

# Corso di Robotica (ROB)



## C. Modulo di Percezione

### **Visione artificiale retinica**

Cecilia Laschi

Istituto di BioRobotica, Scuola Superiore Sant'Anna

[cecilia.laschi@sssup.it](mailto:cecilia.laschi@sssup.it)

050-883486

# Sommario della lezione



- Principi di base della visione retinica
- Alcune proprietà delle immagini retiniche
- Le relazioni matematiche tra immagini retiniche e cartesiane
- La foveazione
- Una testa robotica antropomorfa
- Esempi di applicazione in robotica

*Riferimenti bibliografici:*

*G. Sandini, G. Metta, "Retina- like sensors: motivations, technology and applications". in Sensors and Sensing in Biology and Engineering. T.W. Secomb, F. Barth, and P. Humphrey, Editors. Springer-Verlag. 2002.*

# Principi di base della visione retinica

Standard image



Retina-like image



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



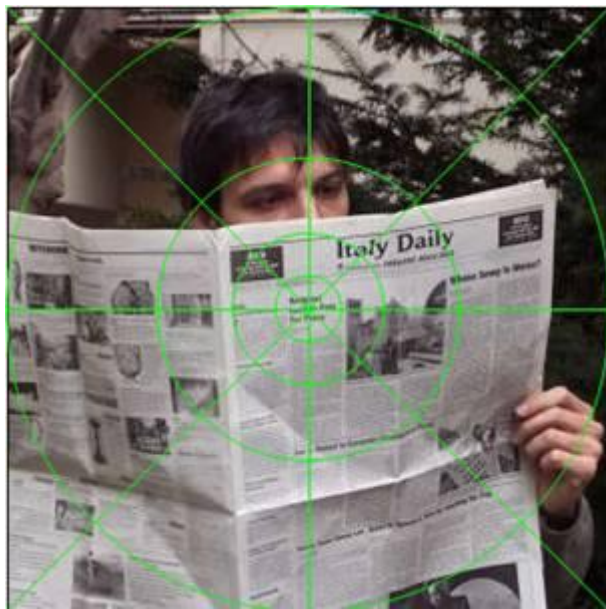
Log-polar projection



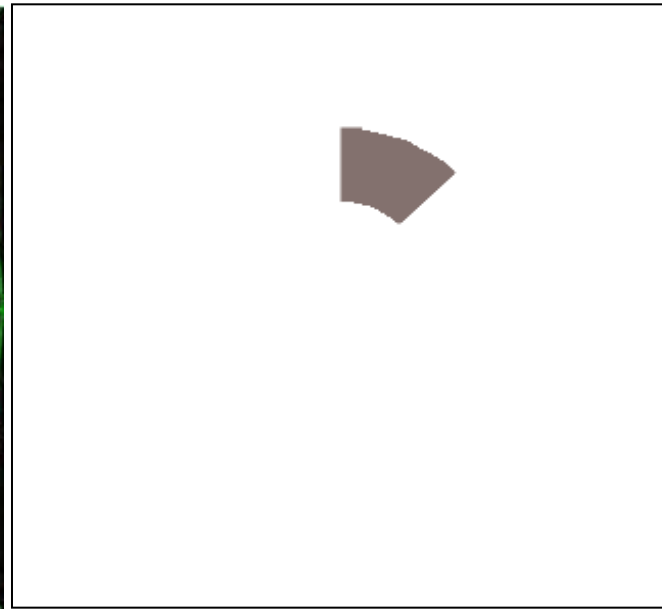
# Costruzione di un'immagine retinica



Immagine cartesiana  
tradizionale



Divisione in  
circonferenze e spicchi



Calcolo del valore  
medio di un settore

# Costruzione di un'immagine retinica



Copia del valore medio di un settore  
in un pixel di un'immagine polare

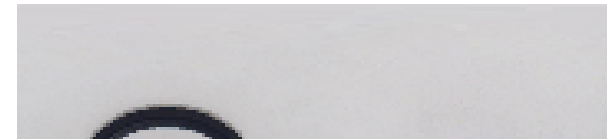
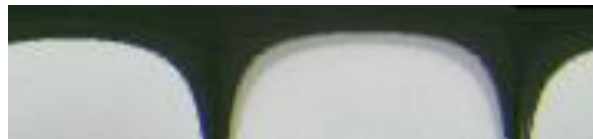
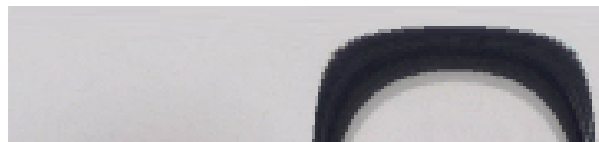
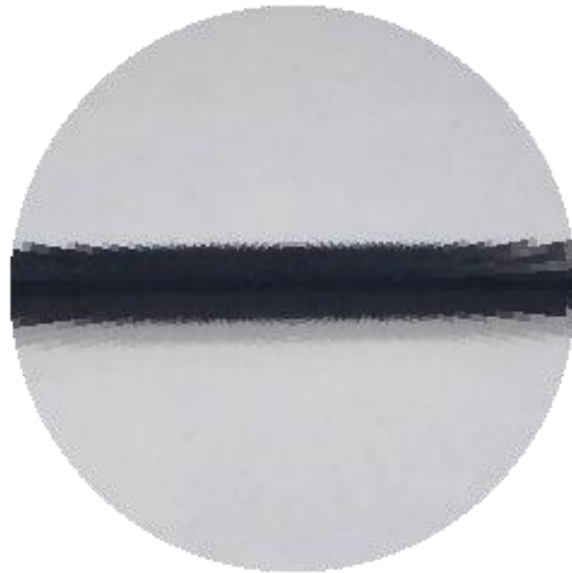


Immagine polare risultante

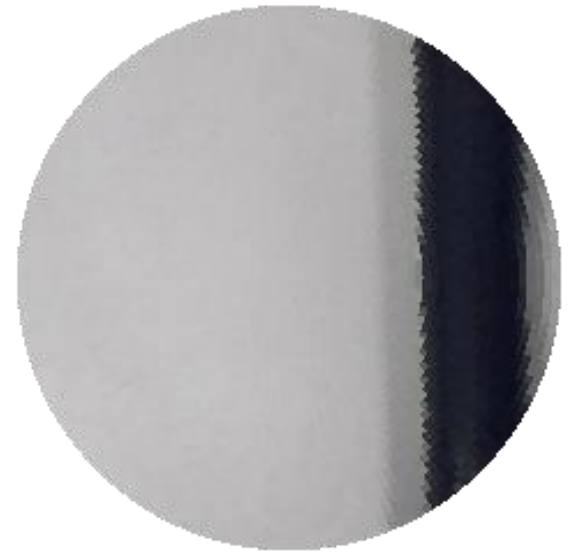
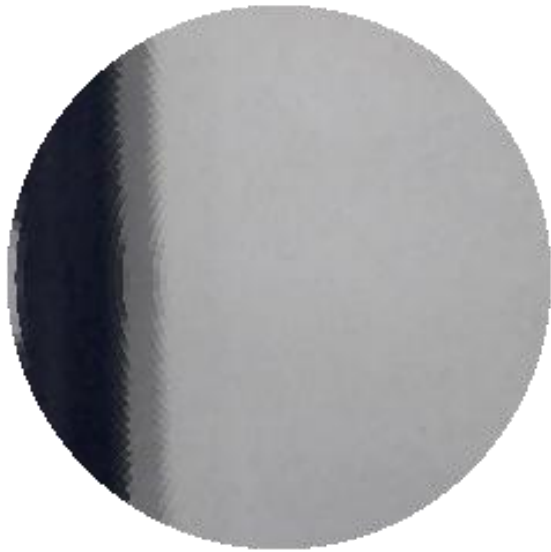


Immagine cartesiana  
ricostruita dalla polare

# An example of pattern translation



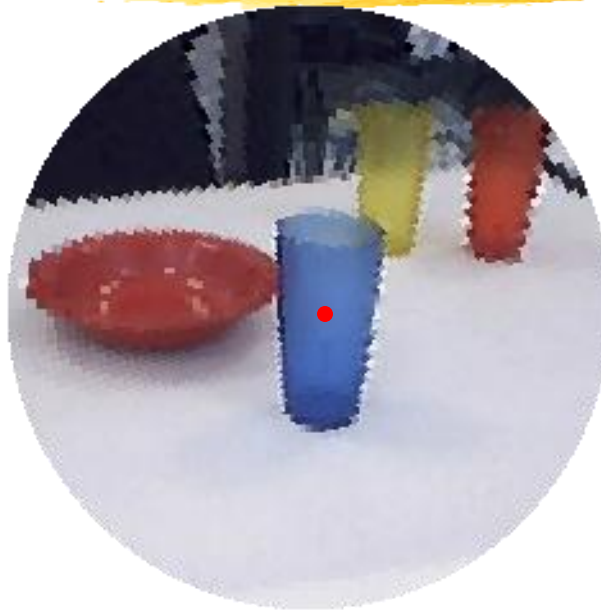
# An example of pattern translation



# An example of simulated foveation



Object detection  
in the periphery



Object foveation



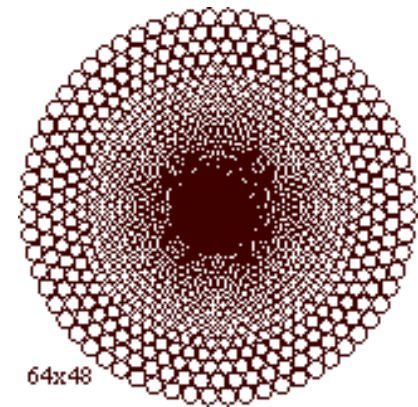
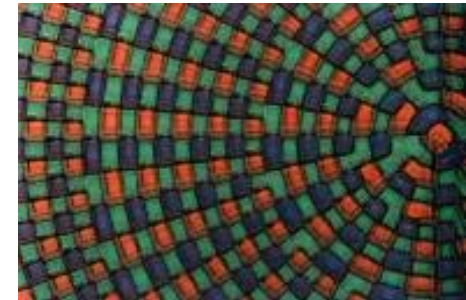
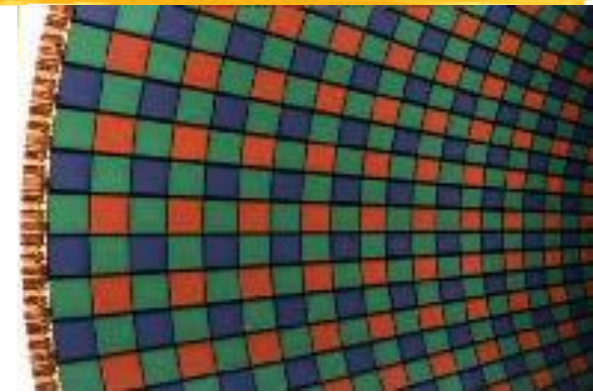
Foveation of a  
point of interest  
(edge)





# 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



# Le relazioni matematiche

## From standard image to log-polar image

$$\rho(x, y) = \begin{cases} (F - 1) + \log_{\lambda} \left[ \left( F - \frac{1}{2} - \sqrt{x^2 + y^2} \right) (1 - \lambda) + \lambda \right] & \text{if } \sqrt{x^2 + y^2} > (F - \frac{1}{2}) \\ \left( \sqrt{x^2 + y^2} + \frac{1}{2} \right) & \text{if } \sqrt{x^2 + y^2} < (F - \frac{1}{2}) \end{cases}$$

$$r(\rho) = \left[ \left( F - \frac{1}{2} \right) + \lambda \frac{1 - \lambda^{\rho - F}}{1 - \lambda} \right] \text{ if } \rho > F$$

$$\theta(x, y) = \frac{\Theta}{2\pi} \cdot \arctan\left(\frac{y}{x}\right) + \frac{\Theta}{2} + \text{Shift Factor}$$

$F=42$   
 $P=152$   
 $\Theta=252$   
 $X=545$   
 $Y=545$   
 $\lambda=1.02314422608633$

$F$  = size of the fovea in rings.

$P$  = total number of rings.

$\Theta$  = maximum # of pixels in each ring.

$2X$  = horizontal size of the cartesian image.

$2Y$  = vertical size of the cartesian image.

$\rho$  = ring number in the log polar image.

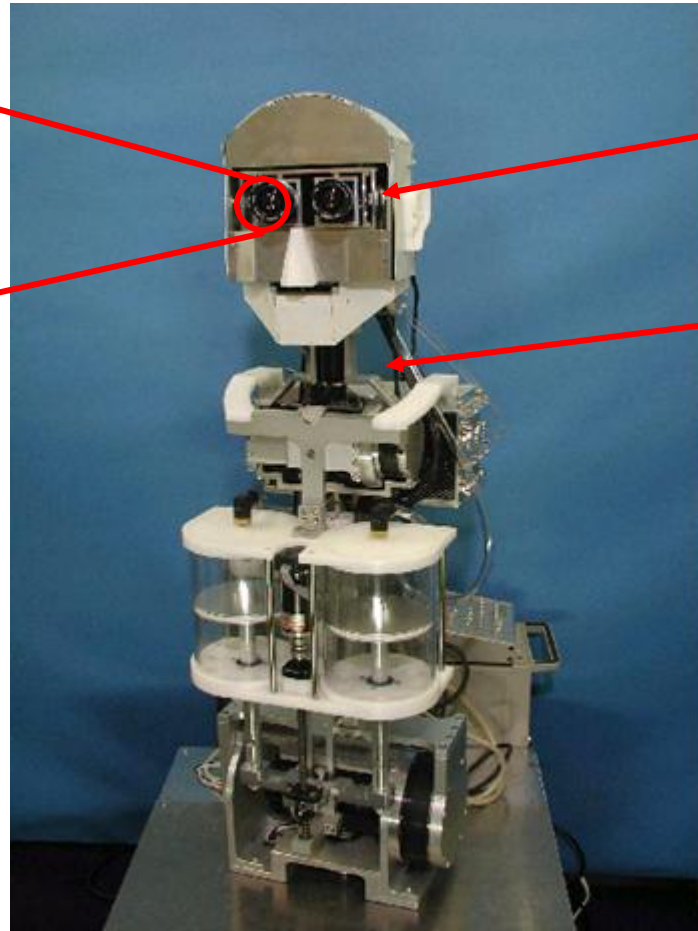
$\theta$  = angular polar coordinate.

# Retina-like vision for visuo-motor co-ordination of a robot head

## WE-4 robotic head with Giotto cameras



*Retina-like  
Giotto cameras  
by the  
University of  
Genova, Italy*



3 dof for eye  
movements

4 dof for  
neck  
movements

*WE-4 robotic head by  
Takanishi Lab, Waseda  
University, Tokyo, Japan*

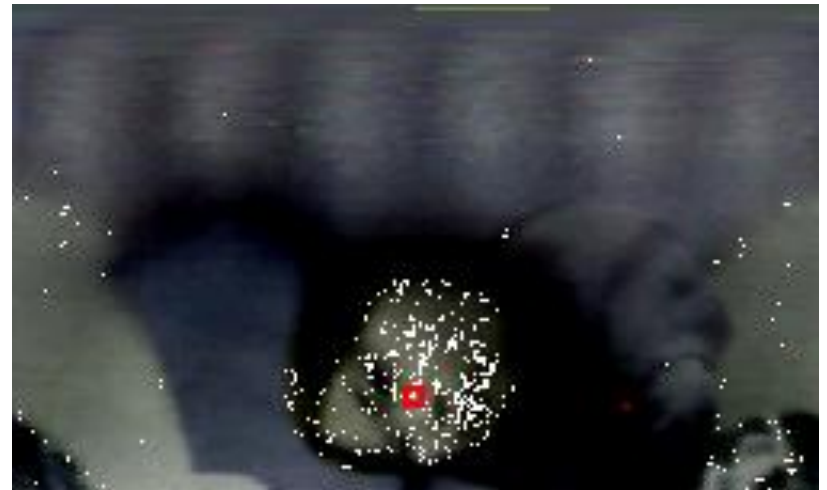
# Face detection by hue

Hue = information on the color

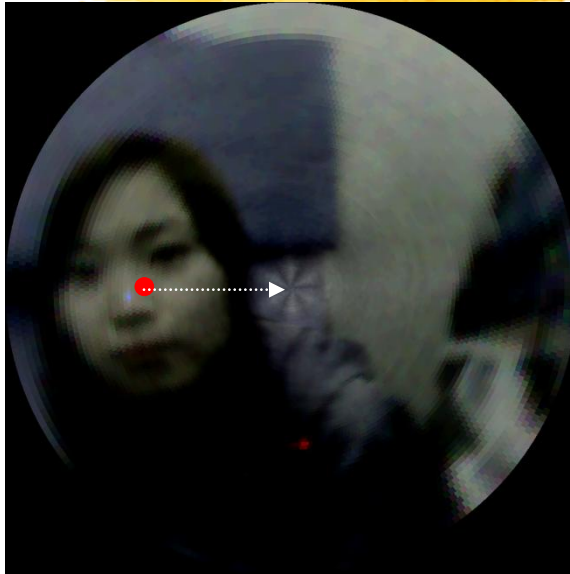
$$\text{Hue} = \cos^{-1} \left( \frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \right)$$

if  $B > G$  then  $\text{Hue} = 2\pi - \text{Hue}$

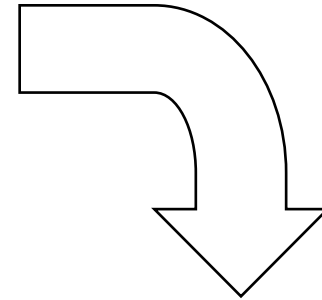
R, G, B = RED, GREEN, BLUE components, respectively



# An example of foveation

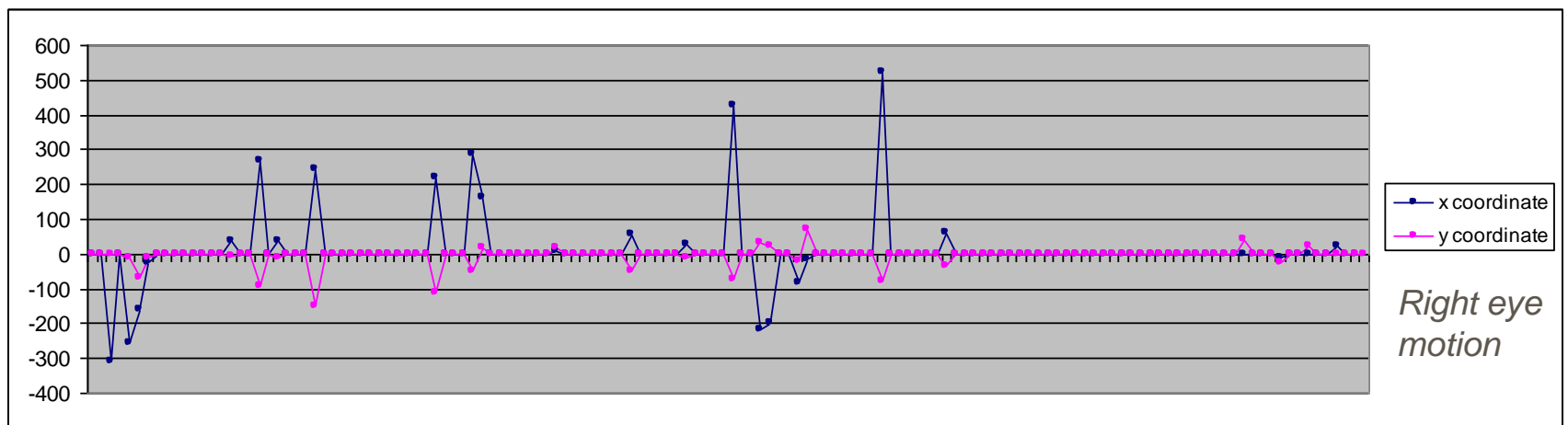


Eye/neck movements



*Proportions are rescaled for display purposes*

# Experimental trials



[Cecilia Laschi, Hiroyasu Miwa, Atsuo Takanishi, Eugenio Guglielmelli, Paolo Dario, 2002]

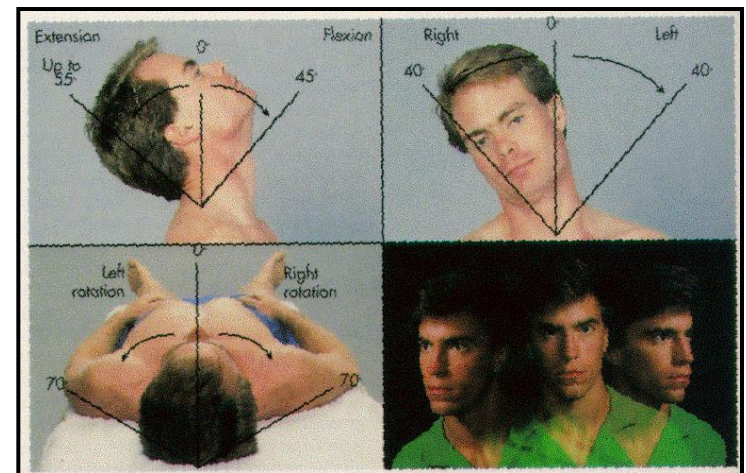
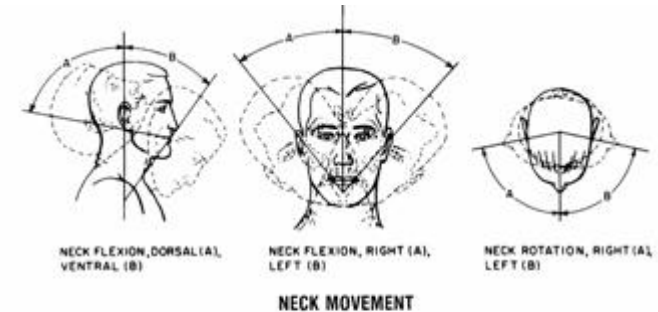
Example of design and development of  
a human-like robotic head



**The ARTS humanoid robot  
head**

# Synthesis of characteristics of the human oculo-motor system

- Eye movements:
  - Saccades
  - Vergence
  - Pursuit
- Ranges of motion:
  - 120° for the tilt eye movements
  - 60° for the pan eye movements
- Eye speed:
  - Up to 900°/sec (in saccades)
- Inter-ocular distance: between 60 and 80 mm



[Thibodeau & Patton, 1996]



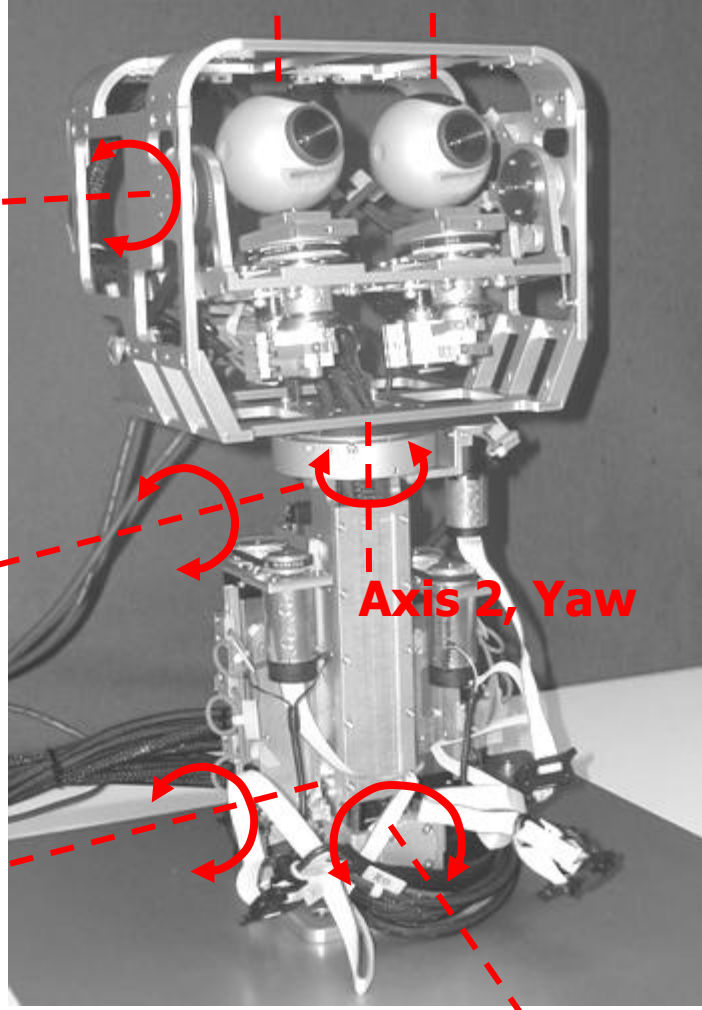
# Kinematic structure of the SSSA Robot Head

Axis 5, Right Eye Yaw      Axis 6, Left Eye Yaw

Axis 4, Eye Pitch

Axis 3, Upper Pitch

Axis 0, Lower Pitch



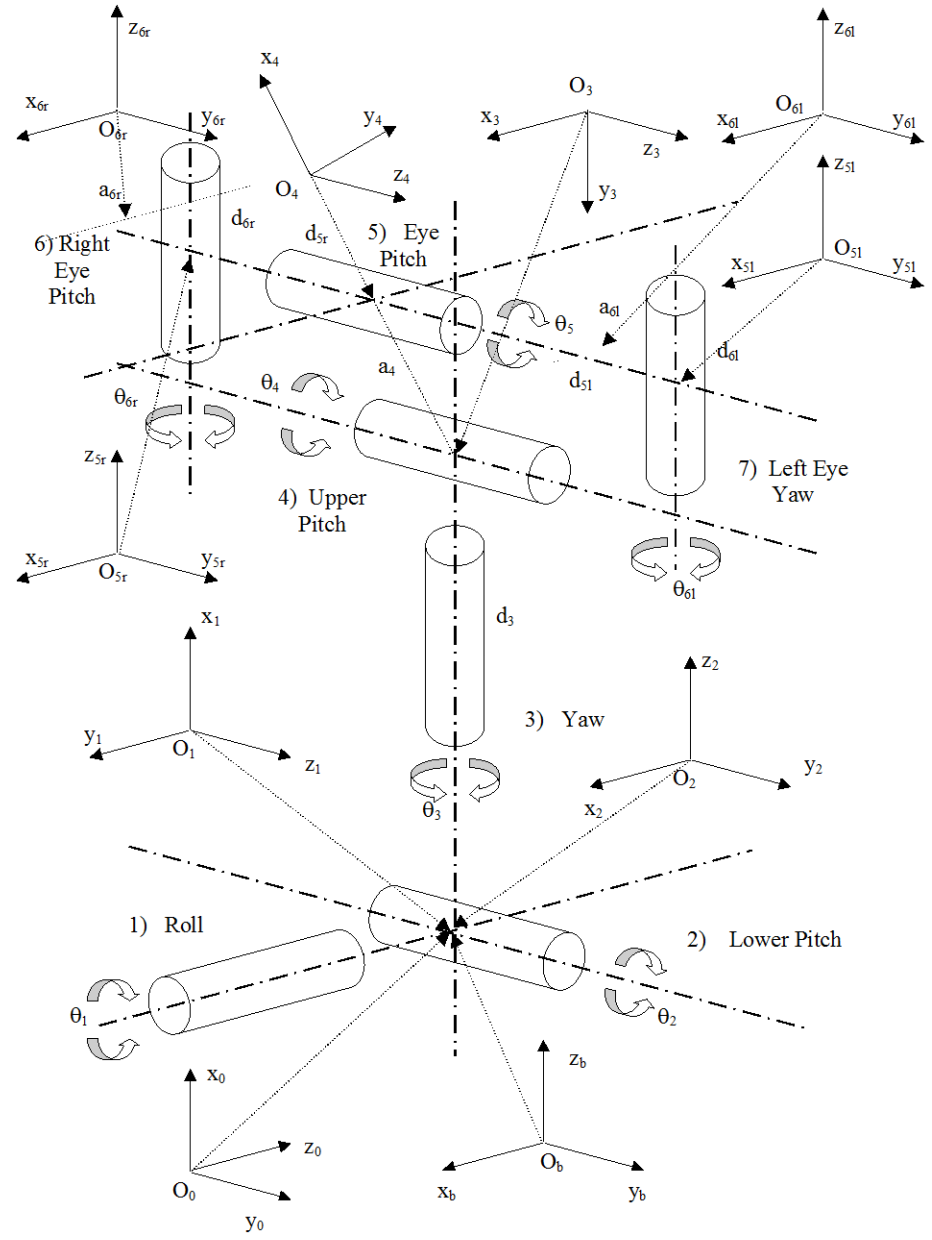
Axis 2, Yaw

Axis 1, Roll

- Eye Pitch Axis:  $\pm 47^\circ$ ,  $600^\circ/\text{s}$
- Eye R/L Yaw Axis:  $\pm 45^\circ$ ,  $1000^\circ/\text{s}$
- Yaw:  $\pm 100^\circ$ ,  $170^\circ/\text{s}$
- Roll:  $\pm 30^\circ$ ,  $25^\circ/\text{s}$
- Upper Pitch:  $\pm 30^\circ$ ,  $120^\circ/\text{s}$
- Lower Pitch:  $\pm 25^\circ$ ,  $20^\circ/\text{s}$

# Head kinematic chain and Denavit-Hartenberg parameters

Joint	$a_i$ (mm)	$d_i$ (mm)	$\alpha_i$ (rad)
J1	0	0	$-\pi/2$
J2	0	0	$\pi/2$
J3	0	195	$-\pi/2$
J4	137.5	0	0
J5 <sub>r</sub>	0	-30 ÷ -50	$\pi/2$
J5 <sub>l</sub>	0	30 ÷ 50	$\pi/2$
J6 <sub>l</sub>	$a_{6l}$	$d_{6l}$	0
J6 <sub>r</sub>	$a_{6r}$	$d_{6r}$	0



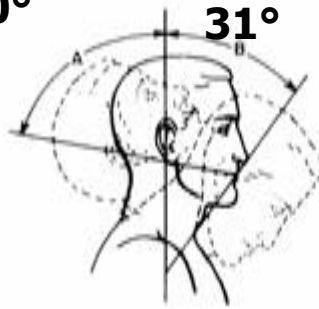
# Comparison of performances between human and robotic head

## Neck:

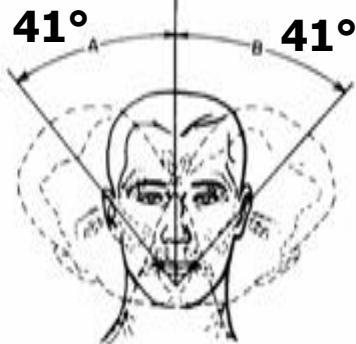
## Eye:

Human

60°

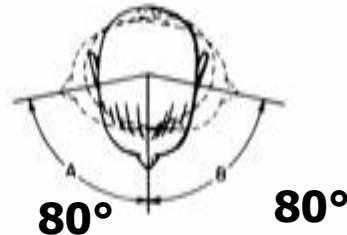


NECK FLEXION, DORSAL (A), VENTRAL (B)



NECK FLEXION, RIGHT (A), LEFT (B)

NECK MOVEMENT



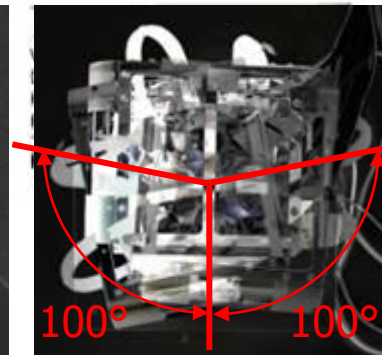
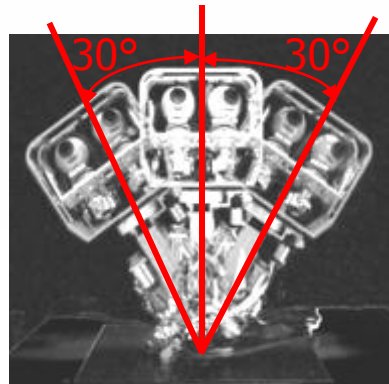
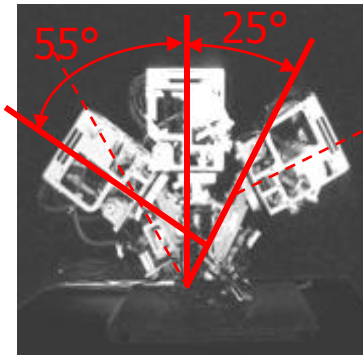
NECK ROTATION, RIGHT (A), LEFT (B)

[Hamill et al., 1995]

Pitch:  $\pm 60^\circ$ ,  $600^\circ/\text{s}$

Yaw:  $\pm 30^\circ$ ,  $600^\circ/\text{s}$

Robot



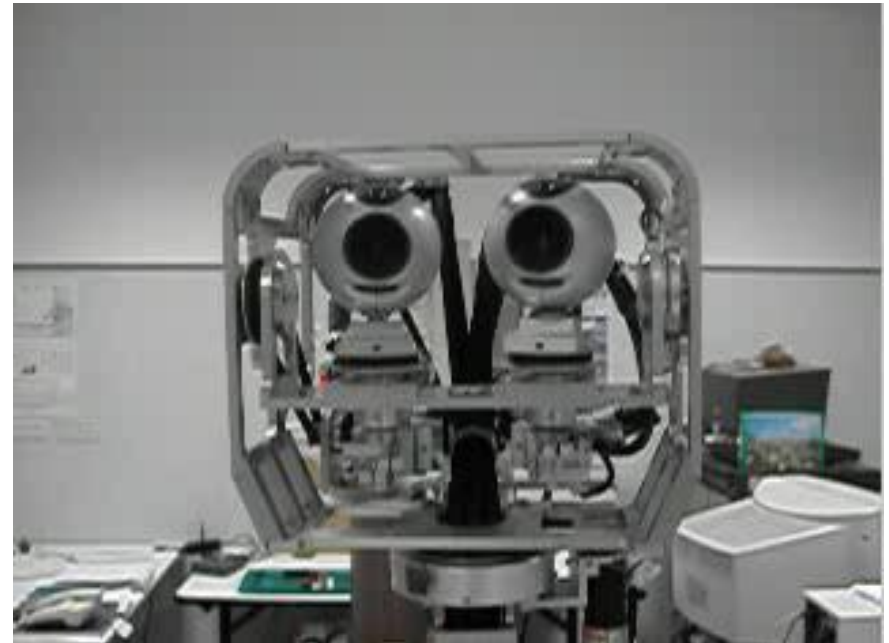
Pitch:  $\pm 47^\circ$ ,  $600^\circ/\text{s}$

Yaw:  $\pm 45^\circ$ ,  $1000^\circ/\text{s}$

# The movements of the 7 dofs of the robotic head



Neck Movements

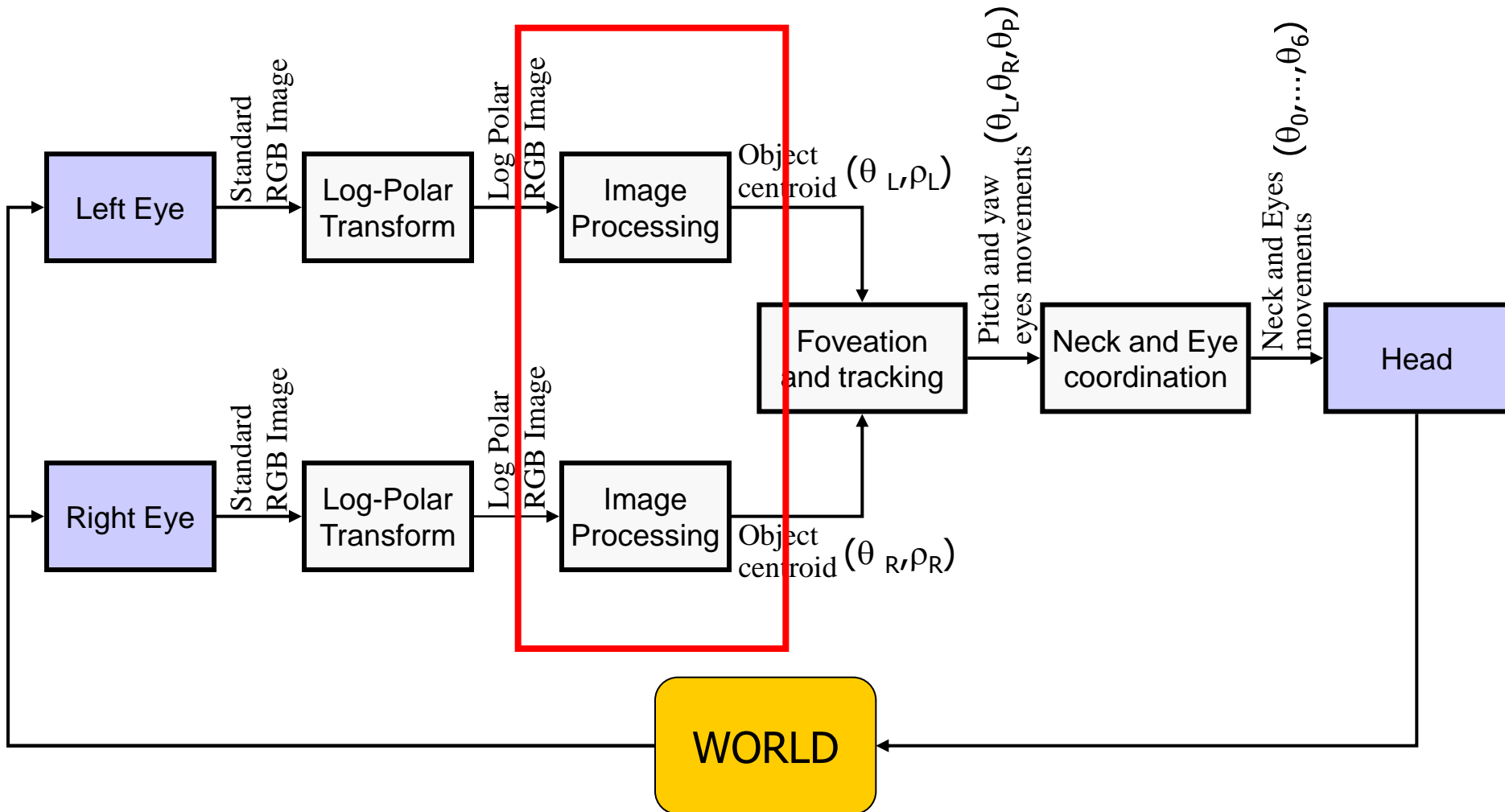


Saccades,  $400^\circ/\text{sec}$



saccade 300° sec.avi

# Overall sensory-motor scheme of the visual apparatus



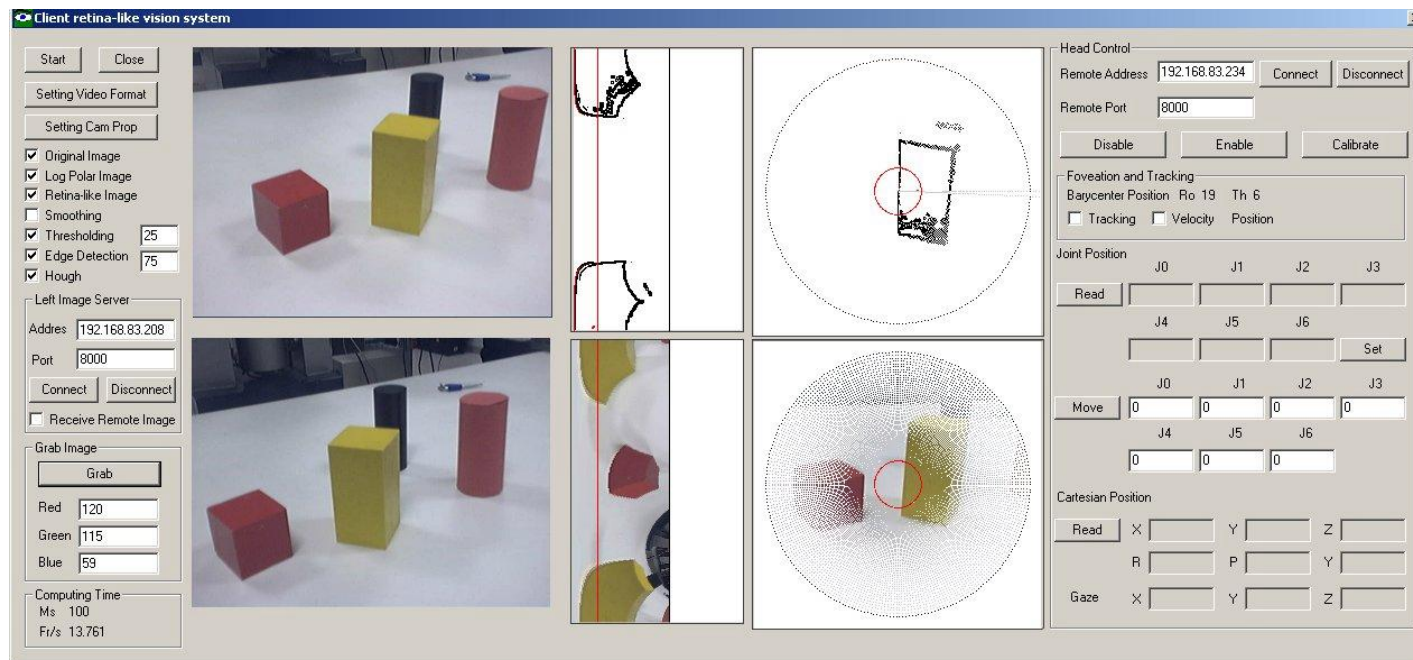
# Examples of algorithms developed for retina-like image processing



- 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
- Computation of the centroid of a thresholded area
- Edge detection
- Line detection

# 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 screenshot displays the PALOMA Robotic Artefact Control Panel interface. The main window is divided into several sections:

- Left Panel:** Contains control buttons (Start, Close, Setting Video Format, Setting Cam Prop) and a list of image processing options: Original Image (checked), Log Polar Image (checked), Retina-like Image (unchecked), Smoothing (unchecked), Thresh (100, checked), Edge Detection (75, unchecked), and Hough (unchecked). It also includes a Left Image Server section with Address (PALOMA1), Port (8000), and Connect/Disconnect buttons, and a Grab Image section with a Grab button and a table of image data.
- Image Processing Results:** A central area showing a sequence of images: the original scene with colored blocks, a vertical strip of the image, and two circular regions of interest (ROIs) around a yellow block, one in the original image and one in a processed version.
- Right Panel:** Contains control sections for Head Control (Remote Address: HEAD, Remote Port: 8000, STATUS: CONNECTED), Foveation and Tracking (Bar. Pos. Right Ro 5 Th 164 Left Ro 4 Th 62, Tracking unchecked, Velocity checked, Prop. Par 70, Velocity 0.50), Joint Position (Read and Move buttons for joints J0-J6 and Vel/T), Cartesian Position (Read buttons for X, Y, Z, R, P, Y), Gaze (X, Y, Z), and Head Neurocontroller (X: 85, Y: 0, Z: 10, Clamped Joints unchecked, Value Joints unchecked, Sym Move button).

R	G	B	H	S	V
124	109	66	45.57	0.468	99.66



# Edge Detection (gradient based method)

**PALOMA Robotic Artefact Control Panel**

Start Close

Setting Video Format

Setting Cam Prop

Original Image  
 Log Polar Image  
 Retina-like Image  Or.  
 Smoothing  
 Thresh 100  HSV  
 Edge Detection 75  
 Hough

Left Image Server

Address PALOMA1

Port 8000 Con.

Connect Disconnect

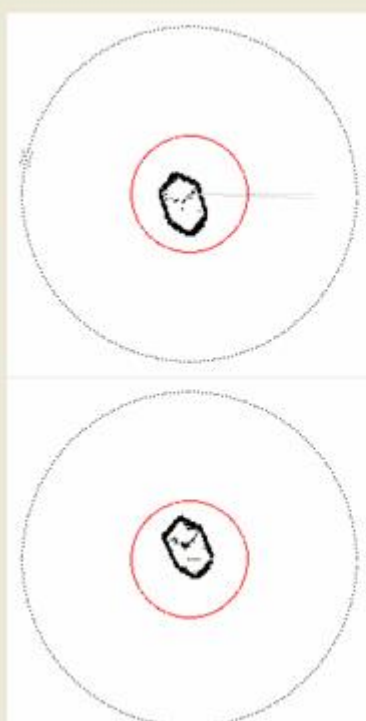


Receive Remote Image

Grab Image

Grab

R 124 H 45.57  
G 109 S 0.468  
B 66 V 99.66

Computing Time  
Ms 47  
Fr/s 6.347



Head Control

Remote Address HEAD Connect Disconnect

Remote Port 8000 STATUS: CONNECTED

Disable Enable Calibrate

Foveation and Tracking

Bar. Pos. Right Ro 5 Th 164 Left Ro 4 Th 63

Tracking  Velocity Position

Prop. Par 70

Velocity 0.50

Joint Position

Read	J0	J1	J2	J3	
	0	0	0	0	
Move	J4	J5	J6	Vel	T
	0	0	0	0.5	Set

Cartesian Position

Read	X	Y	Z
	R	P	Y
Gaze	X	Y	Z

Head Neurocontroller

X	Y	Z	<input type="checkbox"/> Clamped Joints
85	0	10	
<input type="checkbox"/> Sym	Move		<input type="checkbox"/> Value Joints

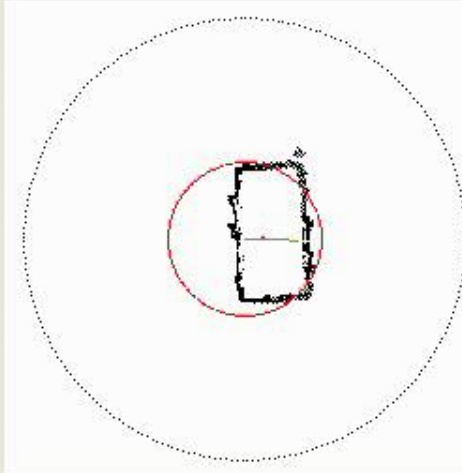
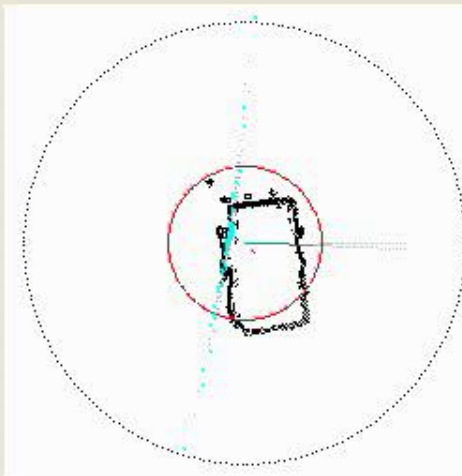
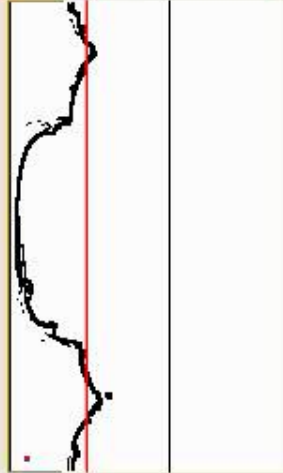
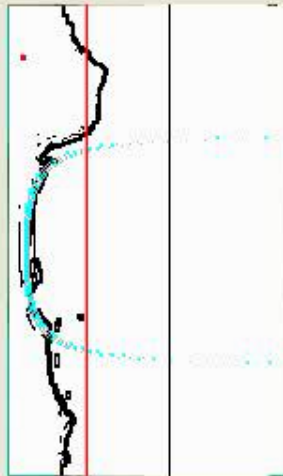
# Line detection (Hough method)

The screenshot displays the PALOMA Robotic Artefact Control Panel software interface. The main window is divided into several sections:

- Left Panel:** Contains control buttons (Start, Close, Setting Video Format, Setting Cam Prop) and a list of image processing options: Original Image, Log Polar Image, Retina-like Image, Smoothing, Thresh (100), Edge Detection (75), and Hough. It also includes a 'Left Image Server' section with address (PALOMA1) and port (8000) settings, and a 'Grab Image' section with a 'Grab' button and RGB data (R: 124, G: 109, B: 66; H: 45.57, S: 0.468, V: 99.66). A 'Computing Time' section shows 78 ms and 6.349 Fr/s.
- Image Processing Pipeline:** A sequence of four images showing the detection process: the original camera view, the image after log-polar transformation, the edge detection result, and the final Hough transform space with detected lines highlighted in cyan.
- Tracking View:** Two circular views showing the detected line (a red circle) being tracked by a blue dashed line.
- Right Panel:** Contains control sections for 'Head Control' (Remote Address: HEAD, Remote Port: 8000, STATUS: CONNECTED), 'Foveation and Tracking' (Bar. Pos. Right Ro 5 Th 164, Left Ro 4 Th 63, Tracking: unchecked, Velocity: checked, Position: unchecked, Prop. Par: 70, Velocity: 0.50), 'Joint Position' (Read and Move buttons for J0-J6, Vel, T), 'Cartesian Position' (Read buttons for X, Y, Z, R, P, Y), and 'Head Neurocontroller' (X: 85, Y: 0, Z: 10, Clamped Joints, Value Joints, Sym, Move).

- Applied only to pixels belonging to the fovea

Or.  
HSV  
75  
on.  
connect  
Image  
2.49  
626  
12.0



### Head Control

Remote Address: HEAD  
Remote Port: 8000

#### Foveation and Tracking

Bar. Pos. Right Ro 8 Th 223 Left

Tracking  Velocity Position -0

Prop. Par  70  
Velocity  0.35

#### Joint Position

<input type="button" value="Read"/>	J0	J1	J2
	0	10	0
<input type="button" value="Move"/>	J4	J5	J6
	-15	-4	8

#### Cartesian Position

<input type="button" value="Read"/>	X	Y
	R	P
Gaze	X	Y

#### Head Neurocontroller

X	Y	Z
85	0	10

Sym

### Arm Control Panel

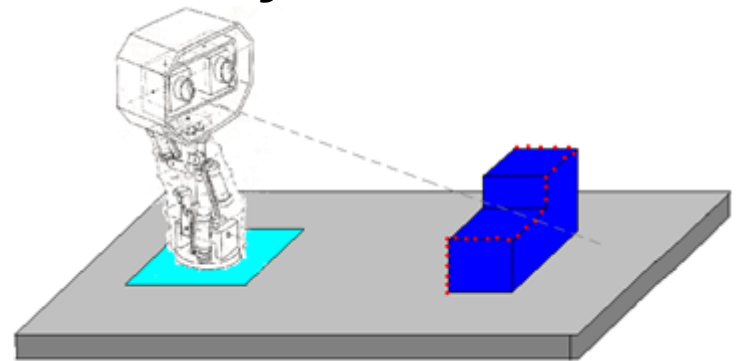
<input type="button" value="Move"/>	J0	J1	J2	J3	J4	J5	J6	J7	POS
	90.0	0.0	135.0	0.0	-90.0	0.0	0.0	0.0	

<input type="button" value="Read"/>	X	Y	Z	Roll	Pitch	Yaw	J0	Elb

Block  Compliant

# Preliminary activities

Foveation and tracking of borders of object and reconstruction of the geometry of the object



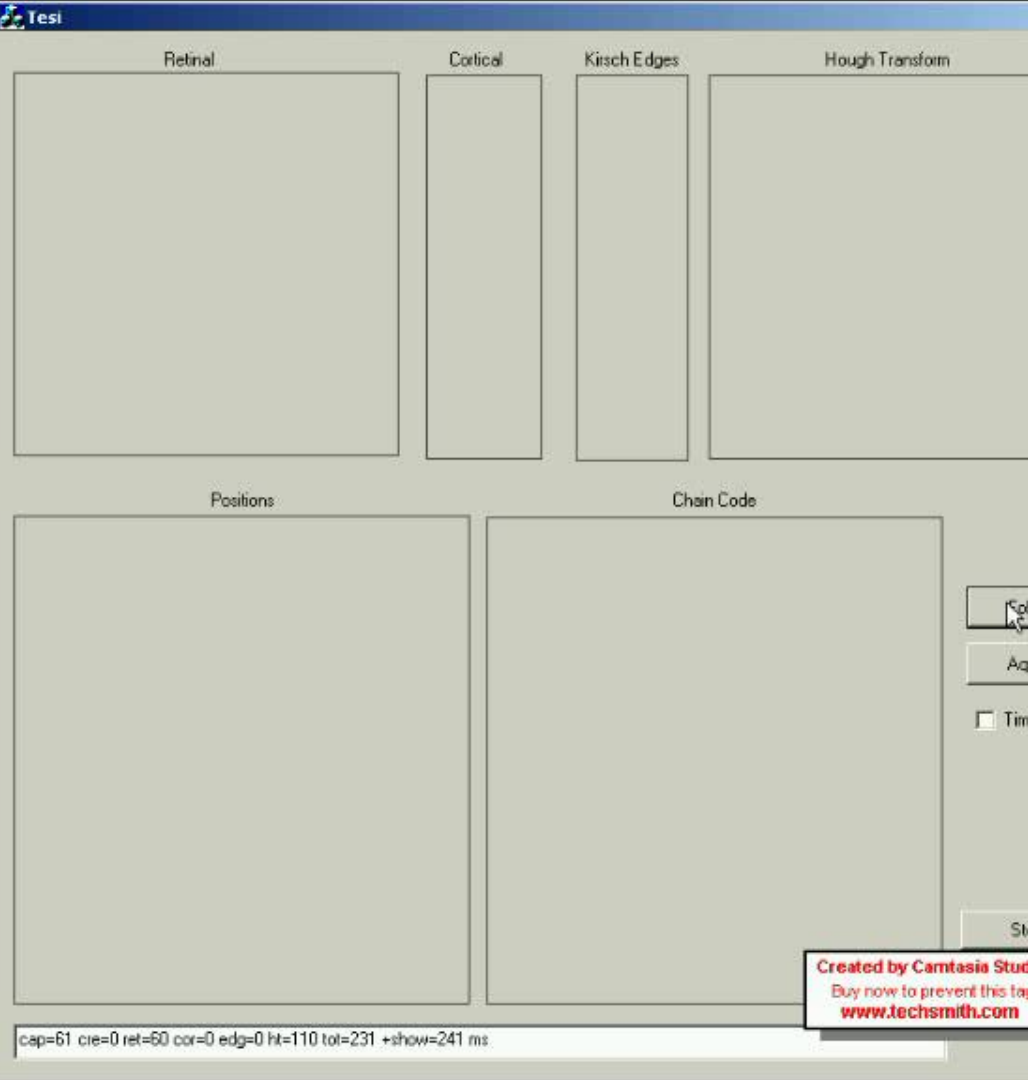
Retina Like image



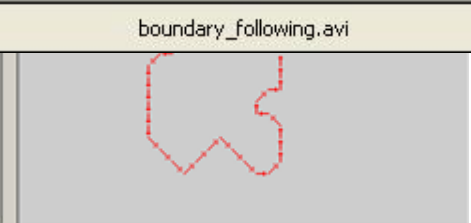
Log Polar Image



Edge of log polar image

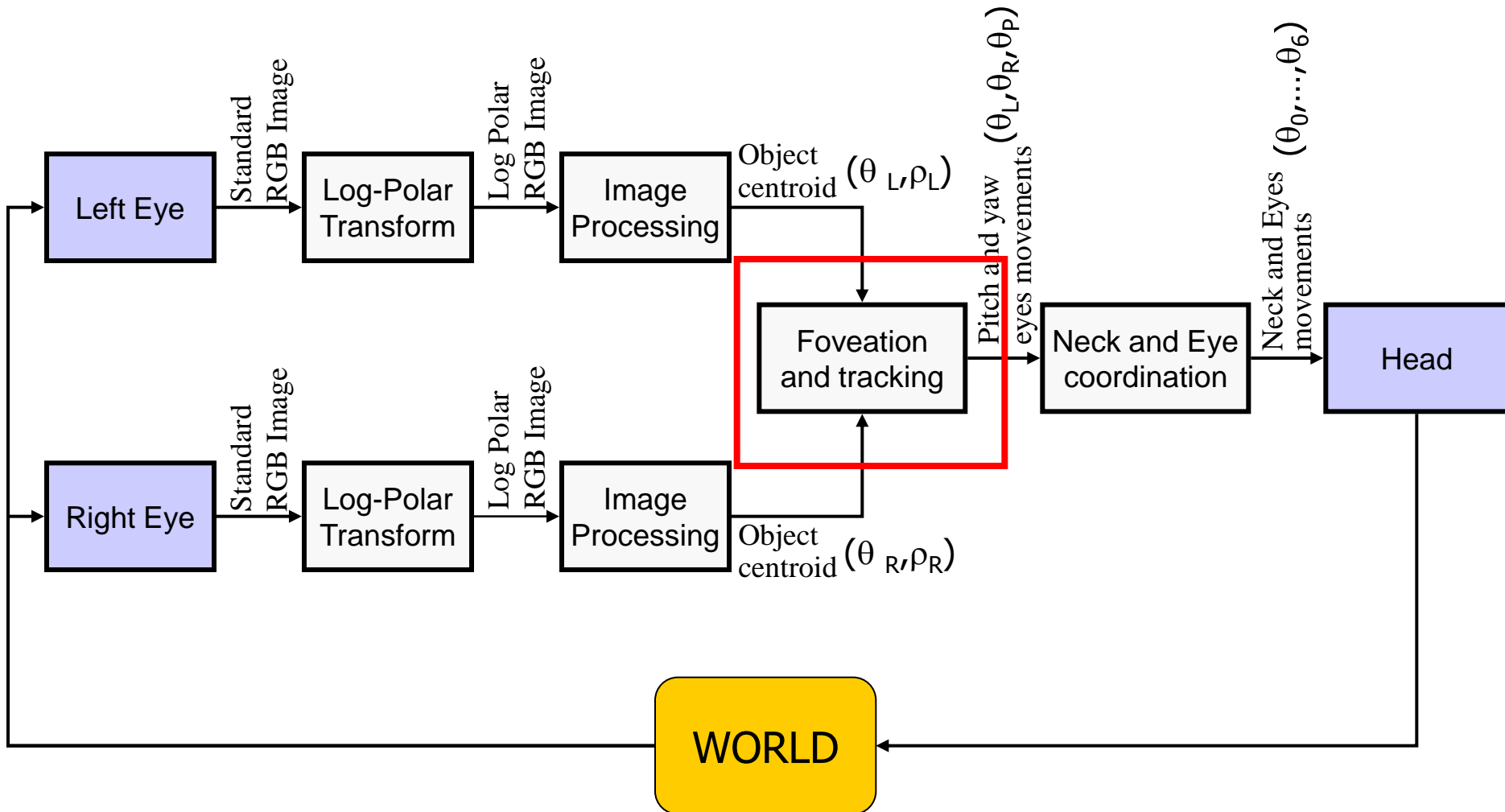


Detected lines (Boundaries)



Boundary reconstruction based on eye positions

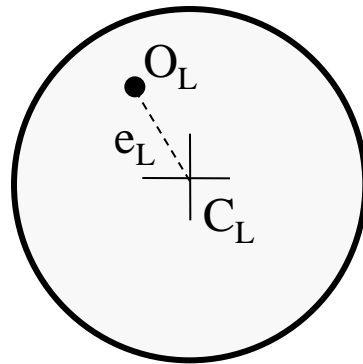
# Overall sensory-motor scheme of the visual apparatus



# Foveation of the object centroid

Proportional control based on the visual error

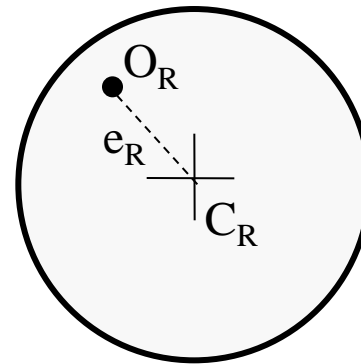
Left Image



$$O_L = (\rho_L, \theta_L)$$

$$e_L = \rho_L / M_{r0}$$

Right Image

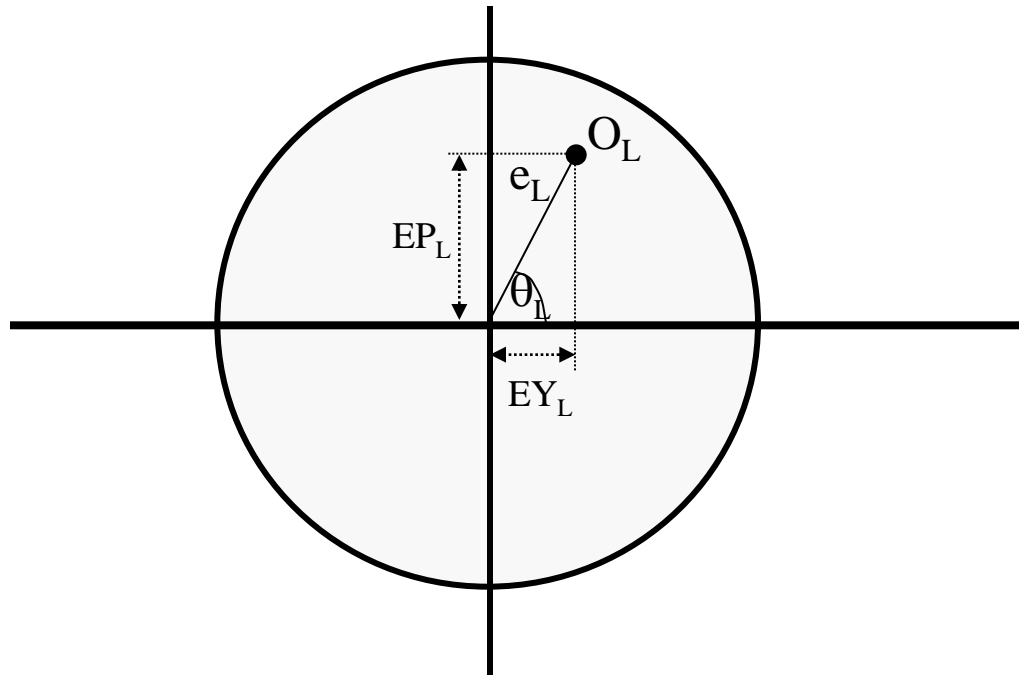


$$O_R = (\rho_R, \theta_R)$$

$$e_R = \rho_R / M_{r0}$$

$M_{r0}$  is the maximum  $\rho$  value (i.e. 152)

# Computation of yaw and pitch eye movements



$$EY_L = e_L * \cos(\theta_L) * P_L$$

$$EP_L = e_L * \sin(\theta_L) * P_L$$

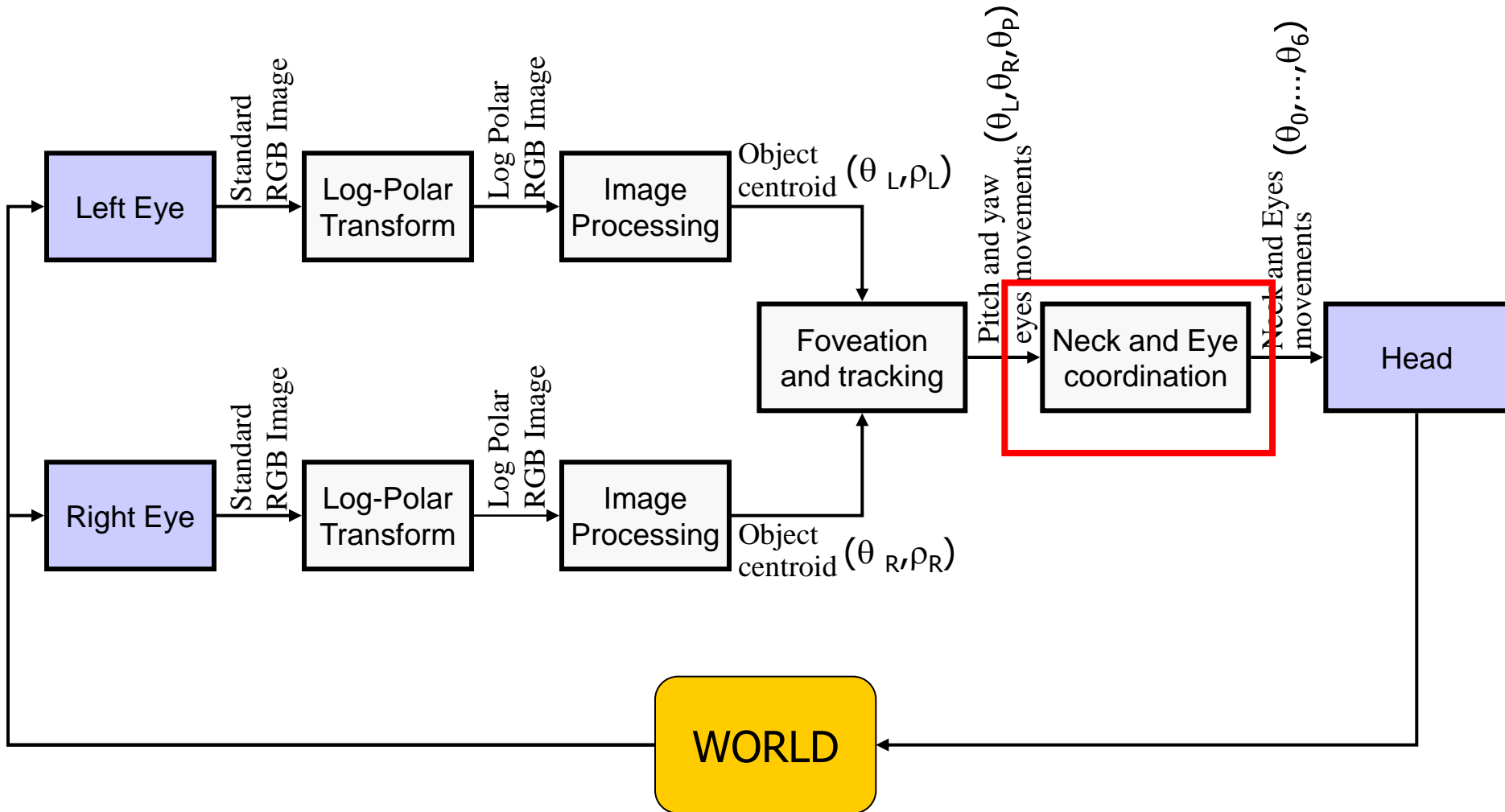
$$EY_R = e_R * \cos(\theta_R) * P_R$$

$$EP_R = e_R * \sin(\theta_R) * P_R$$

$$EP = (EP_L + EP_R) / 2$$

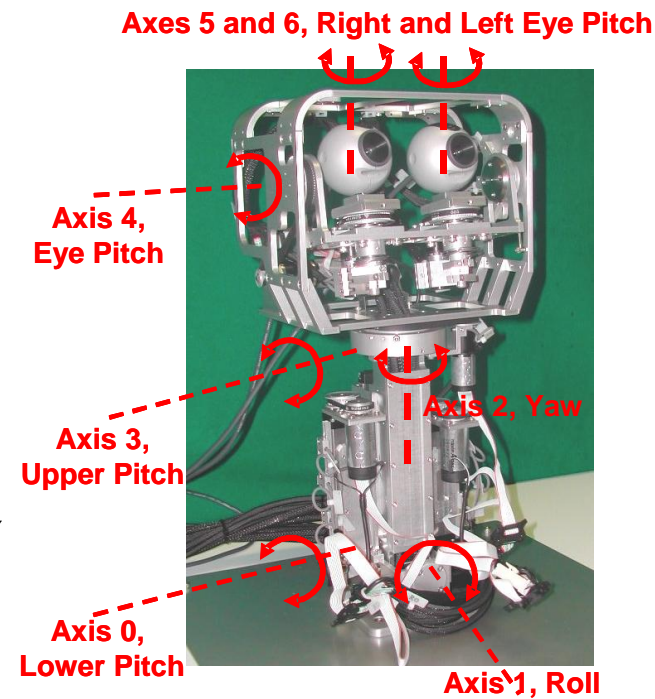
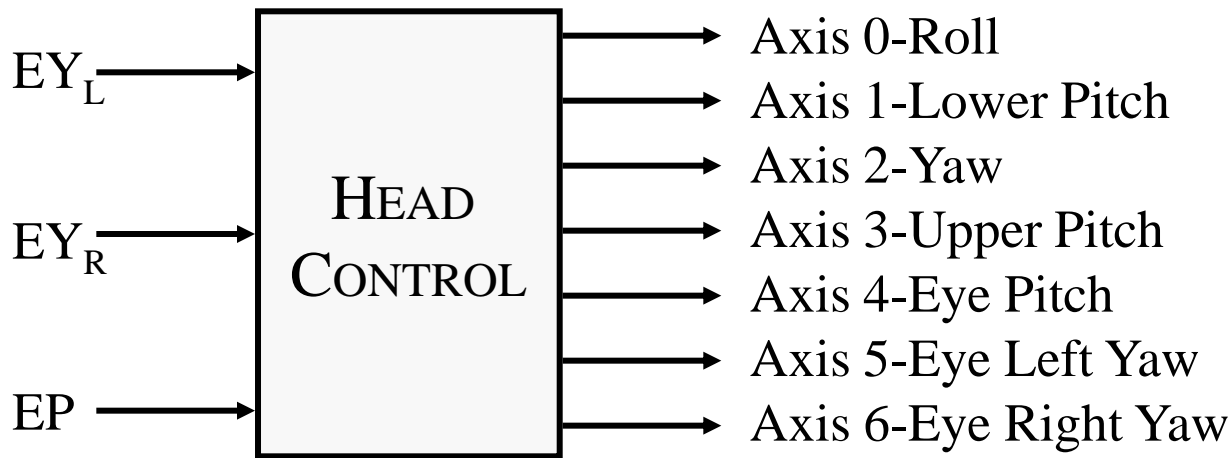
$P_L$  and  $P_R$  are the proportional parameters for left and right eye, respectively.

# Overall sensory-motor scheme of the visual apparatus





# Eye-neck coordination



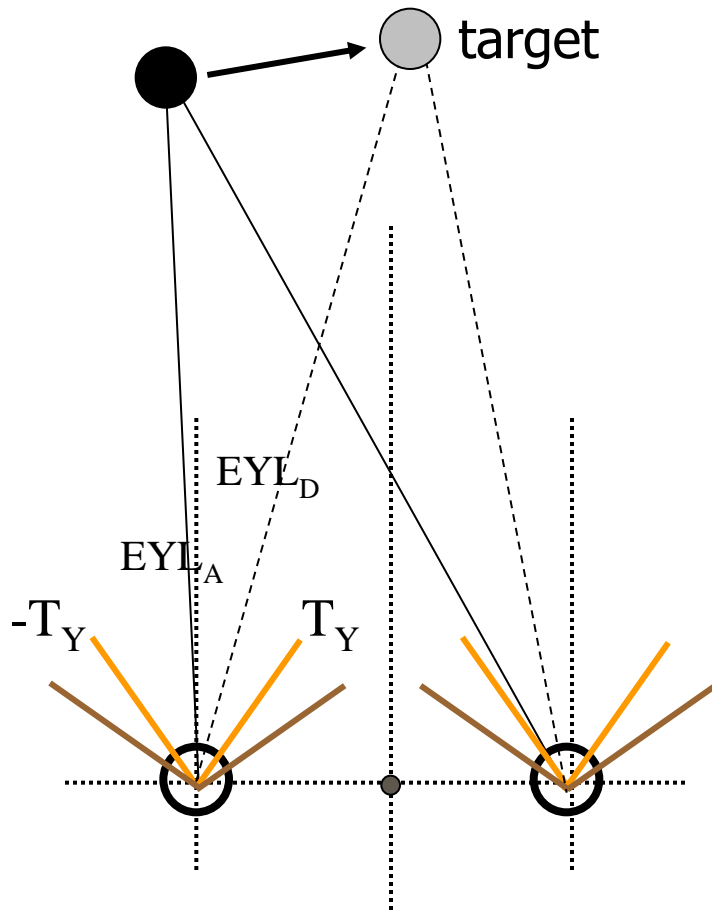
# Solution 1



**Distribution of the  
movements between the  
neck and eye DOF**

# Strategy for the coordination of neck and eye movement (yaw)

- If the movement is small, it is executed by the eyes, only



$$EYL_A + EYL_L < T_Y$$

and

$$EYL_R + EYL_R < T_Y$$



*Left Eye Yaw*

$$EYL_D = EYL_A + EYL_L$$

*Right Eye Yaw*

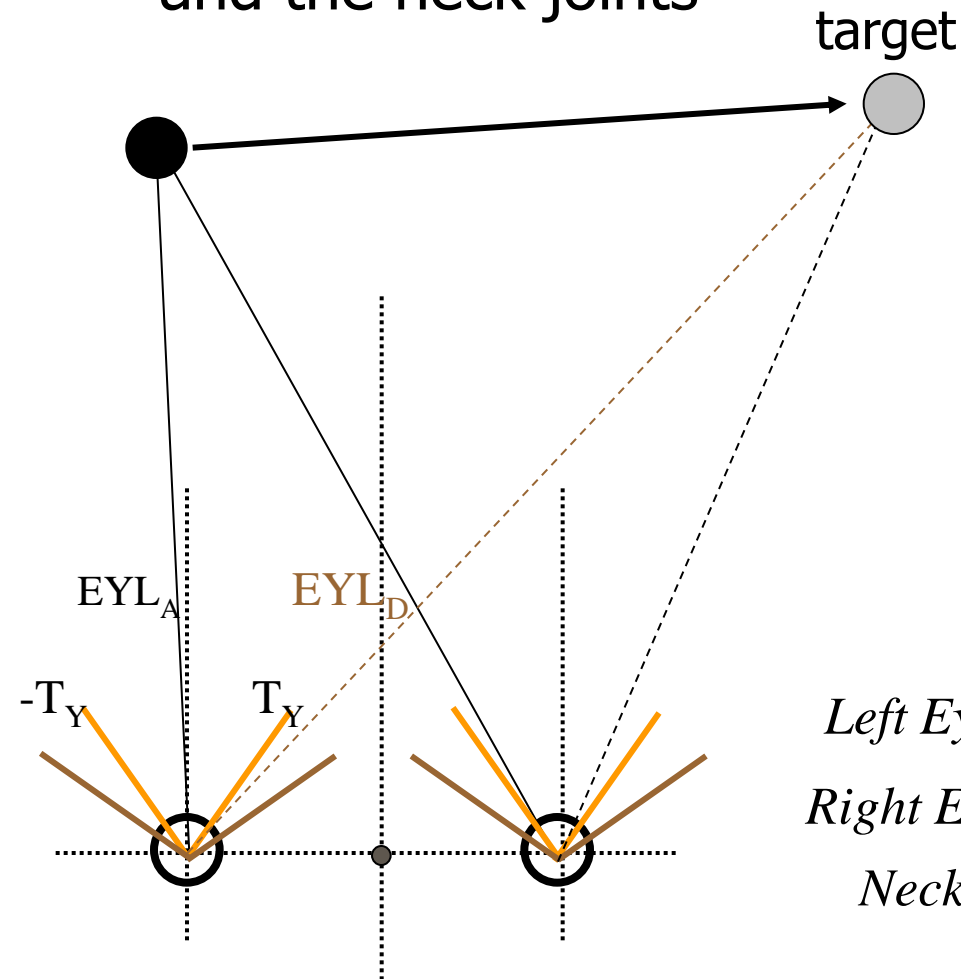
$$EYL_D = EYL_R + EYL_R$$

*Neck Yaw*

$$YAW_D = YAW_A$$

# Strategy for the coordination of neck and eye movement (yaw)

- If the movement is larger, it is distributed among the eyes and the neck joints



$$EYL_A + EY_L > T_Y$$

or

$$EYR_A + EY_R > T_Y$$



Eyes and neck

$$\theta = \text{atan}((\tan(EY_L) + \tan(EY_R))/2)$$

*Left Eye Yaw*  $EYL_D = EYL_A + EY_L - \theta$

*Right Eye Yaw*  $EYR_D = EYR_A + EY_R - \theta$

*Neck Yaw*  $YAW_D = YAW_A + \theta$

# Strategy for the coordination of neck and eye movement (pitch)

Eye, upper and lower pitch of the head are calculated as a percentage (proportional to the available range) of EP.

$$K1 = EP * EYP_{AV} / P_{AV}$$

$$K2 = EP * UP_{AV} / P_{av}$$

$$K3 = EP * LP_{AV} / P_{av}$$

$$EYP_D = EYP_A + EP * K1$$

$$EUP_D = EUP_A + EP * K2$$

$$ELP_D = ELP_A + EP * K3$$

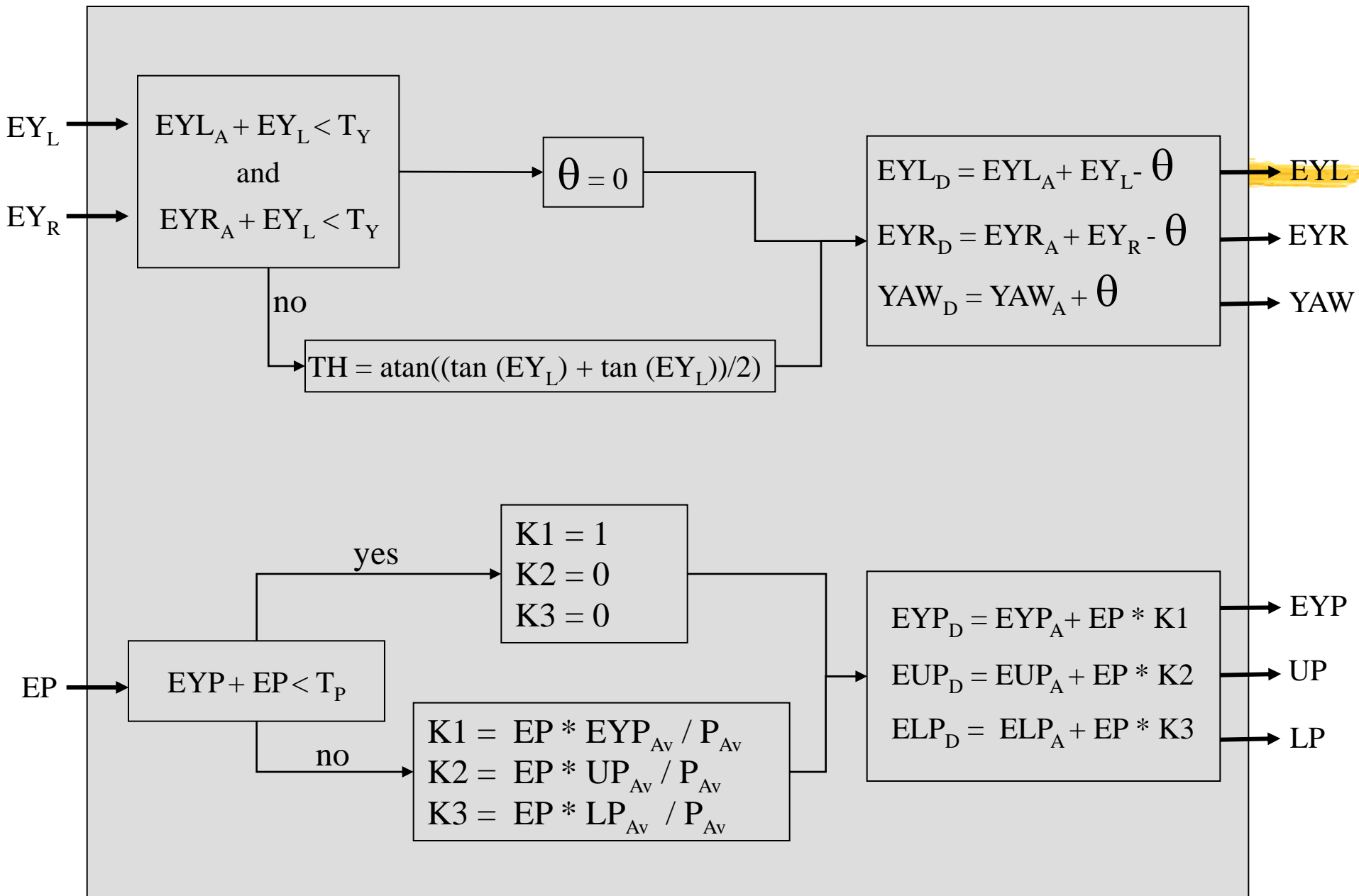
$$EYP_{AV} = EYP_M - EYP_A$$

$$UP_{AV} = UP_M - UP_A$$

$$LP_{AV} = LP_M - UP_A$$

$$P_{AV} = EYP_{AV} + UP_{AV} + LP_{AV}$$

$EYP_M$ ,  $UP_M$  and  $LP_M$  are the range limits respectively for eye pitch, upper pitch and lower pitch axis



**PALOMA Robotic Artifact Control Panel**

Start Close

Setting Video Format

Setting Cam Prop

Original Image  
 Log Polar Image  
 Retina-like Image  Dr.  
 Smoothing  
 Thresh 80  HSV  
 Edge Detection 75  
 Hough

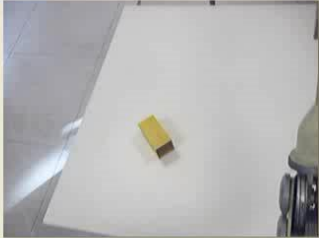
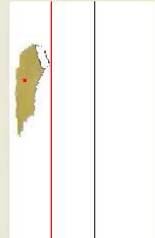
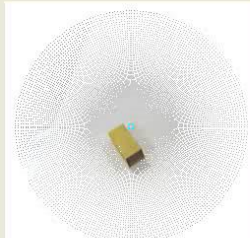
Left Image Server  
 Address: PALOMA1  
 Port: 8000 Con.  
   
 Receive Remote Image

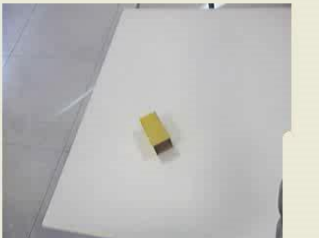
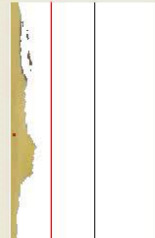
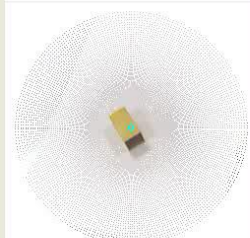
Grab Image

R	176	H	53.62
G	164	S	0.563
B	77	V	139.0

Computing Time  
 Ms: 47  
 Fr/s: 0.015

Hand Control Panel

Head Control

Remote Address: HEAD

Remote Port: 8000 STATUS: CONNECTED

Foveation and Tracking

Bar. Pos. Right Ro 15 Th 171 Left Ro 4 Th 117

Tracking  Velocity Position 0.002-0.009:0.010

Prop. Par \_\_\_\_\_ 70

Velocity \_\_\_\_\_ 0.68

Joint Position

Read	J0	J1	J2	J3		
	0	0	0	0	0	0
Move	J4	J5	J6	Vel	T	
	0	0	0	0.5	Set	

Cartesian Position

Read	X	Y	Z		
R		P	Y		
Gaze	X	Y	Z		

Head Neurocontroller

X	Y	Z	<input type="checkbox"/> Clamped Joints
85	0	10	<input type="checkbox"/> Value Joints
<input type="checkbox"/> Sym	<input type="button" value="Move"/>		

# Pursuit Movement



Frame rate: 10 fps for both images  
 Head Control loop: 100 ms