Corso di Robotica (ROB)

C. Modulo di Robotica Bioispirata

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

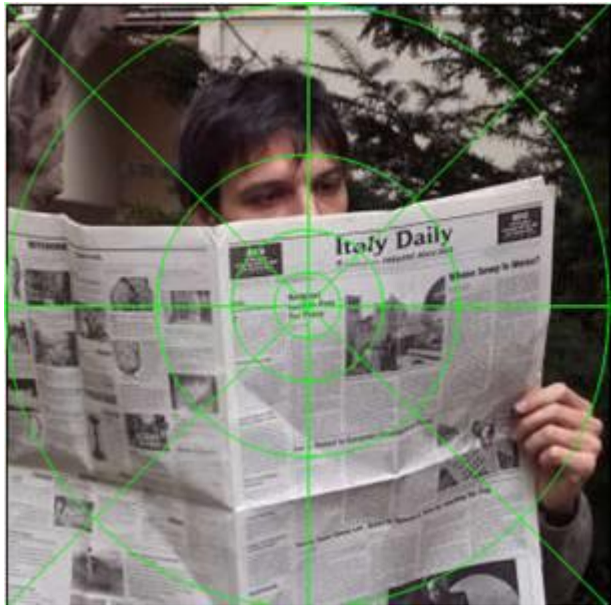
Retina-like image

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

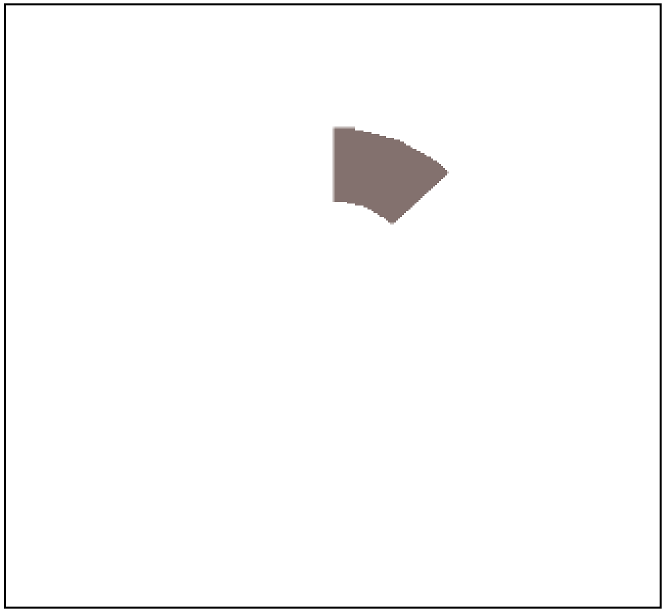
Log-polar projection

Costruzione di un'immagine retinica

Immagine cartesiana tradizionale



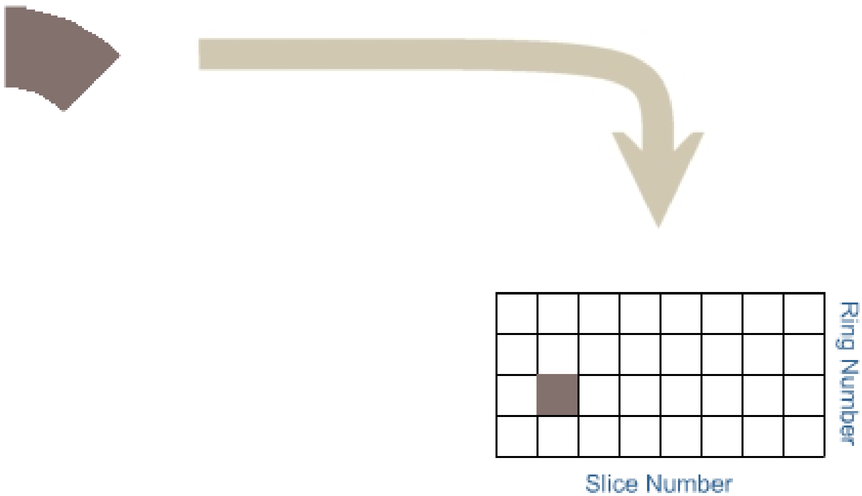
Divisione in circonferenze e spicchi



Calcolo del valore medio di un settore

http://www.retinica.com/

Costruzione di un’immagine retinica

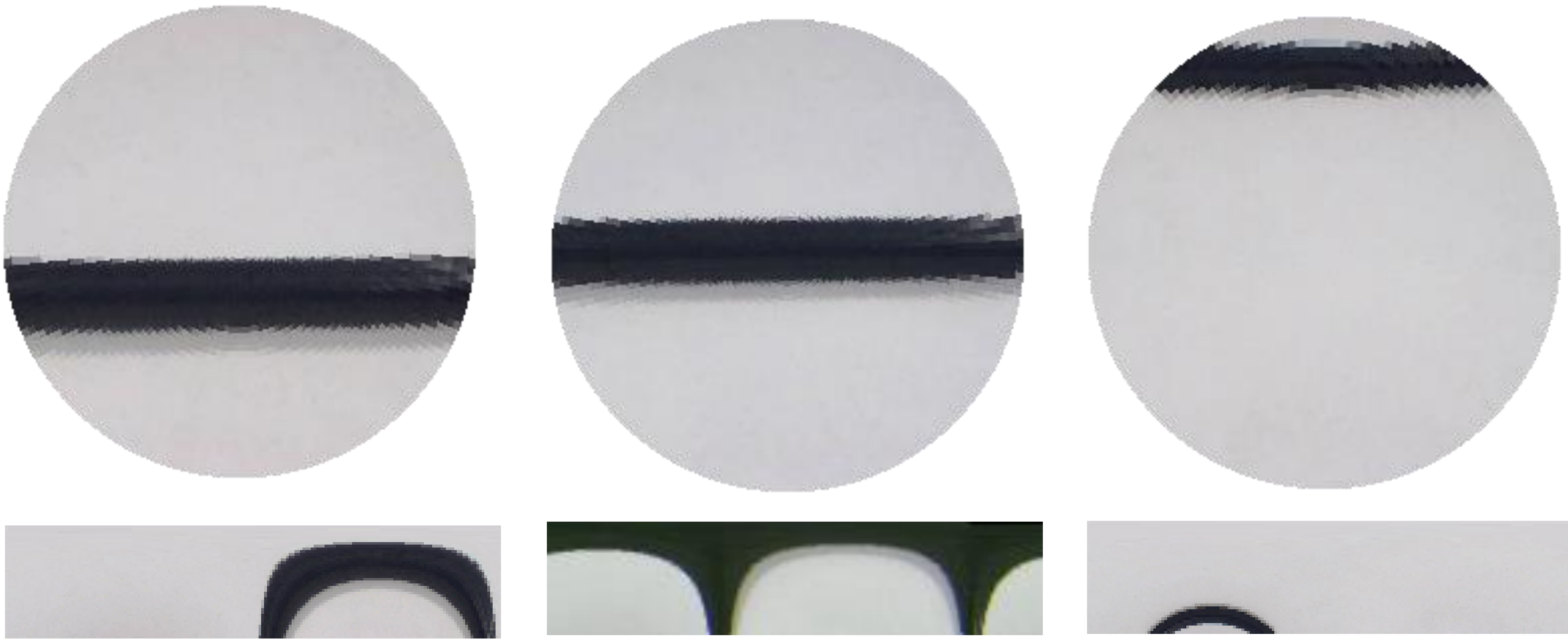


Copia del valore medio di un settore in un pixel di un’immagine polare

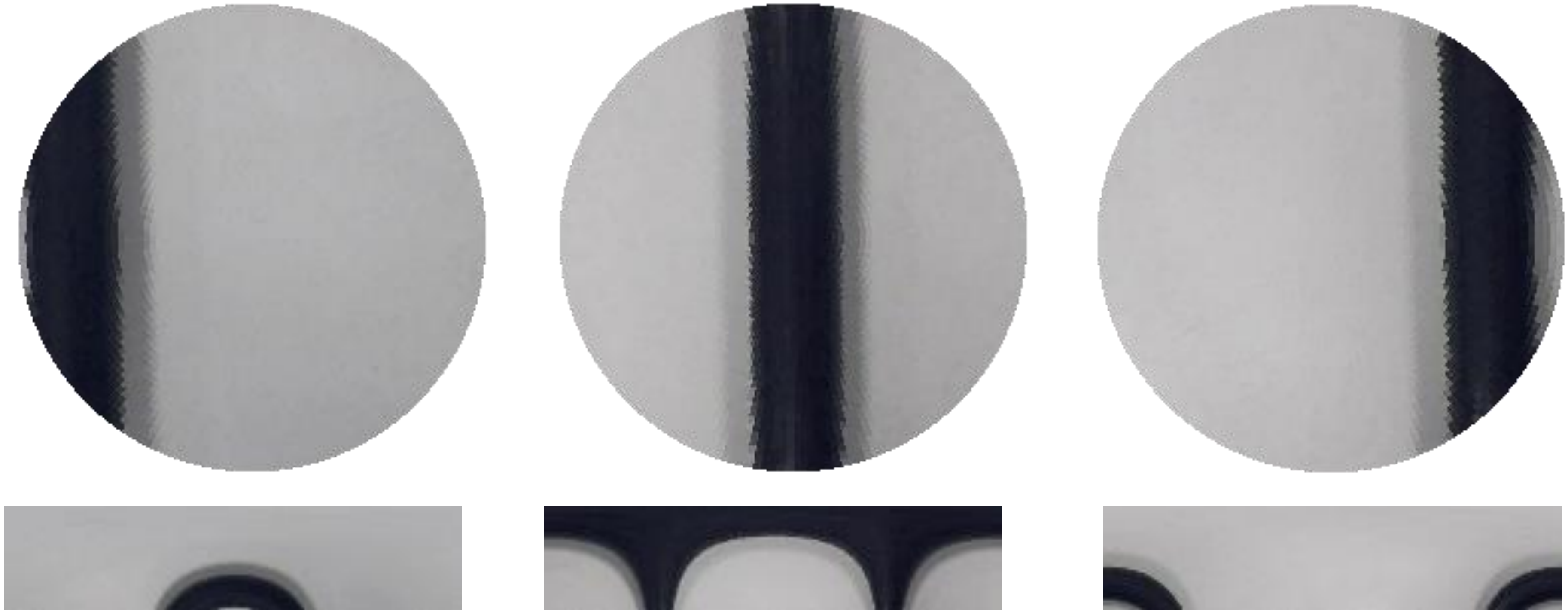
Immagine polare risultante

Immagine cartesiana ricostruita dalla polare

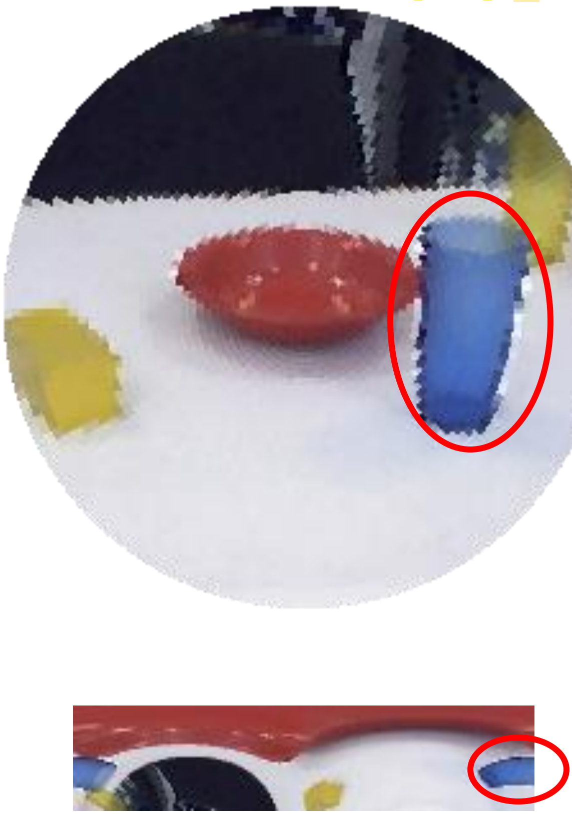
An example of pattern translation



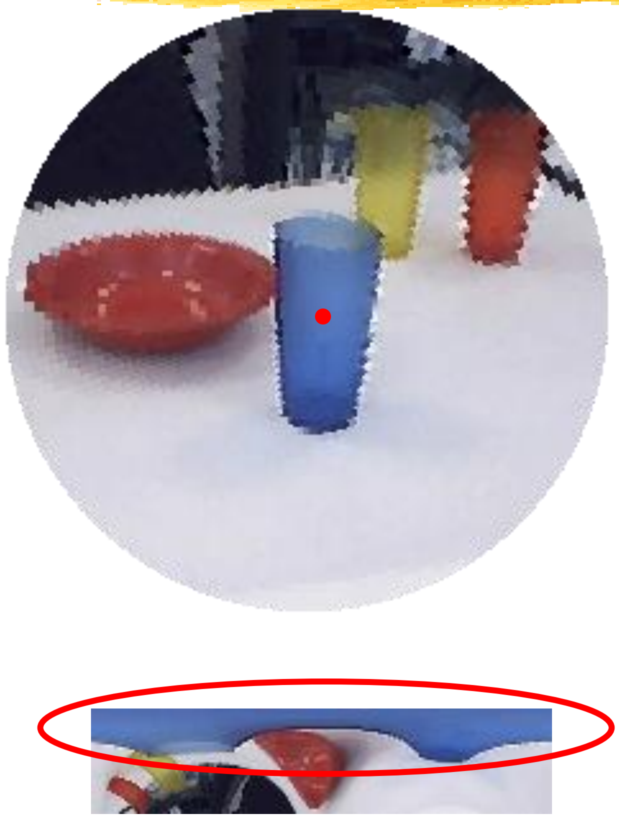
An example of pattern translation



An example of simulated foveation



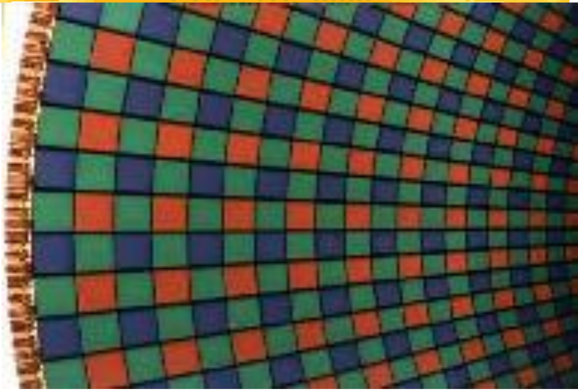
Object detection
in the periphery



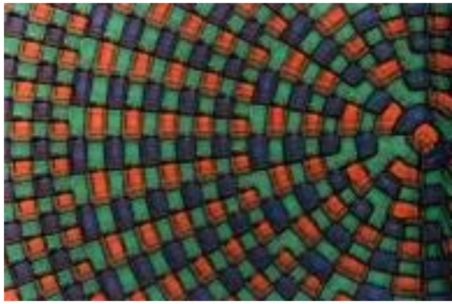
Object foveation

Foveation of a
point of interest
(edge)

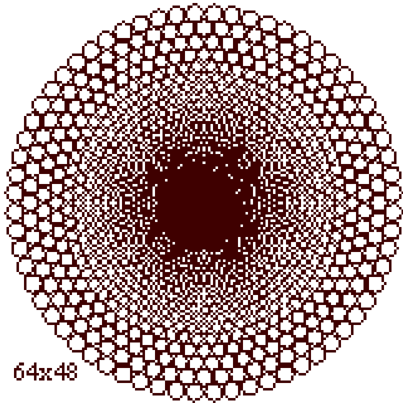
The Retina-like Giotto cameras



- 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

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

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

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

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

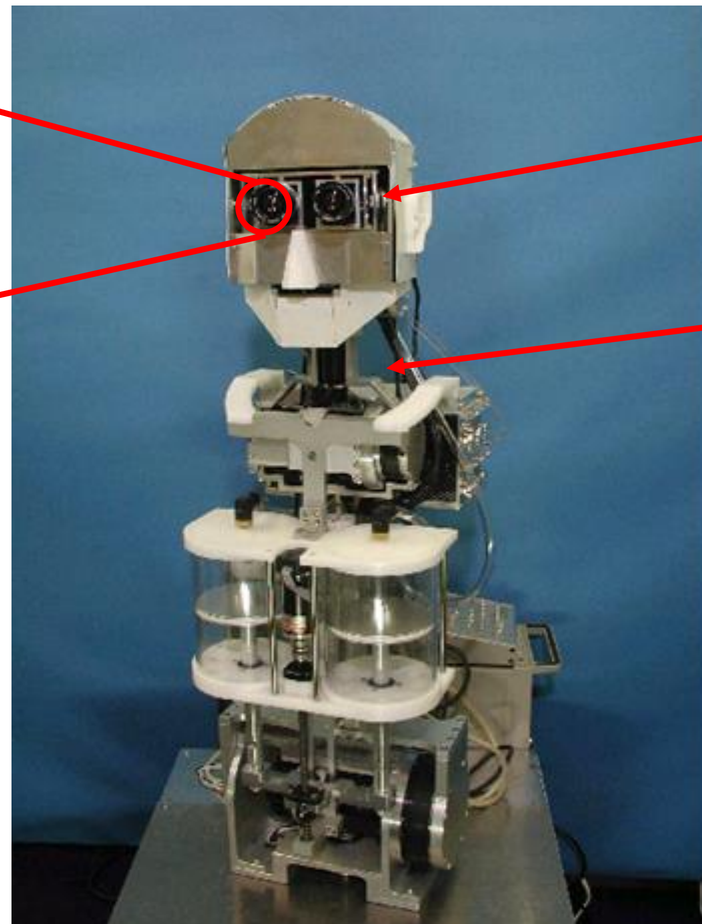
F = size of the fovea in rings.
P = total number of rings.
Θ = maximum # of pixels in each ring.

$2X$ = 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

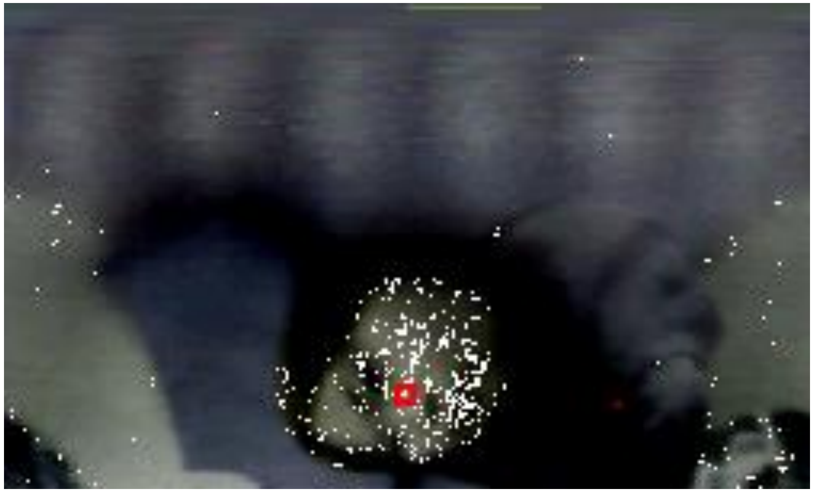
Face detection by hue

Hue = information on the color

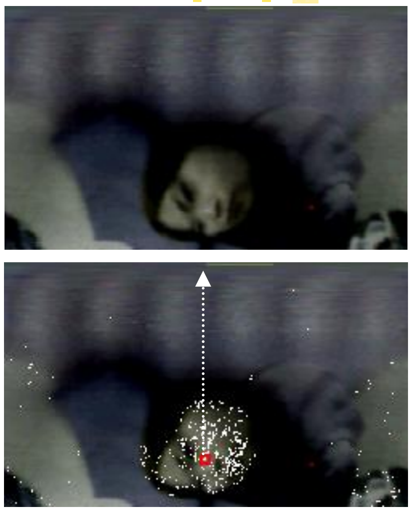
$$\text{Hue} = \cos^{-1}\left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right)$$

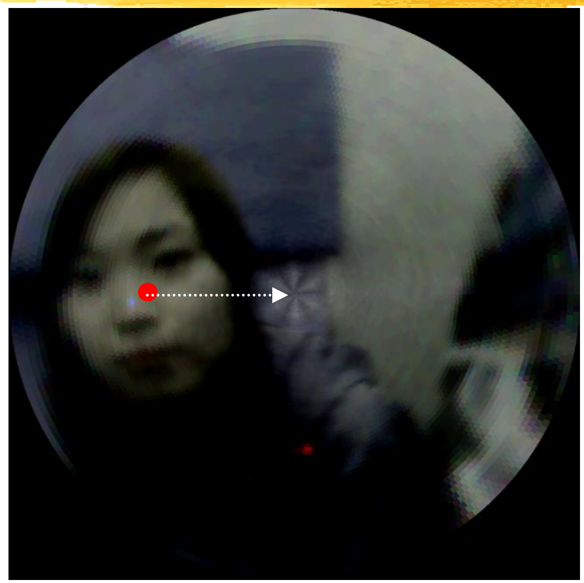
if B>G then Hue = 2π - Hue

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

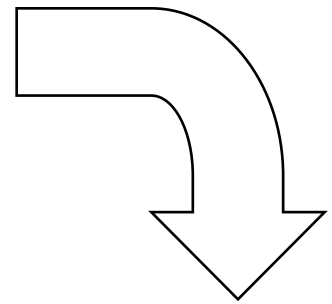


An example of foveation



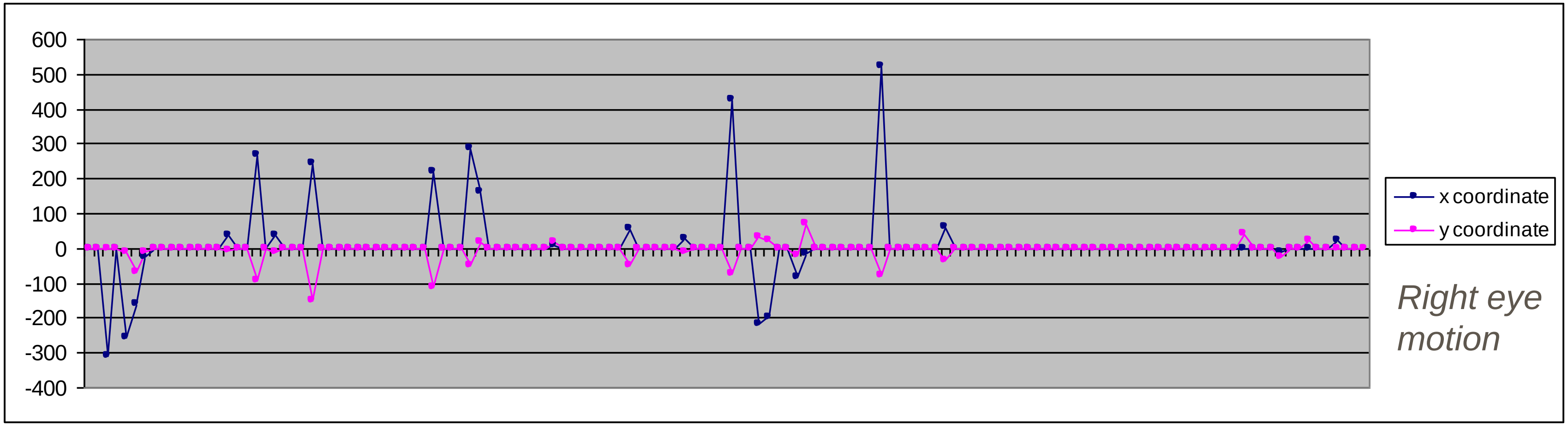


Eye/neck movements



Proportions are rescaled for display purposes

Experimental trials



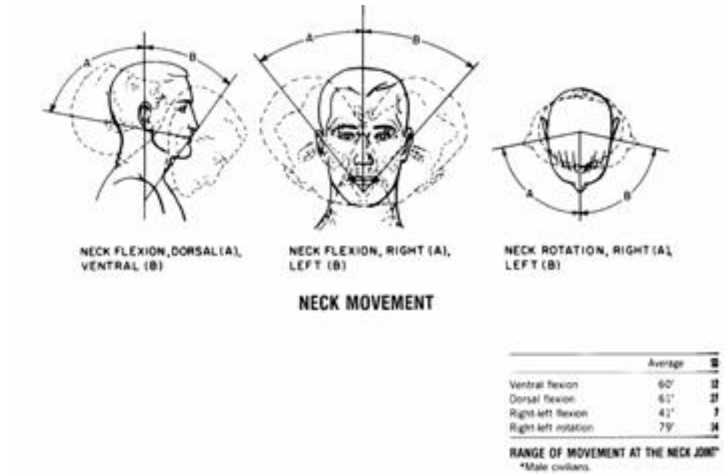
[Cecilia Laschi, Hiroyasu Miwa, Atsuo Takanishi, Eugenio Guglielmelli, Paolo Dario, 2002]

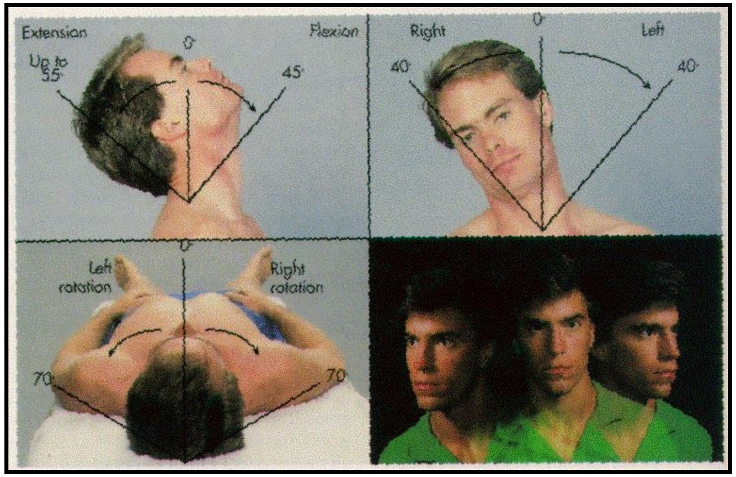
Example of design and development of a human-like robotic head

The ARTS humanoid robot head

Synthesis of characteristics of the human oculo-motor system

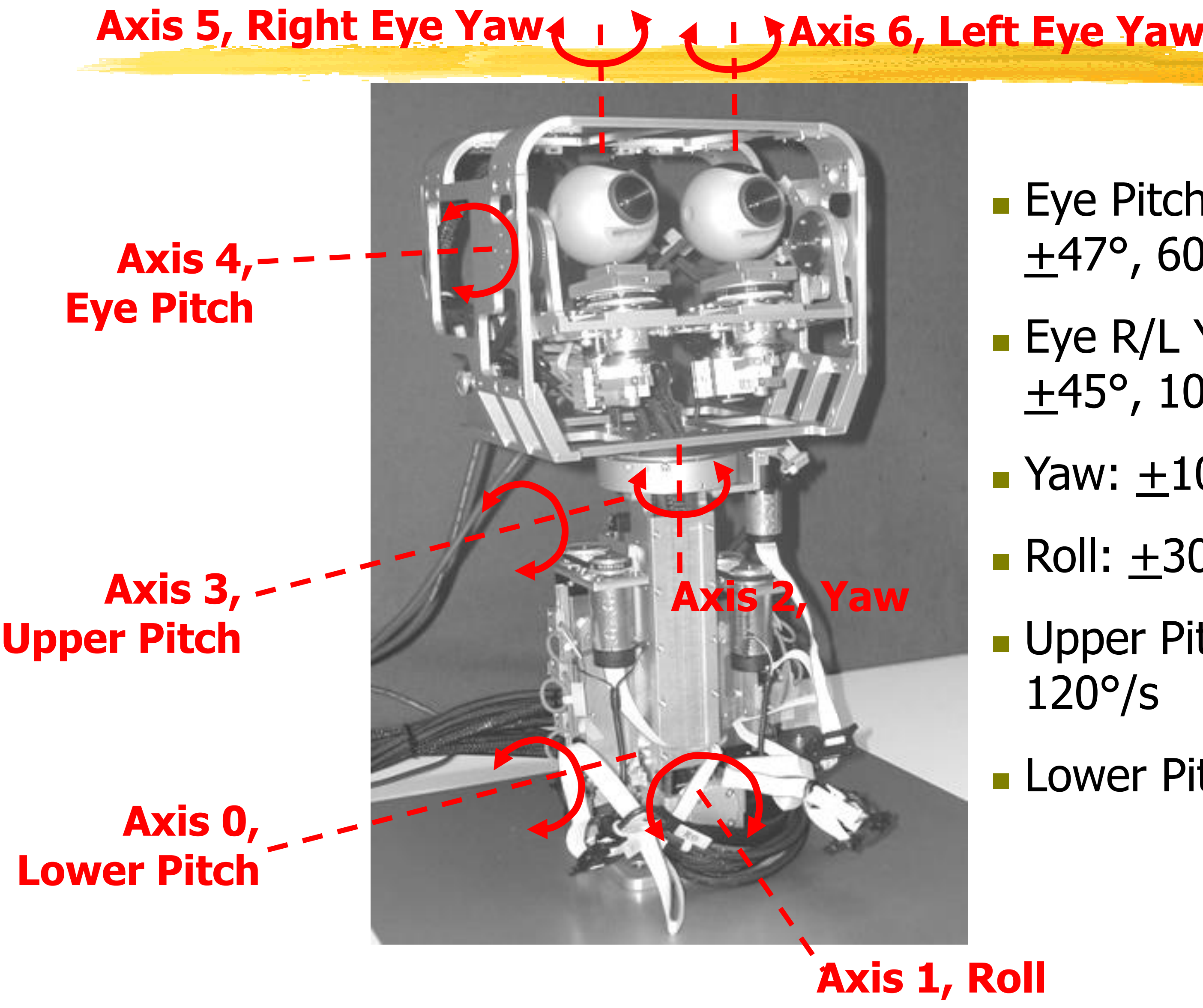
- Eye movements:
  - Saccades
  - Vergence
  - Pursuit
- Ranges of motion:
  - 120° for the tilt eye movements
  - 60° for the pan eye movements
- Eye speed:
  - Up to 900°/sec (in saccades)
- Inter-ocular distance: between 60 and 80 mm





[Thibodeau & Patton, 1996]

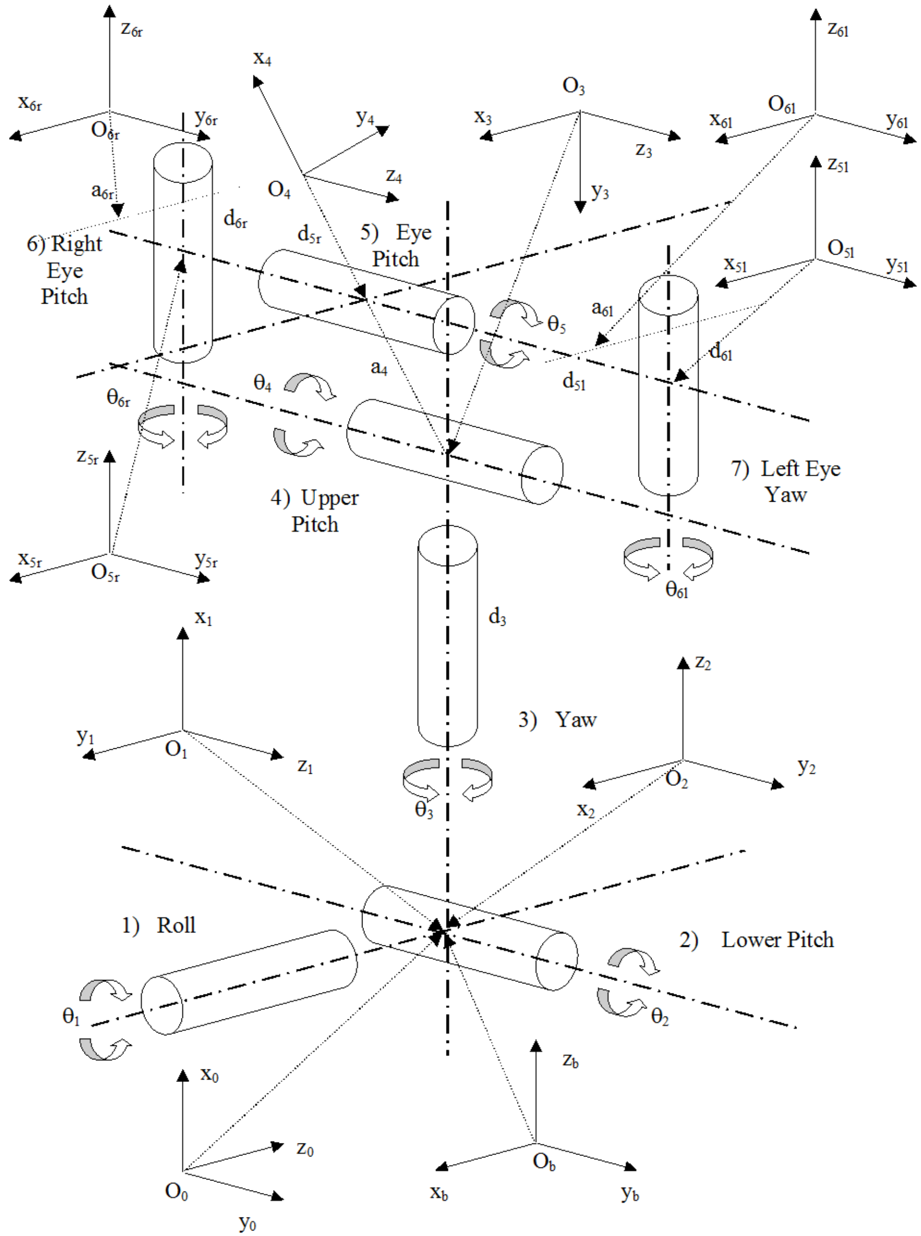
Kinematic structure of the SSSA Robot Head



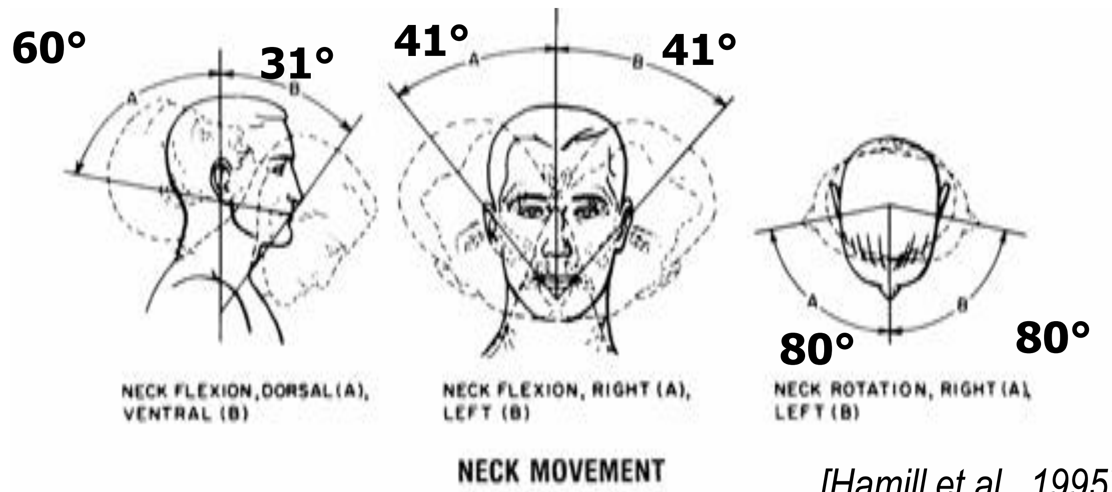
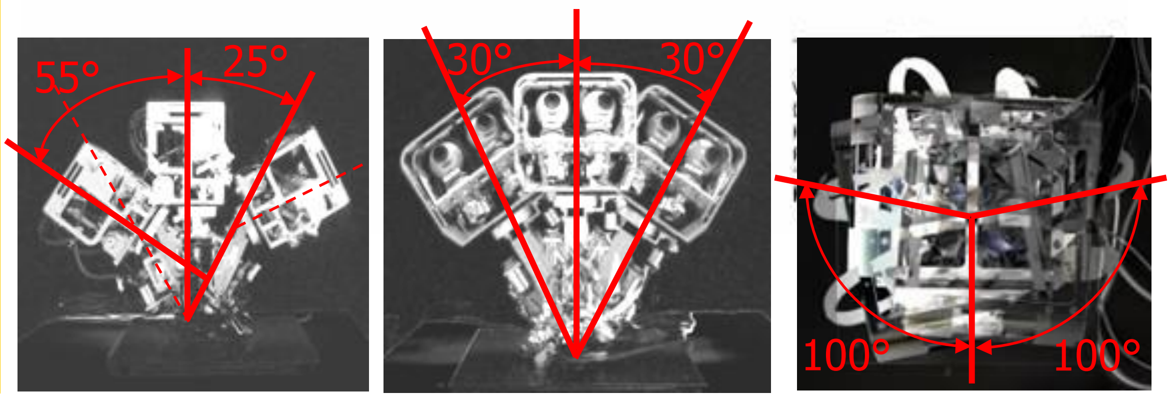
- Eye Pitch Axis: $\pm 47°$, 600°/s
- Eye R/L Yaw Axis: $\pm 45°$, 1000°/s
- Yaw: $\pm 100°$, 170°/s
- Roll: $\pm 30°$, 25°/s
- Upper Pitch: $\pm 30°$, 120°/s
- Lower Pitch: $\pm 25°$, 20°/s

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: $\underline{+}$ 60°, 600°/s Yaw: $\underline{+}$ 30°, 600°/s
Robot		Pitch: $\underline{+}$ 47°, 600°/s Yaw: $\underline{+}$ 45°, 1000°/s

The movements of the 7 dofs of the robotic head

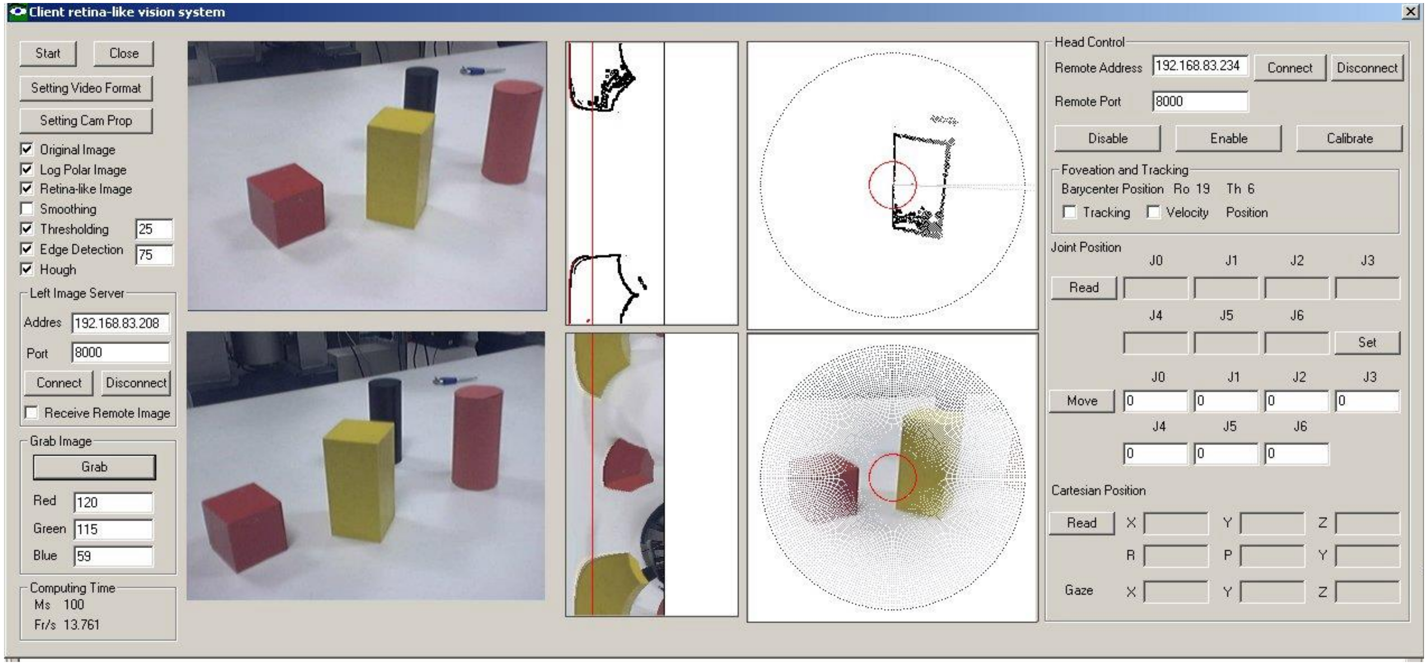
head_performances.avi

Examples of algorithms developed for retina-like image processing

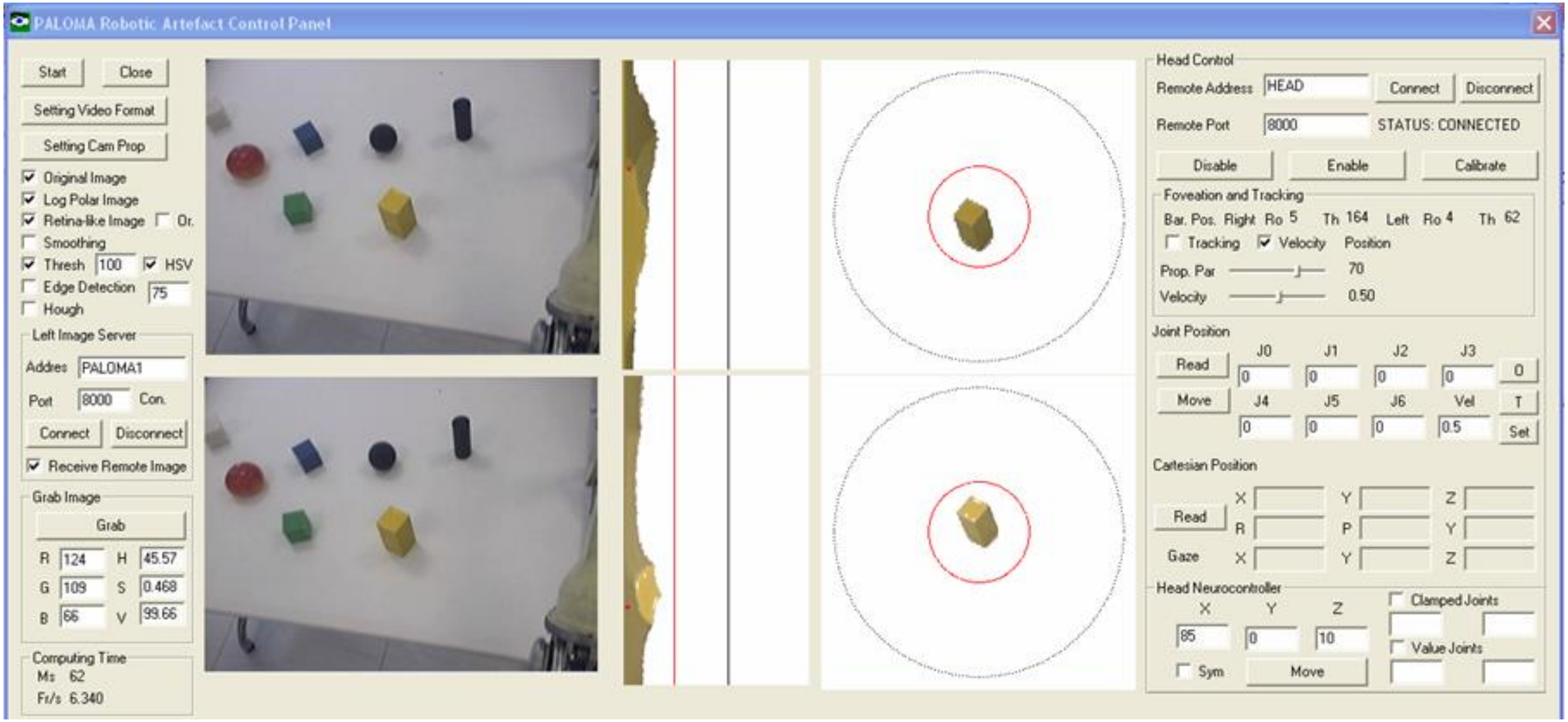
- Acquiring standard image
- Creating log-polar image from standard image
- Creating retina-like image from log-polar image
- Thresholding of image based on RGB and HUE
- Computation of the centroid of a thresholded area
- Edge detection
- Line detection

Simulation of retina-like cameras and basic image processing

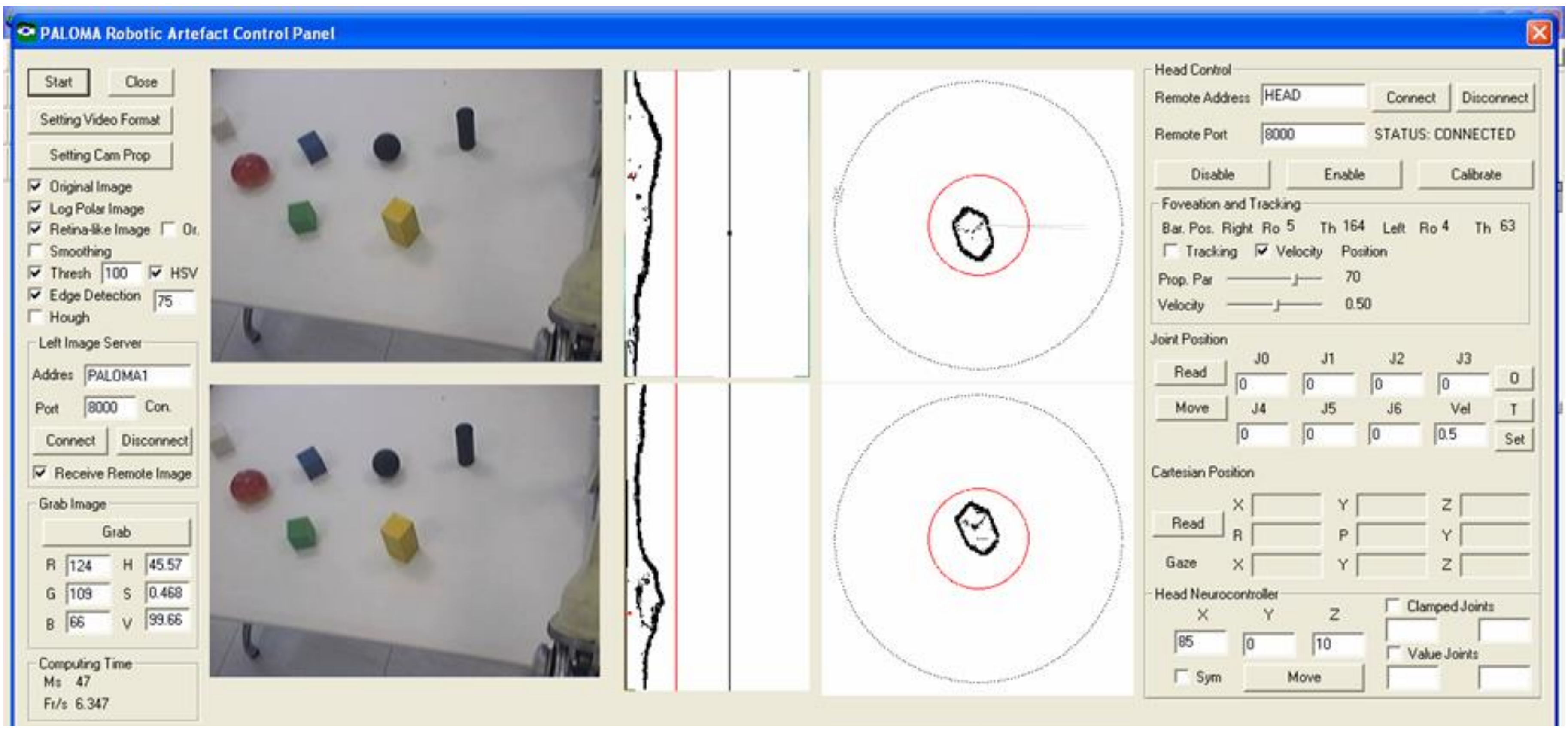
- Acquiring standard image
- Creating log-polar image from standard image
- Creating retina-like image from log-polar image



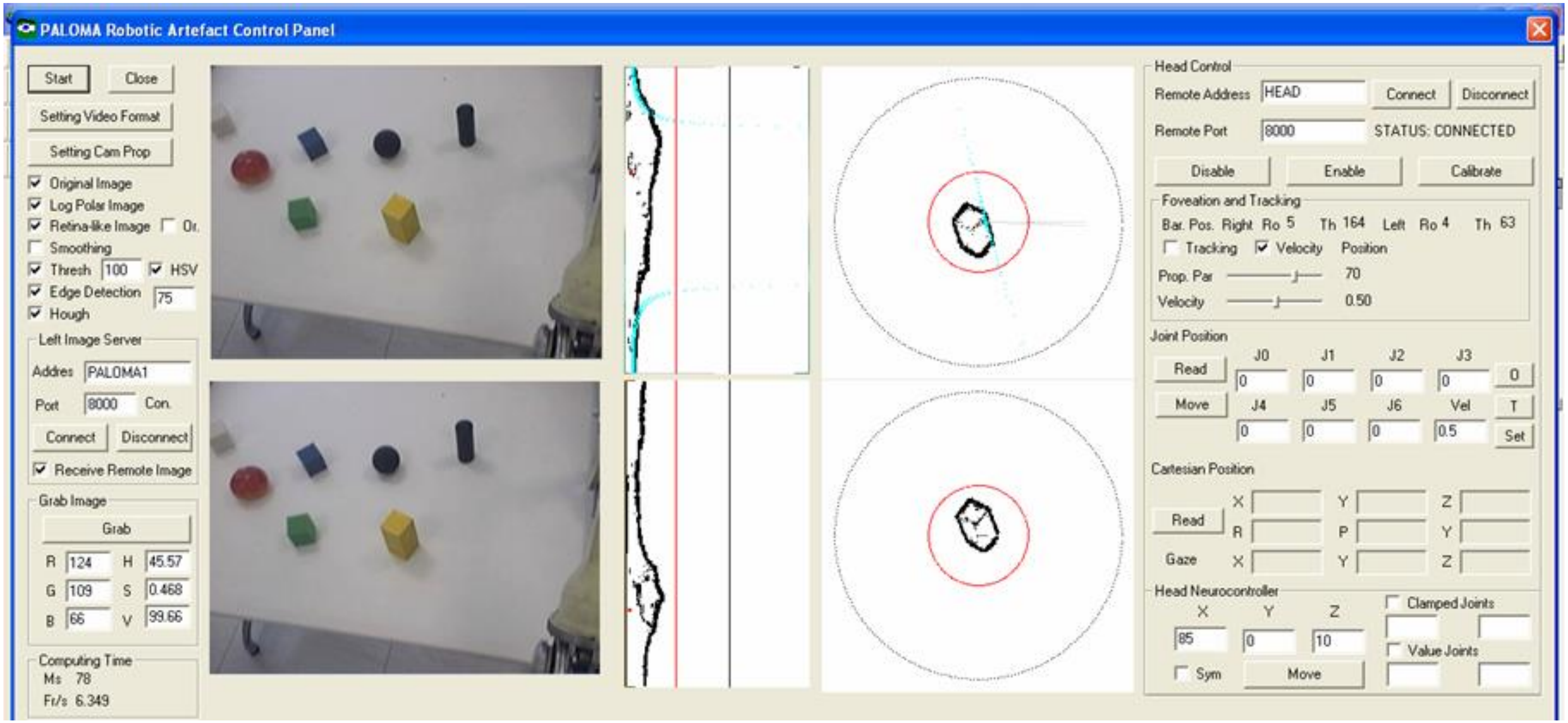
Thresholding of image based on RGB and HUE



Edge Detection (gradient based method)

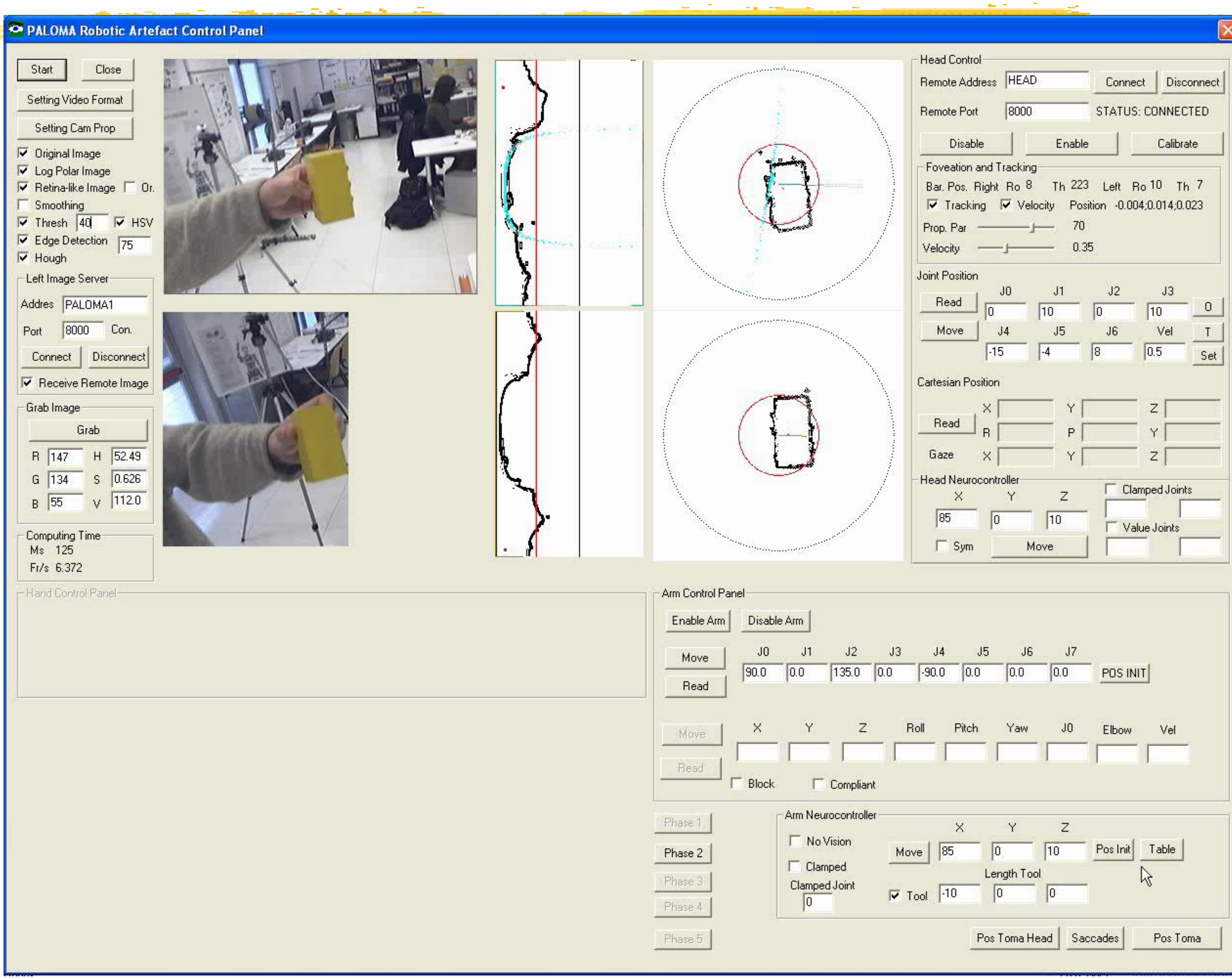


Line detection (Hough method)

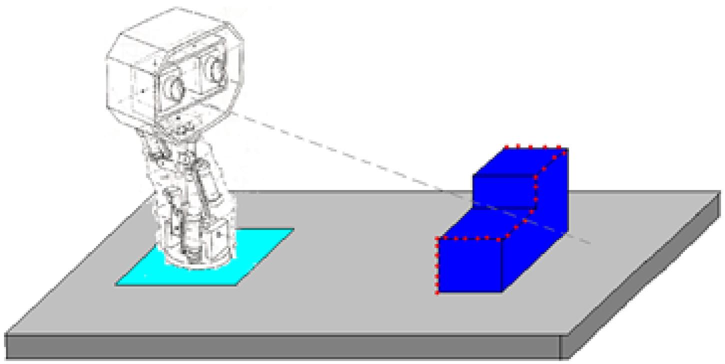


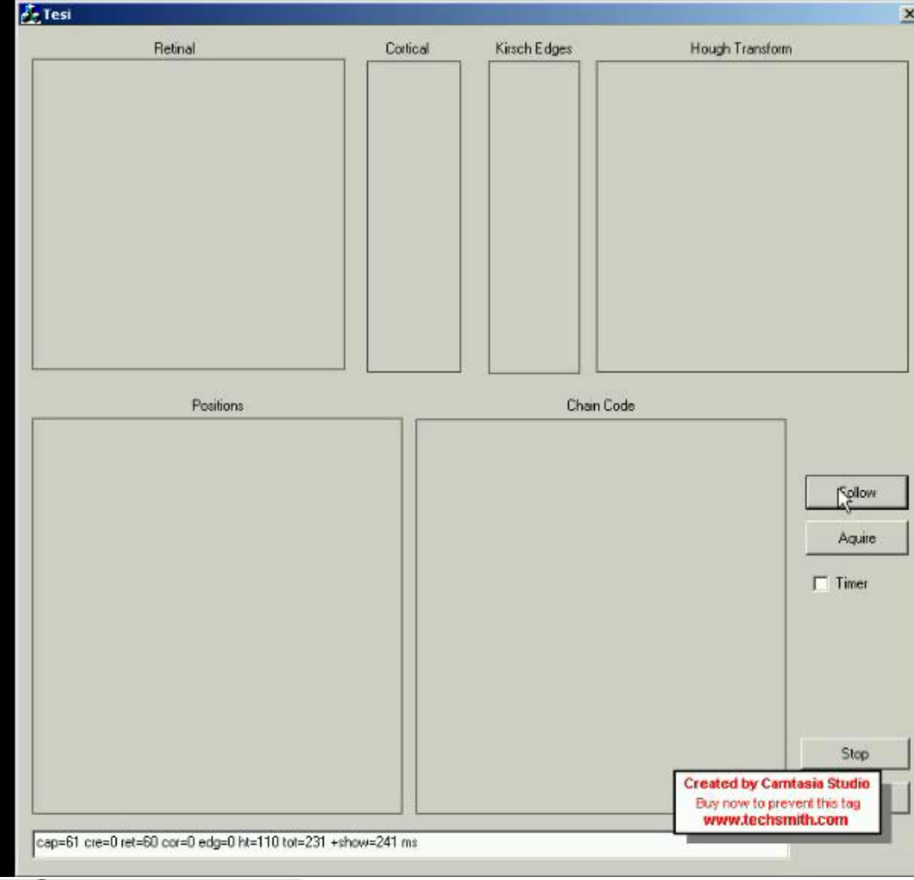
- Applied only to pixels belonging to the fovea

Line detection



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

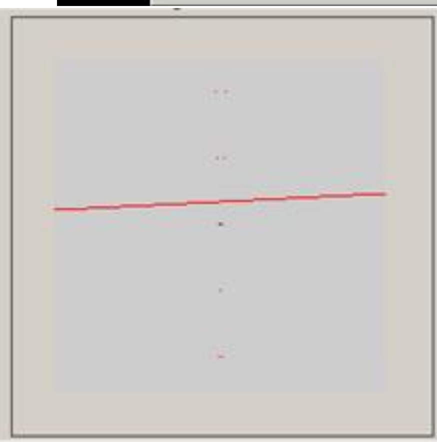




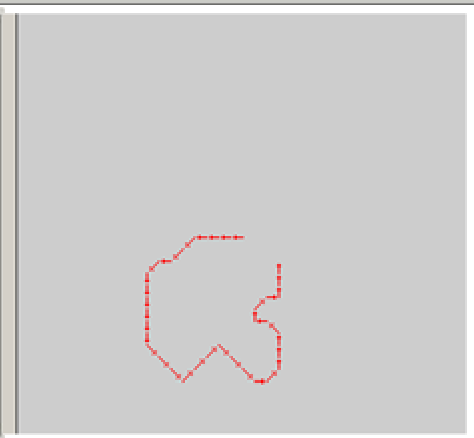
Retina Like image

Log Polar Image

Edge of log polar image

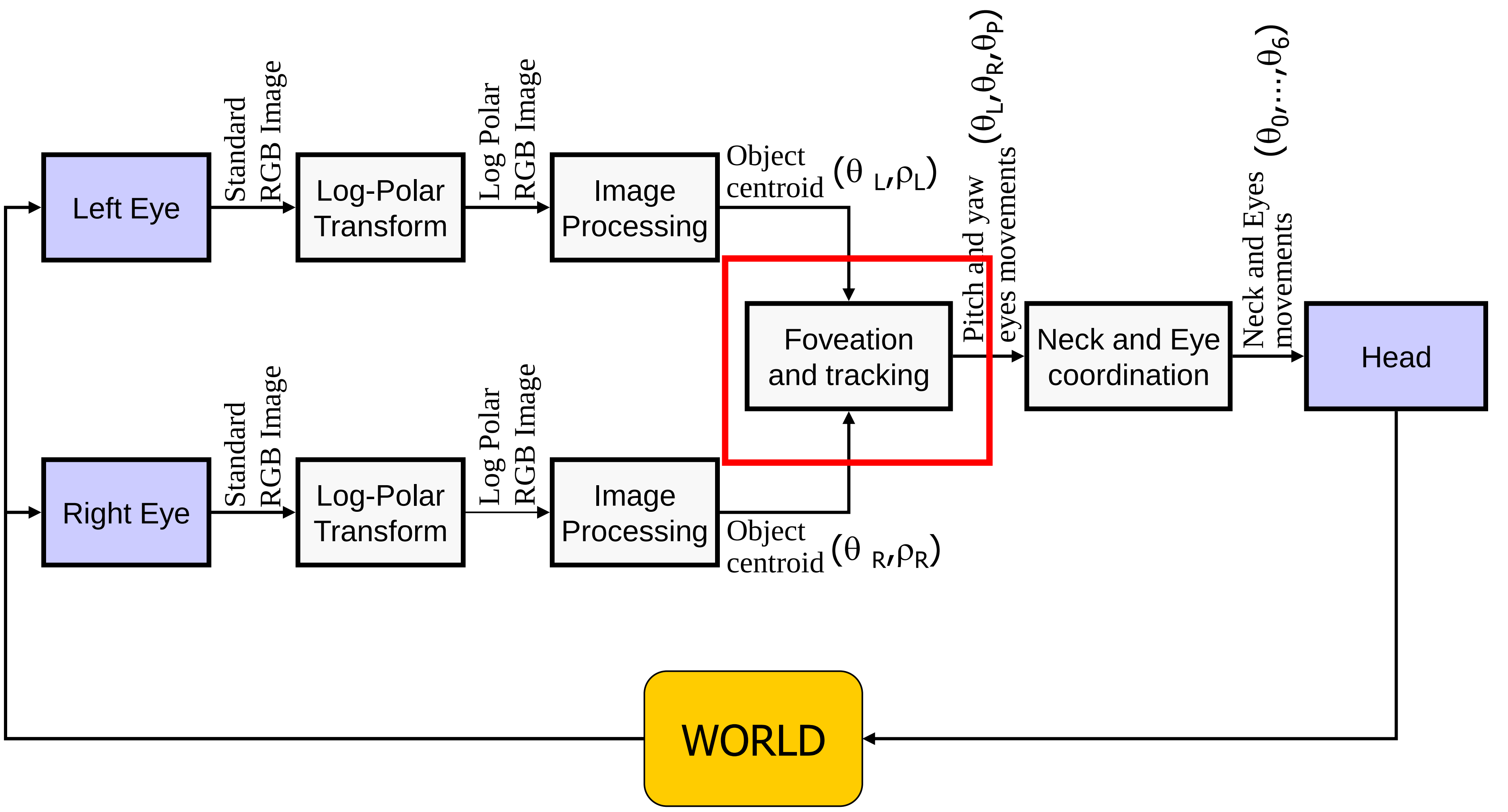


Detected lines (Boundaries)



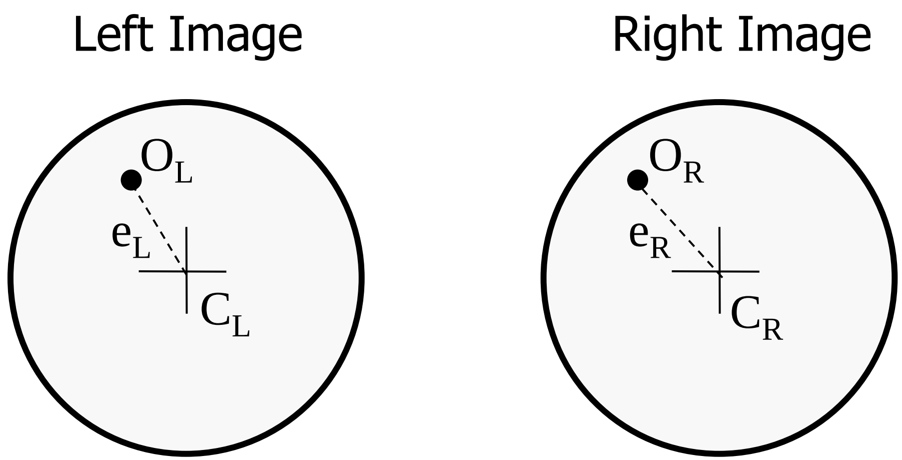
Boundary reconstruction based on eye positions

Overall sensory-motor scheme of the visual apparatus



Foveation of the object centroid

Proportional control based on the visual error

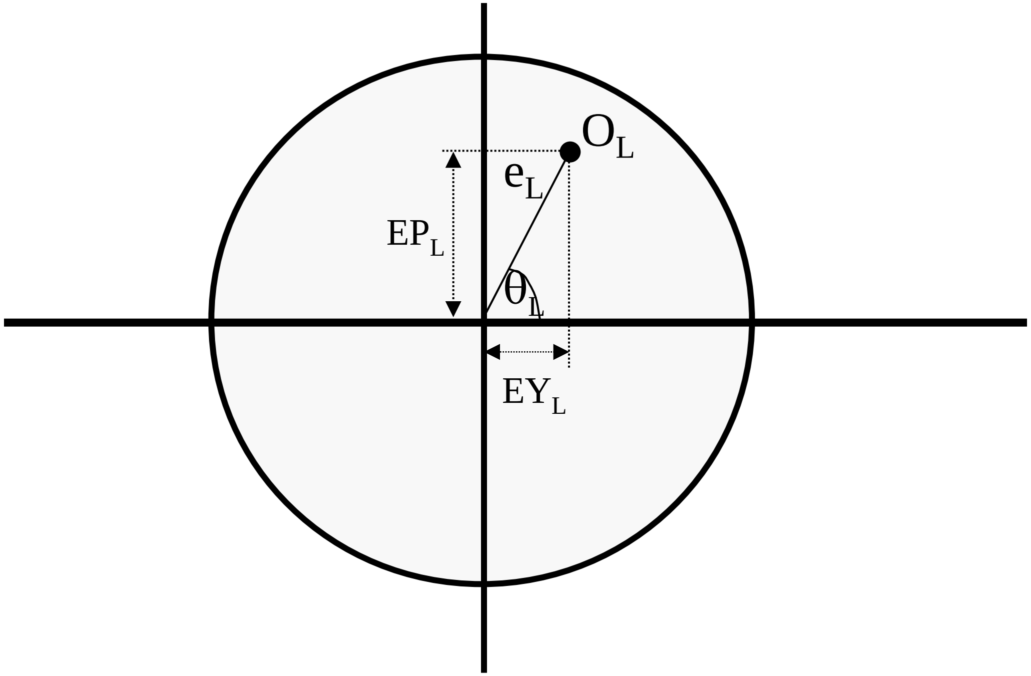


$O_L = (\rho_L, \theta_L)$ $O_R = (\rho_R, \theta_R)$

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

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

Computation of yaw and pitch eye movements



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

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

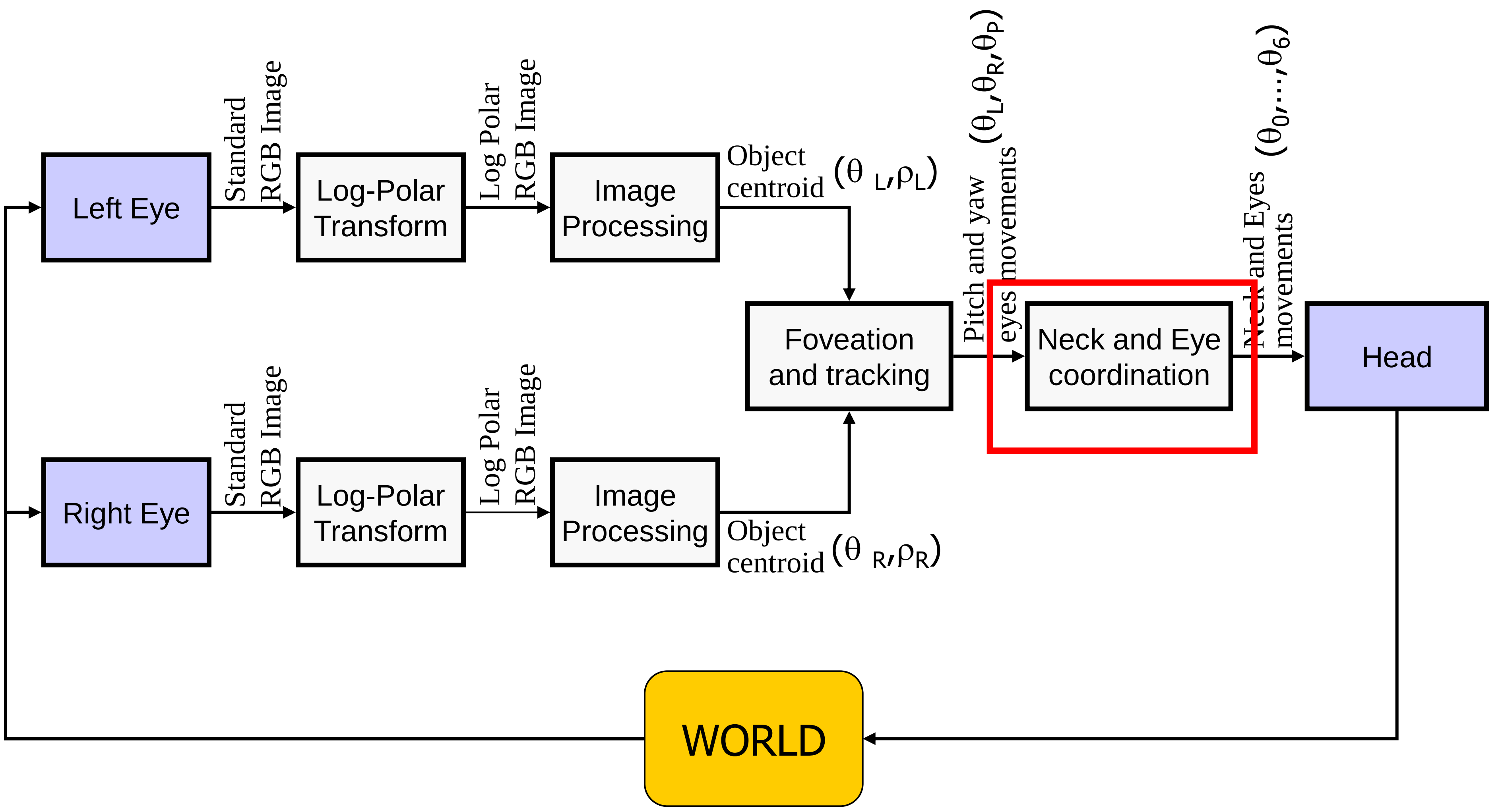
$EP_L = e_L * \sin(\theta_L) * P_L$

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

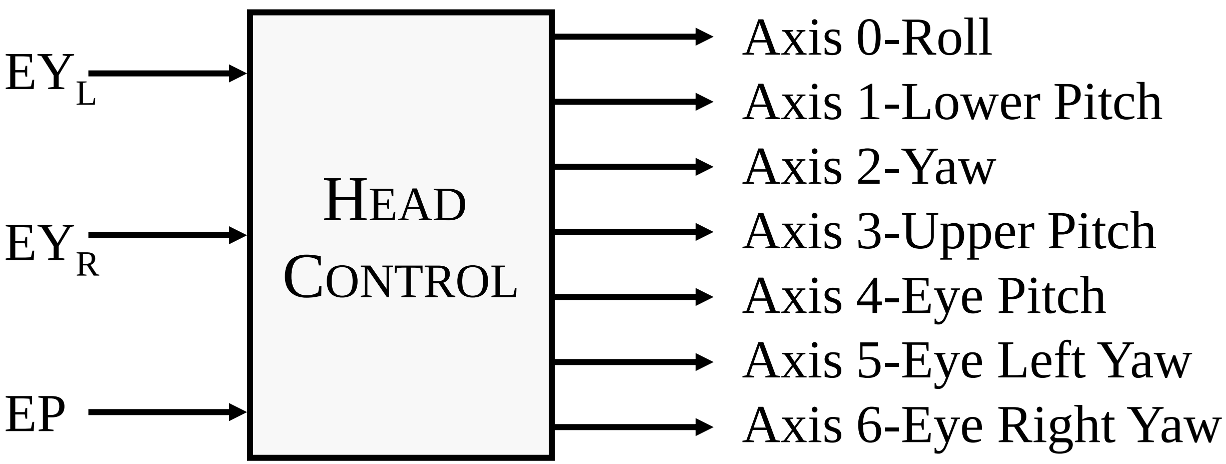
$EP=(EP_L+EP_R)/2$

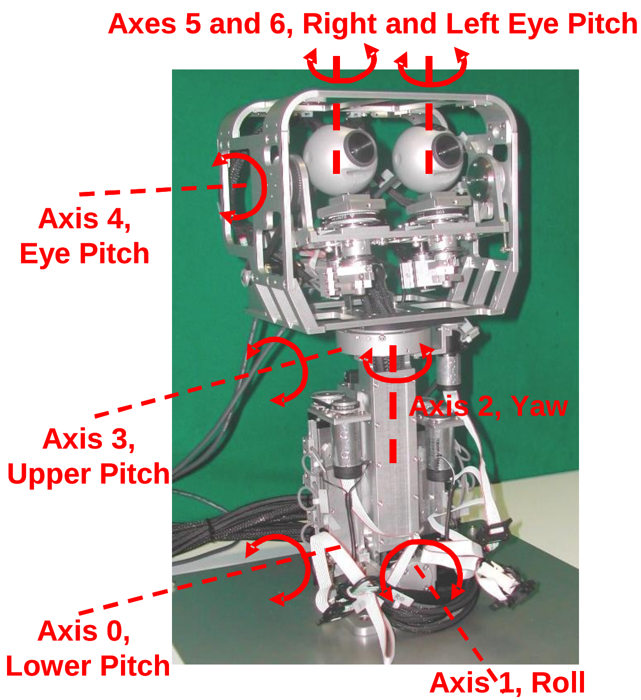
P_L and P_R are the proportional parameters for left and right eye, respectively.

Overall sensory-motor scheme of the visual apparatus



Eye-neck coordination



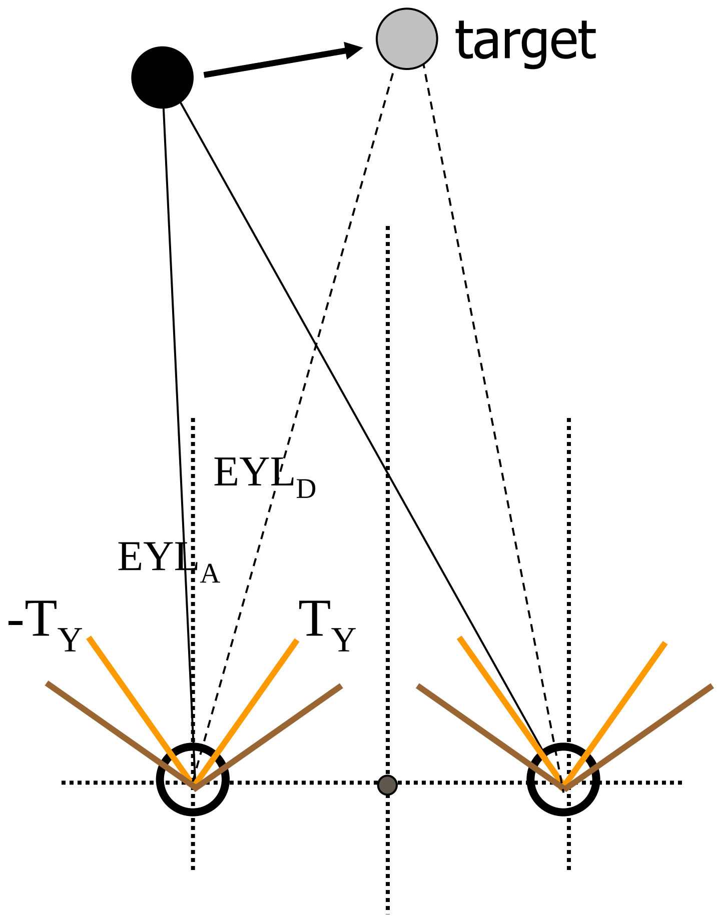


Solution 1

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

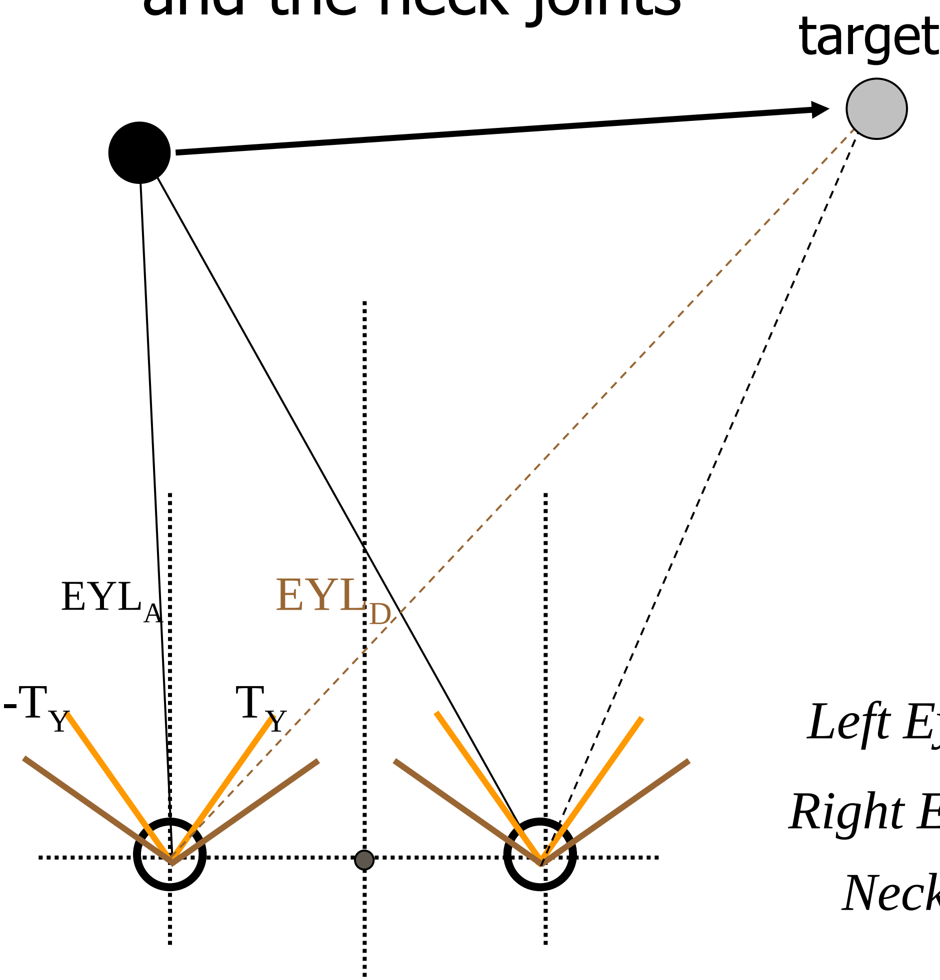
$$EYR_A + EY_R < T_Y$$

⇩

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

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

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.

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

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

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

$EYP_D = EYP_A + EP * K1$

$EUP_D = EUP_A + EP * K2$

$ELP_D = ELP_A + EP * K3$

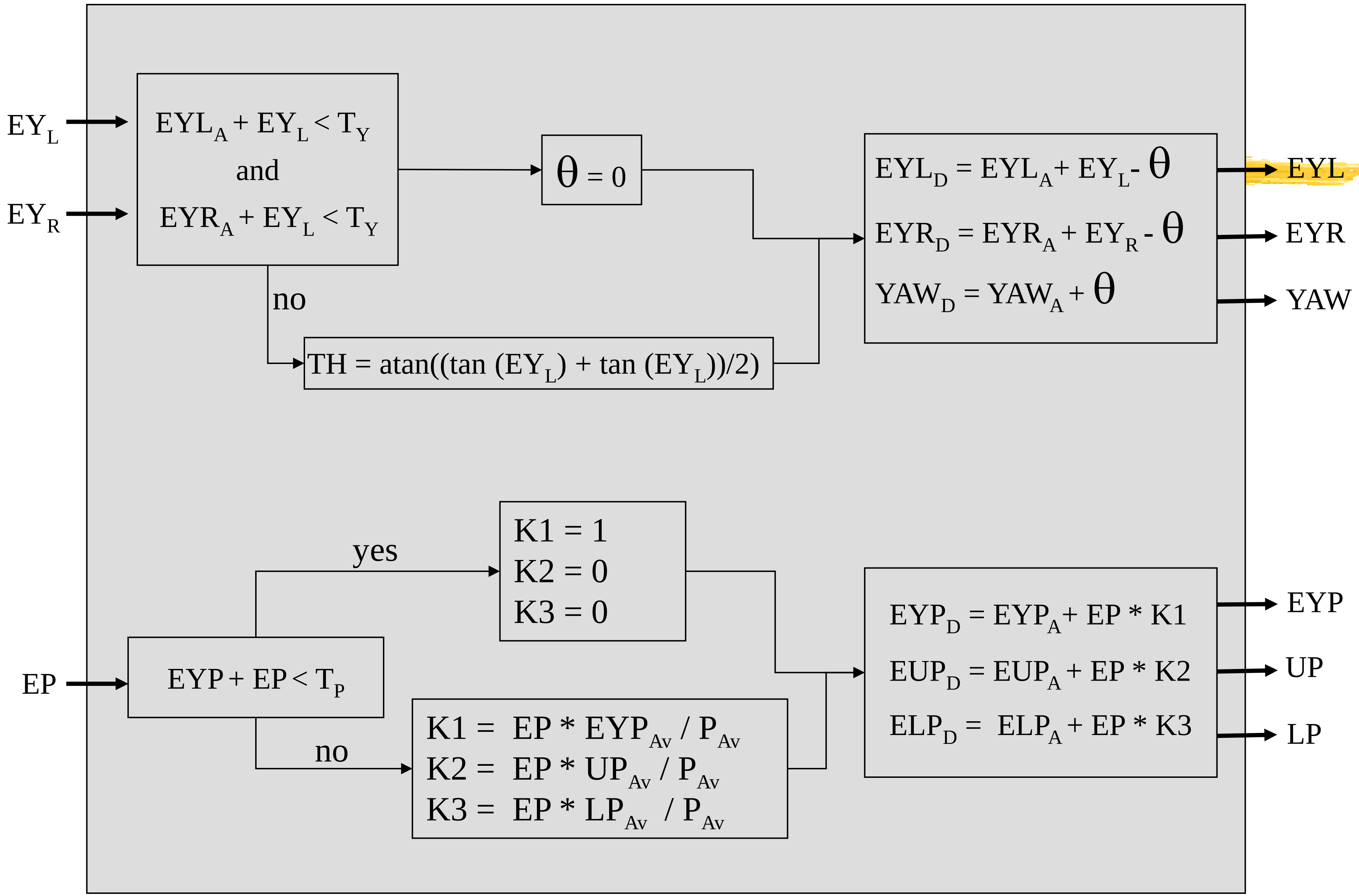
$EYP_{AV} = EYP_M - EYP_A$

$UP_{AV} = UP_M - UP_A$

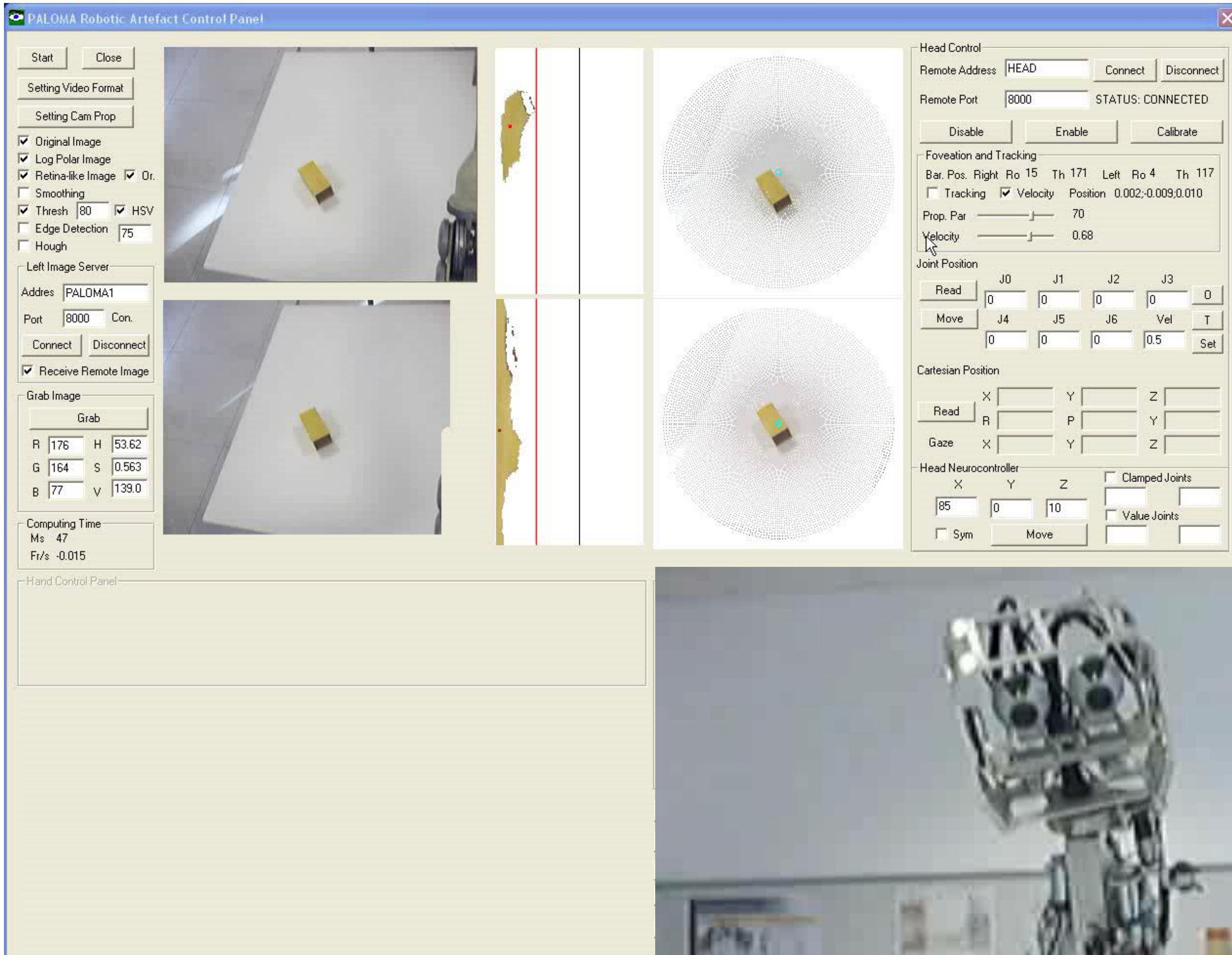
$LP_{AV} = LP_M - UP_A$

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

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



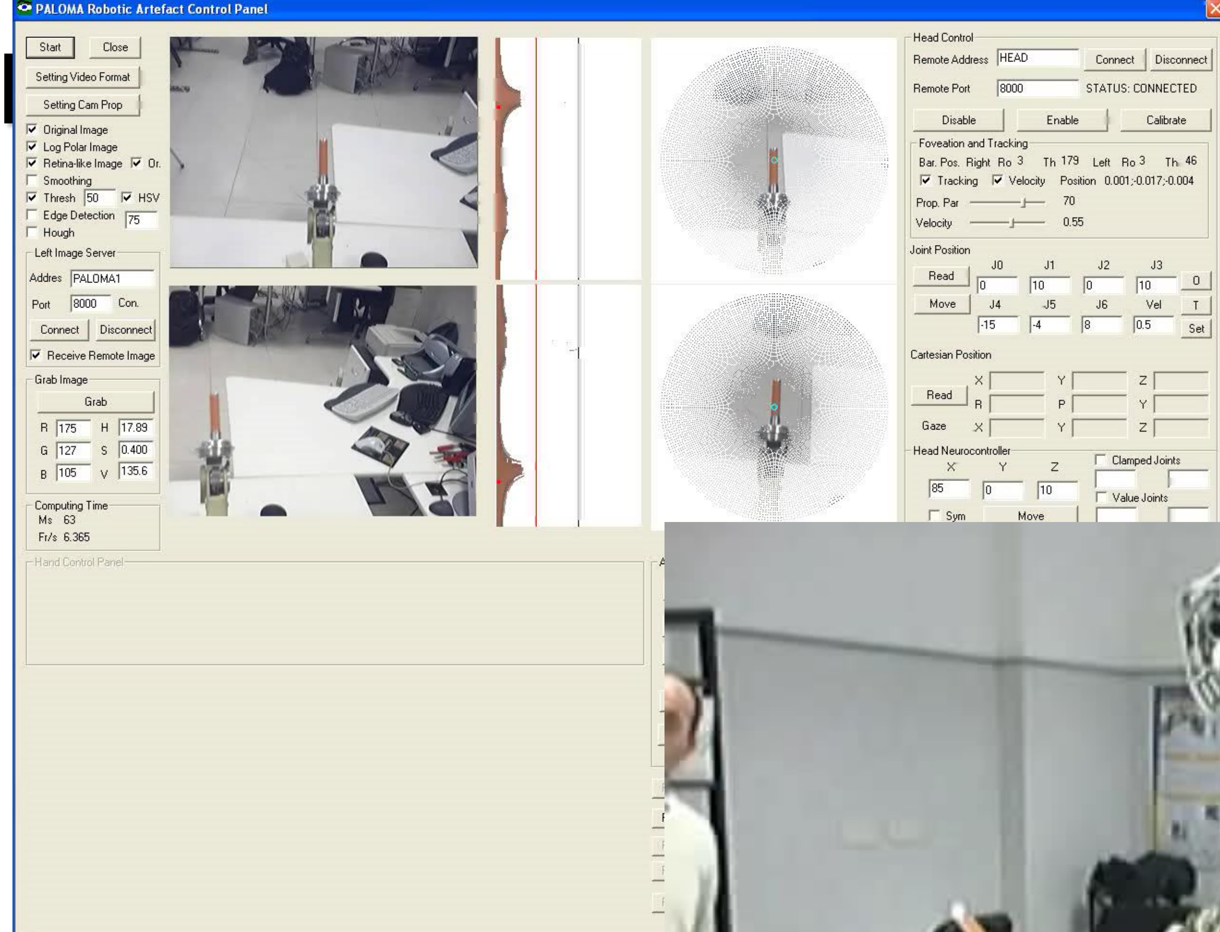
EY_L
EY_R
EYL_A + EY_L < T_Y
and
EYR_A + EY_L < T_Y
θ = 0
no
TH = atan((tan (EY_L) + tan (EY_L))/2)
EYL_D = EYL_A + EY_L - θ
EYR_D = EYR_A + EY_R - θ
YAW_D = YAW_A + θ
EYL
EYR
YAW
EP
EYP + EP < T_P
yes
K1 = 1
K2 = 0
K3 = 0
no
K1 = EP * EYP_Av / P_Av
K2 = EP * UP_Av / P_Av
K3 = EP * LP_Av / P_Av
EYP_D = EYP_A + EP * K1
EUP_D = EUP_A + EP * K2
ELP_D = ELP_A + EP * K3
EYP
UP
LP



Pursuit Movement

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

Hand Tracking



Frame rate: 10 fps for both images

Head Control loop: 100 ms

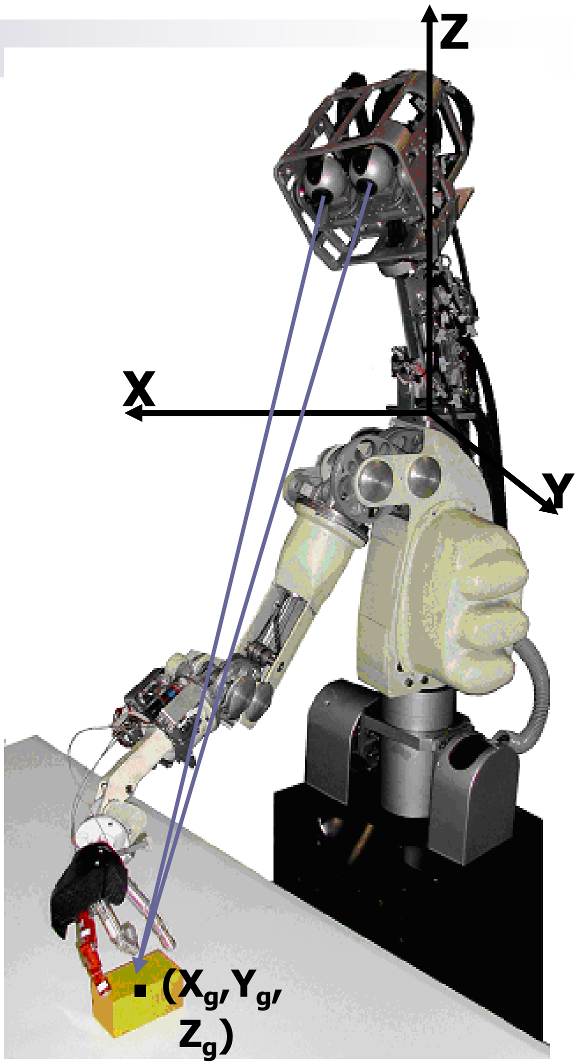
Arm movement 0.2 m/s

Solution 3

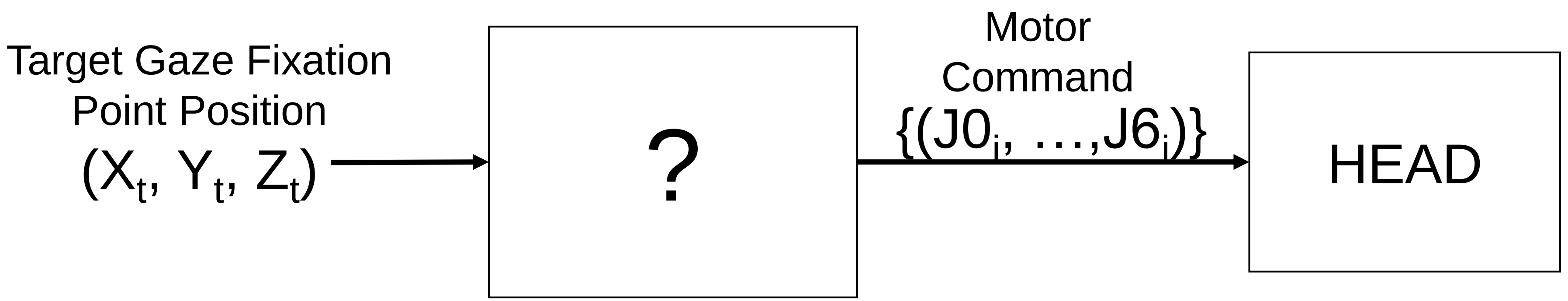
Implementation of a bioinspired model of head-eye coordination based on learning

G. Asuni, G. Teti, C. Laschi, E. Guglielmelli, P. Dario, “A Robotic Head Neuro-controller Based on Biologically-Inspired Neural Models”, *IEEE International Conference on Robotics and Automation – ICRA 2005*, Barcelona, Spain, April 18-22, 2005, pp.2373-2378.

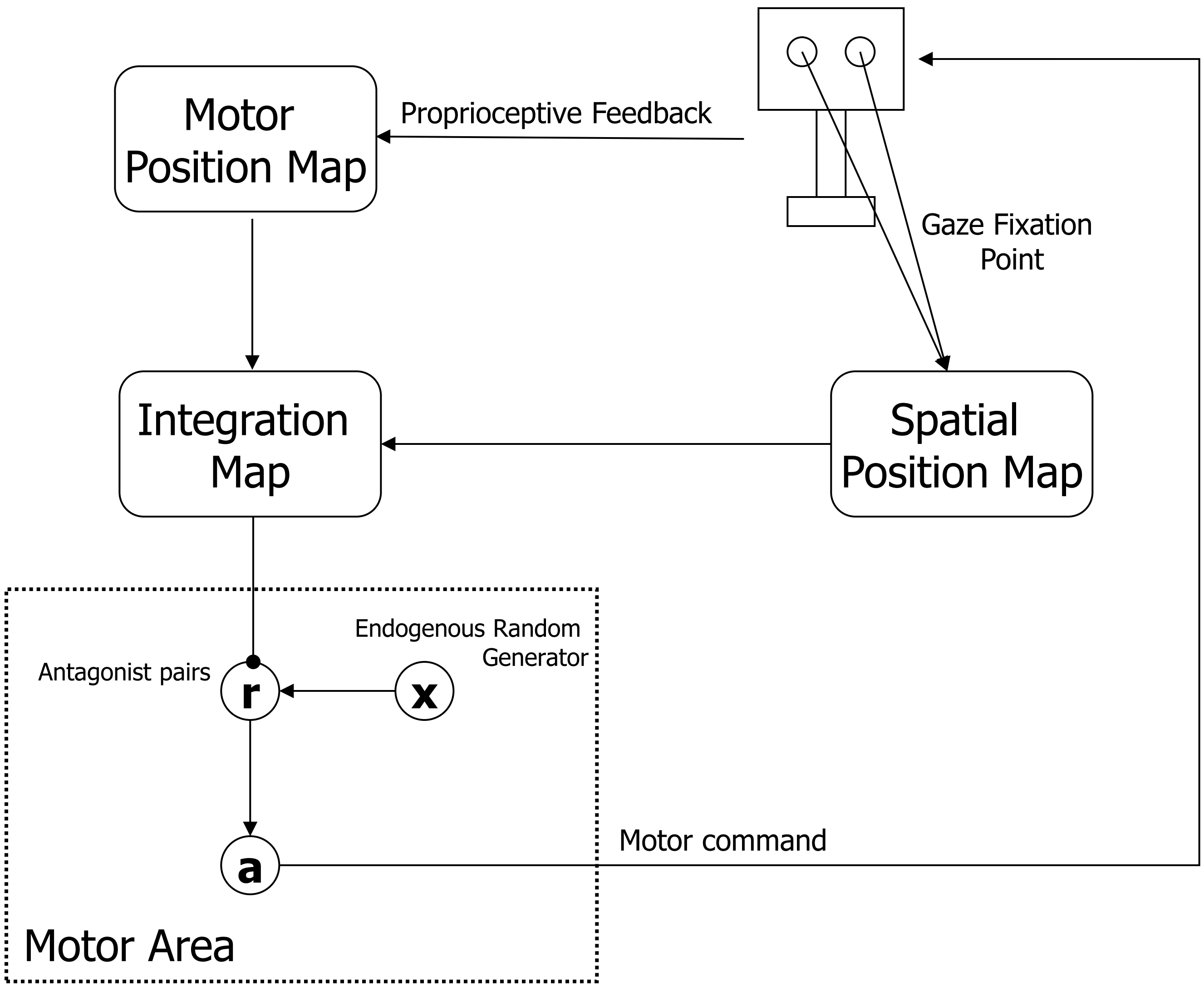
Addressed Problem



To develop a control module that receives in input a target gaze position and provides in output a command sequence able to reach it

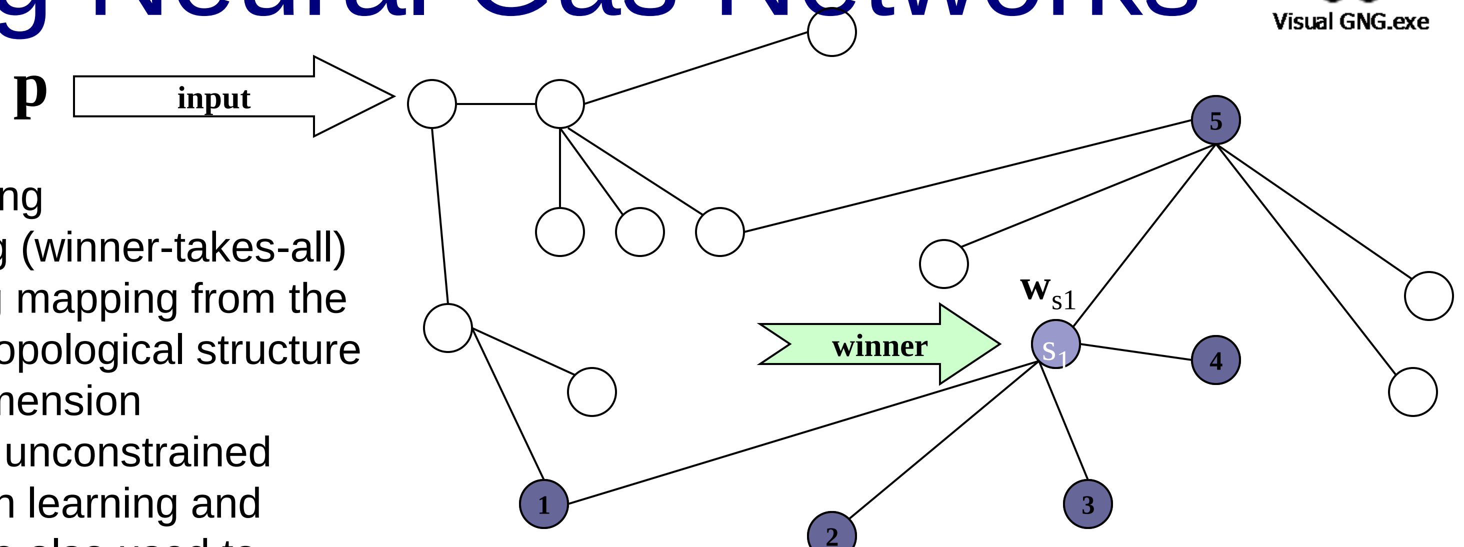


The proposed neural model



Implementation tools: Growing Neural Gas Networks

Visual GNG.exe



- Unsupervised learning
- Competitive learning (winner-takes-all)
- Topology-preserving mapping from the input space onto a topological structure of equal or lower dimension
- Network topology is unconstrained
- Competitive Hebbian learning and connection aging are also used to generate the topology
- Growth mechanism (the network size need not be predefined)
- The growth process can be interruped when a user defined performance criterion has been fulfilled

$\mathbf{w}_i$ is the weight vector associated to the unit i

Set of direct topological neighbors of the winner unit (S_1)

$$N_{s1} = \{ 1 \quad 2 \quad 3 \quad 4 \quad 5 \}$$

Updating rules:

$$\mathbf{w}_{s1} = \mathbf{w}_{s1} + \varepsilon_b (\mathbf{p} - \mathbf{w}_{s1})$$

$$\mathbf{w}_i = \mathbf{w}_i + \varepsilon_n (\mathbf{p} - \mathbf{w}_i) \; (\forall i \in N_{s_1})$$

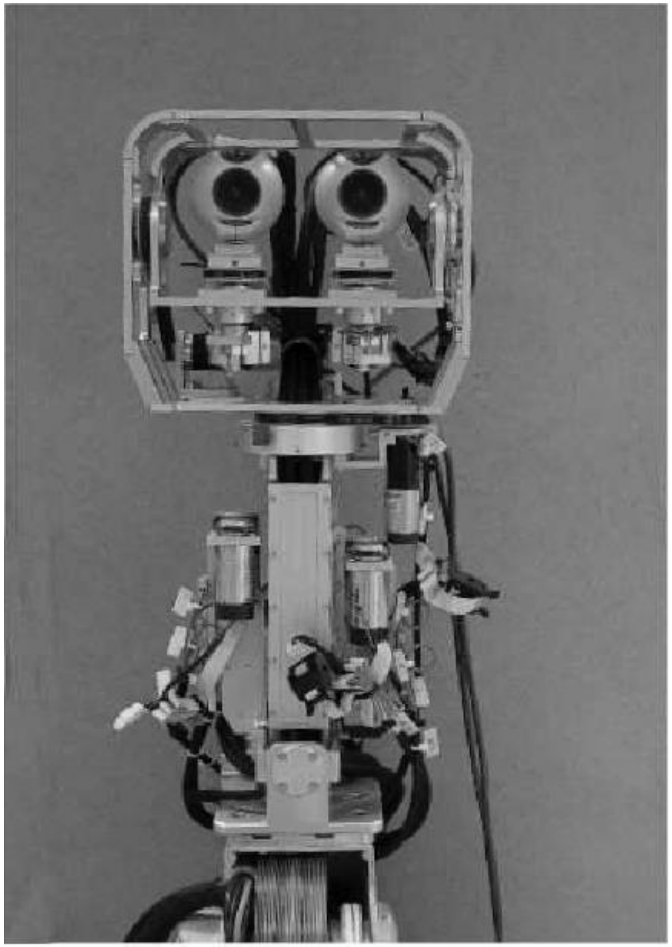
Bernd Fritzke, “Growing Cell Structures - A Self-organizing Network for Unsupervised and Supervised Learning”. ICSI TR-93-026, 1993. *Neural Networks* 7(9):1441-1460, 1994a

Testing phase

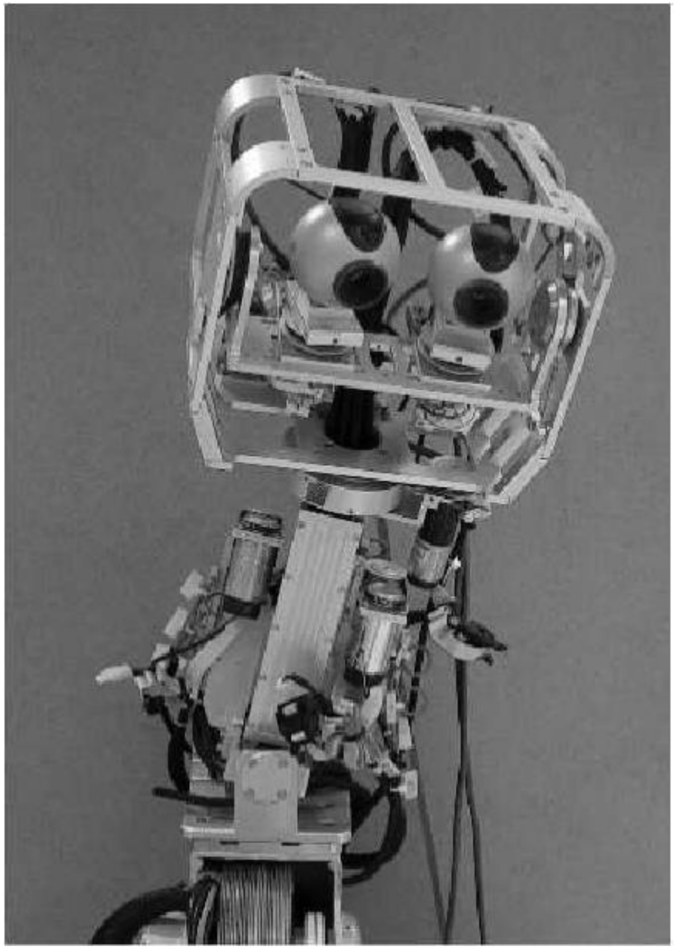
- After the training phase, given a target fixation point the system provides the joint rotations that drives the current gaze fixation point in the target point
- Three different modalities:
  1. Normal (without any constraint)
  2. With a clamped joint 0
  3. With symmetric angles for eye joints

All trials have been executed without additional learning

Experimental results: normal gazing



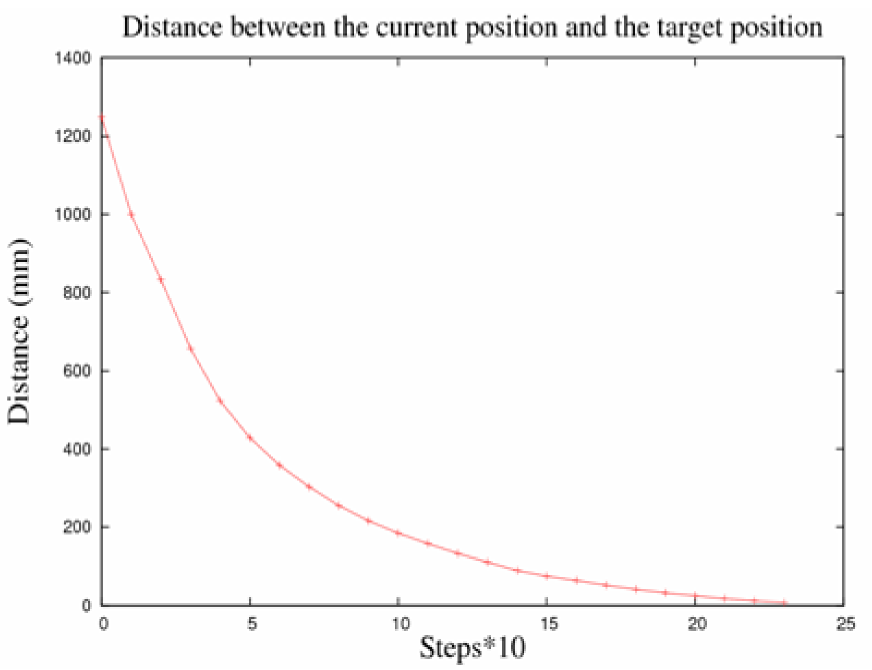
Initial posture



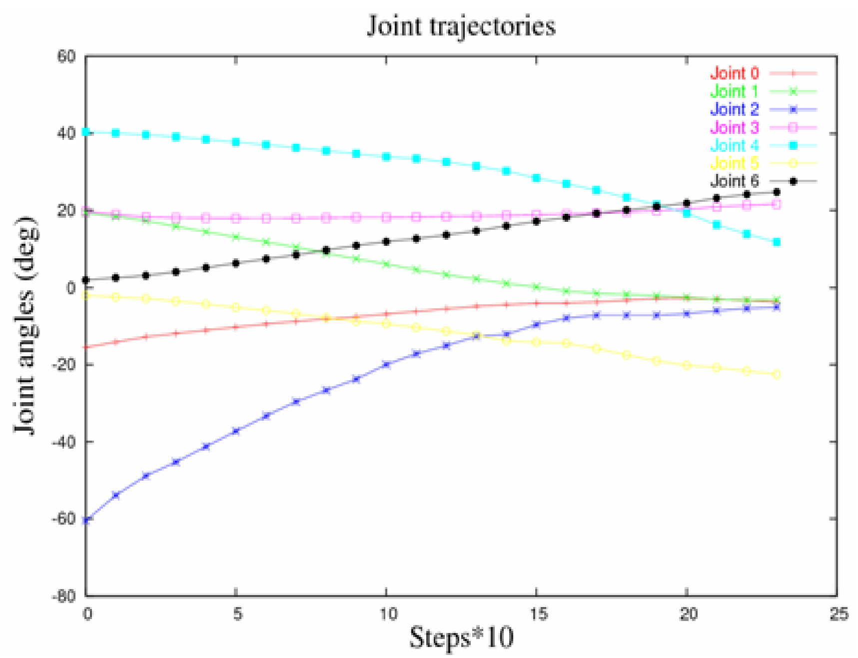
Final posture (normal)

Experimental results: robotic head (7 d.o.f)

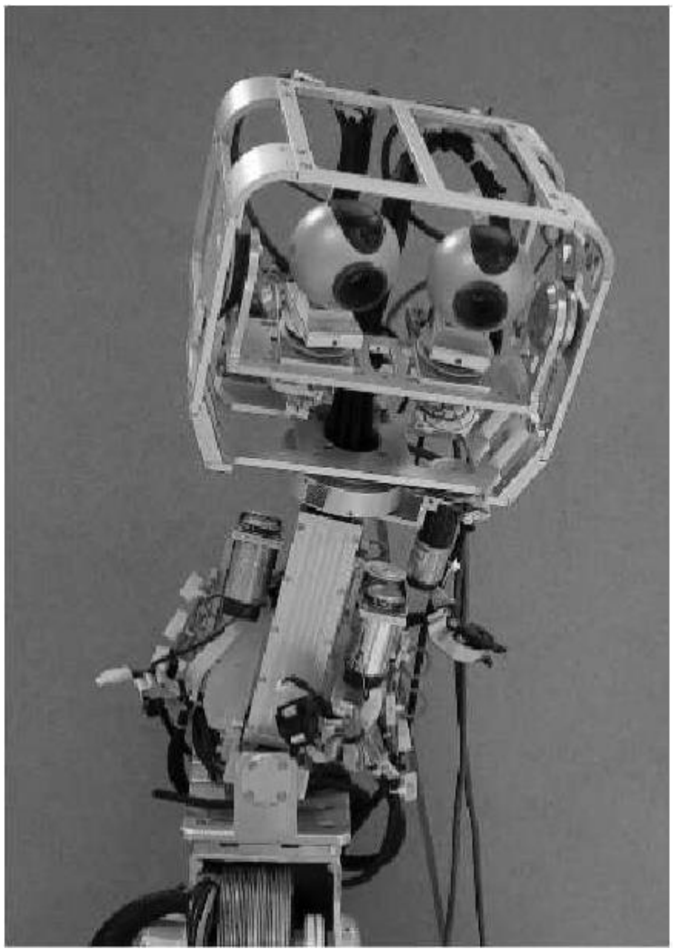
Distance between the current gaze fixation point and the target: monotonic trend



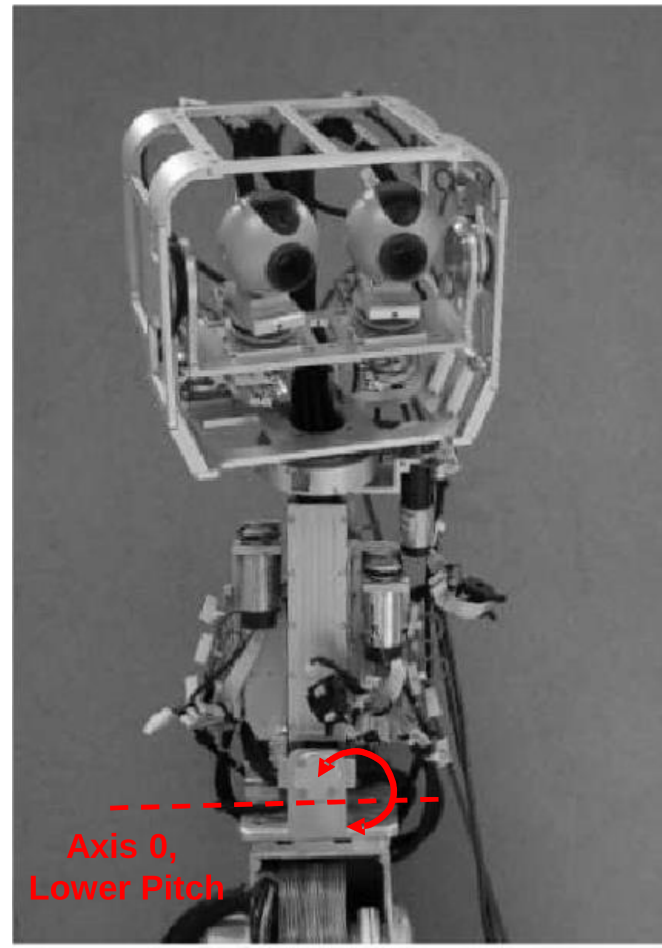
Joint trajectory



Experimental results: gazing with a clamped joint

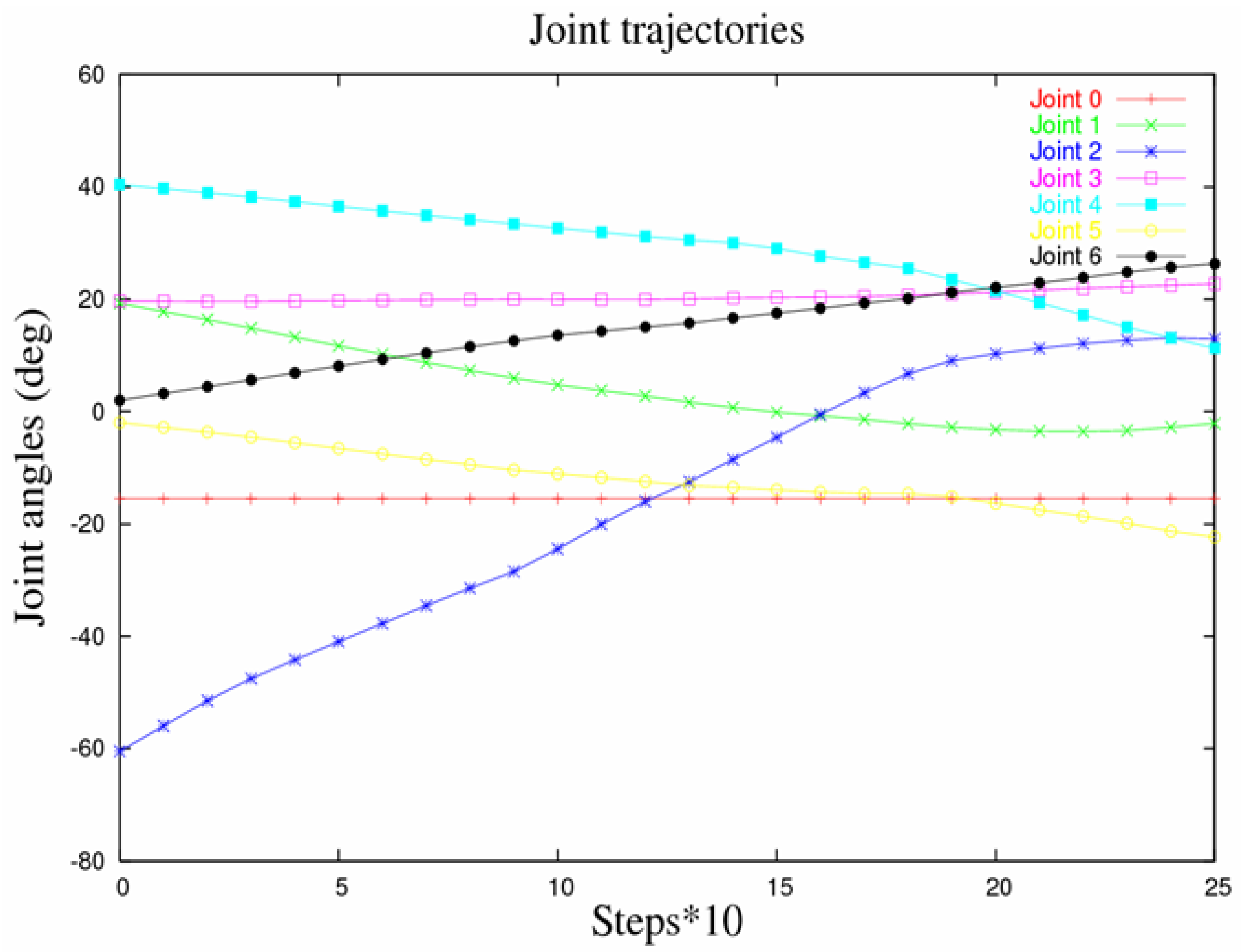


Final posture in normal mode



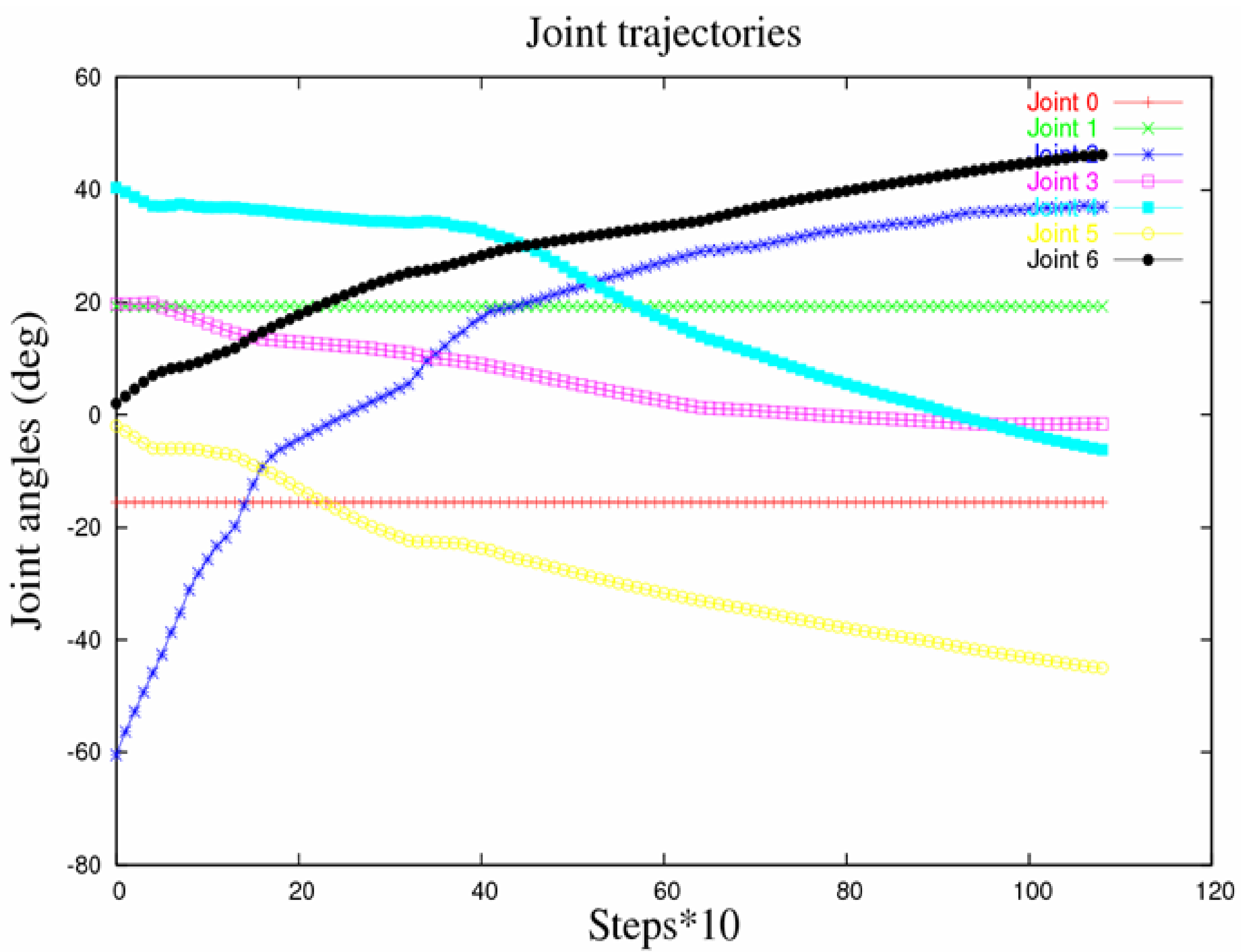
Final posture (clamped joint 0)

Experimental results: robotic head (7 d.o.f)



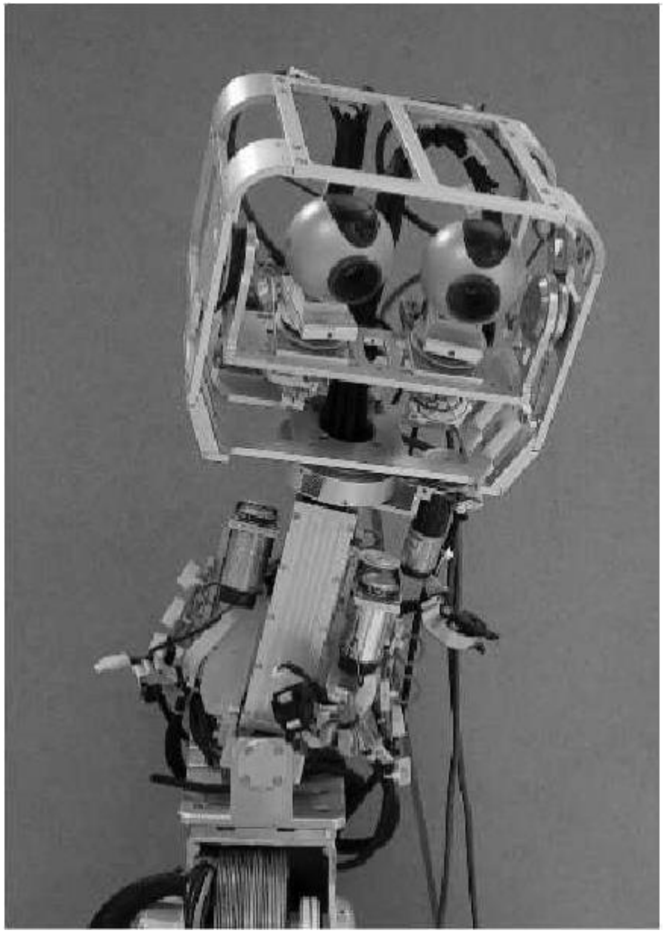
Joint trajectory: clamped joint 0

Experimental results: robotic head (7 d.o.f)

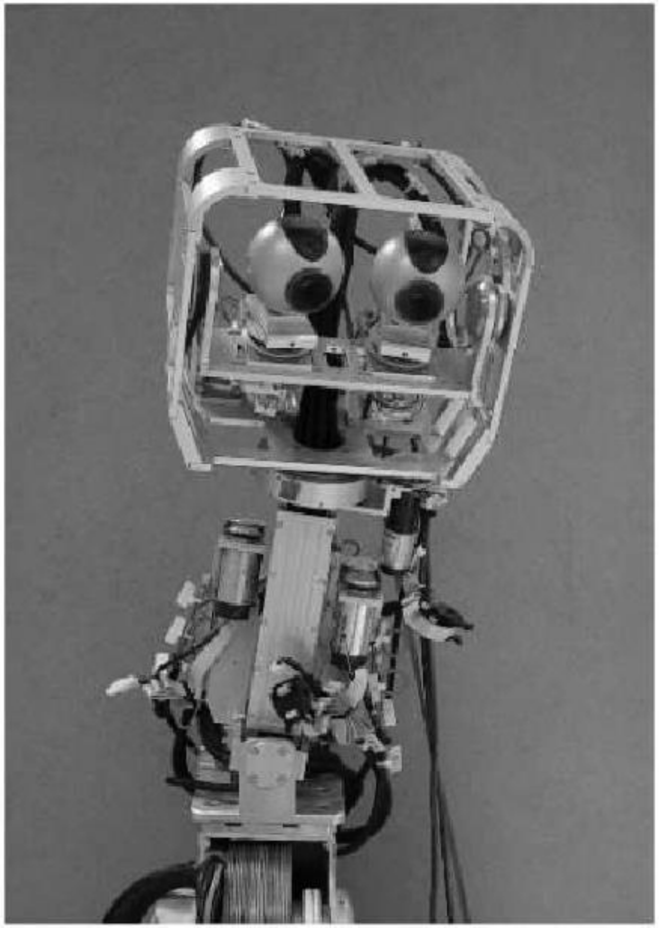


Joint trajectory: clamped joint 0 and joint 1

Experimental results: gazing with symmetric eye angles

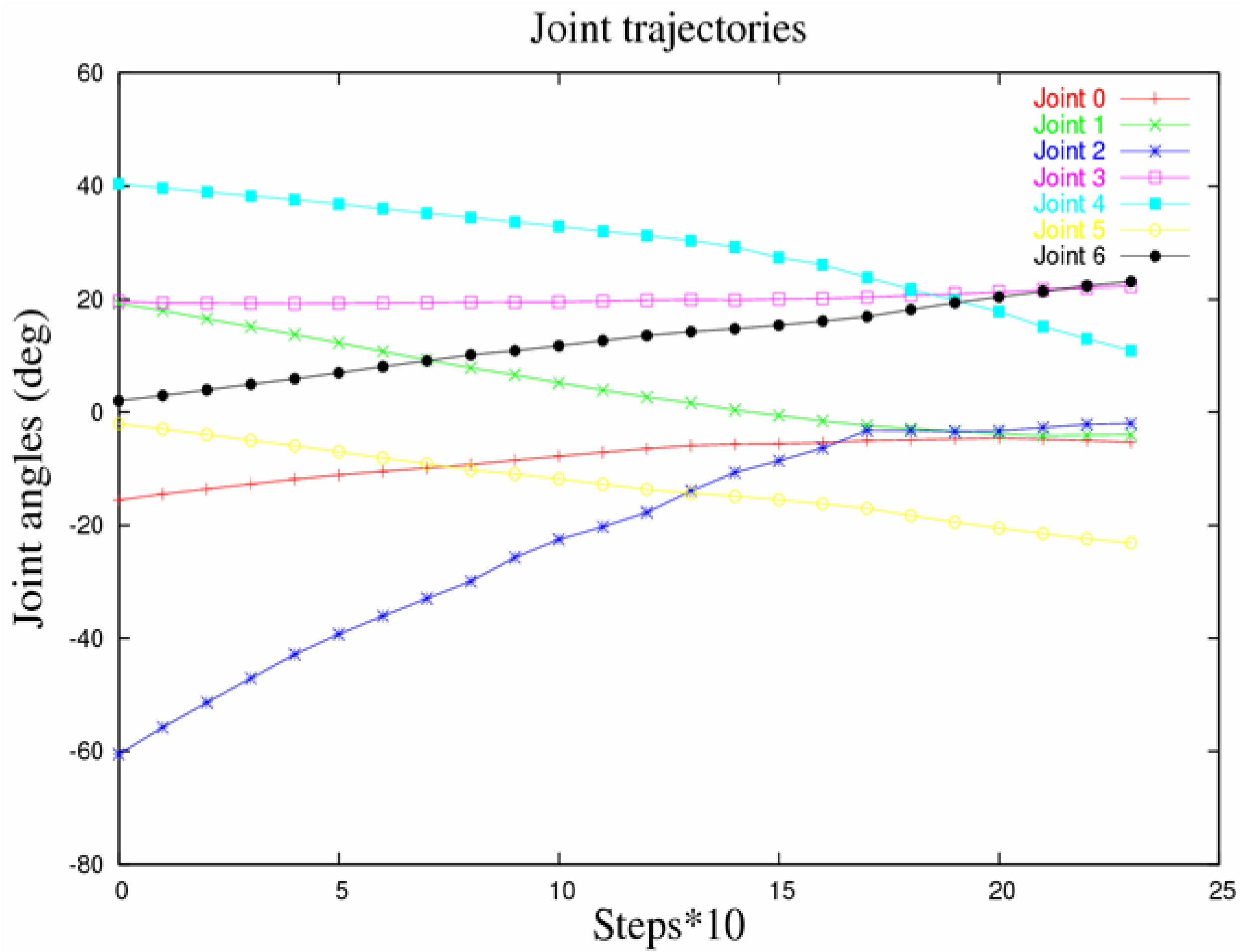


Final posture in normal mode



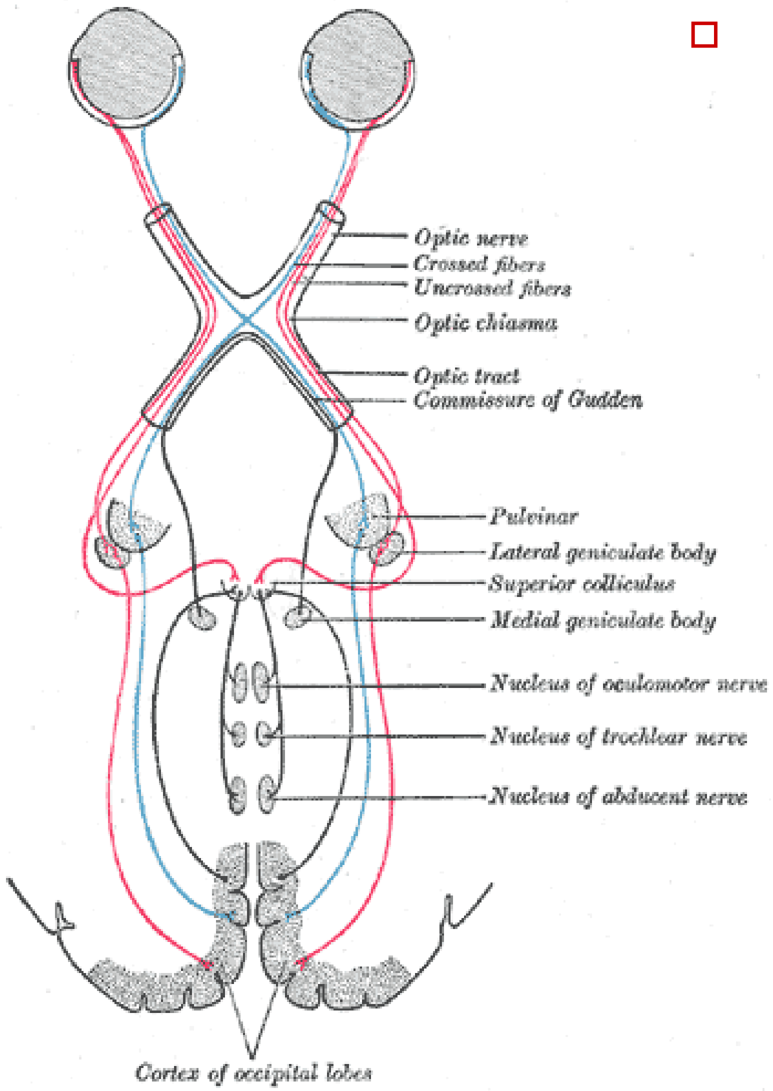
Final posture with symmetric angles for eye joints

Experimental results: robotic head (7 d.o.f)

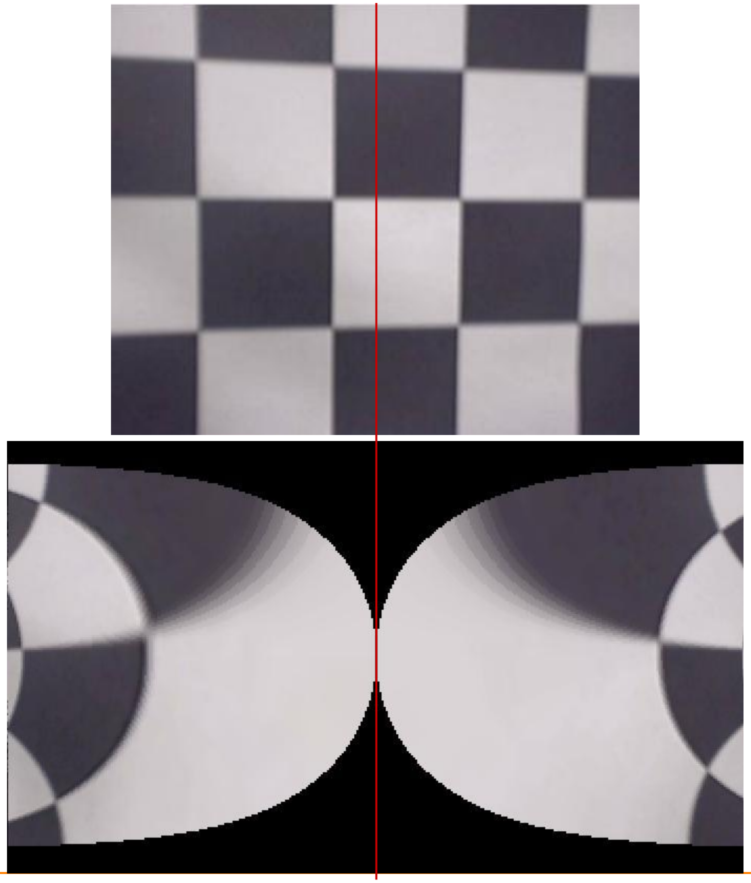


Joint trajectory: symmetric angles for eye joints (vergence)

Validation of a model of gaze control (by Prof. Alain Berthoz, College de France, Paris)



- Implementation of the mapping from the polar coordinates in visual space to the superior colliculus coordinate system, according to the model



L. Manfredi, C. Laschi, E.S. Maini, B. Girard, N. Tabereau, A. Berthoz, "Implementation of a neurophysiologic model of saccadic movements on an anthropomorphic robotic head", accepted for Humanoids 2006, Genova, Italy, Dec.4-6, 2006.

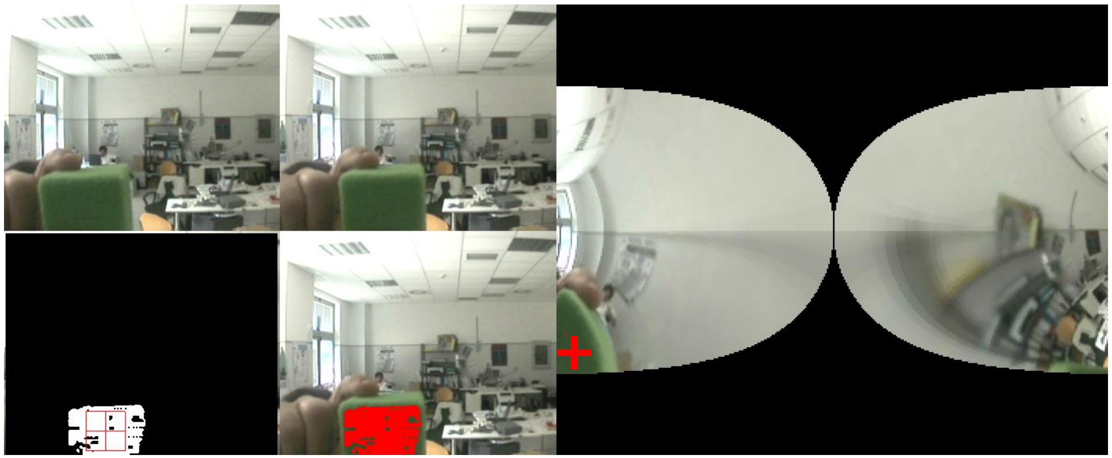
Validation of a model of gaze control (by Prof. Alain Berthoz, College de France)

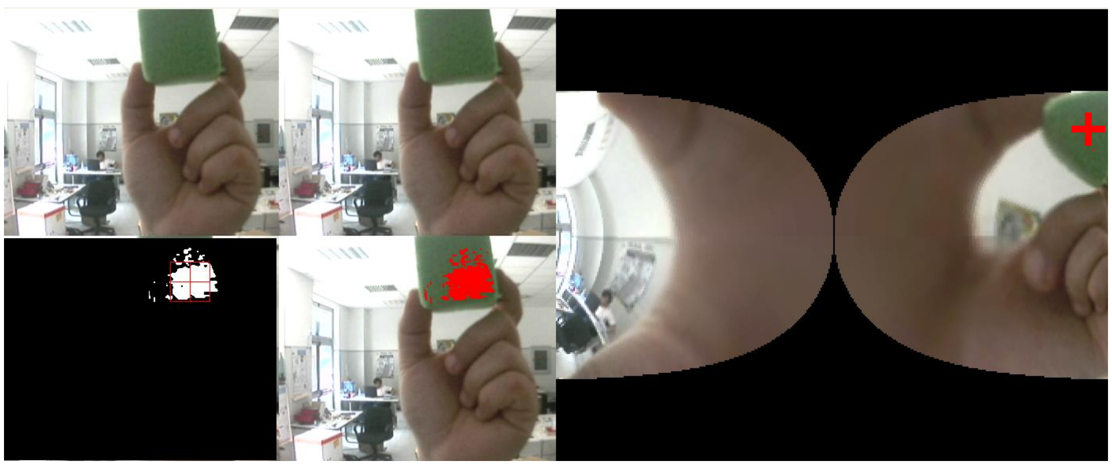
- Implementation of the mapping from the polar coordinates in visual space to the superior colliculus coordinates system, according to the model
- Generation of saccade movements:
  - A stimulus of a given colour can be detected in the map and the coordinates calculated in the superior colliculus, in real time
  - These coordinates are sent to the gaze control model to calculate the velocity profile for gaze control
  - The velocity profiles are used to control the robot head to generate the saccade movements of the eyes

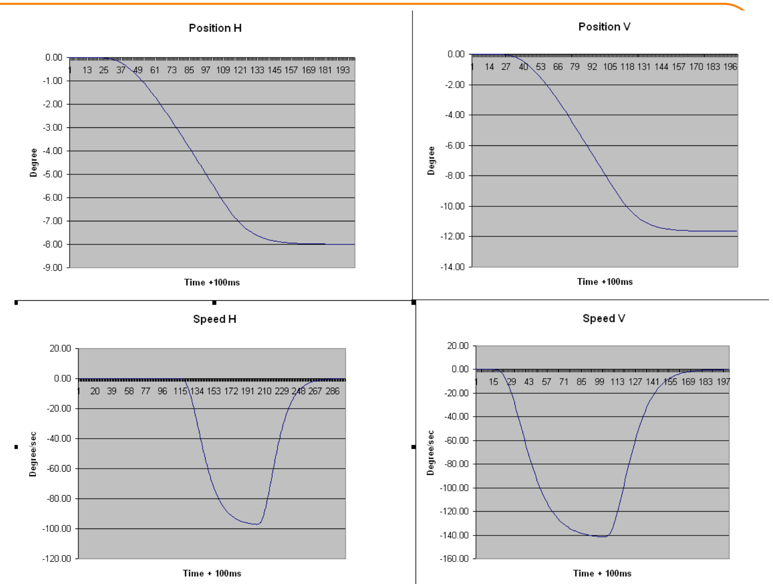
Original images

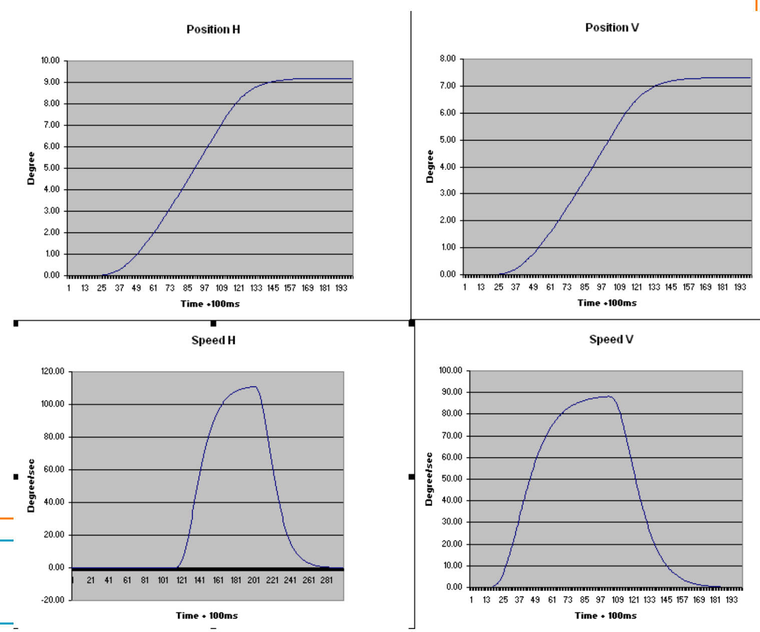
Collicular mapping (red point: stimulus coordinates)

Generation of saccade movements









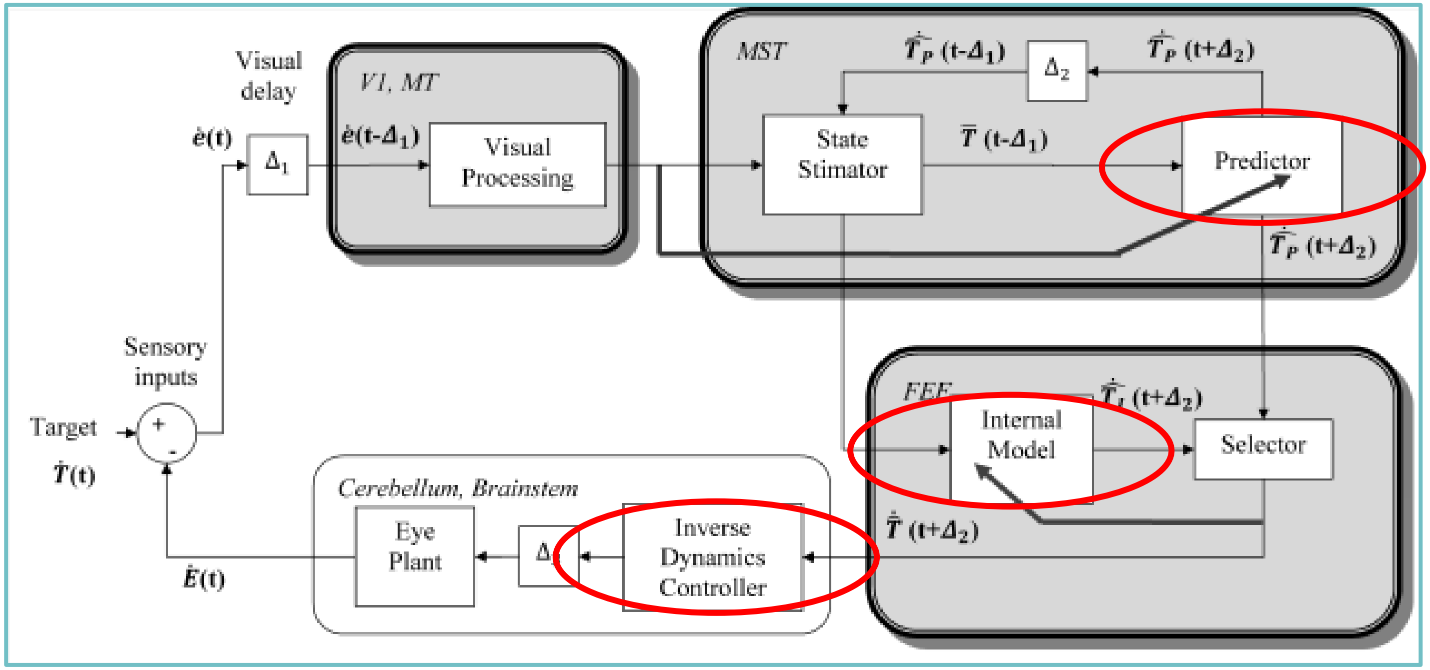
Generation of saccade movements

Stimulus #1

Stimulus #2

Saccades executed by the right eye

A predictive model for smooth pursuit

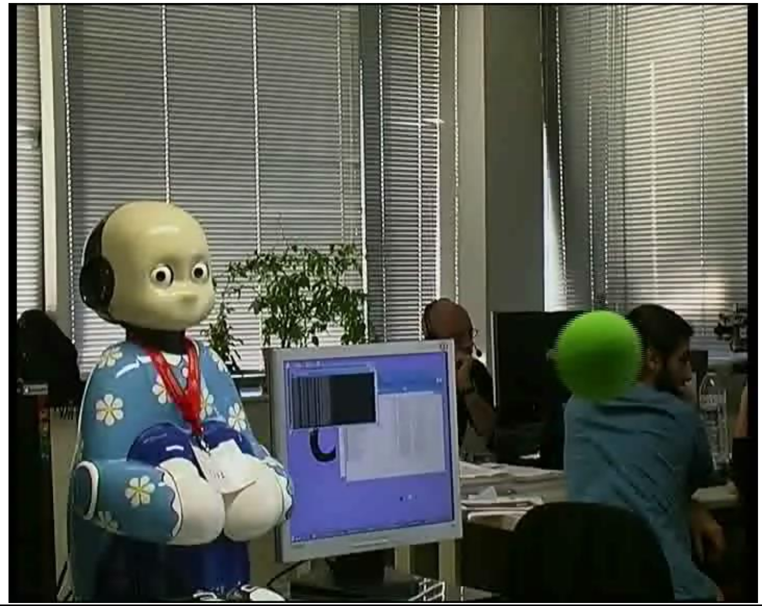


This circuit is based on Shibata and Schaal's model (*Shibata 2005*) of smooth pursuit and consists of **three subsystems**:

1. a **recurrent neural network** (RNN) mapped onto medial superior temporal area (MST), which receives the retinal slip with delays and ***predicts*** the current target motion,
2. an **inverse dynamics controller** (IDC) of the oculomotor system, mapped onto the cerebellum and the brainstem,
3. and **a memory block** that recognizes the target dynamics and provides the correct weights values before the RNN.

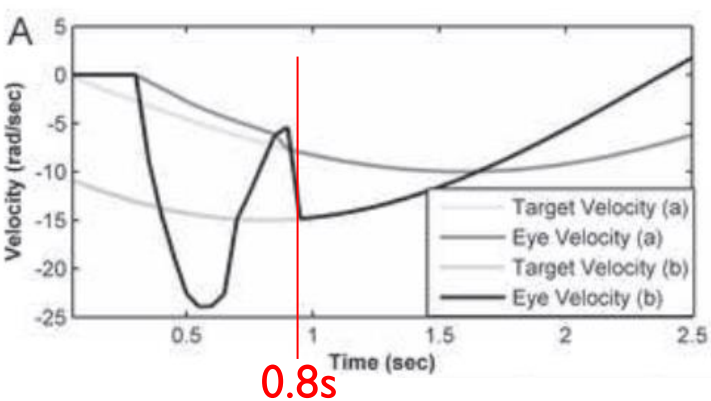
Zambrano D, Falotico E, Manfredi L, and Laschi C. (2010). "A model of the smooth pursuit eye movement with prediction and learning". *Applied Bionics and Biomechanics*

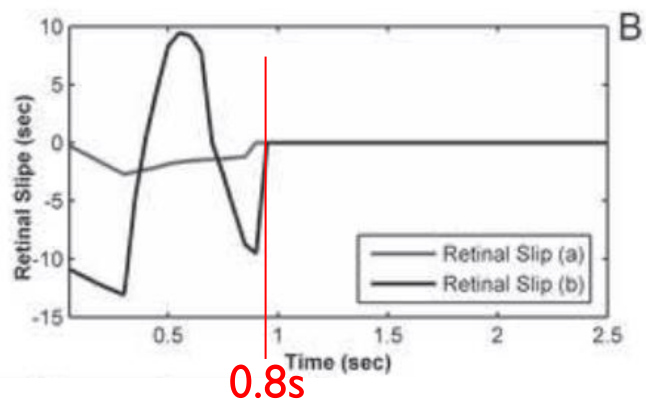
The predictive smooth pursuit on a robot head



iCub platform
head, 6 dof:
3 for the eyes
3 for the neck

The *retinal slip* (target velocity onto the retina) reaches zero after that the algorithm converges. When the target is unexpectedly stopped, the system goes on tracking the target for a short time.





Sinusoidal dynamics:

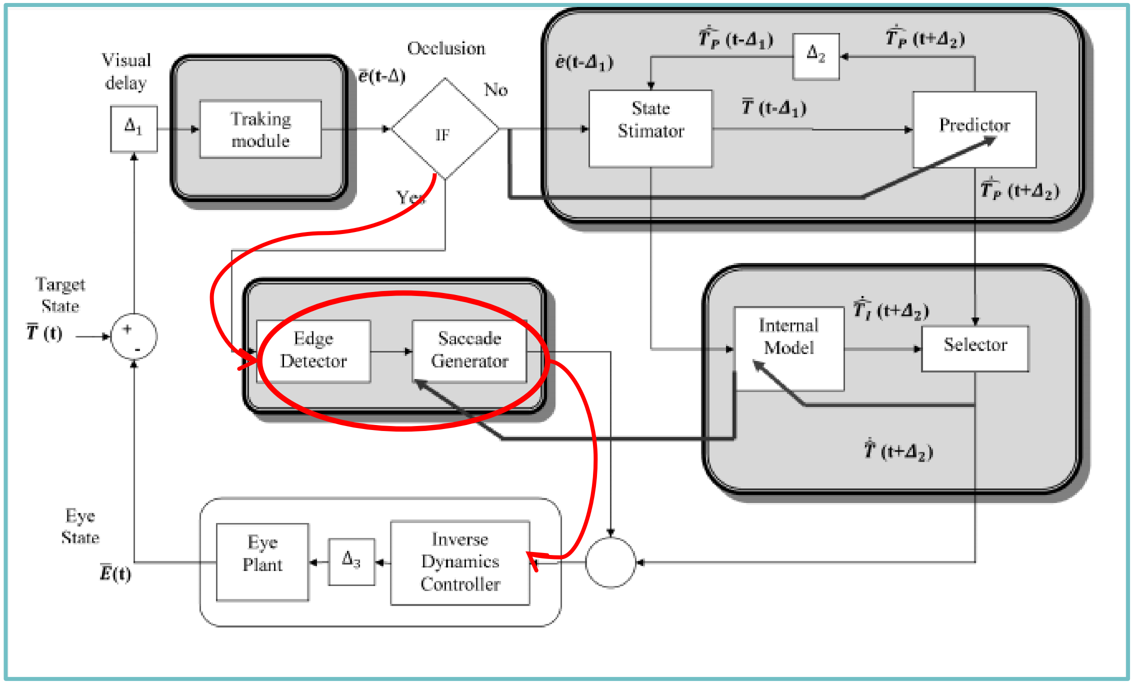
a) angular frequency: 1 rad/s, amplitude: 10 rad, phase: π/2

b) angular frequency: 1 rad/s, amplitude: 15 rad, phase of ¾ π

Smooth pursuit and occlusions

- Tracking across occlusions is not made with continuous smooth pursuit (von Hofsten, 2006)
- Humans are able to successfully track moving targets across occlusions by combining:
  - **Smooth pursuit** while the object remains visible;
  - One **saccade** to the predicted point where the object reappears;
- The saccade is elicited slightly before the target reappearance

A model of smooth pursuit and occlusions

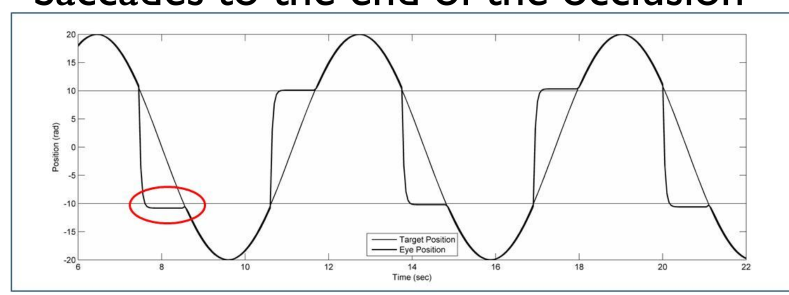


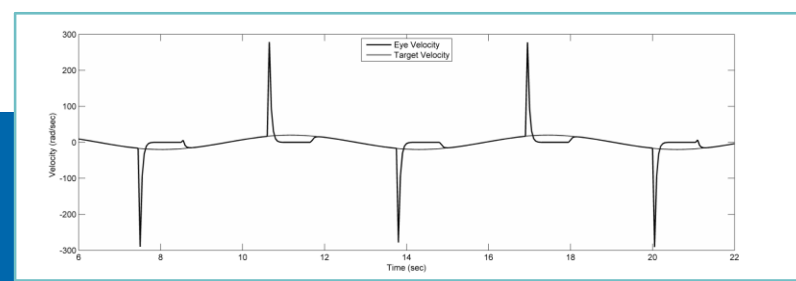
If the object disappears behind the occluder an event of occlusion is noticed and another module starts to detect the edges in the image to find where the object will reappear.
At this point the **saccade generator** module repeats the prediction of the target dynamics until the predicted position is equal to the edge detected from the previous module.

Implementation of smooth pursuit with occlusions

The tracking algorithm based on particle filtering detects the position of the target on the image and sends the results directly to the smooth pursuit system. When the target reappears, **the gaze points to the position of the target reappereance**, so the tracking algorithm is able to find the ball at the center of the image.

Saccades to the end of the occlusion



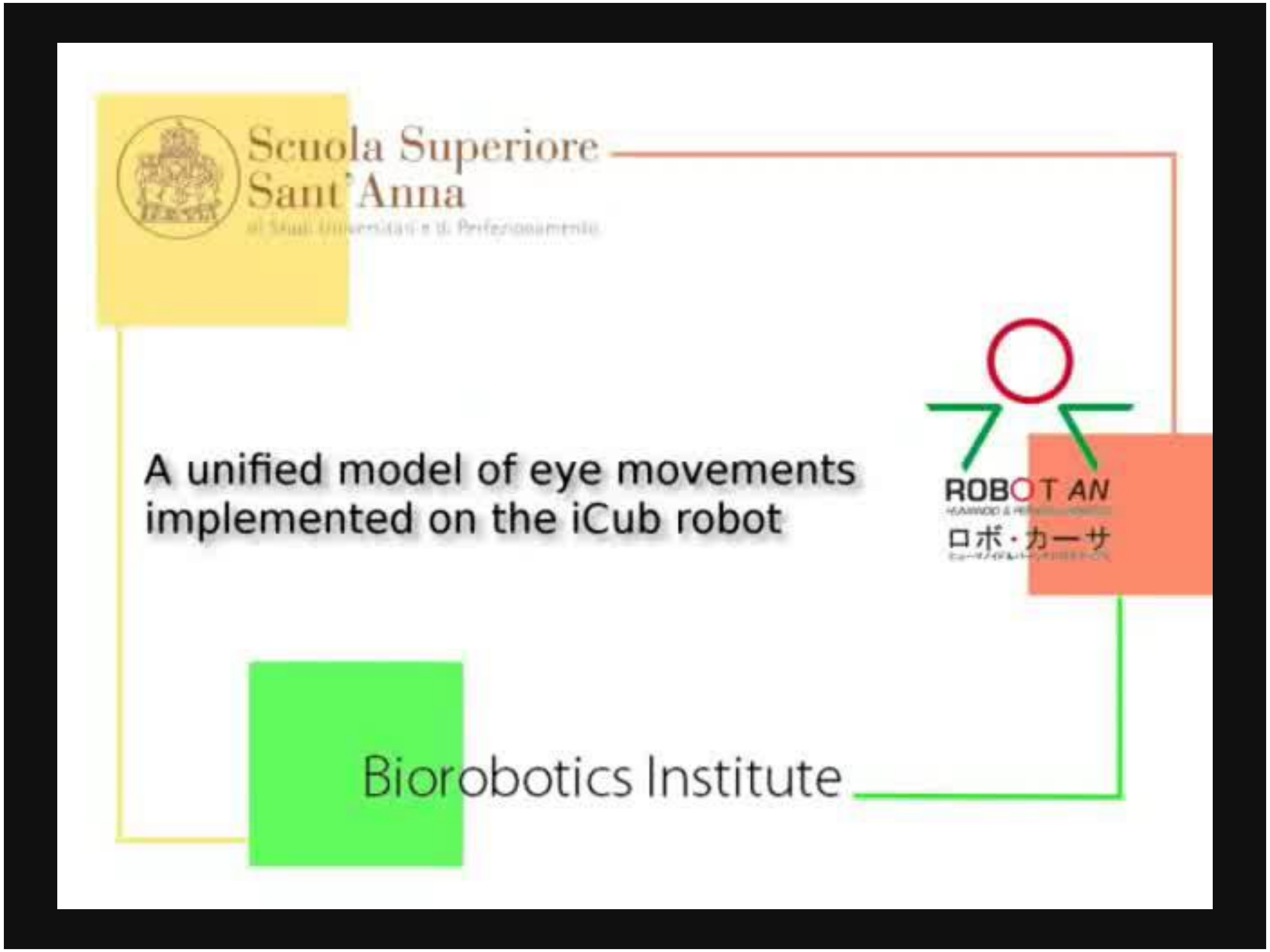


E. Falotico, M. Taiana, D. Zambrano, A. Bernardino, J. Santos Victor, P. Dario, and C. Laschi. “Predictive Tracking Across Occlusions in the iCub Robot.” *9th IEEE-RAS International Conference on Humanoid Robots*, December 7th-10th, 2009,

Punching a moving target - robot experiments

The prediction is iterated ahead 0.5 seconds
As the predicted target is inside the arm workspace, the robot executes a movement to punch the ball in the ***predicted position***

Robotic implementation of gaze control, integrating different eye movements



E. Falotico. D. Zambrano, C. Laschi, P. Dario, "Bioinspired integrated eye movements in a humanoid robot", (in preparation) Autonomous Robots

D. Zambrano, E. Falotico. C. Laschi, P. Dario, "A model of basal ganglia for robotic eye movement control", (in preparation) Autonomous Robots