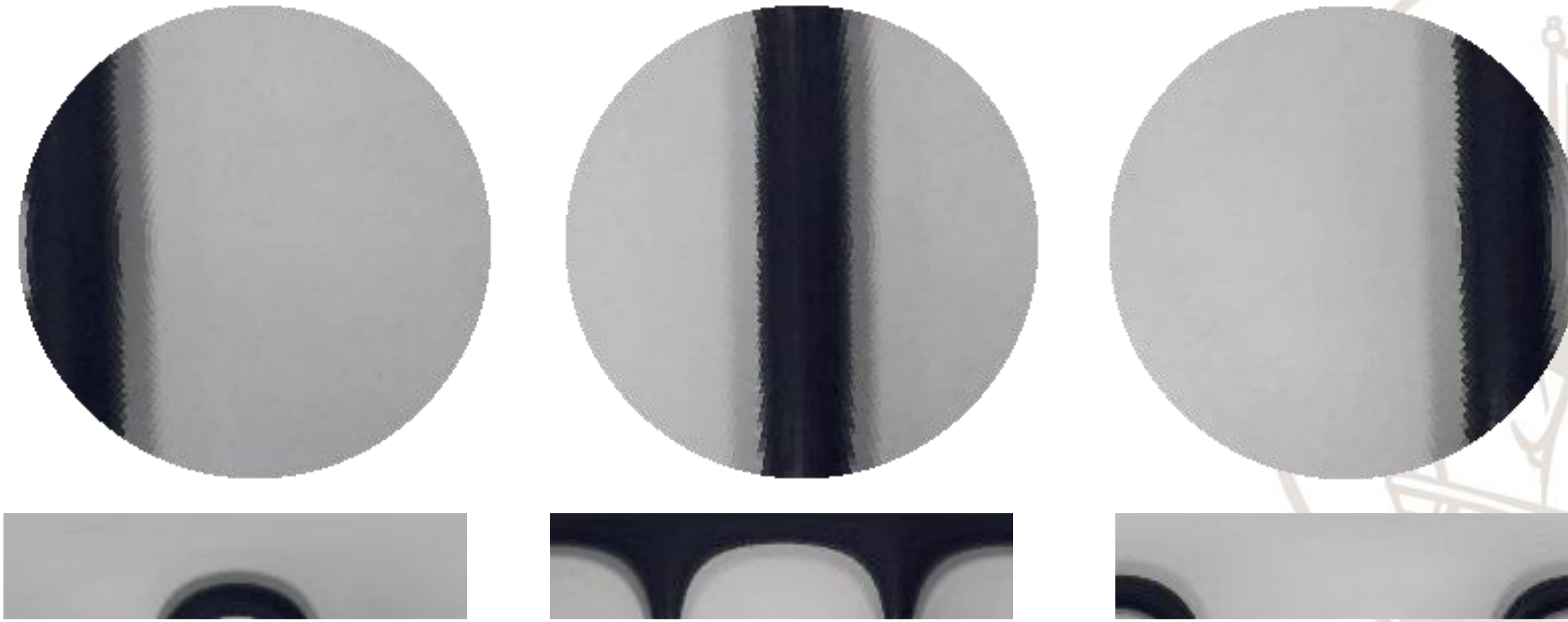


- Technology: 0.35 micrometer CMOS
- Total Pixels: 33193
- Geometry:
  - 110 rings with 252 pixels
  - 42 rings with a number of pixels decreasing toward the center with a "sunflower" arrangement
- Tessellation: pseudo-triangular
- Pixels: direct read-out with logarithmic response
- Size of photosensitive area: 7.1mm diameter
- Constant resolution equivalent: 1090x1090
- On-chip processing: addressing, A/D, output amplifier

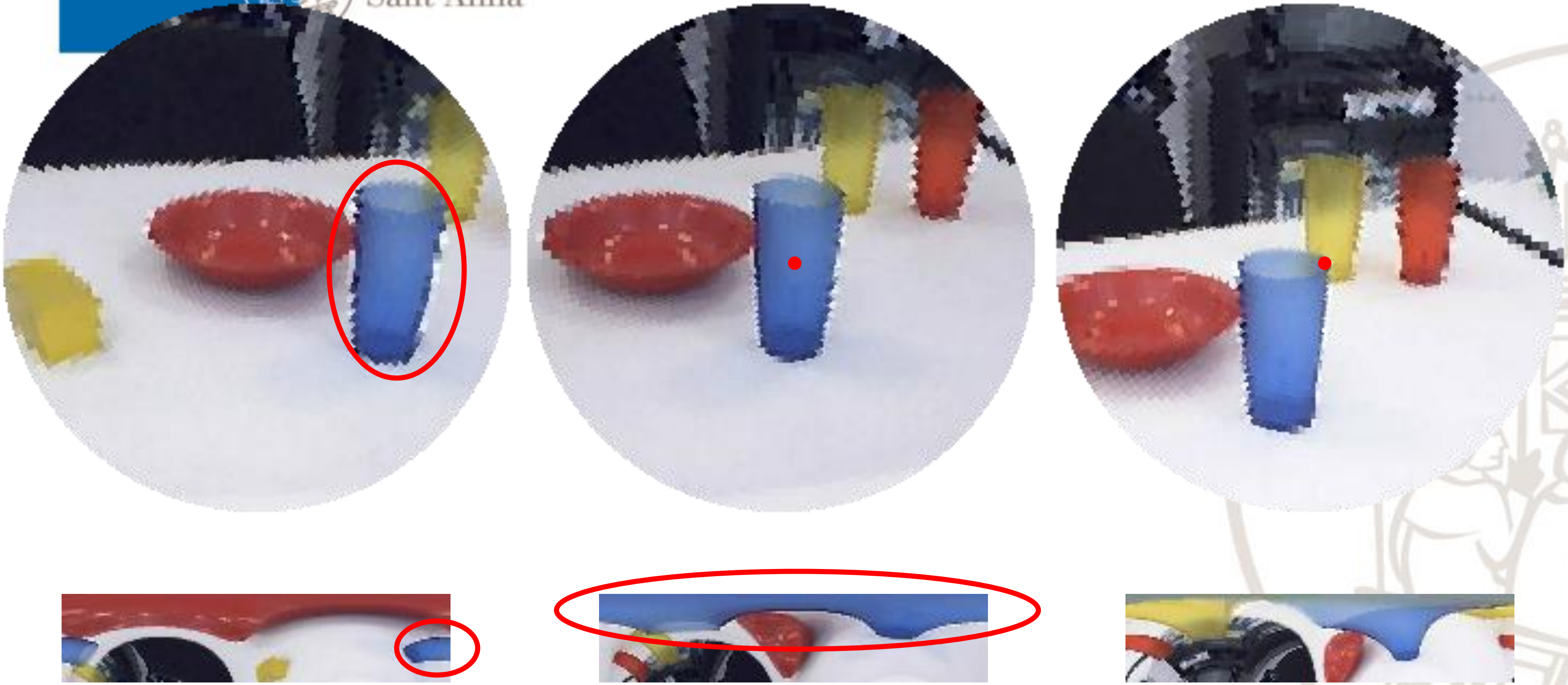
An example of pattern translation



An example of pattern translation



An example of simulated foveation

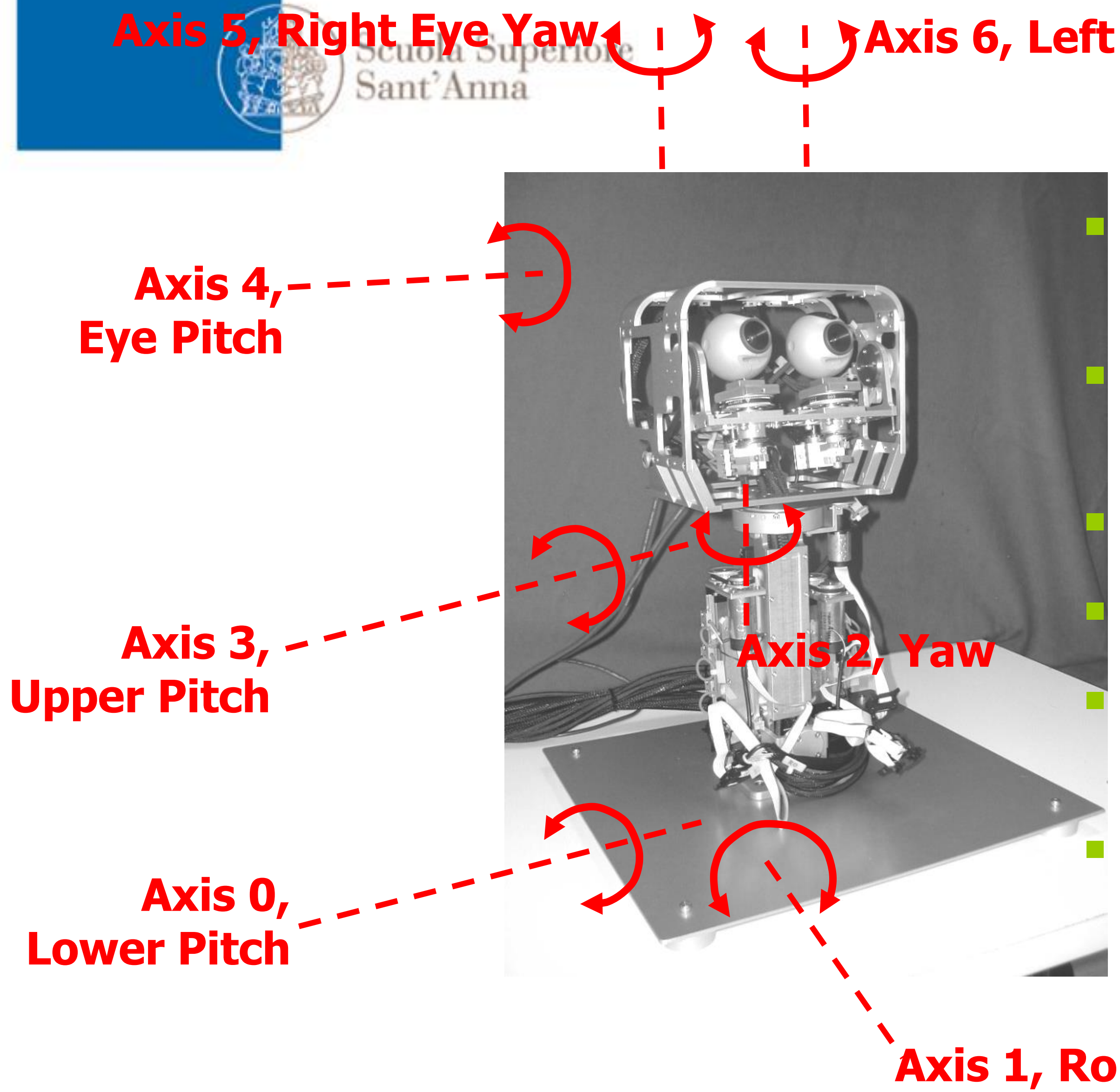


Object detection in the periphery

Object foveation

Foveation of a point of interest (edge)

Kinematic structure of the SSSA Robot Head



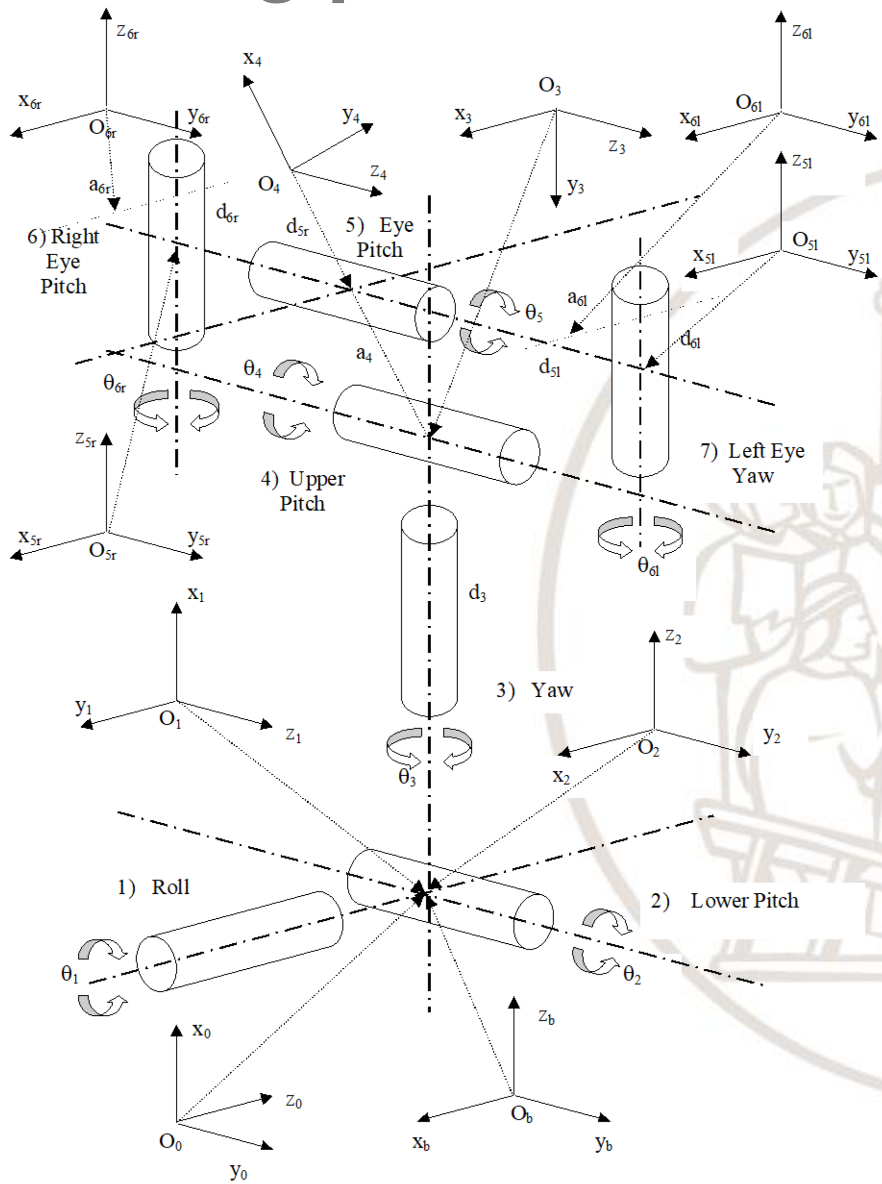
- Eye Pitch Axis: ±47°, 600°/s
- Eye R/L Yaw Axis: ±45°, 1000°/s
- Yaw: ±100°, 170°/s
- Roll: ±30°, 25°/s
- Upper Pitch: ±30°, 120°/s
- Lower Pitch: ±25°, 20°/s

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

Head kinematic chain and Denavit-Hartenberg parameters

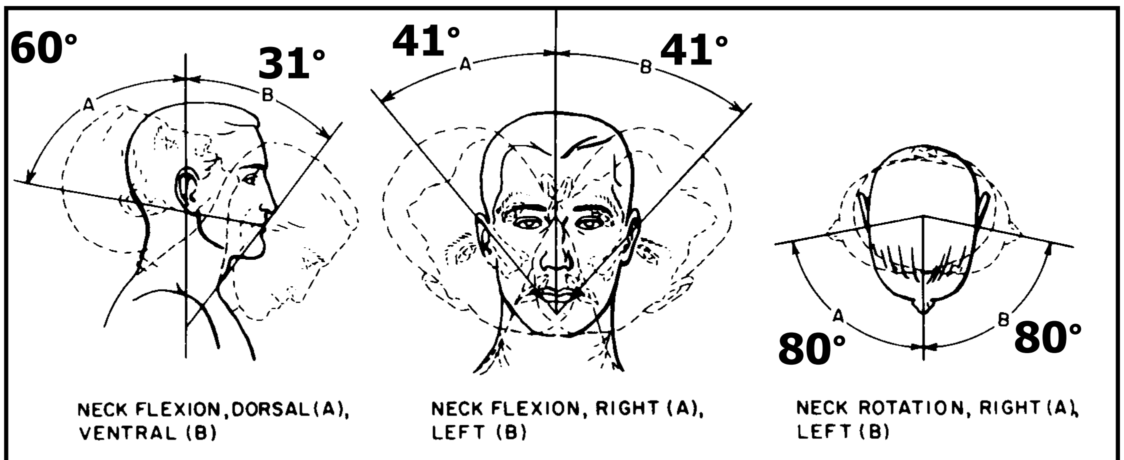
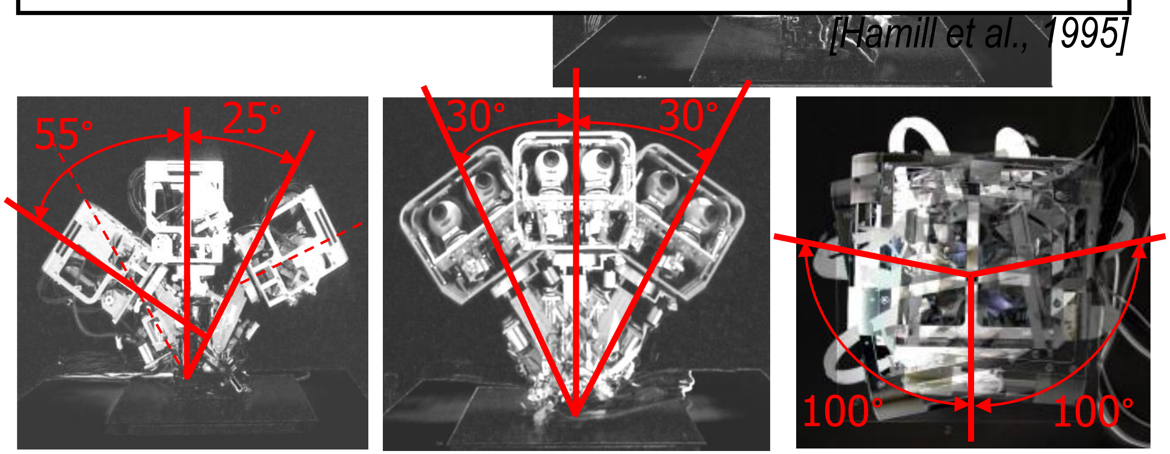
Joint	a_i (mm)	d_i (mm)	α_i (rad)
J1	0	0	$-\pi/2$
J2	0	0	$\pi/2$
J3	0	195	$-\pi/2$
J4	137.5	0	0
$J5_r$	0	$-30 \div -50$	$\pi/2$
$J5_l$	0	$30 \div 50$	$\pi/2$
$J6_l$	a_{6l}	d_{6l}	0
J6r	a_{6r}	d_{6r}	0



Comparison of performances between human and robotic head

THE BIOROBOTICS INSTITUTE

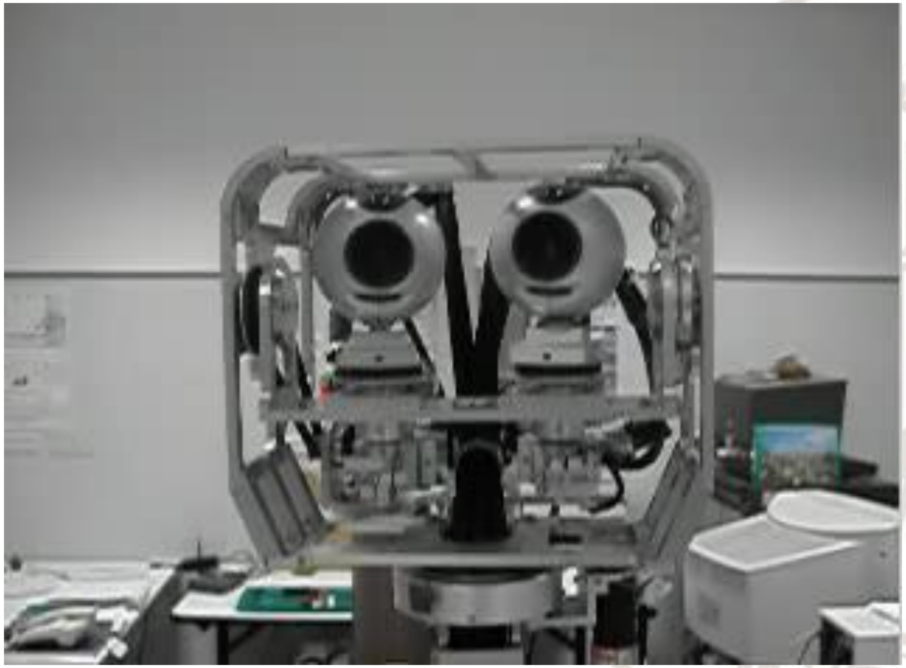
Scuola Superiore Sant'Anna

	Neck:	Eye:
Human		Pitch: ± 60°, 600°/s Yaw: ± 30°, 600°/s
Robot		Pitch: ± 47°, 600°/s Yaw: ± 45°, 1000°/s

[Hamill et al., 1995]

The movements of the 7 dofs of the robotic head

Neck Movements



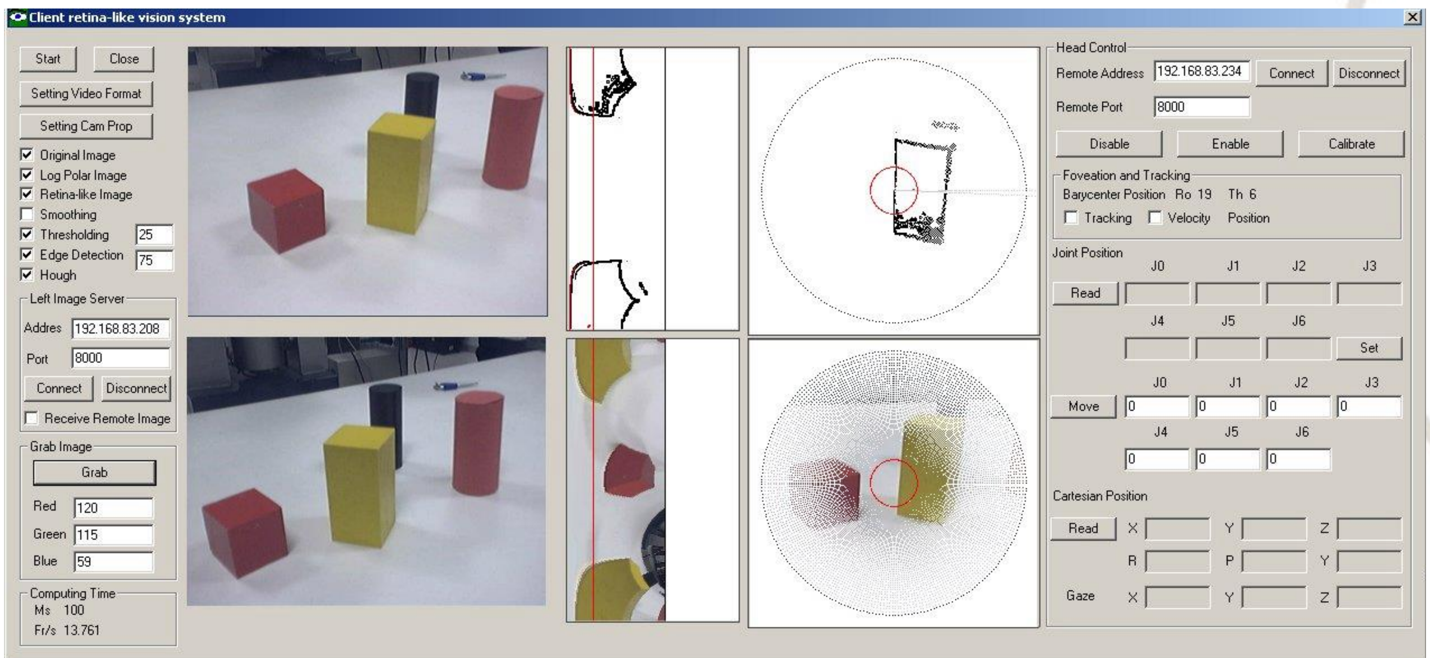
Saccades, 400°/sec

THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

Simulation of retina-like cameras and basic image processing

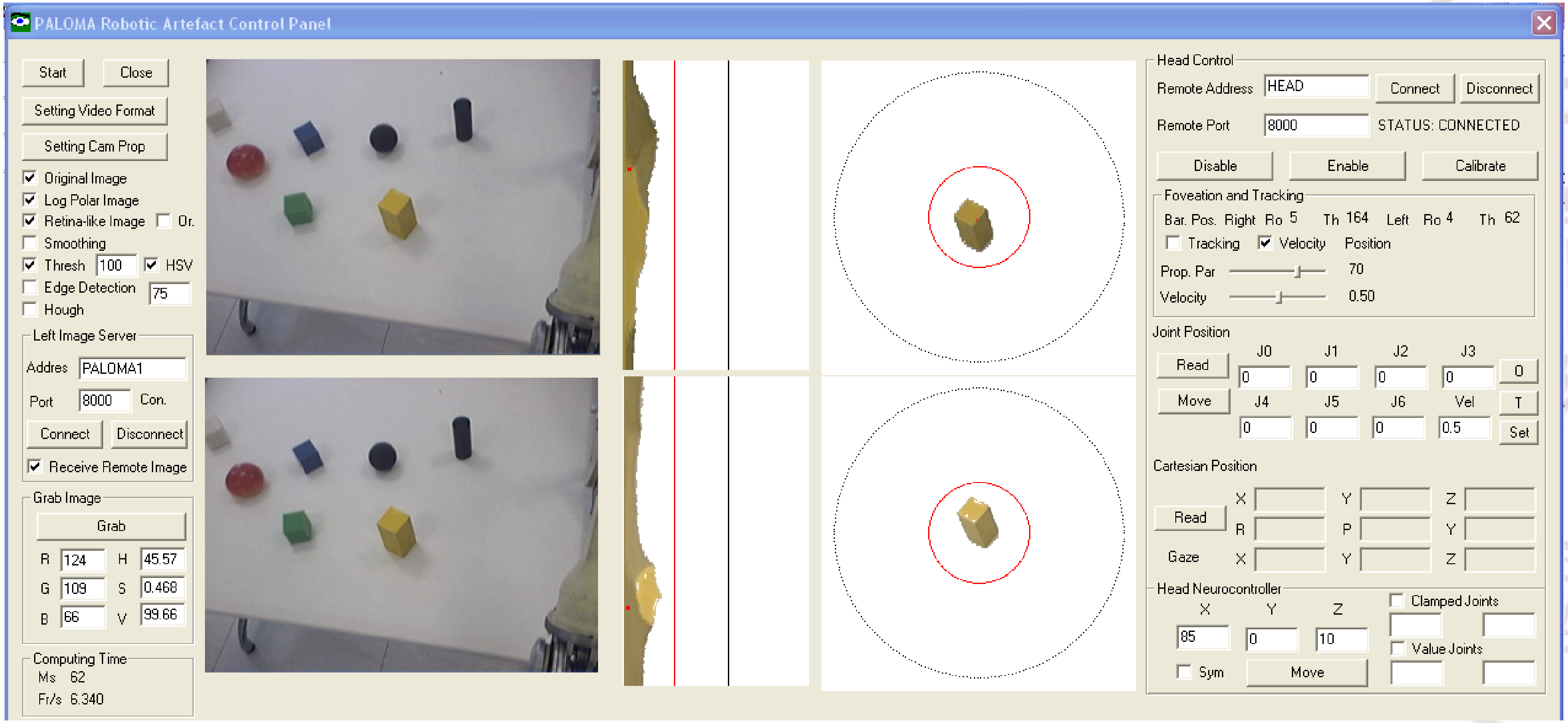
- Acquiring standard image
- Creating log-polar image from standard image
- Creating retina-like image from log-polar image



Thresholding of image based on RGB and HUE

THE BIOROBOTICS INSTITUTE

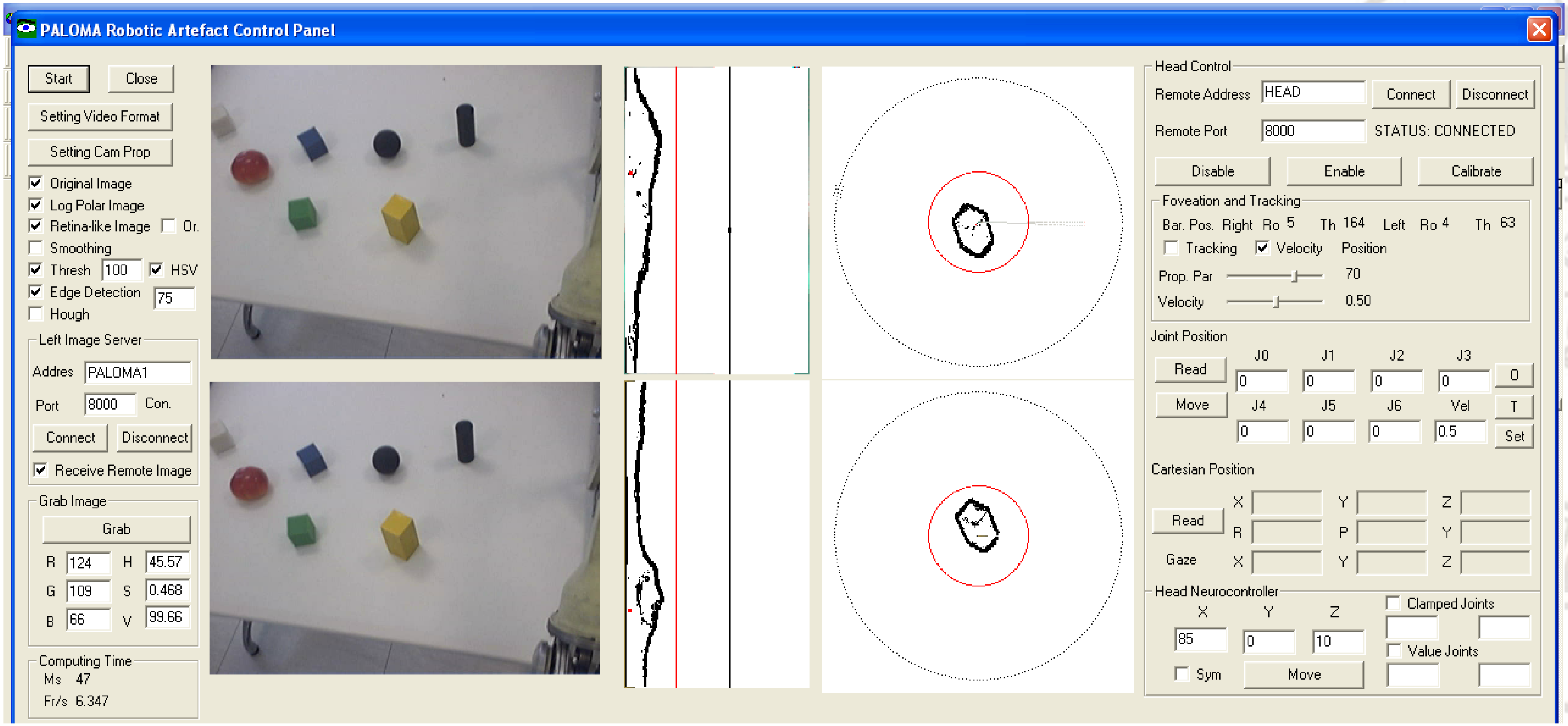
Scuola Superiore Sant'Anna



THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

Edge Detection (gradient based method)

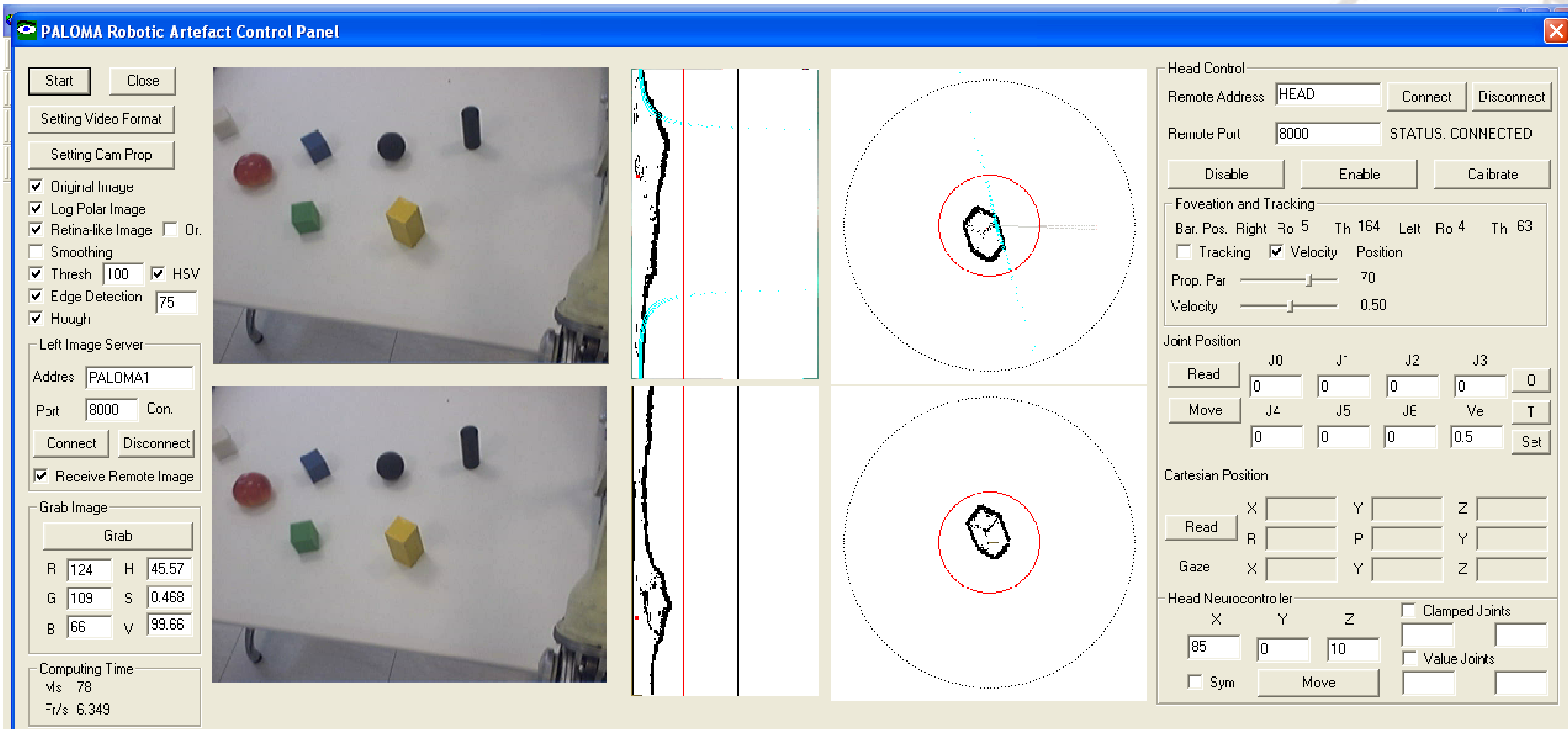


THE BIOROBOTICS INSTITUTE

Scuola Superiore Sant'Anna

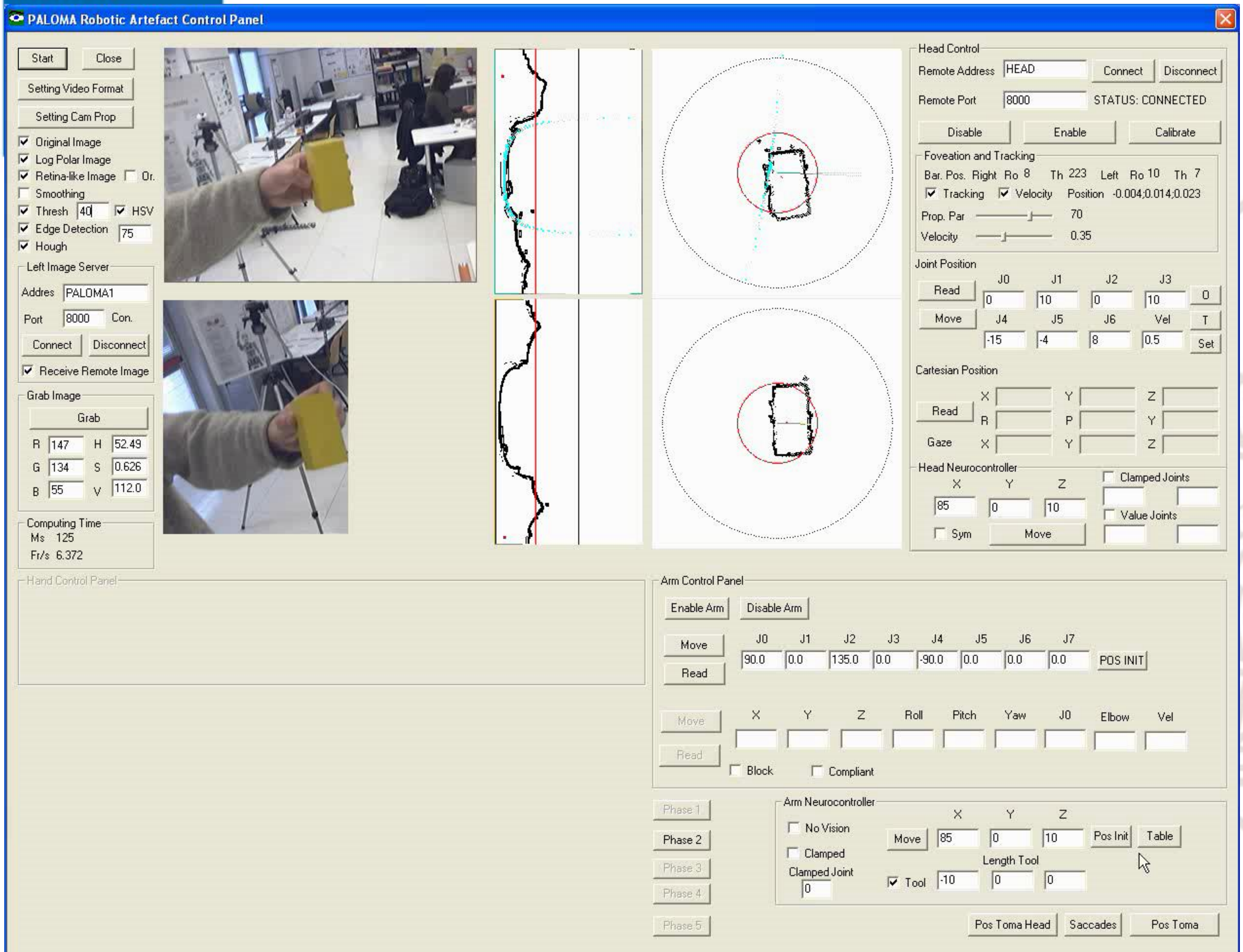
Line detection (Hough method)

- Applied only to pixels belonging to the fovea



THE BIOROBOTICS INSTITUTE

Line detection



Summary

- In robotics, retina-like vision can reduce image data amount and speed up image processing
- This strategy requires eye movements

  -> active perception
- Image processing techniques are applied on a different geometry
  - No difference for local techniques
  - Most global techniques not applicable