Corso di Robotica (ROB)

C. Modulo di Percezione

Visione artificiale retinica

Cecilia Laschi
Istituto di BioRobotica, Scuola Superiore Sant'Anna
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

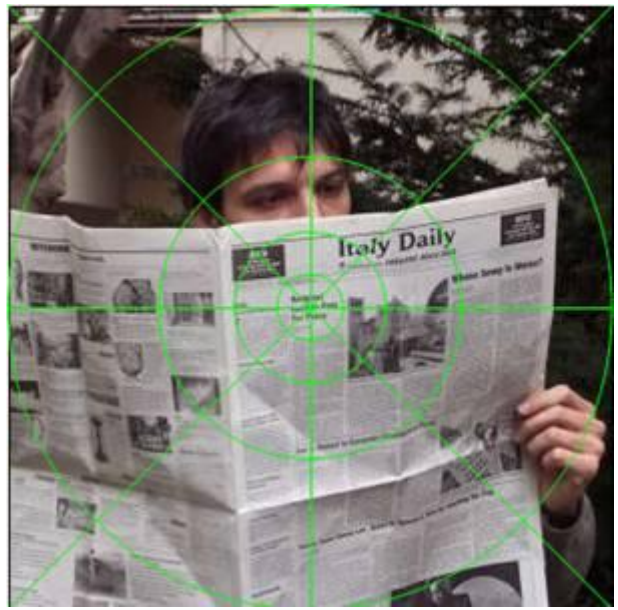
Retina-like image

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

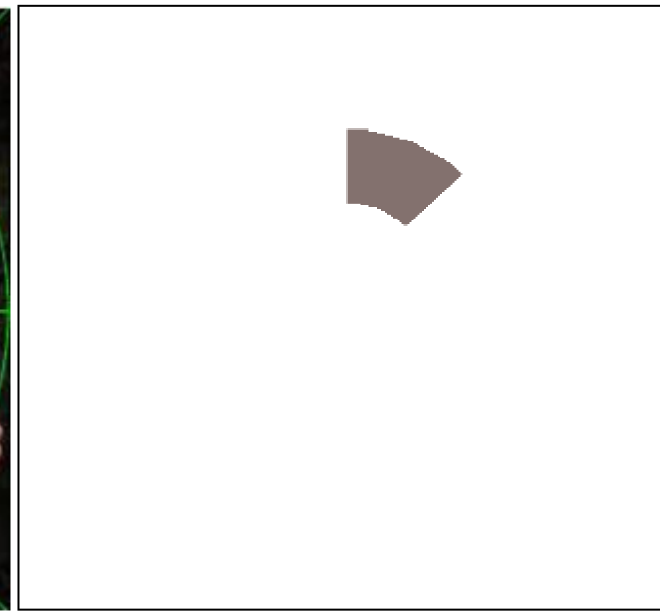
Log-polar projection

Costruzione di un’immagine retinica

Immagine cartesiana tradizionale



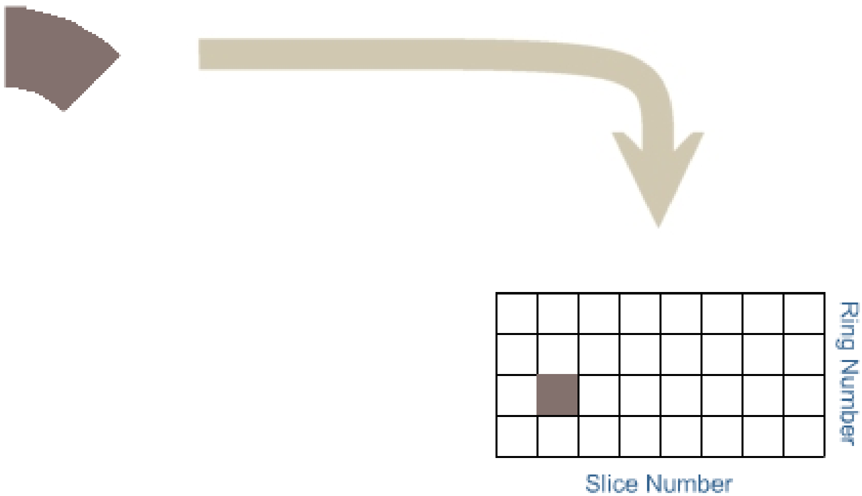
Divisione in circonferenze e spicchi



Calcolo del valore medio di un settore

http://www.retinica.com/

Costruzione di un’immagine retinica

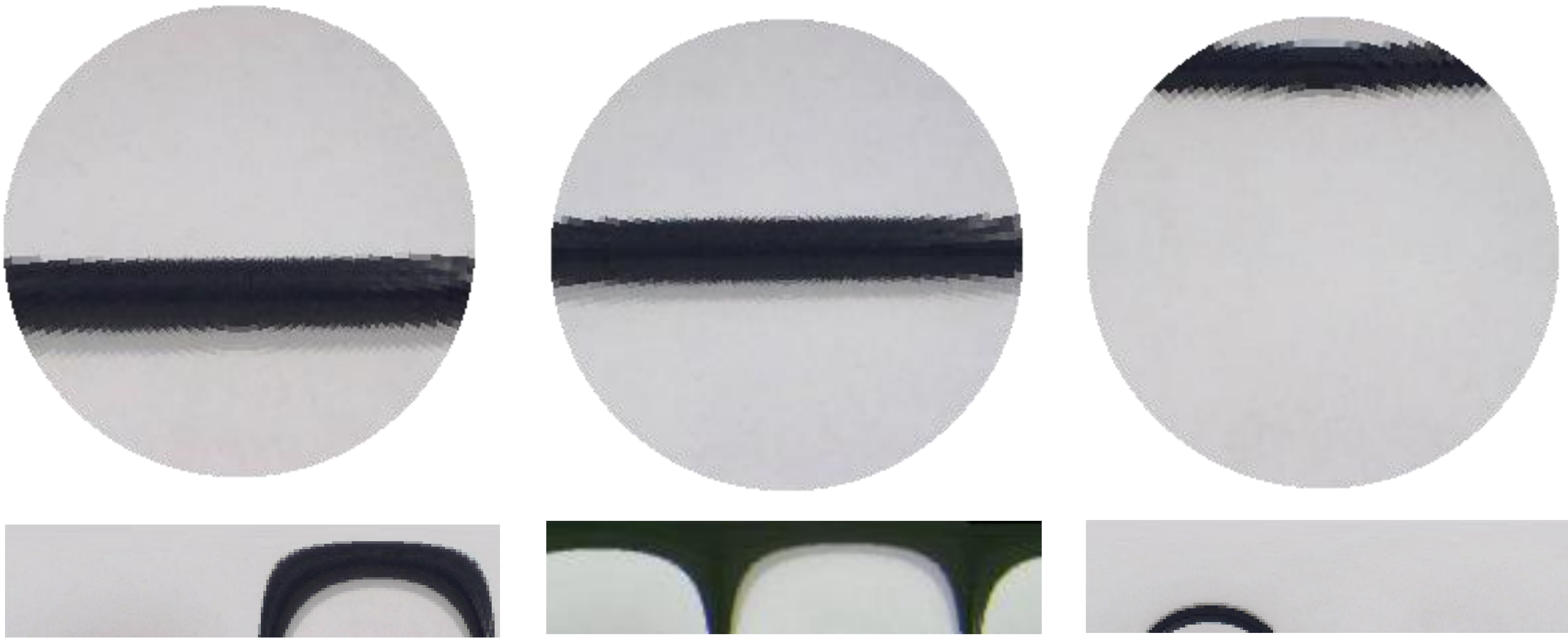


Copia del valore medio di un settore in un pixel di un’immagine polare

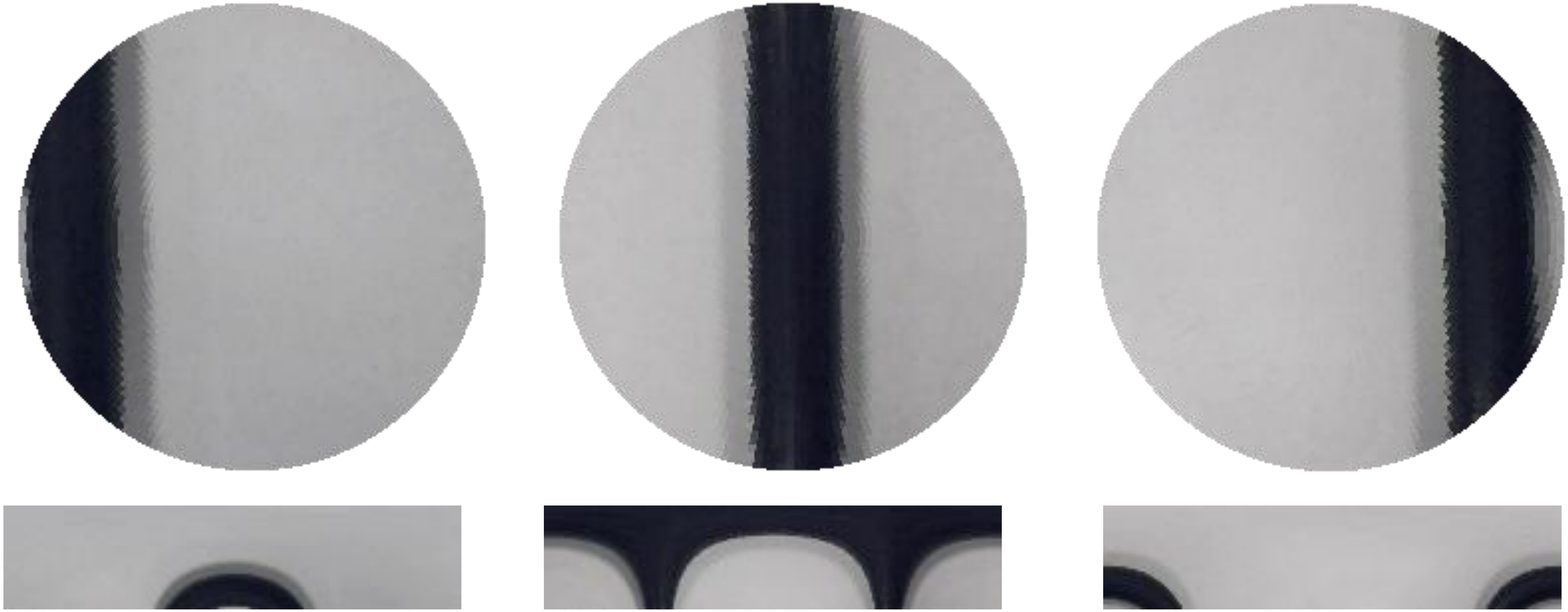
Immagine polare risultante

Immagine cartesiana ricostruita dalla polare

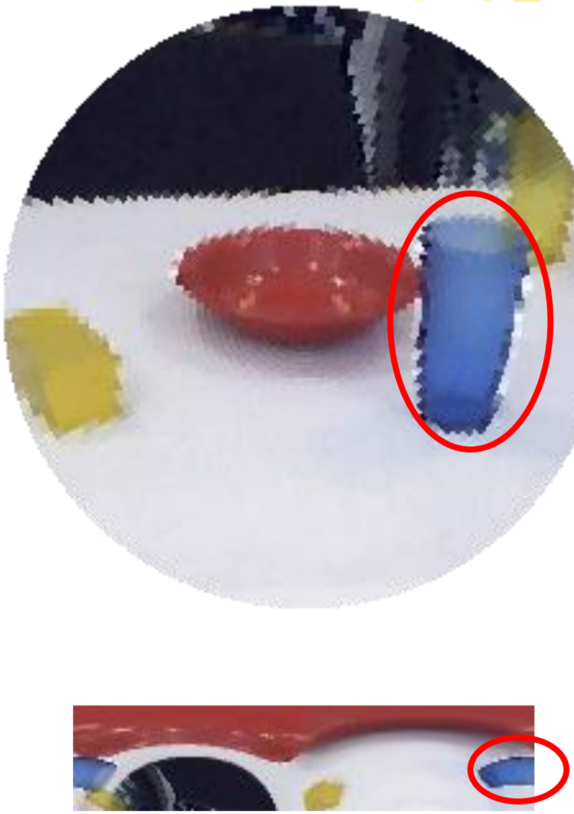
An example of pattern translation



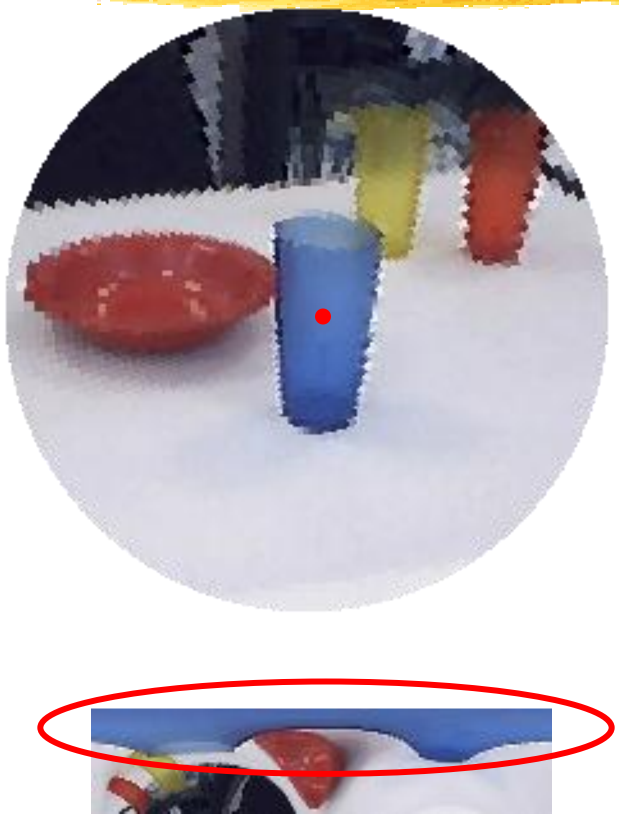
An example of pattern translation



An example of simulated foveation



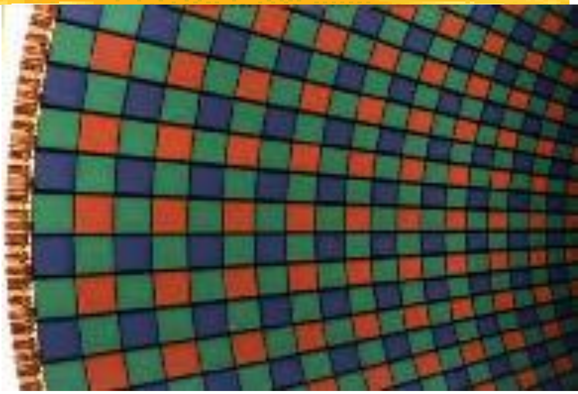
Object detection
in the periphery

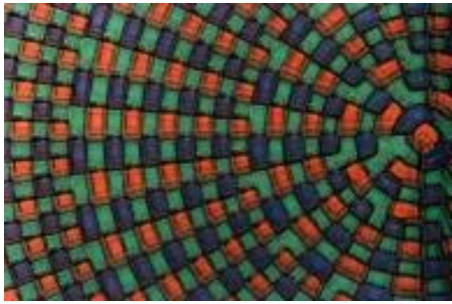


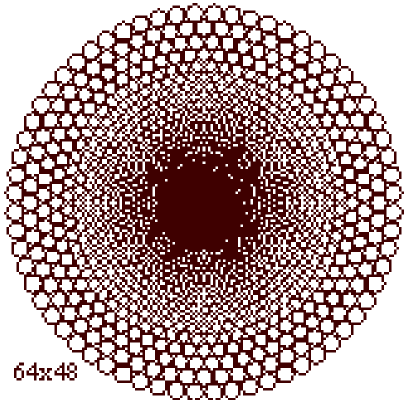
Object foveation

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

$$\rho(x,y)=\begin{cases}(F-1)+\log_\lambda\left[\left(F-½-\sqrt{x^2+y^2}\right)(1-\lambda)+\lambda\right] & \text{if } \sqrt{x^2+y^2}>(F-½)\\ \left(\sqrt{x^2+y^2}+½\right) & \text{if } \sqrt{x^2+y^2}<(F-½)\end{cases}$$

$$r(\rho)=\left[(F-½)+\lambda\frac{1-\lambda^{\rho-F}}{1-\lambda}\right] \text{ if } \rho>F$$

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

F=42
P=152
Θ =252
X=545
Y=545
λ=1.02314422608633

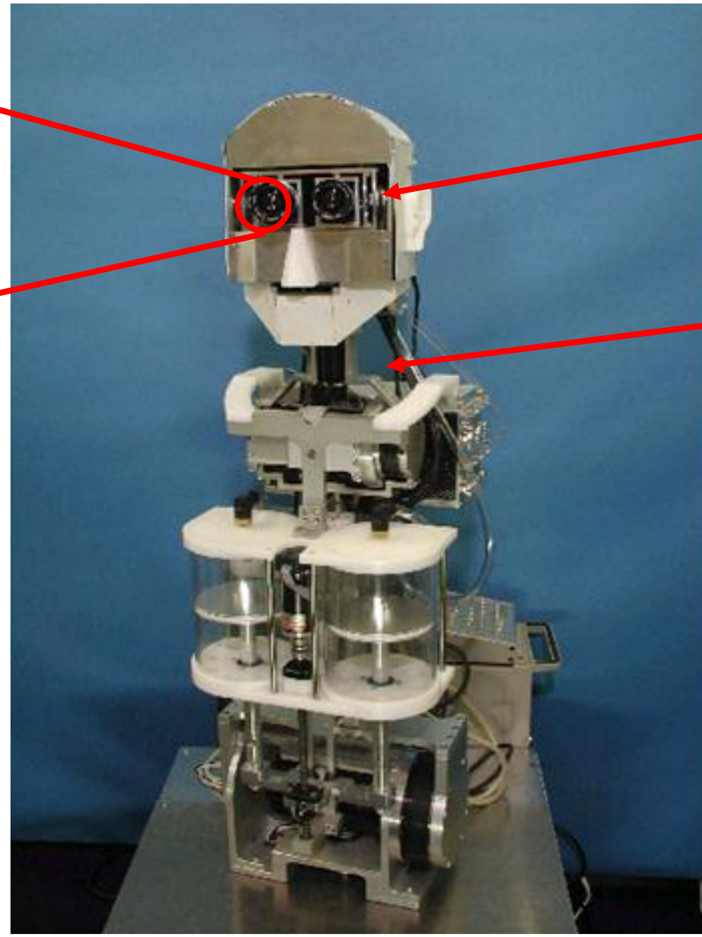
F = size of the fovea in rings.
P = total number of rings.
Θ = maximum # of pixels in each ring.

$2X$ = horizontal size of the cartesian image.
$2Y$ = vertical size of the cartesian image.
ρ = ring number in the log polar image.
θ = 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

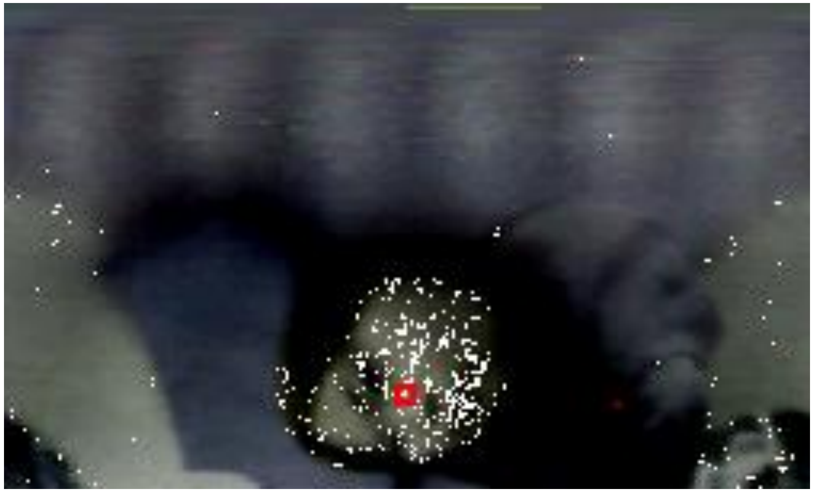
Face detection by hue

Hue = information on the color

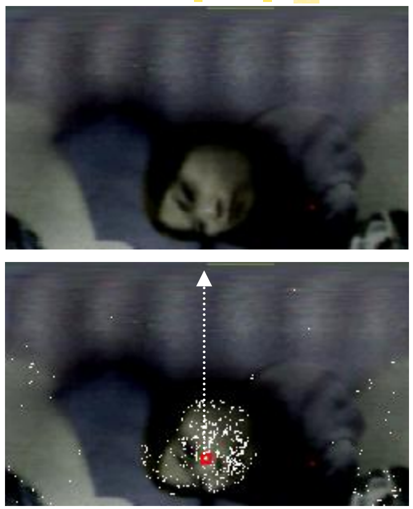
$$\text{Hue} = \cos^{-1}\left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right)$$

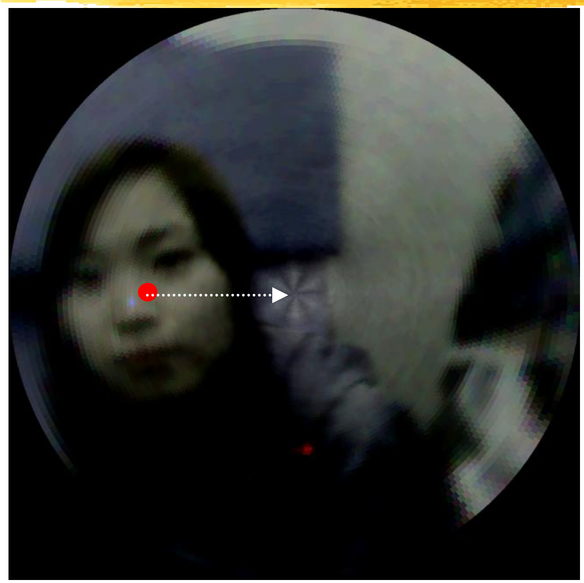
if B>G then Hue = 2π - Hue

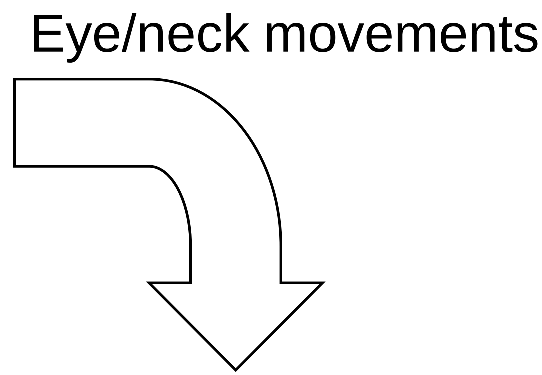
R, G, B = RED, GREEN, BLUE components, respectively



An example of foveation

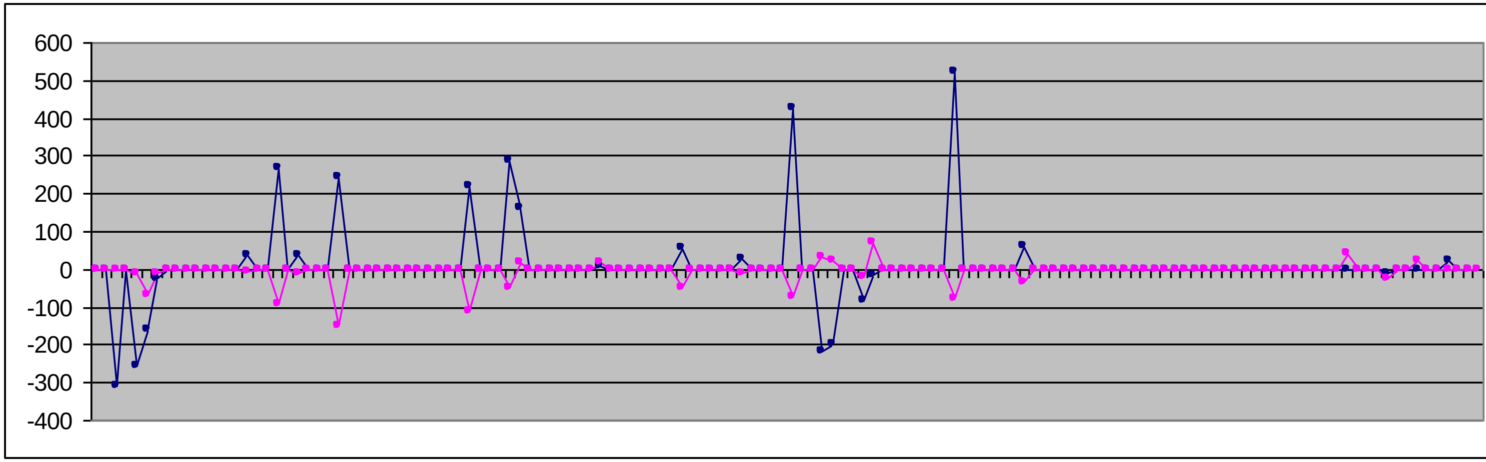






Proportions are rescaled for display purposes

Experimental trials



Right eye motion

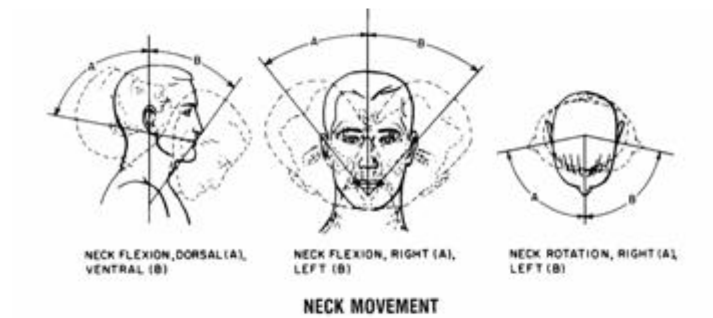
[Cecilia Laschi, Hiroyasu Miwa, Atsuo Takanishi, Eugenio Guglielmelli, Paolo Dario, 2002]

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

The ARTS humanoid robot head

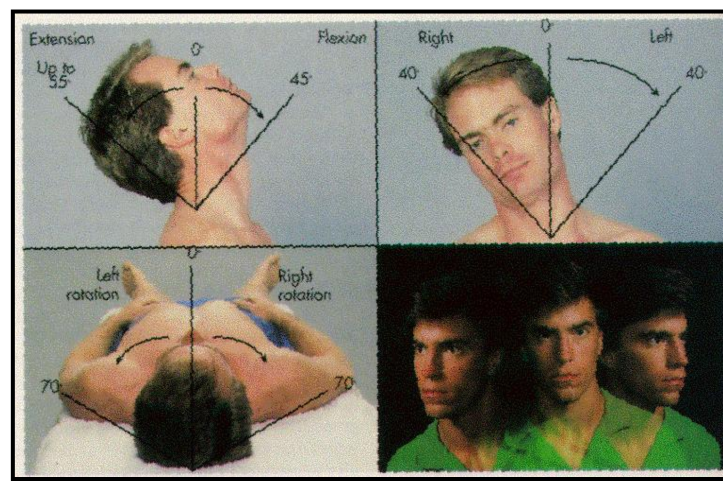
Synthesis of characteristics of the human oculo-motor system

- 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



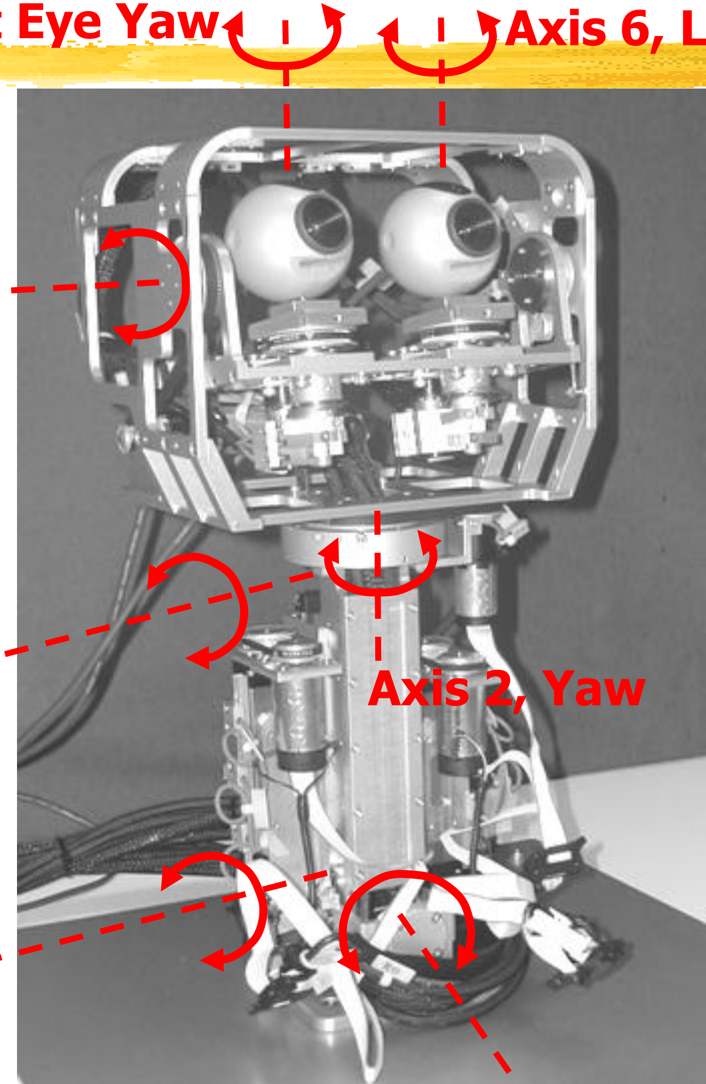
	Average	[illegible]
Ventral flexion	60°	[illegible]
Dorsal flexion	61°	[illegible]
Right-left flexion	41°	[illegible]
Right-left rotation	79°	[illegible]

RANGE OF MOVEMENT AT THE NECK JOINT*
*Male civilians.



[Thibodeau & Patton, 1996]

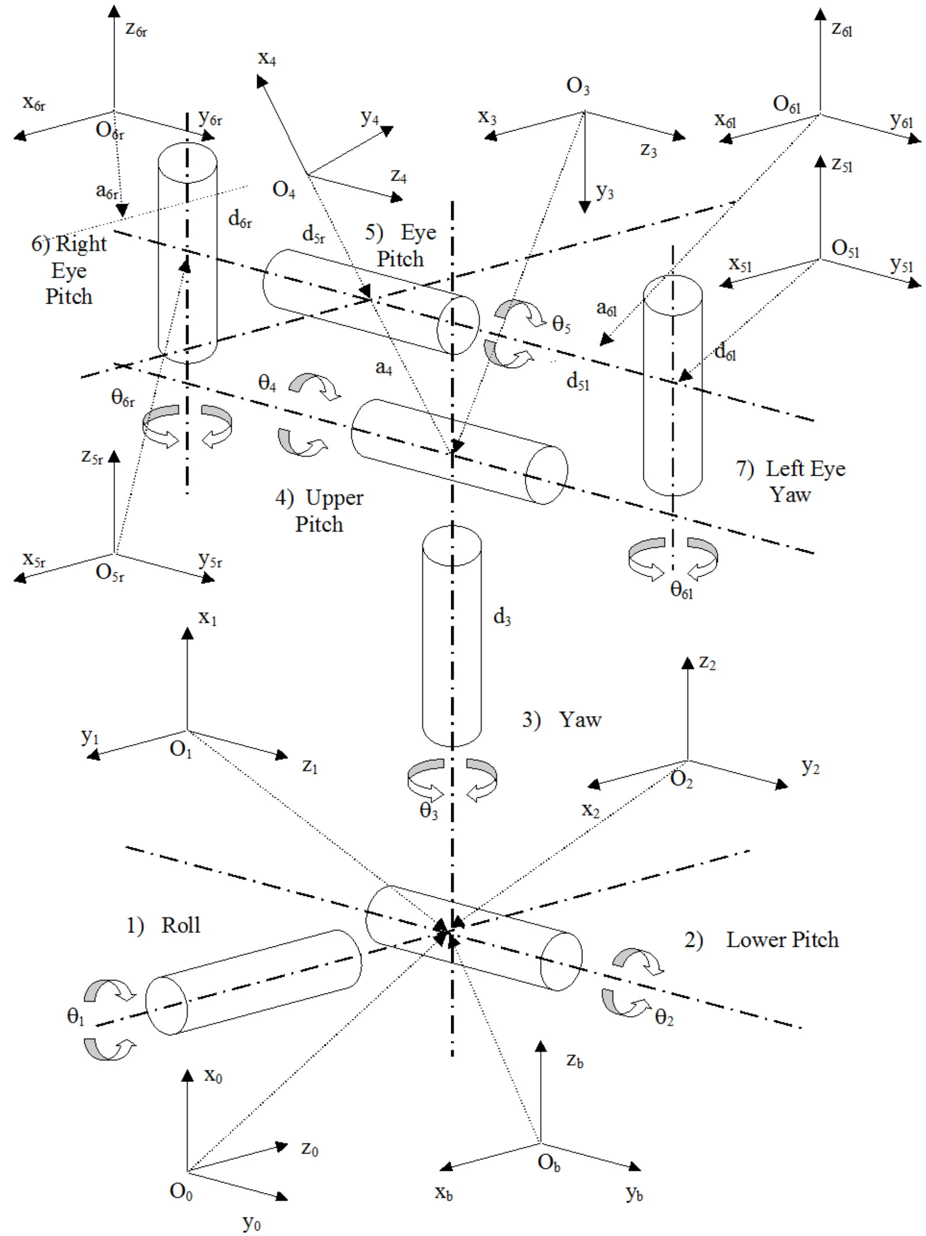
Kinematic structure of the SSSA Robot Head



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

Head kinematic chain and Denavit-Hartenberg parameters

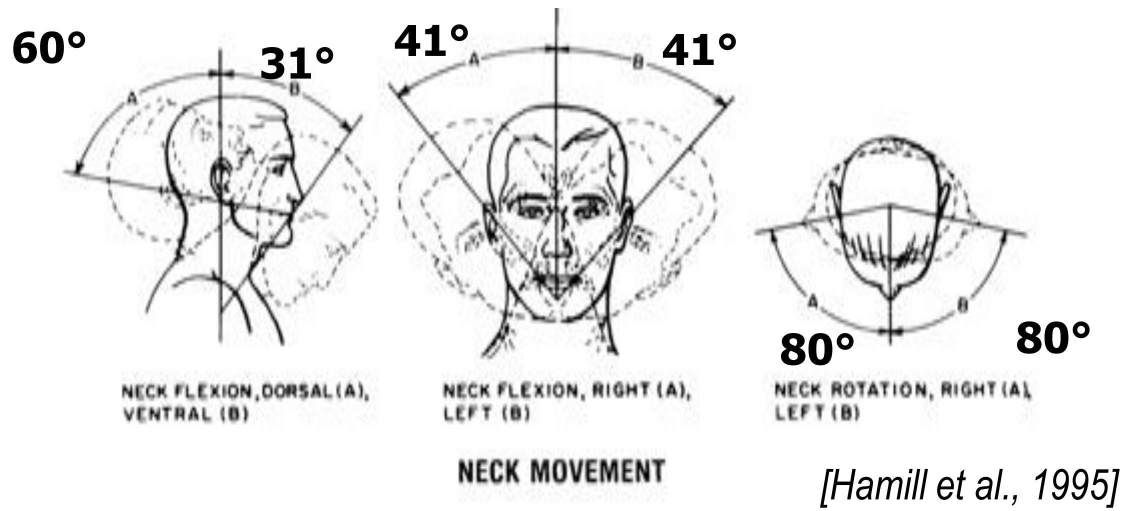
Joint	a_i (mm)	d_i (mm)	α_i (rad)
J1	0	0	$-\pi/2$
J2	0	0	$\pi/2$
J3	0	195	$-\pi/2$
J4	137.5	0	0
$J5_r$	0	$-30 \div -50$	$\pi/2$
$J5_l$	0	$30 \div 50$	$\pi/2$
$J6_l$	a_{6l}	d_{6l}	0
J6r	a_{6r}	d_{6r}	0



Comparison of performances between human and robotic head

Neck: | **Eye:**

Human

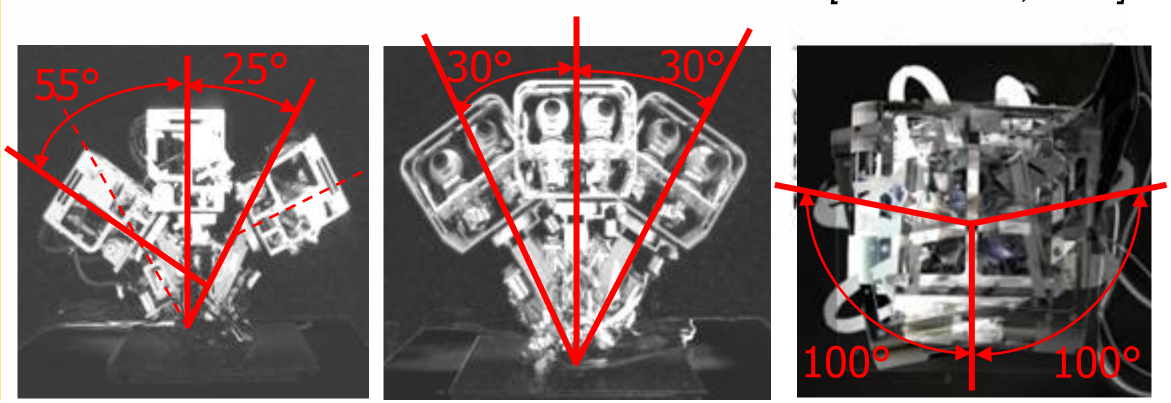


[Hamill et al., 1995]

Pitch: ± 60°, 600°/s

Yaw: ± 30°, 600°/s

Robot



Pitch: ± 47°, 600°/s

Yaw: ± 45°, 1000°/s

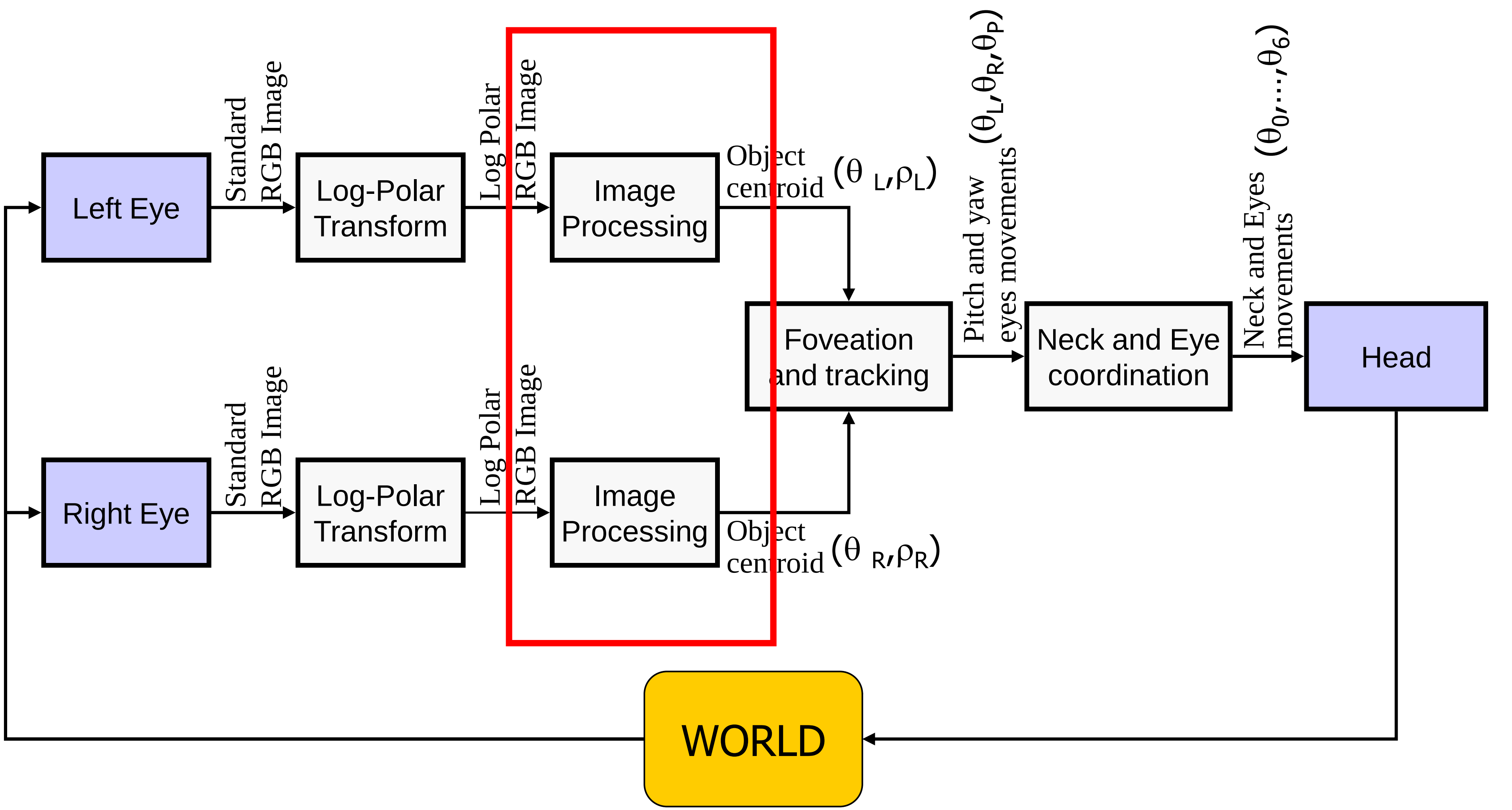
The movements of the 7 dofs of the robotic head

Neck Movements

Saccades, 400°/sec

saccade 300° sec.avi

Overall sensory-motor scheme of the visual apparatus

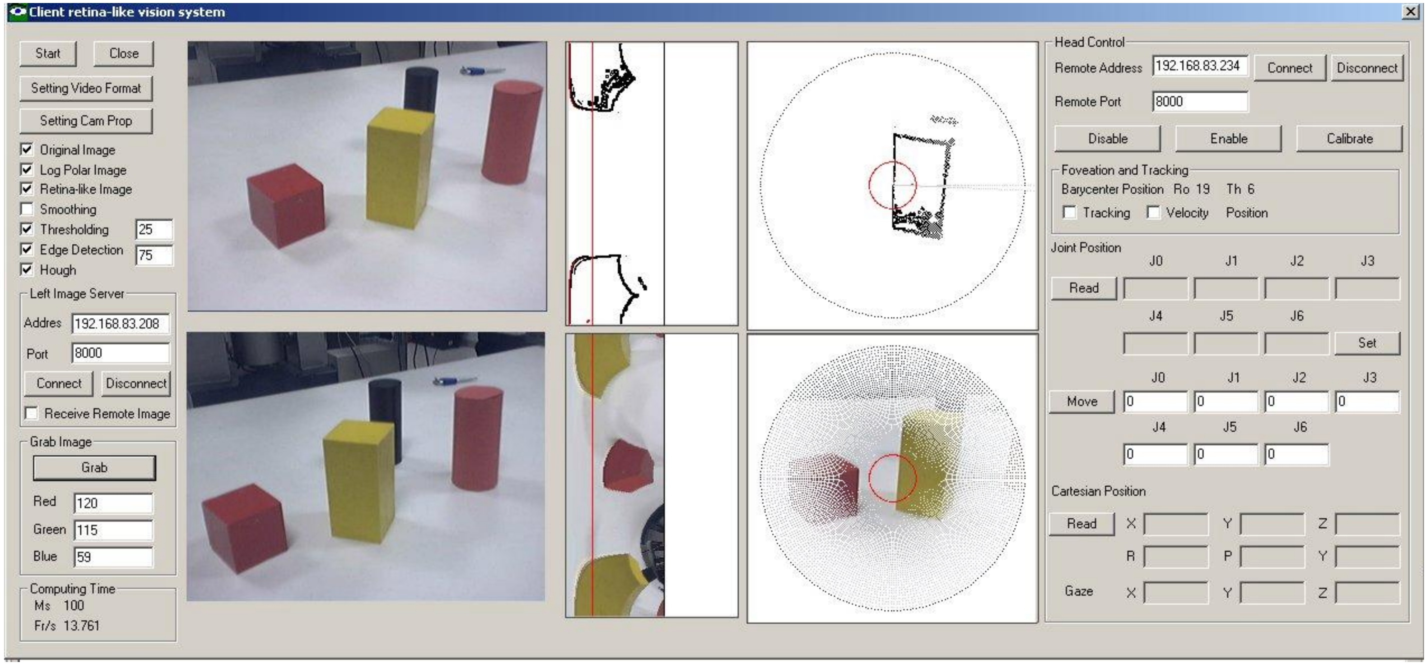


Examples of algorithms developed for retina-like image processing

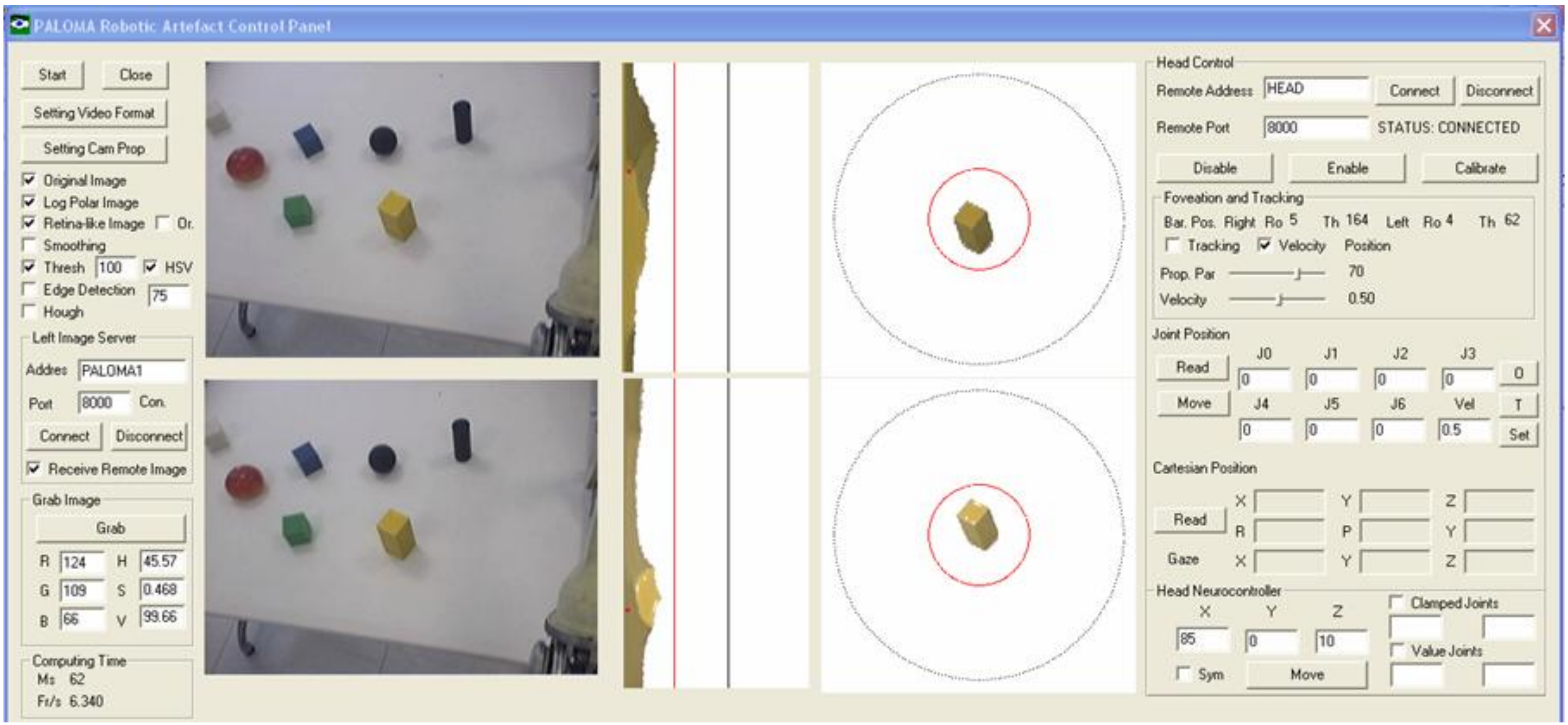
- Acquiring standard image
- Creating log-polar image from standard image
- Creating retina-like image from log-polar image
- Thresholding of image based on RGB and HUE
- Computation of the centroid of a thresholded area
- Edge detection
- Line detection

Simulation of retina-like cameras and basic image processing

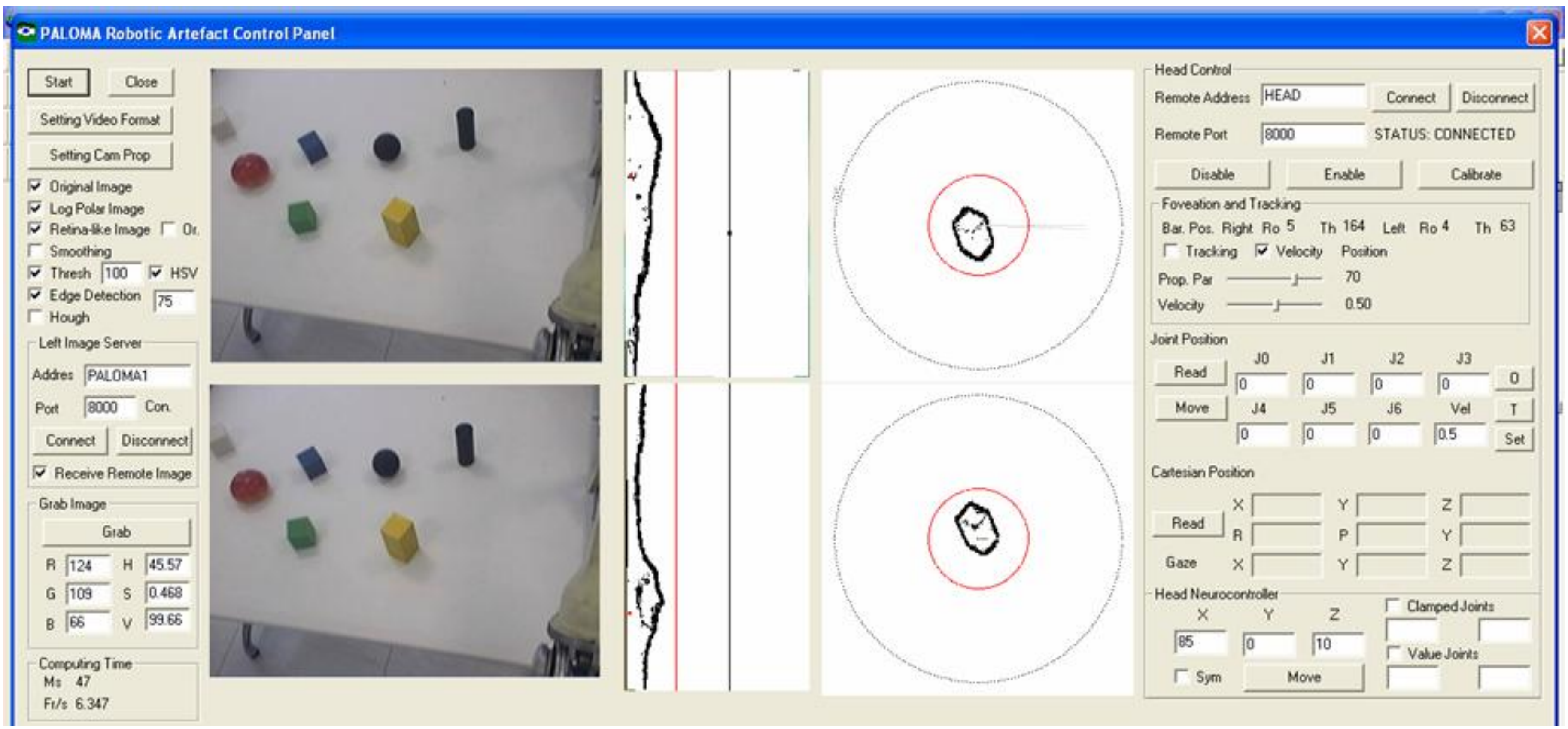
- Acquiring standard image
- Creating log-polar image from standard image
- Creating retina-like image from log-polar image



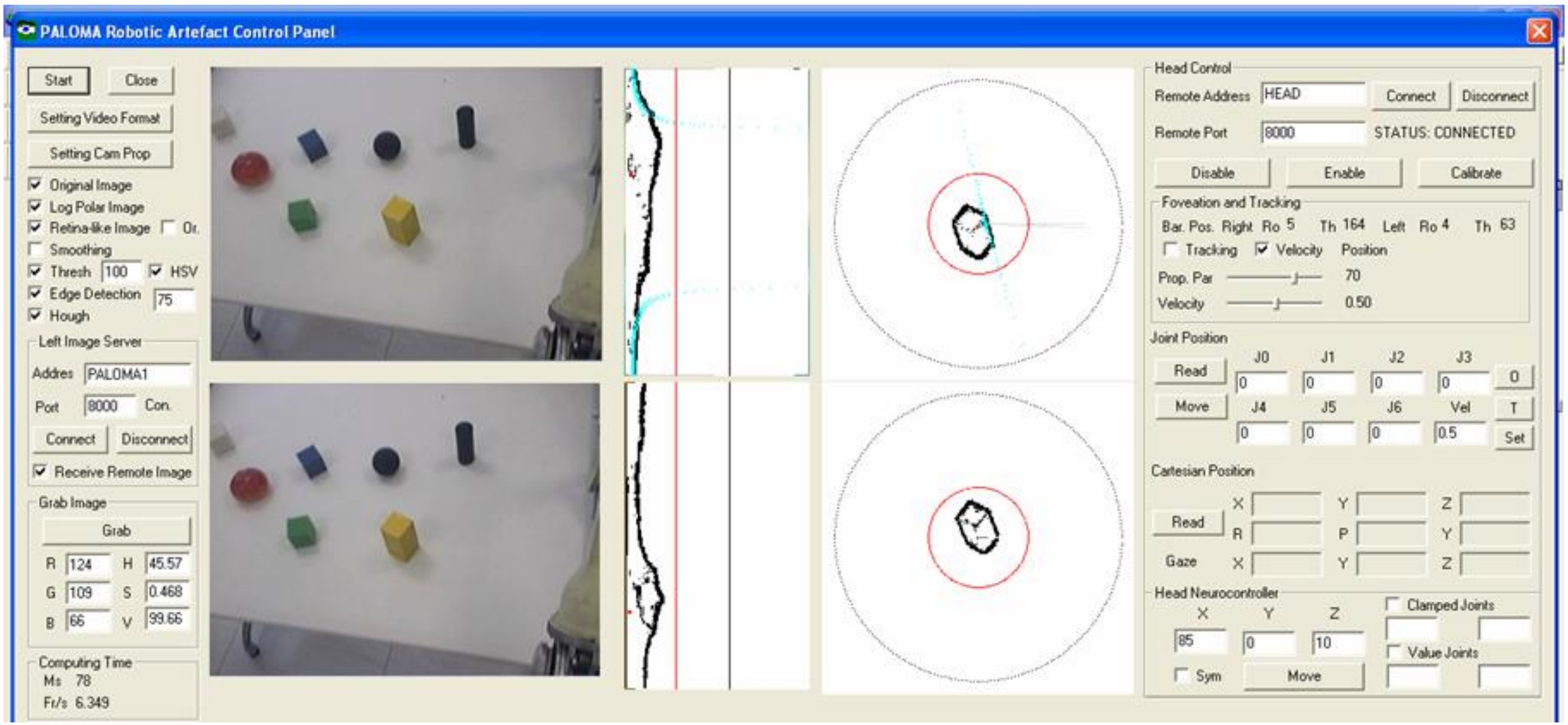
Thresholding of image based on RGB and HUE



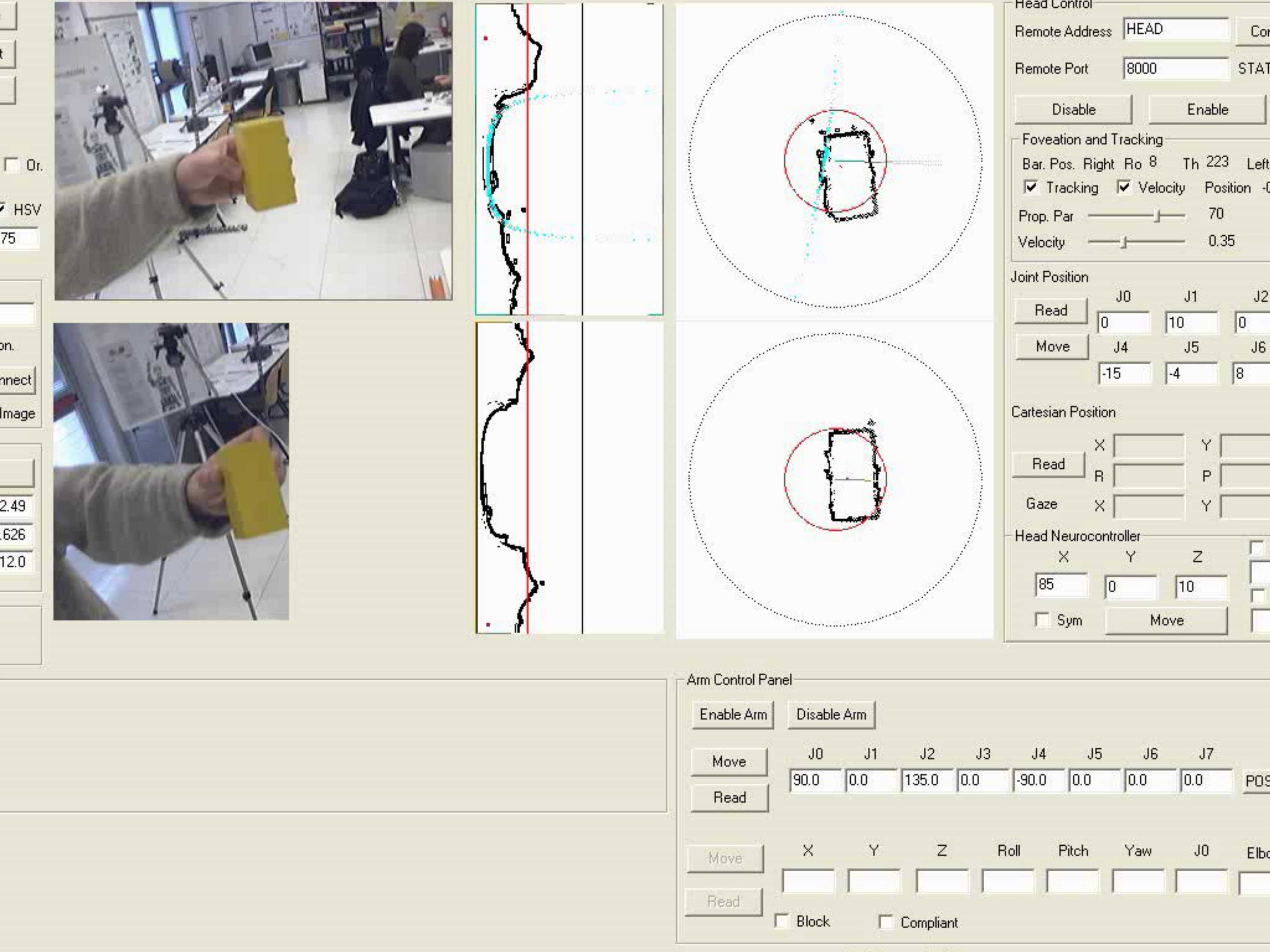
Edge Detection (gradient based method)



Line detection
(Hough method)



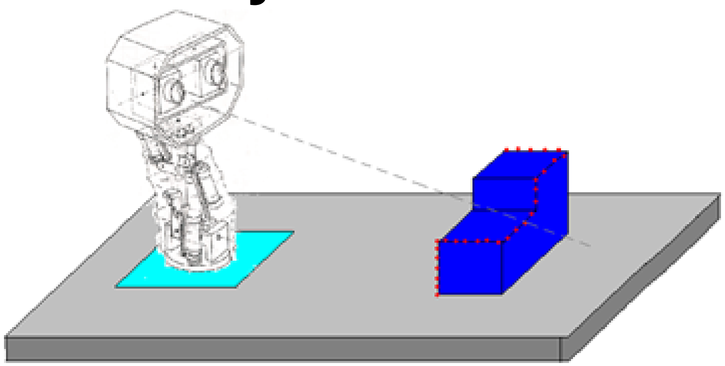
- Applied only to pixels belonging to the fovea

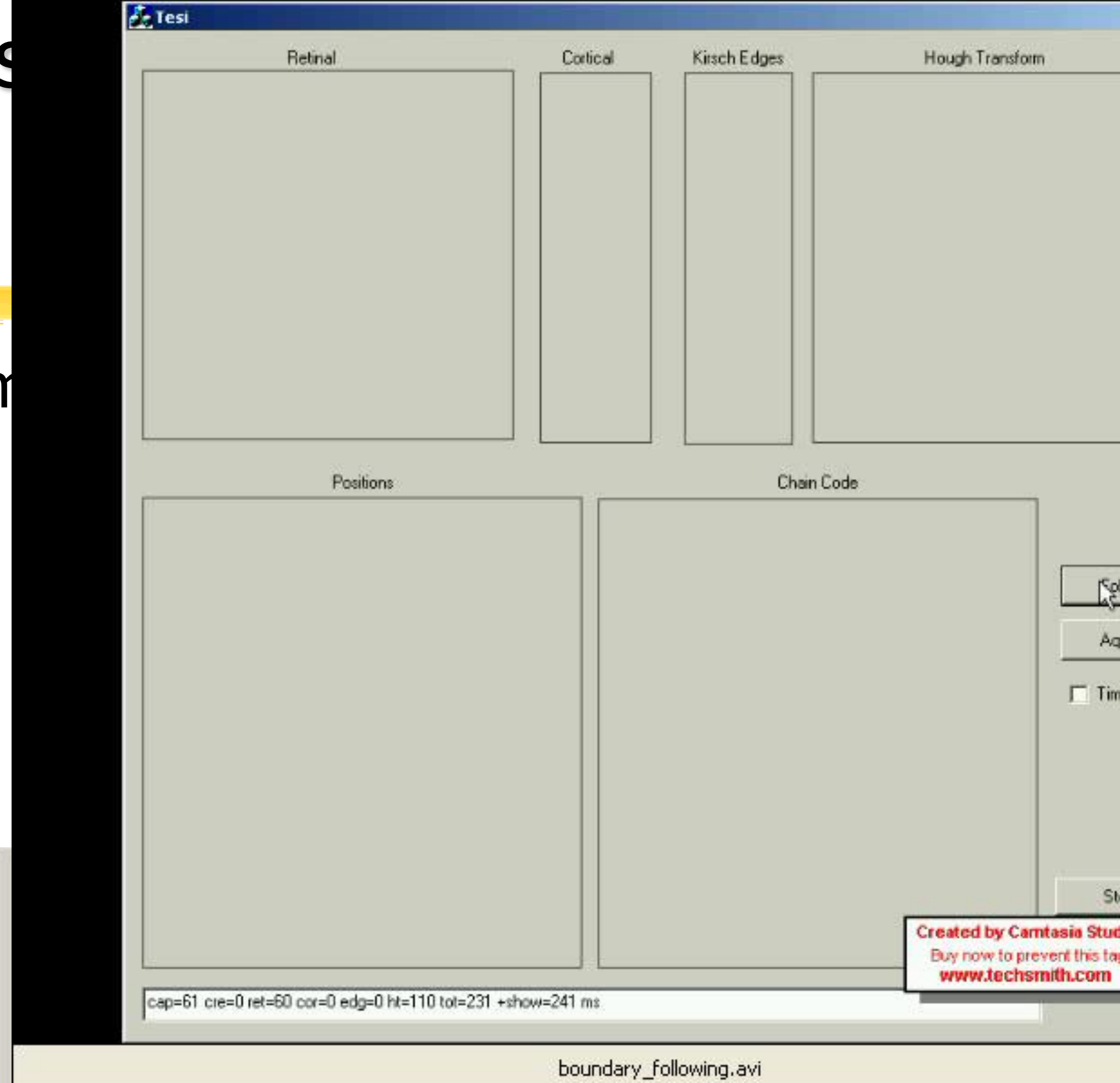


Head Control
Remote Address
HEAD
Remote Port
8000
Disable
Enable
Foveation and Tracking
Bar. Pos. Right Ro 8 Th 223
Tracking
Velocity
Position
Prop. Par
70
Velocity
0.35
Joint Position
Read
Move
J0 J1 J2
0 10 0
J4 J5 J6
-15 -4 8
Cartesian Position
Read
X Y
R P
Gaze X Y
Head Neurocontroller
X Y Z
85 0 10
Sym
Move
Or.
HSV
75
2.49
.626
12.0
Image
Arm Control Panel
Enable Arm
Disable Arm
Move
Read
J0 J1 J2 J3 J4 J5 J6 J7
90.0 0.0 135.0 0.0 -90.0 0.0 0.0 0.0
Move
Read
X Y Z Roll Pitch Yaw J0
Block
Compliant

Preliminary activities

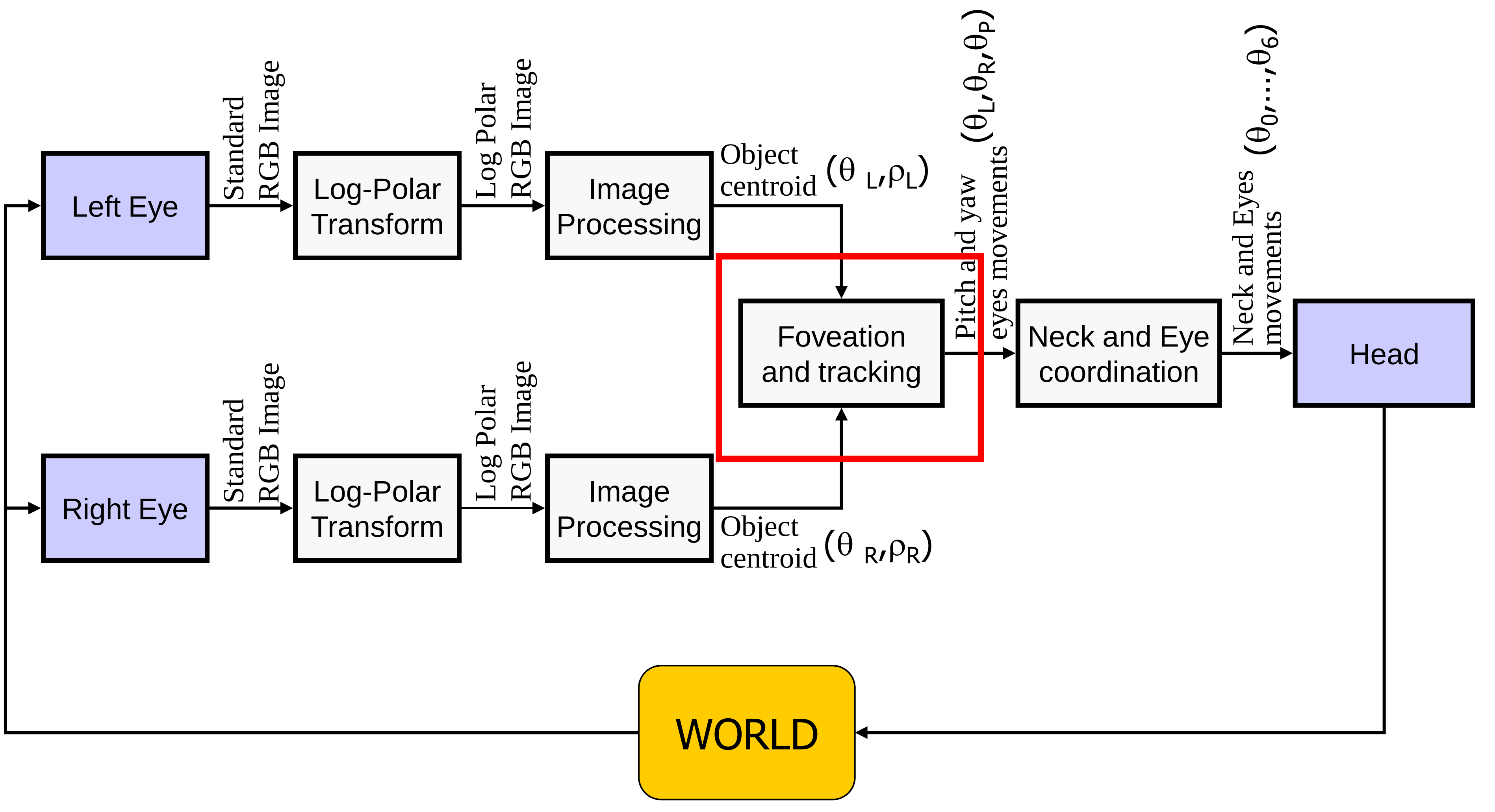
Foveation and tracking of borders of object and reconstruction of the geom of the object





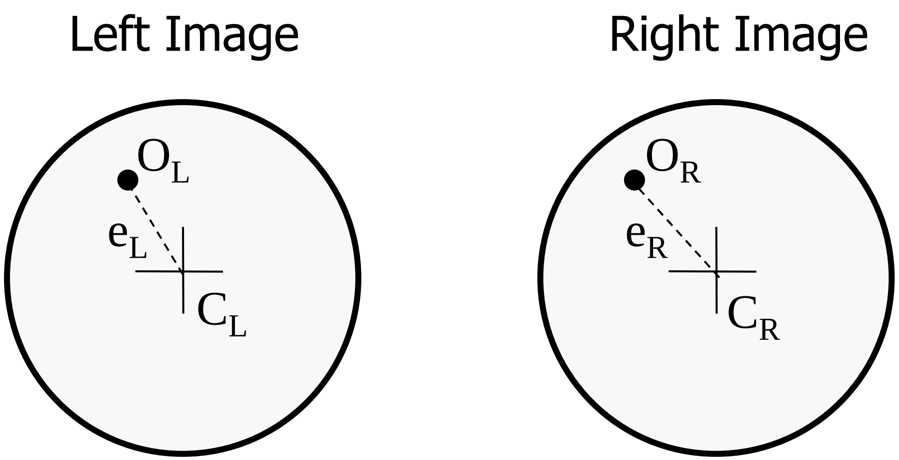
Retina Like image | Log Polar Image | Edge of log polar image | Detected lines (Boundaries) | Boundary reconstruction based on eye positions

Overall sensory-motor scheme of the visual apparatus



Foveation of the object centroid

Proportional control based on the visual error

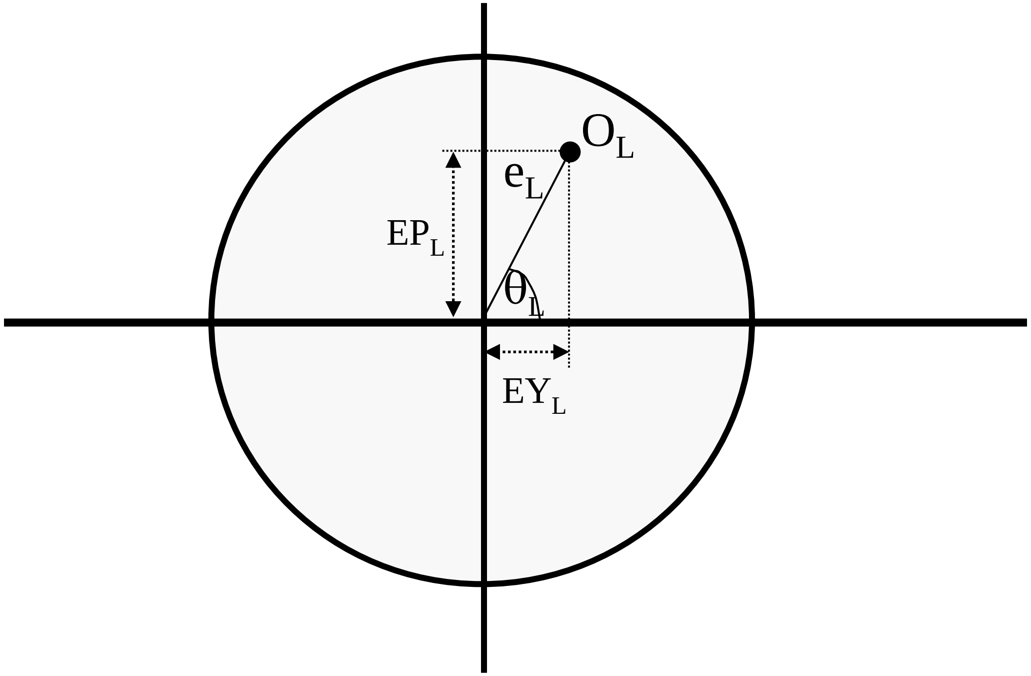


$O_L = (\rho_L, \theta_L)$ $\quad$ $O_R = (\rho_R, \theta_R)$

$e_L = \rho_L / M_{ro}$ $\quad$ $e_R = \rho_R / M_{ro}$

M_{ro} is the maximum ρ value (i.e. 152)

Computation of yaw and pitch eye movements



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

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

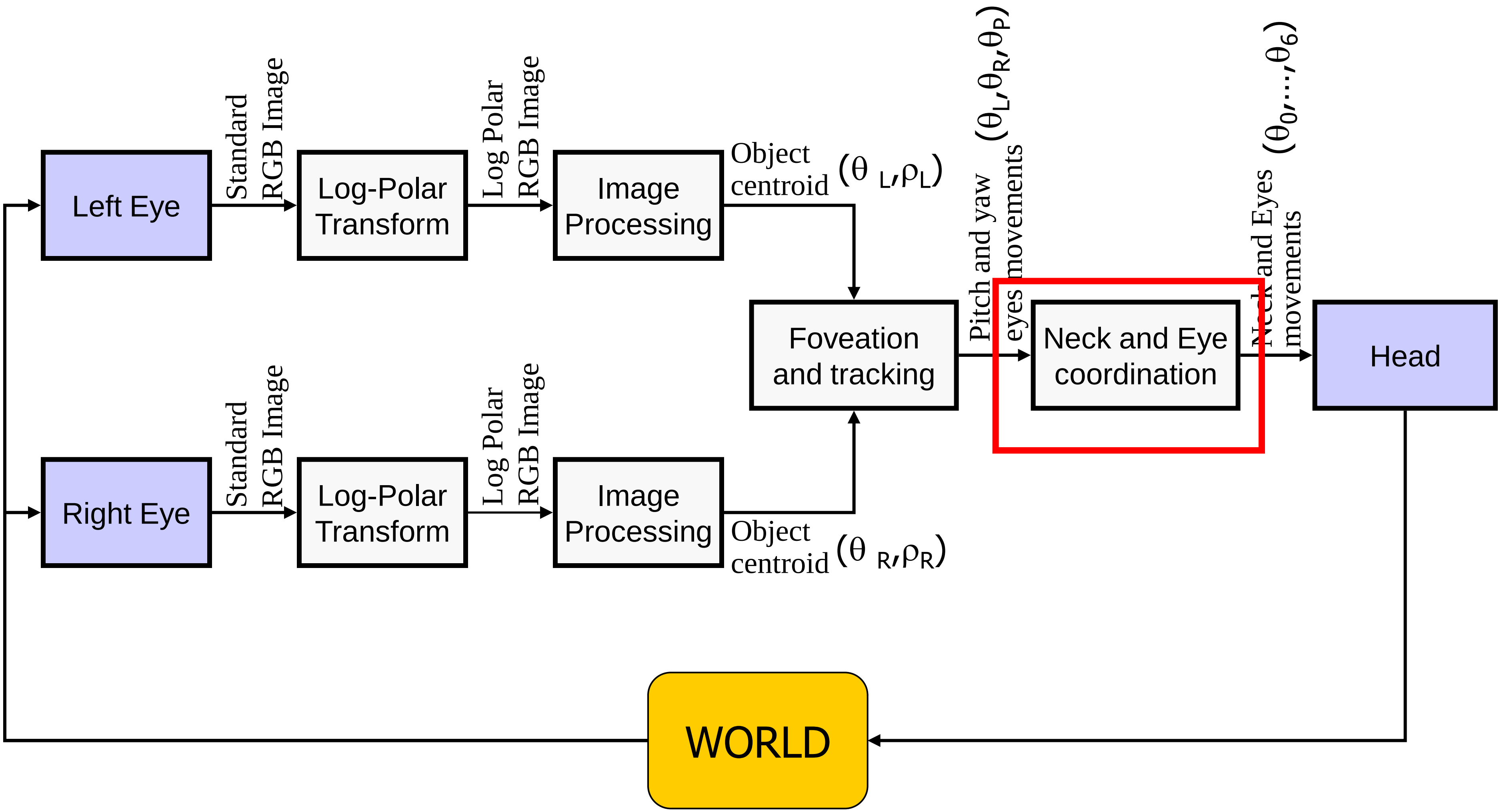
$EP_L = e_L * \sin(\theta_L) * P_L$

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

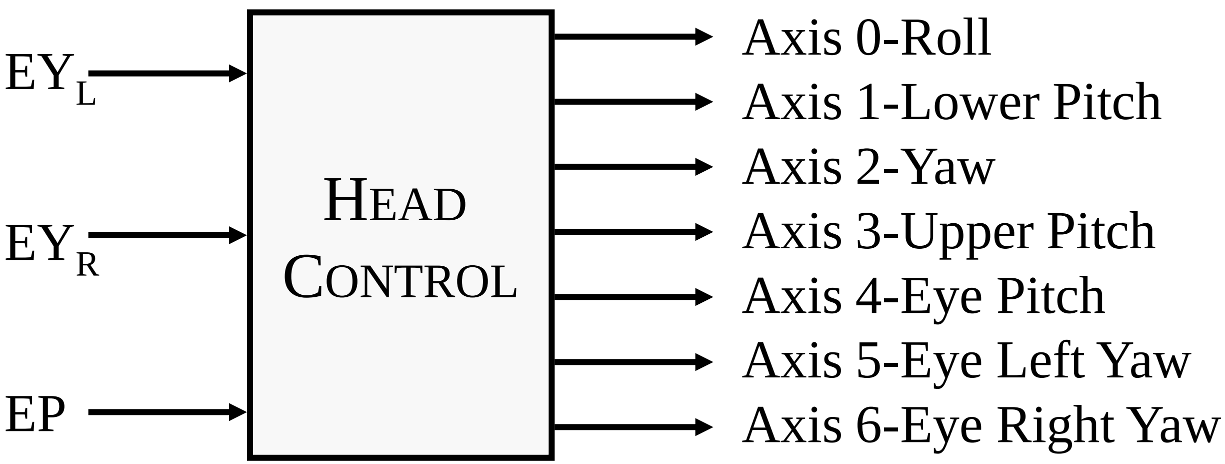
$EP=(EP_L+EP_R)/2$

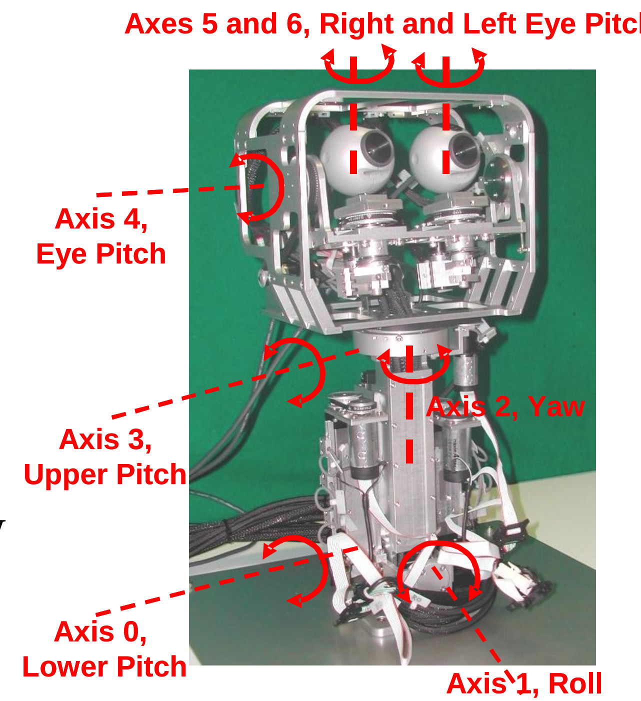
P_L and P_R are the proportional parameters for left and right eye, respectively.

Overall sensory-motor scheme of the visual apparatus



Eye-neck coordination



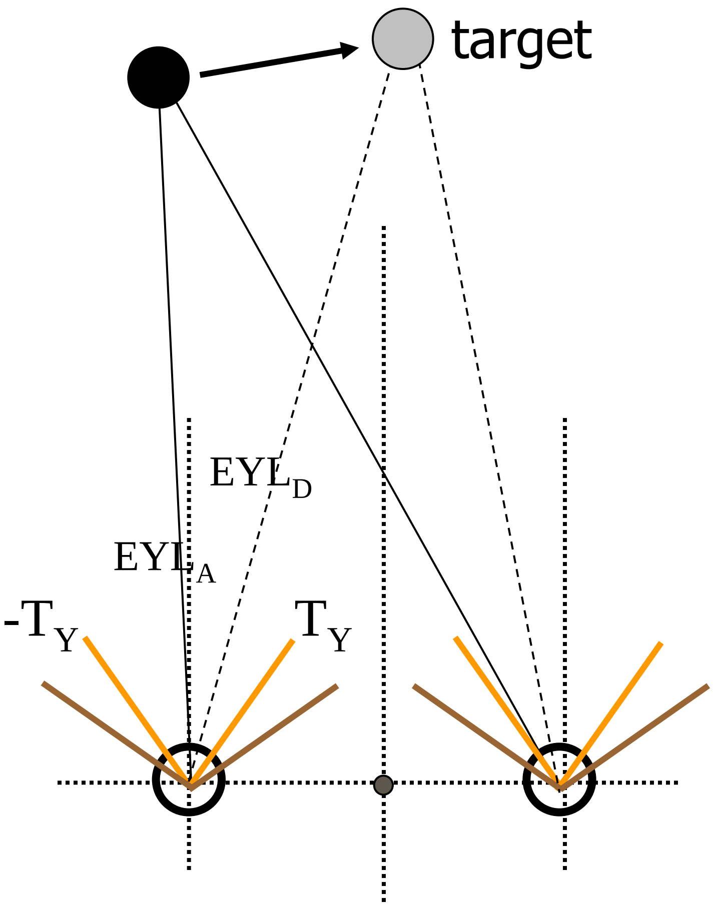


Solution 1

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



$$EYL_A + EY_L < T_Y$$

and

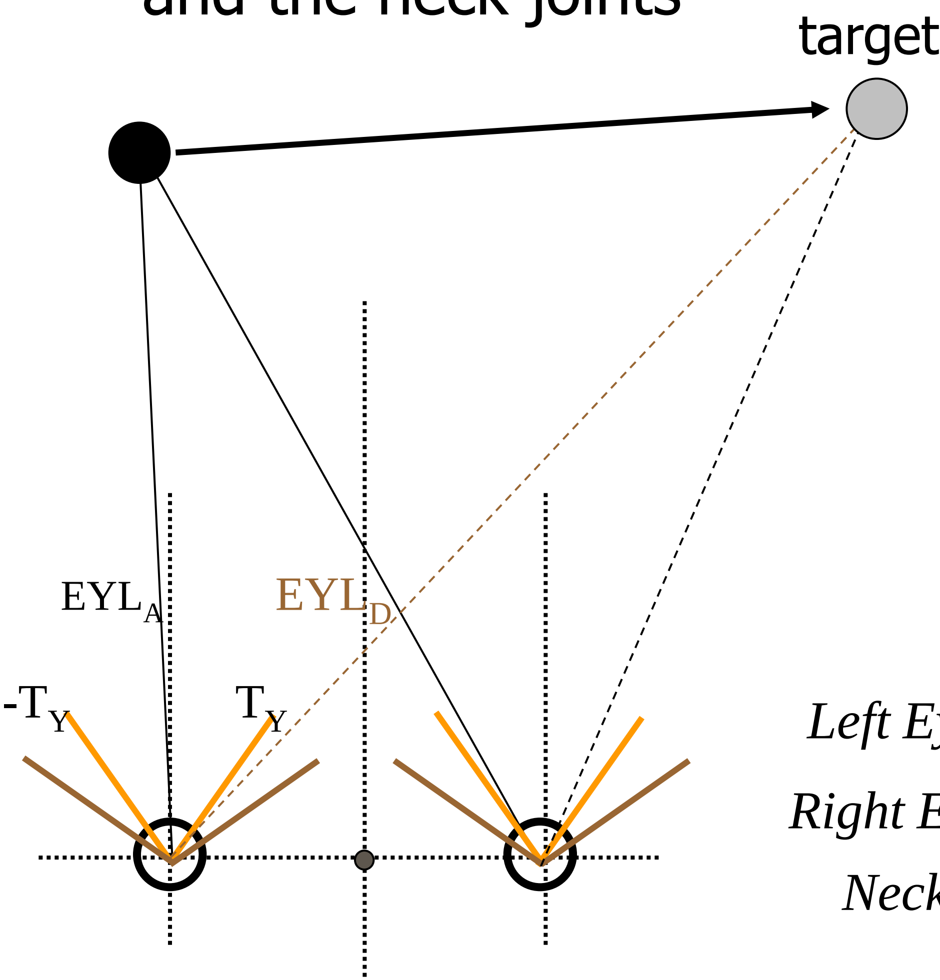
$$EYR_A + EY_R < T_Y$$

⇓

Left Eye Yaw	$EYL_D = EYL_A + EY_L$
Right Eye Yaw	$EYR_D = EYR_A + EY_R$
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.

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

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

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

$EYP_D = EYP_A + EP * K1$

$EUP_D = EUP_A + EP * K2$

$ELP_D = ELP_A + EP * K3$

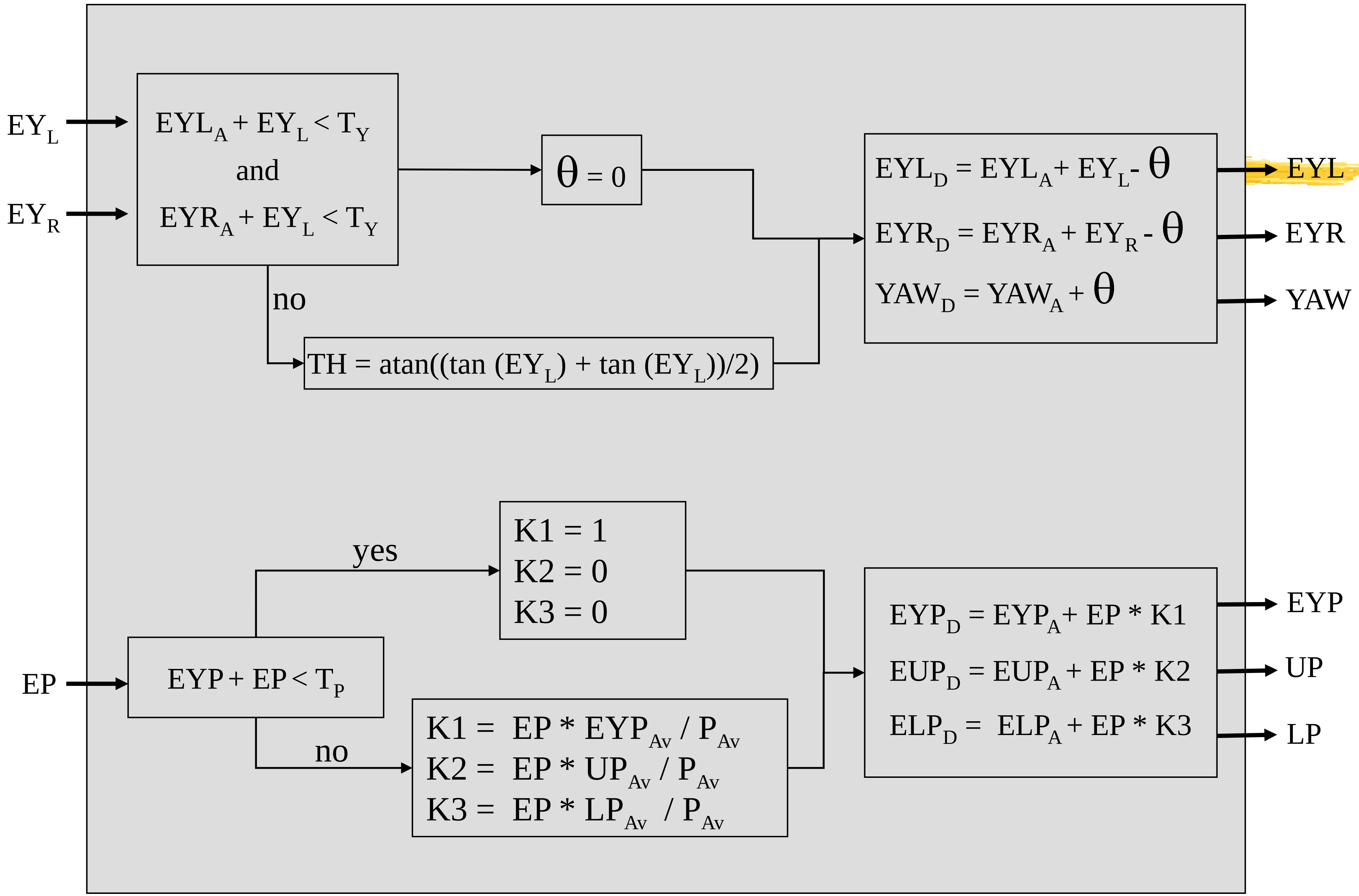
$EYP_{AV} = EYP_M - EYP_A$

$UP_{AV} = UP_M - UP_A$

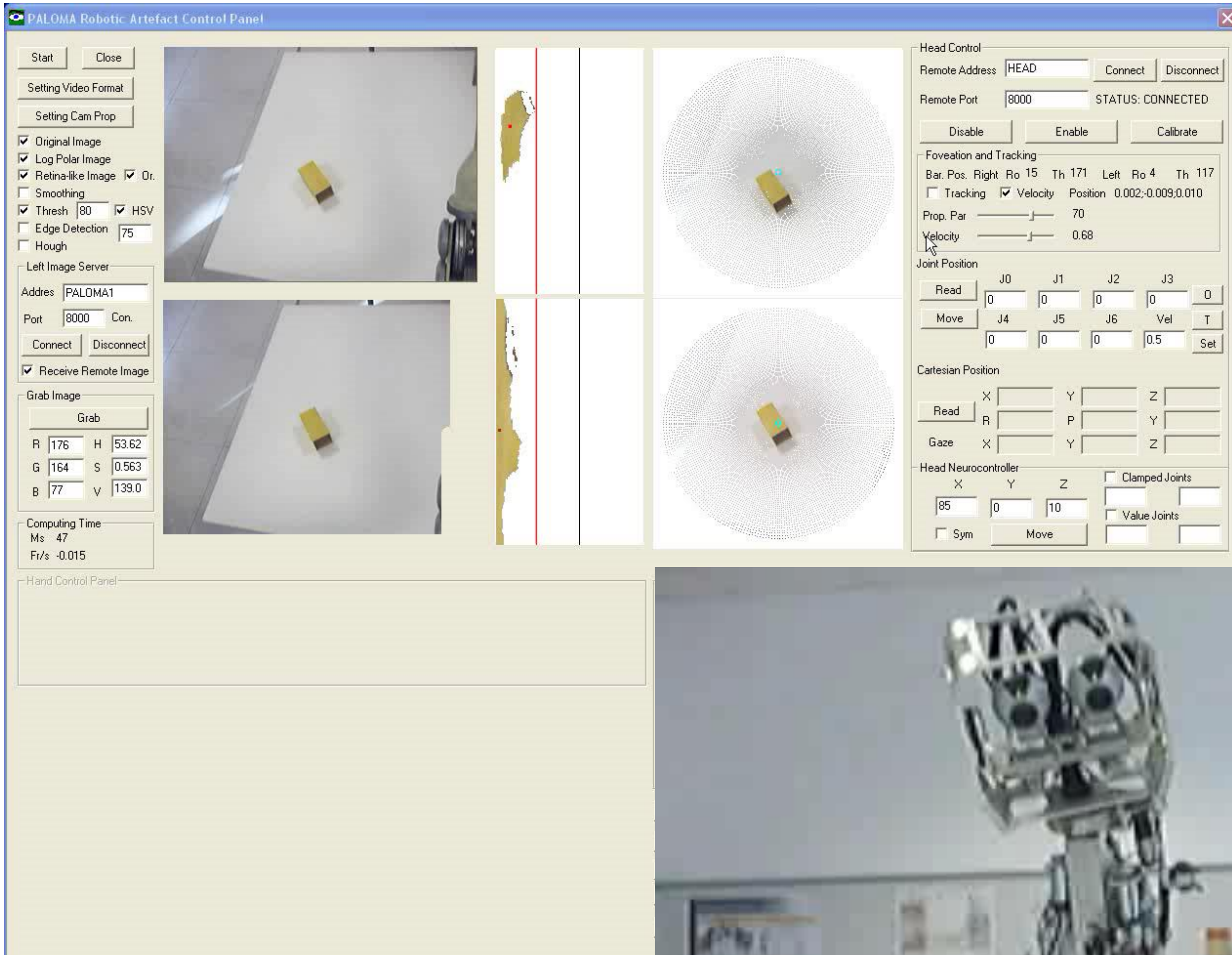
$LP_{AV} = LP_M - UP_A$

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

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



EYL
EYR
EYLA + EYL < TY
and
EYRA + EYL < TY
no
θ = 0
TH = atan((tan (EYL) + tan (EYL))/2)
EYLD = EYLA+ EYL- θ
EYRD = EYRA + EYR - θ
YAWD = YAWA + θ
EYL
EYR
YAW
EP
EYP + EP < TP
yes
no
K1 = 1
K2 = 0
K3 = 0
K1 = EP * EYPAv / PAv
K2 = EP * UPAv / PAv
K3 = EP * LPAv / PAv
EYPD = EYPA+ EP * K1
EUPD = EUPA + EP * K2
ELPD = ELPA + EP * K3
EYP
UP
LP



Pursuit Movement

Frame rate: 10 fps for both images

Head Control loop: 100 ms