

Corso di Percezione Robotica (PRo)



C. Modulo di Percezione

Fondamenti di visione artificiale retinica

Cecilia Laschi

ARTS Lab, Scuola Superiore Sant'Anna

cecilia@arts.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

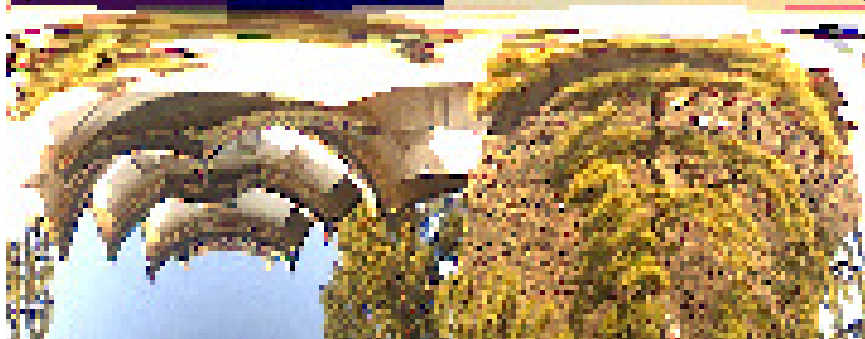
Standard image



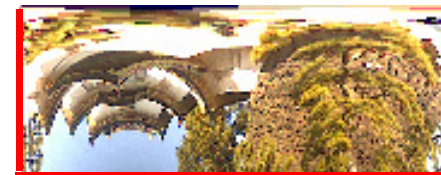
Retina-like image



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



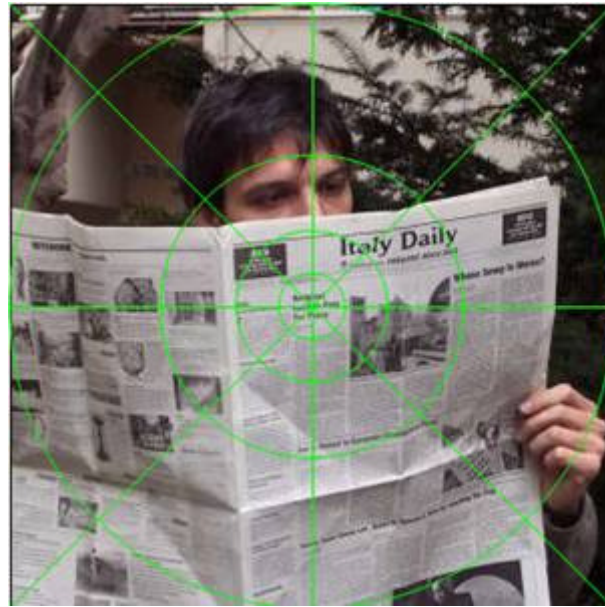
Log-polar projection



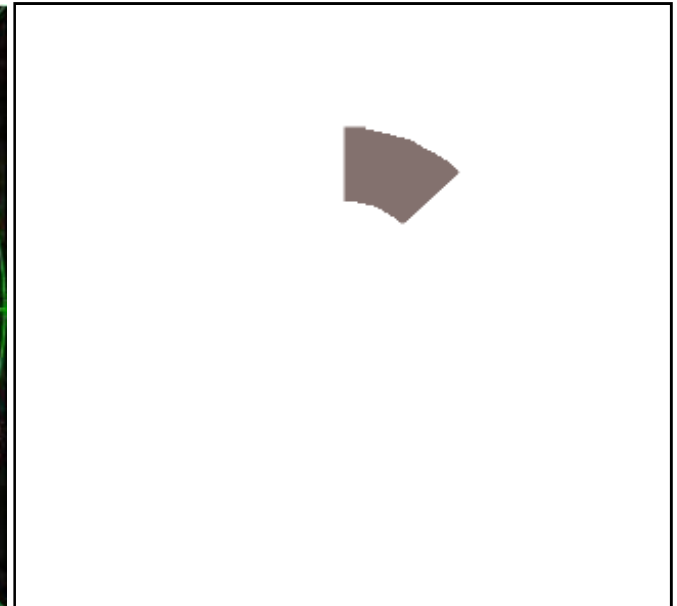
Costruzione di un'immagine retinica



Immagine cartesiana tradizionale



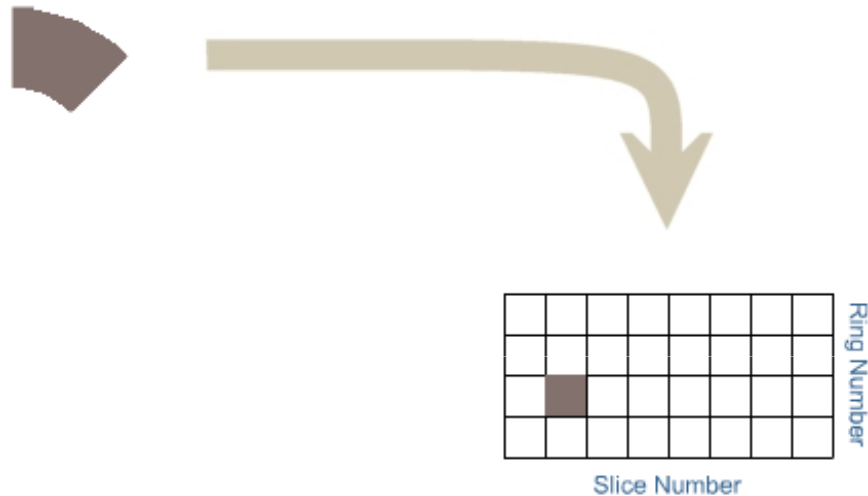
Divisione in circonferenze e spicchi



Calcolo del valore medio di un settore

Costruzione di un'immagine retinica

<http://www.retinica.com/>



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

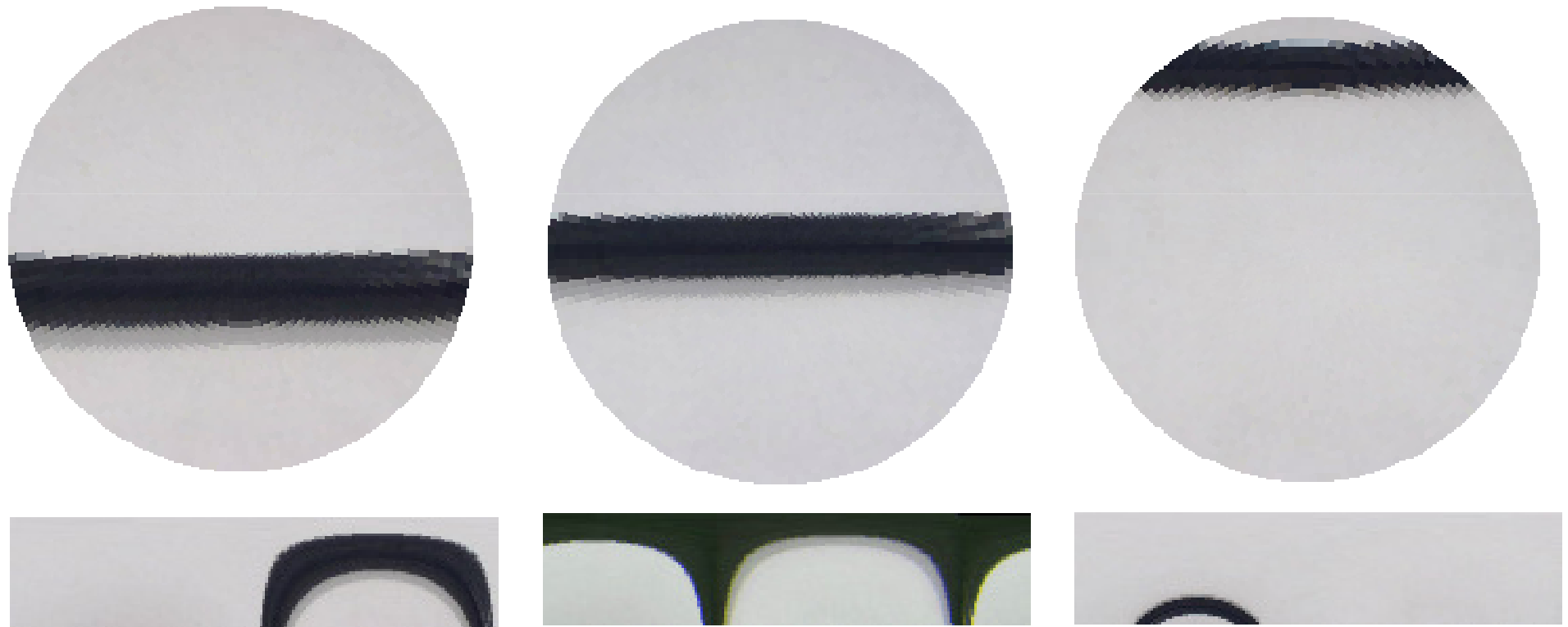


Immagine polare risultante

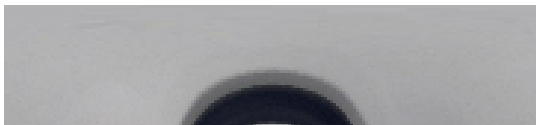
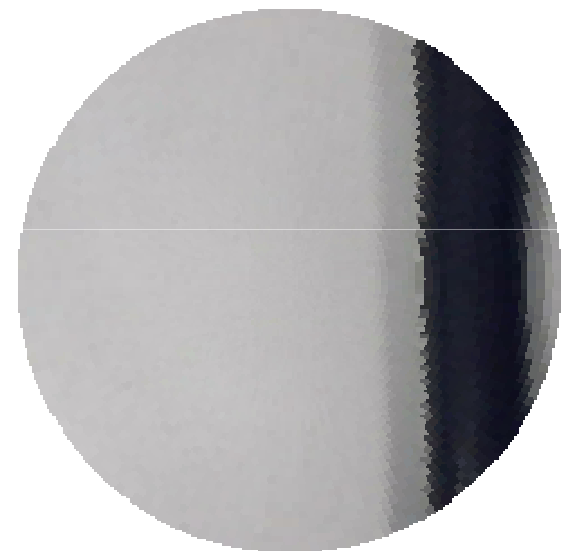
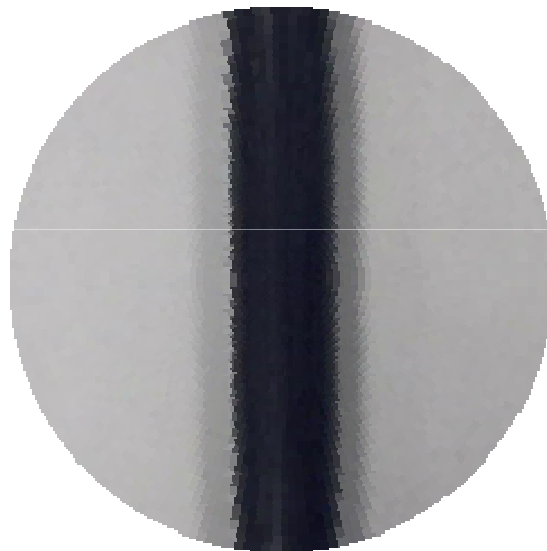
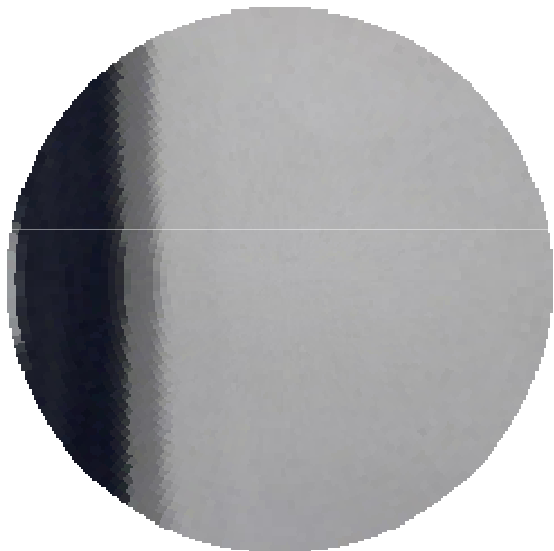


Immagine cartesiana ricostruita dalla polare

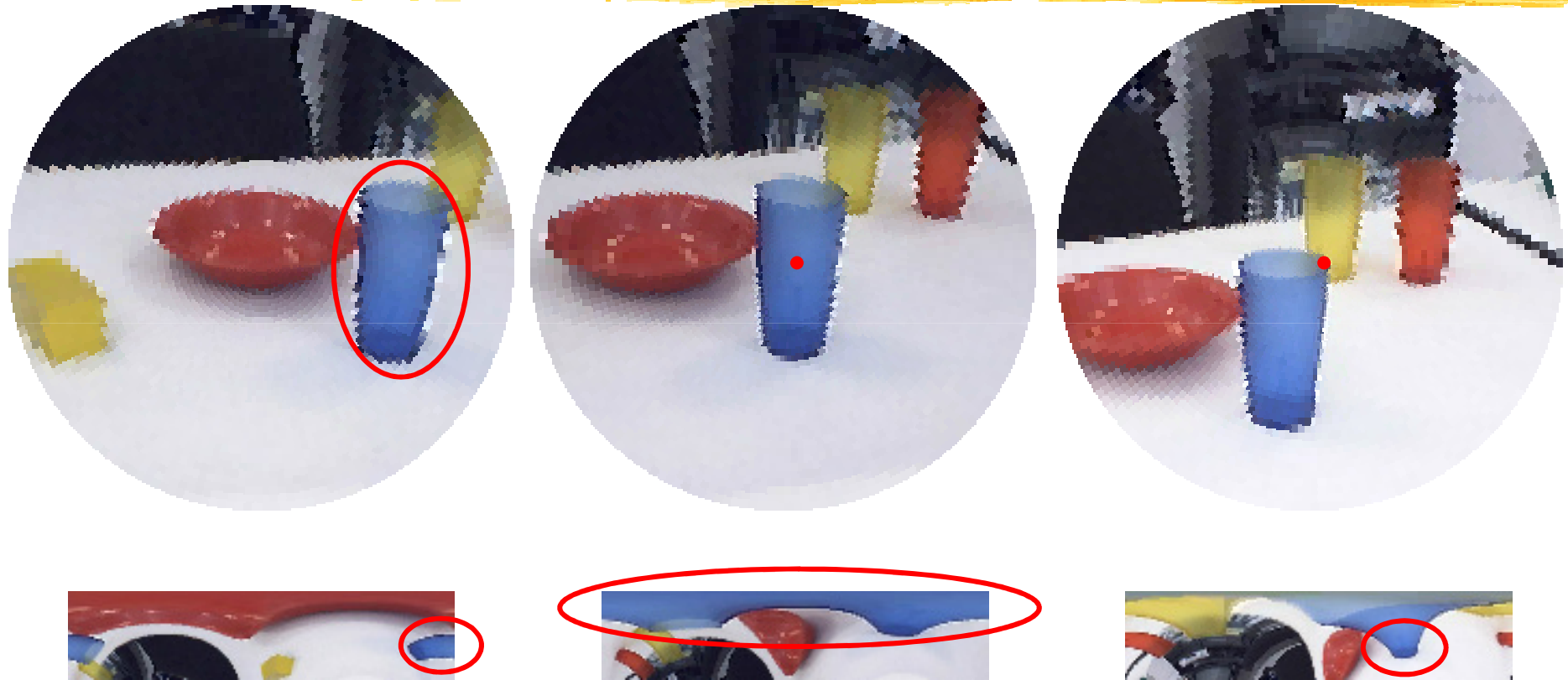
An example of pattern translation



An example of pattern translation



An example of simulated foveation



Object detection
in the periphery

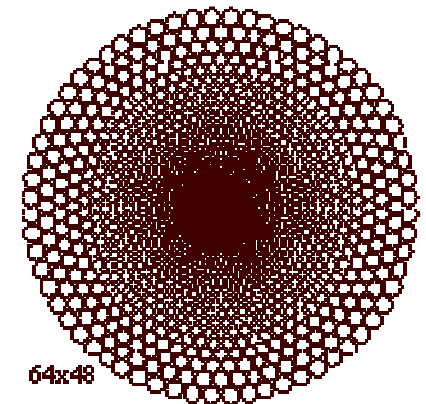
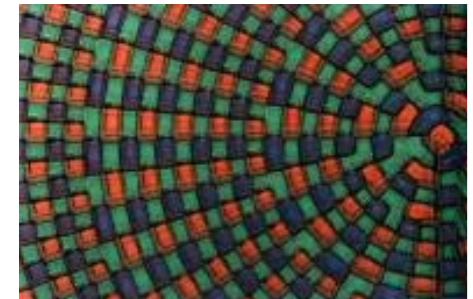
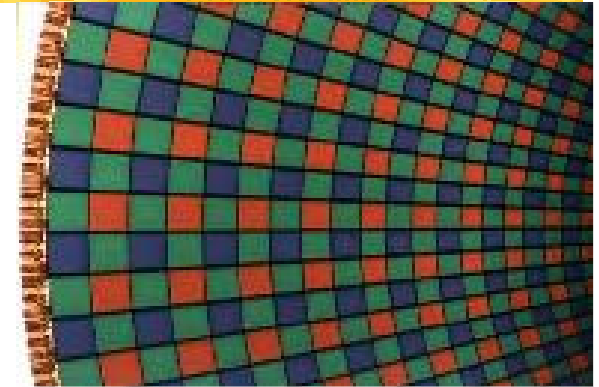
Object foveation

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.

Θ = maximum # of pixels in each ring.

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

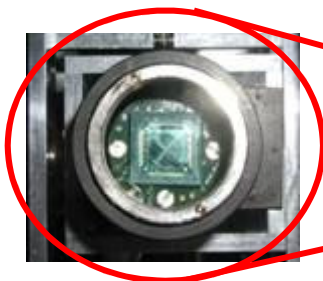
$2Y$ = vertical size of the cartesian image.

ρ = ring number in the log polar image.

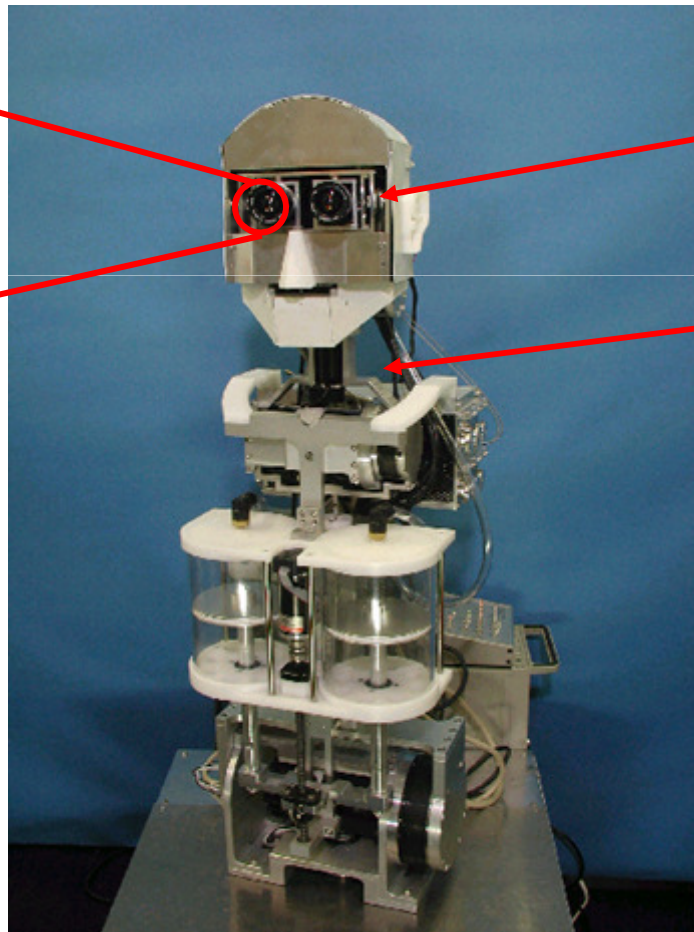
θ = angular polar coordinate.

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*

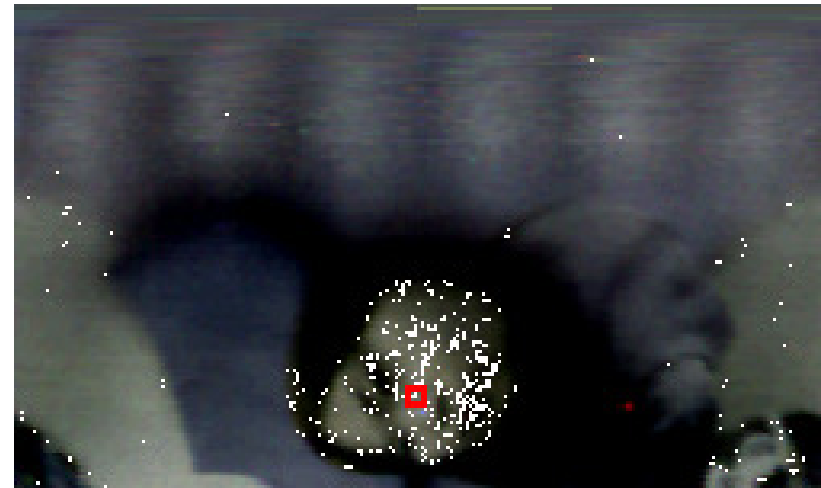
Face detection by hue

Hue = information on the color

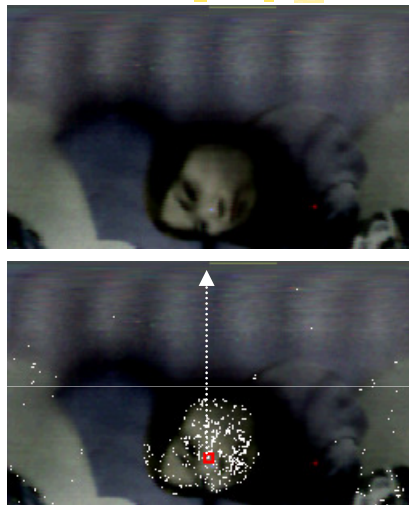
$$\text{Hue} = \cos^{-1} \left(\frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \right)$$

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

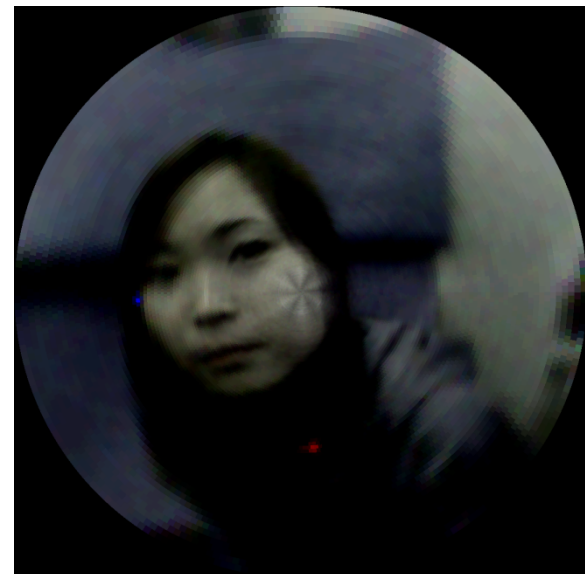
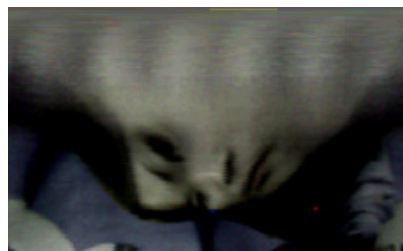
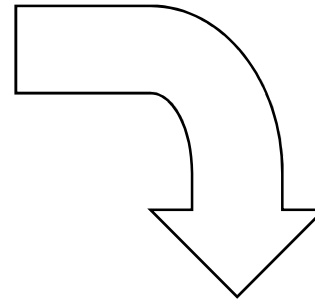
R, G, B = RED, GREEN, BLUE components, respectively



An example of foveation

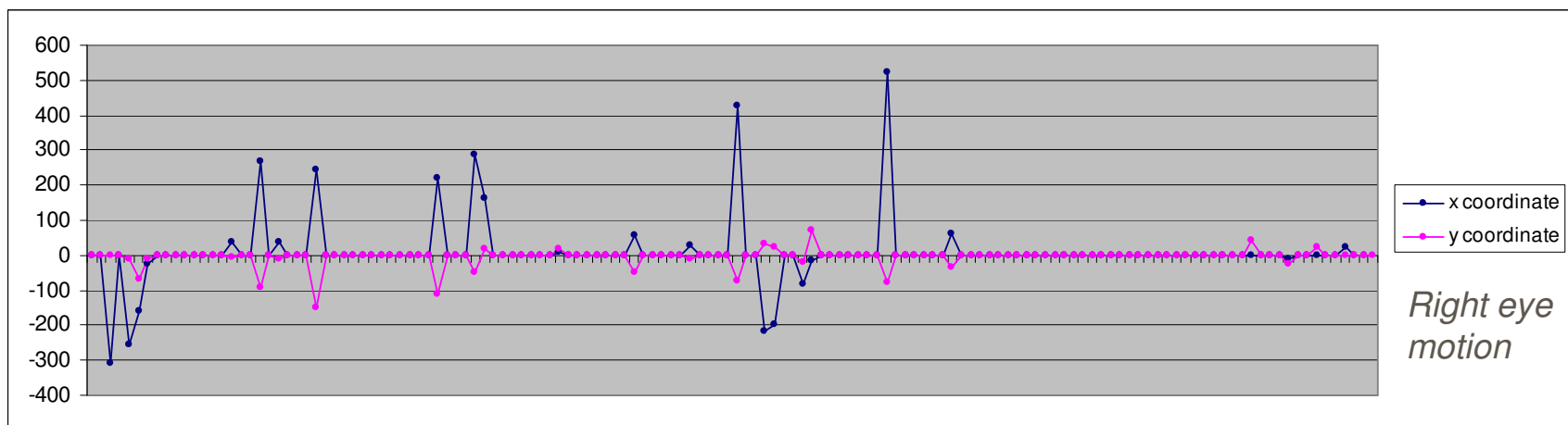
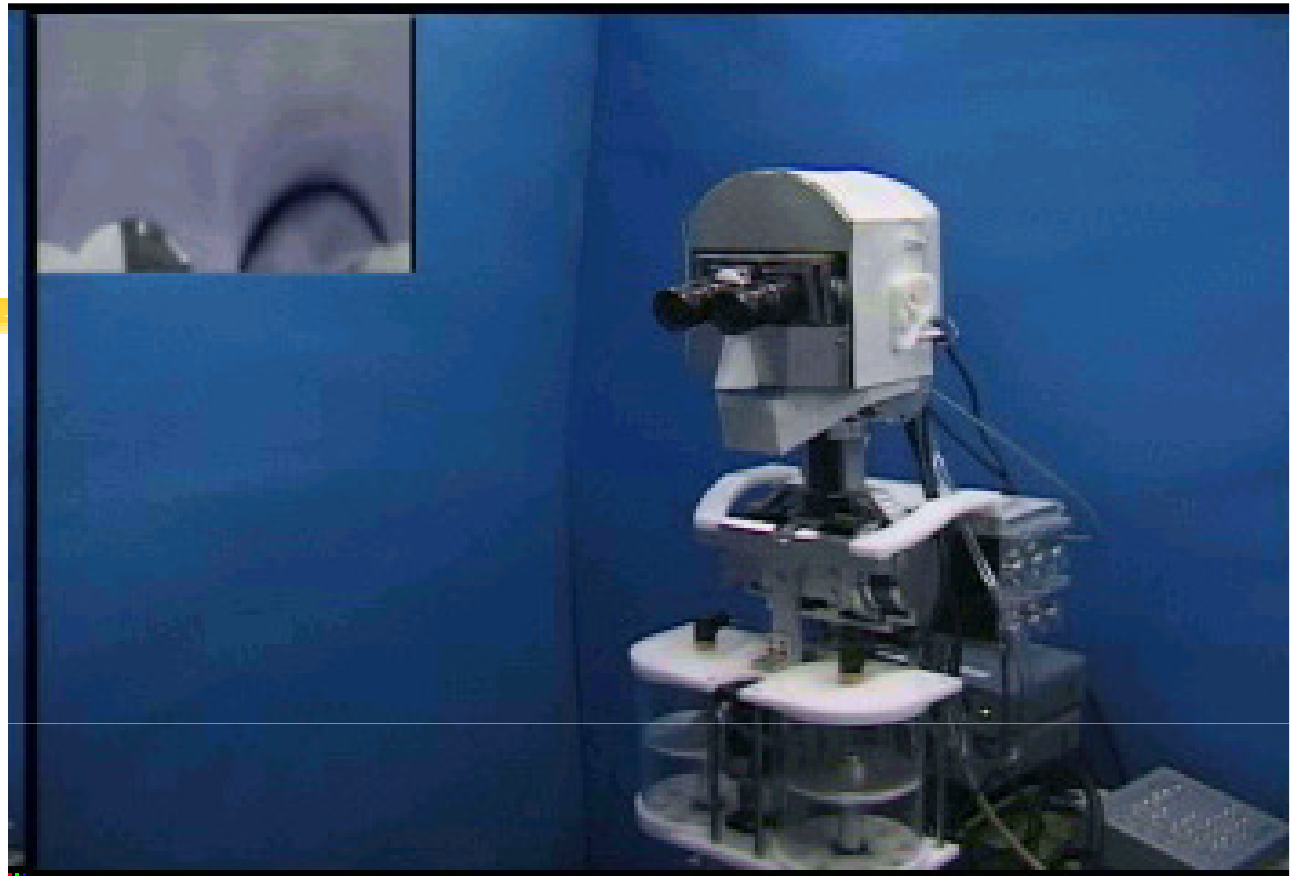


Eye/neck movements



Proportions are rescaled for display purposes

Experimental trials



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

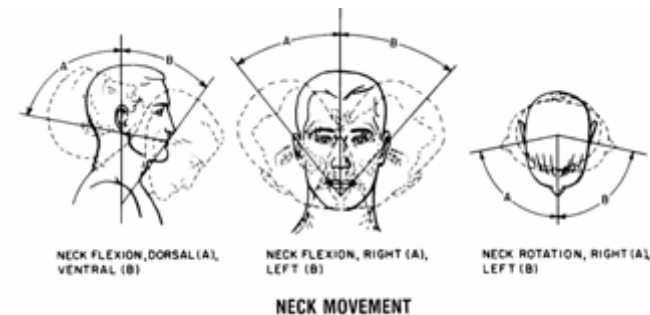
Example of design and development of
a human-like robotic head



**The ARTS humanoid robot
head**

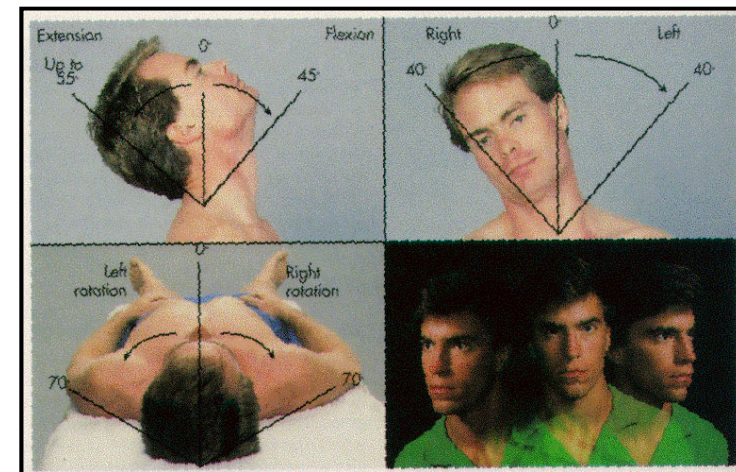
Synthesis of characteristics of the human oculo-motor system

- 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



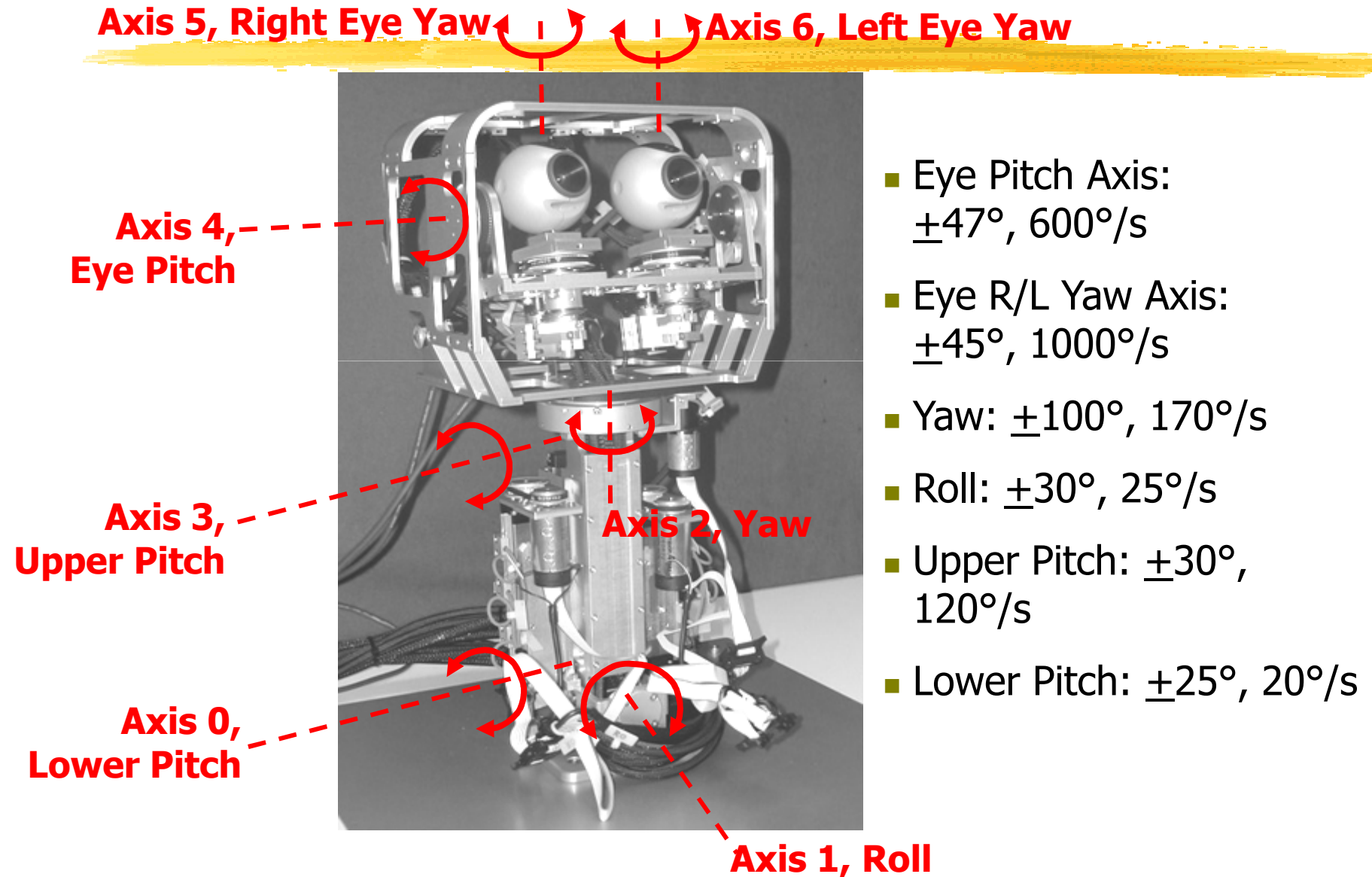
	Average	SD
Ventral flexion	60°	12
Dorsal flexion	51°	17
Right-left flexion	41°	7
Right-left rotation	79°	14

RANGE OF MOVEMENT AT THE NECK JOINT
*Male civilians.



[Thibodeau & Patton, 1996]

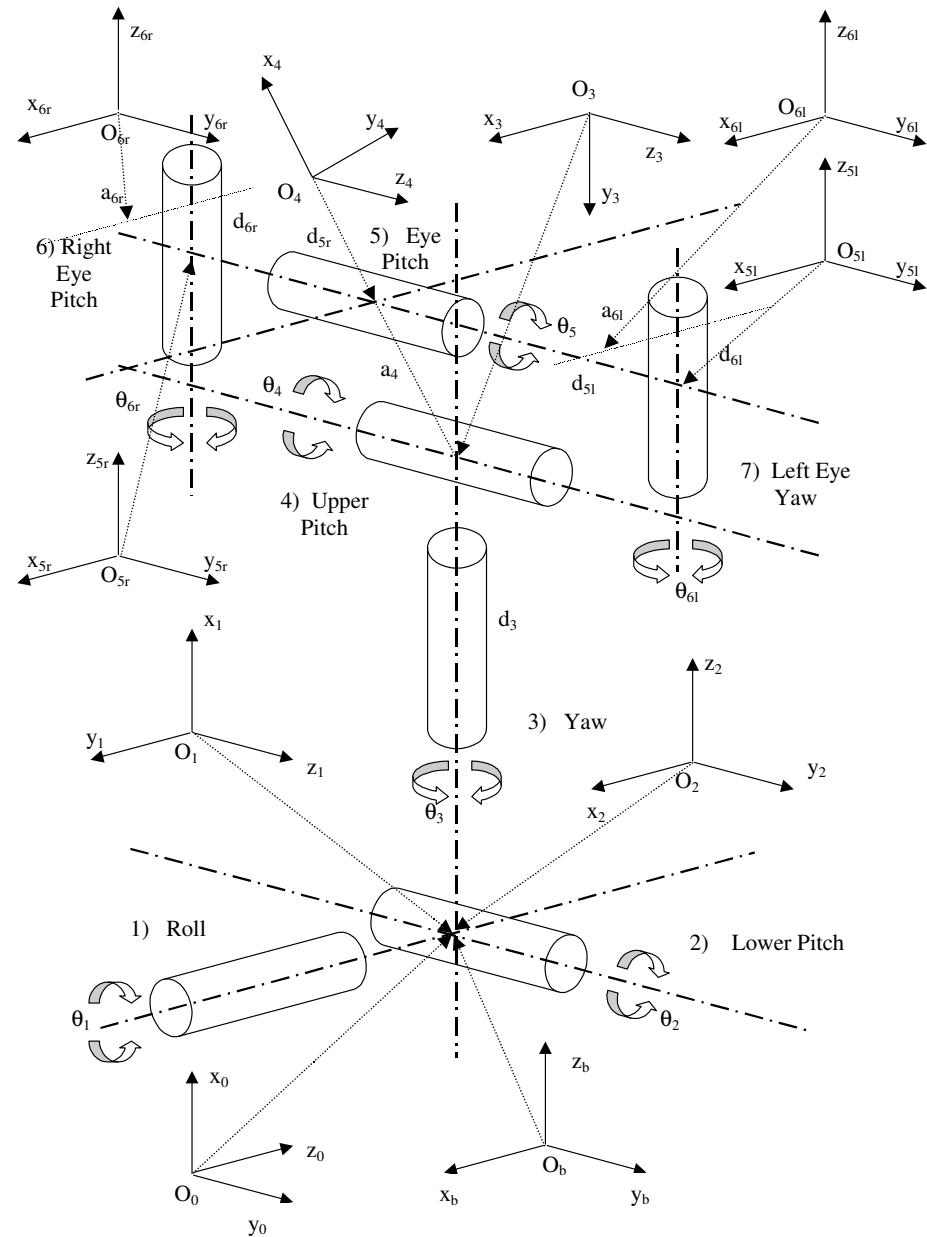
Kinematic structure of the SSSA Robot Head



- 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)	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 ÷ -50	$\pi/2$
J5 _l	0	30 ÷ 50	$\pi/2$
J6 _l	a_{6l}	d_{6l}	0
J6 _r	a_{6r}	d_{6r}	0

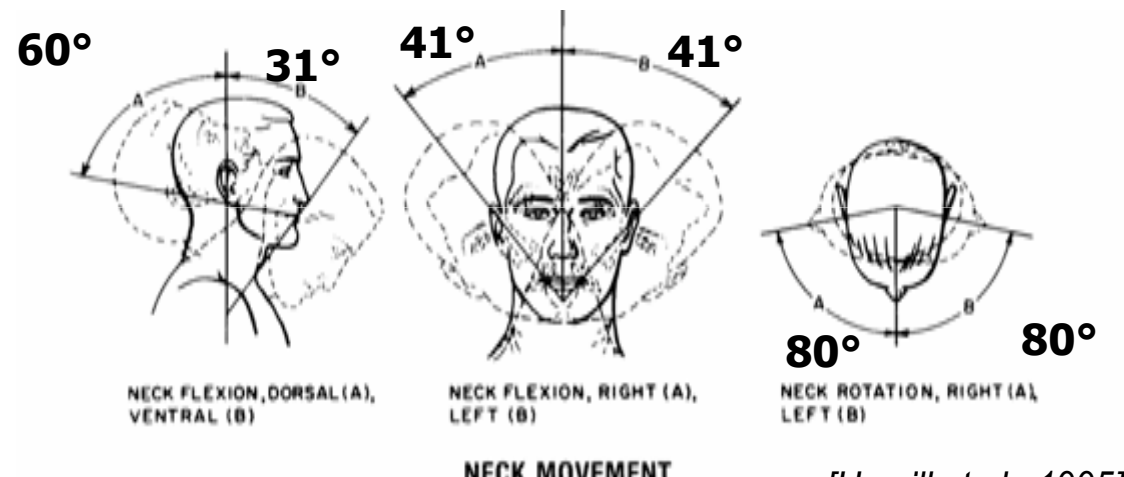


Comparison of performances between human and robotic head

Neck:

Eye:

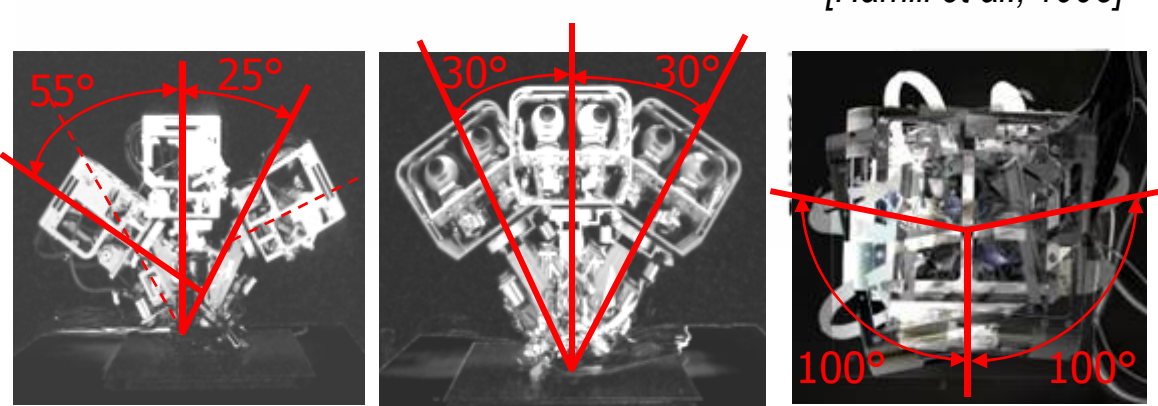
Human



Pitch: $\pm 60^\circ$, $600^\circ/s$
 Yaw: $\pm 30^\circ$, $600^\circ/s$

[Hamill et al., 1995]

Robot



Pitch: $\pm 47^\circ$, $600^\circ/s$
 Yaw: $\pm 45^\circ$, $1000^\circ/s$

The movements of the 7 dofs of the robotic head



Neck Movements

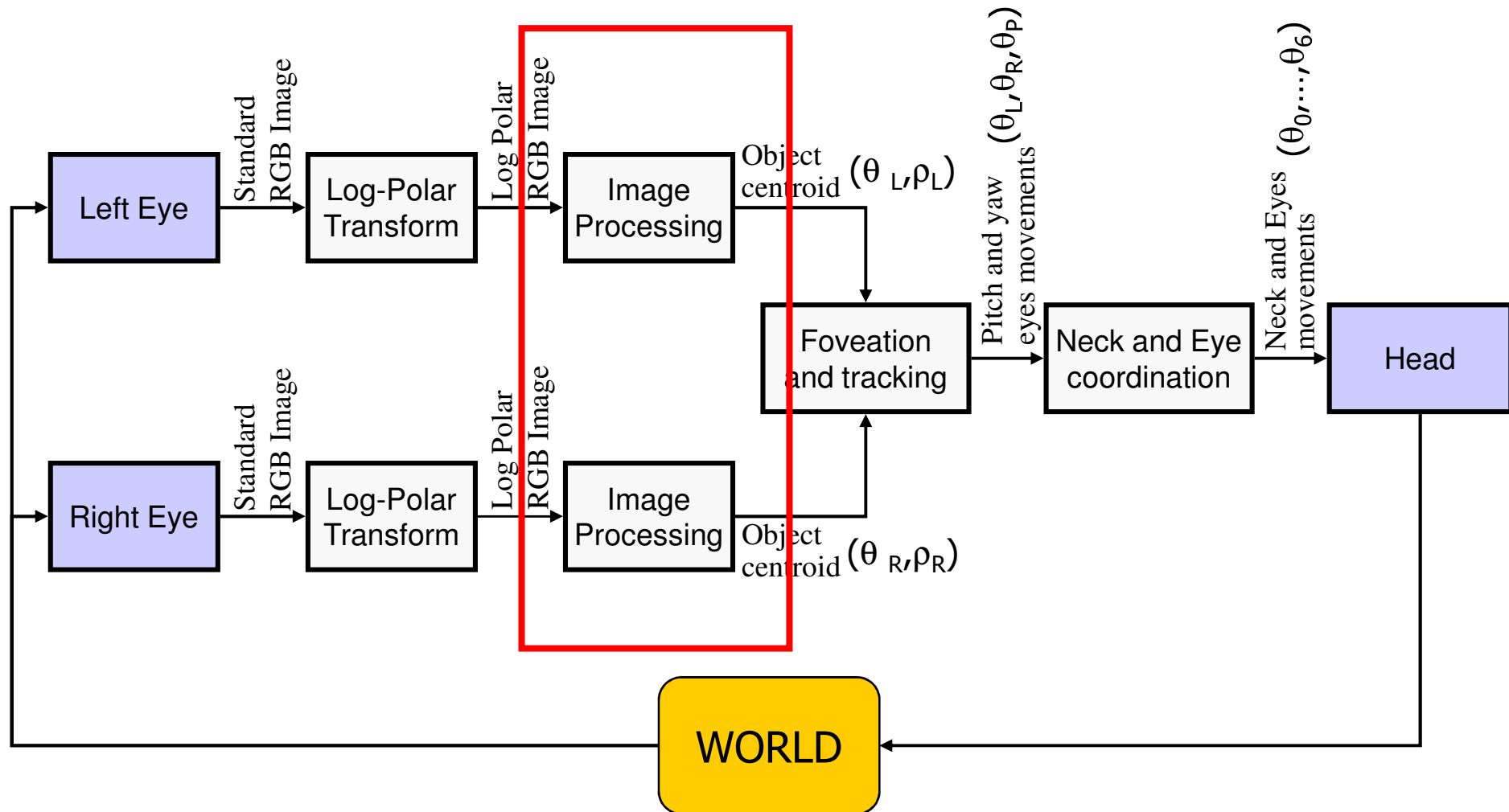


Saccades, 400°/sec



saccade 300° sec.avi

Overall sensory-motor scheme of the visual apparatus



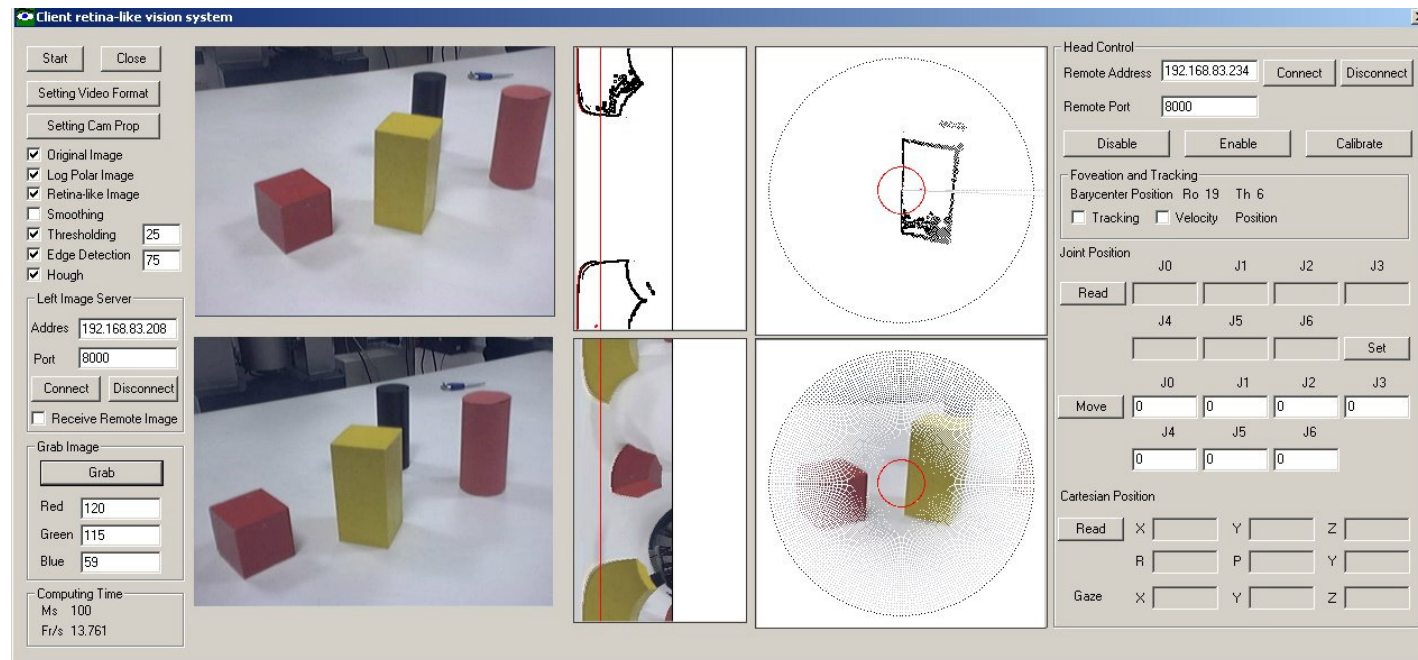
Examples of algorithms developed for retina-like image processing



- 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

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), HSV, Edge Detection (75), and Hough. It also includes a Left Image Server section with fields for Address (PALOMA1) and Port (8000), and a Grab Image section with a Grab button and RGB/HSV values.
- Image Display:** Shows two camera feeds of a table with colored blocks (red, blue, green, yellow, black). The top feed is the original image, and the bottom feed is the processed image. The processed image shows a yellow block highlighted with a red circle, indicating successful thresholding.
- 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, Velocity, Position, Prop. Par 70, Velocity 0.50), Joint Position (Read, Move, J0-J6, Vel, T, Set), Cartesian Position (Read, Gaze, X, Y, Z), and Head Neurocontroller (X, Y, Z, Clamped Joints, Value Joints, Sym, Move).

Computing Time: Ms 62, Fr/s 6.340

Edge Detection (gradient based method)

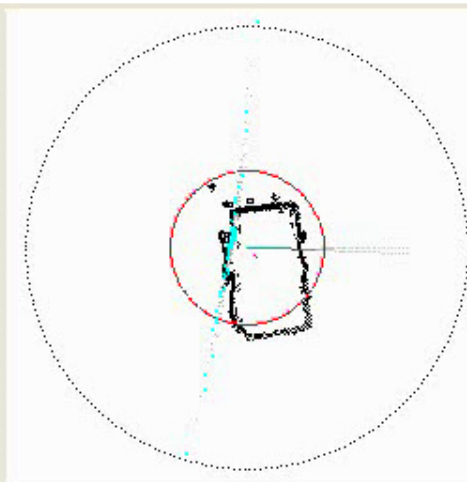
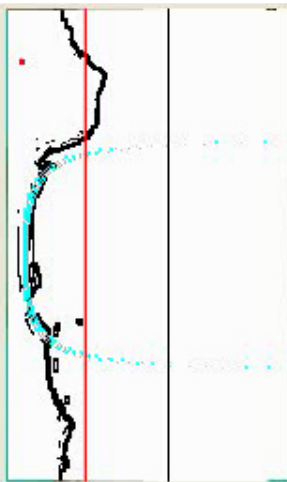
The screenshot displays the PALOMA Robotic Artefact Control Panel software interface. The main window is divided into several sections:

- Left Panel:** Contains control buttons (Start, Close, Setting Video Format, Setting Cam Prop) and a list of processing options: Original Image (checked), Log Polar Image (checked), Retina-like Image (unchecked), Smoothing (unchecked), Thresh (100, checked), HSV (checked), Edge Detection (75, checked), and Hough (unchecked). It also includes a Left Image Server section with address (PALOMA1), port (8000), and connection buttons (Connect, Disconnect), and a Grab Image section with a Grab button and RGB/HSV values.
- Top Left:** A live video feed showing a table with various colored objects (red sphere, blue cube, black cylinder, green cube, yellow cube).
- Top Middle:** A vertical strip showing the edge detection process, with a red vertical line indicating the current position.
- Top Right:** A circular field of view showing the detected edges of the objects, with a red circle highlighting a specific object.
- Bottom Left:** A second live video feed, identical to the top left one.
- Bottom Middle:** A second vertical strip showing the edge detection process.
- Bottom Right:** A second circular field of view showing the detected edges.
- 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/Move buttons for 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, Value Joints, Sym, Move).

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 Left

Tracking Velocity Position -0

Prop. Par: 70

Velocity: 0.35

Joint Position

	J0	J1	J2
Read	0	10	0
Move	J4	J5	J6
	-15	-4	8

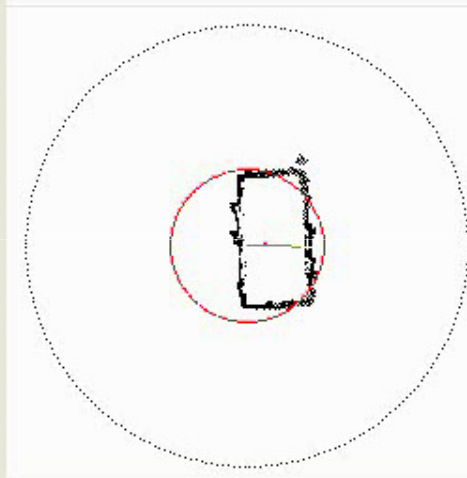
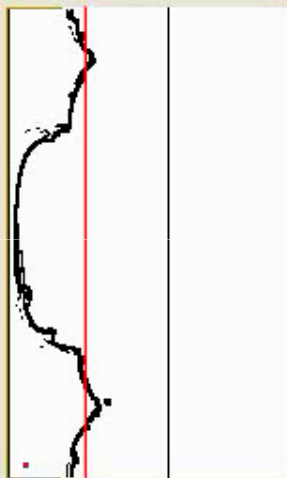
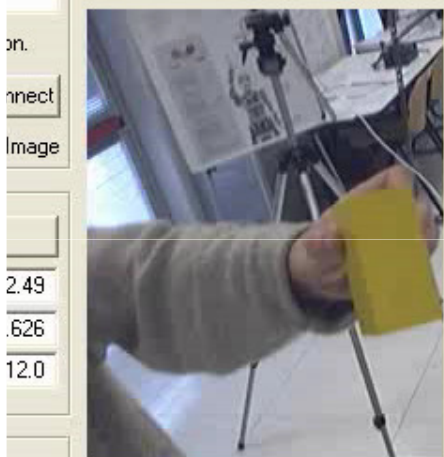
Cartesian Position

Read	X	Y
	R	P
Gaze	X	Y

Head Neurocontroller

	X	Y	Z
	85	0	10

Sym Move



Arm Control Panel

Enable Arm Disable Arm

Move	J0	J1	J2	J3	J4	J5	J6	J7	POS
Read	90.0	0.0	135.0	0.0	-90.0	0.0	0.0	0.0	

Move	X	Y	Z	Roll	Pitch	Yaw	J0	Elb
Read								

Block Compliant

Or

HSV

75

on.

connect

Image

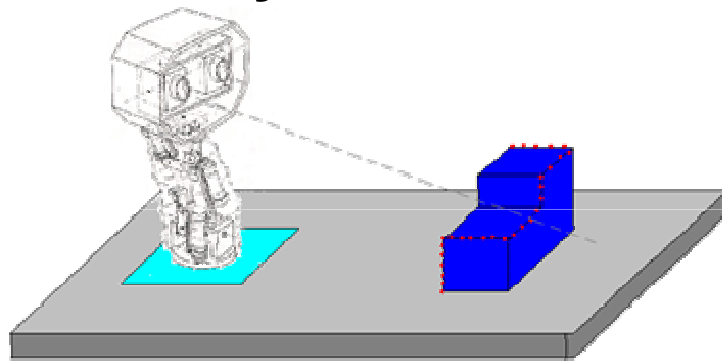
2.49

626

12.0

Preliminary activities

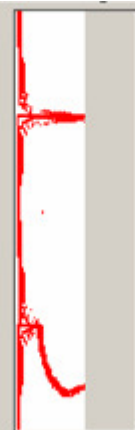
Foveation and tracking of borders of object and reconstruction of the geometry of the object



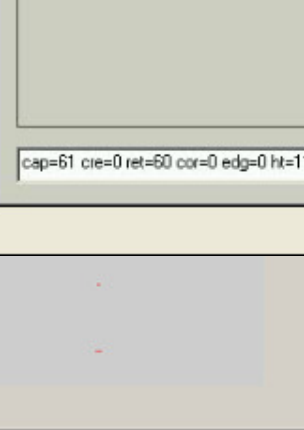
Retina Like image



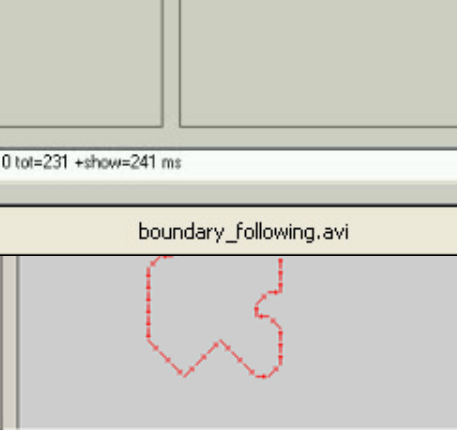
Log Polar Image



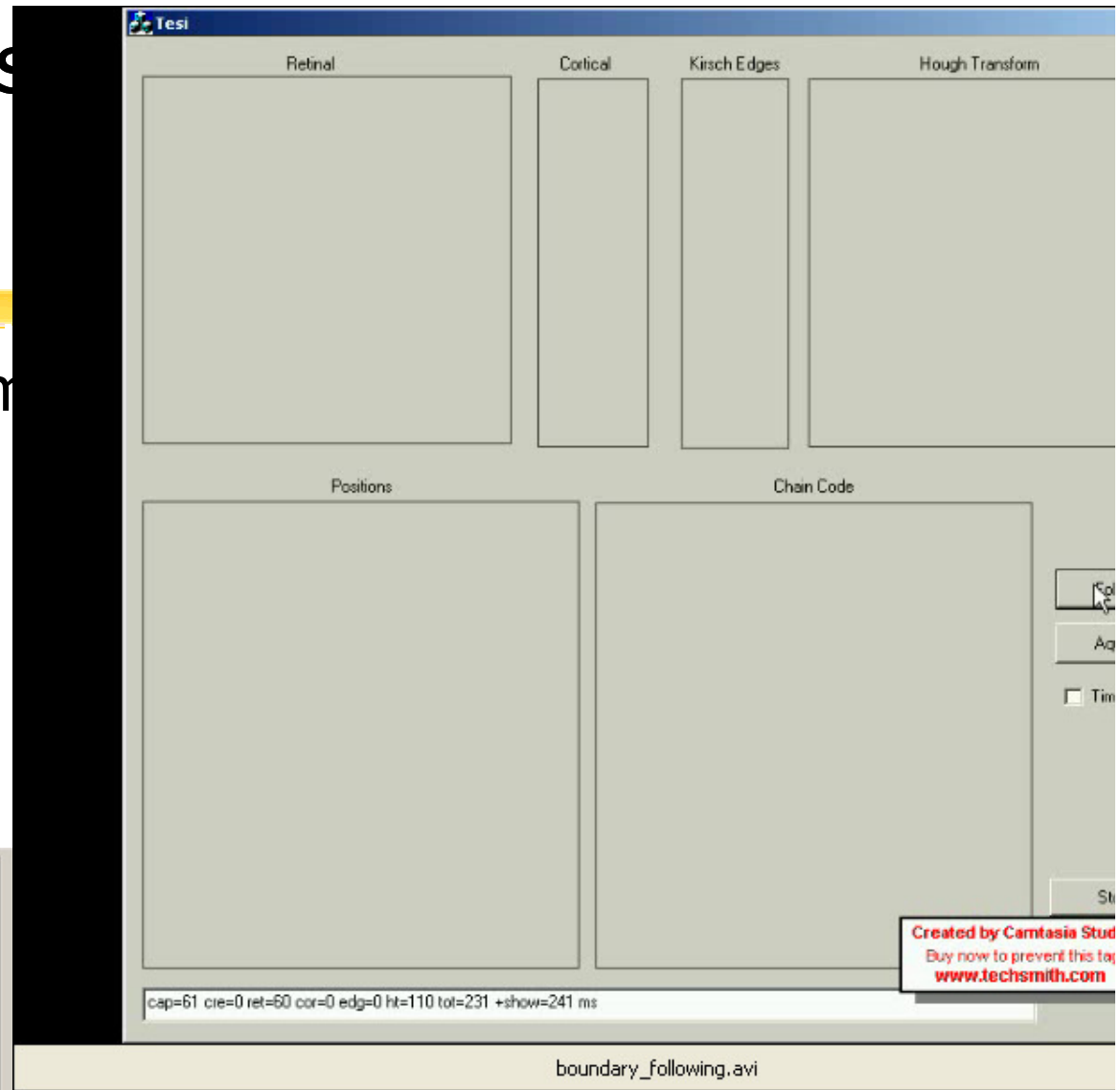
Edge of log polar image



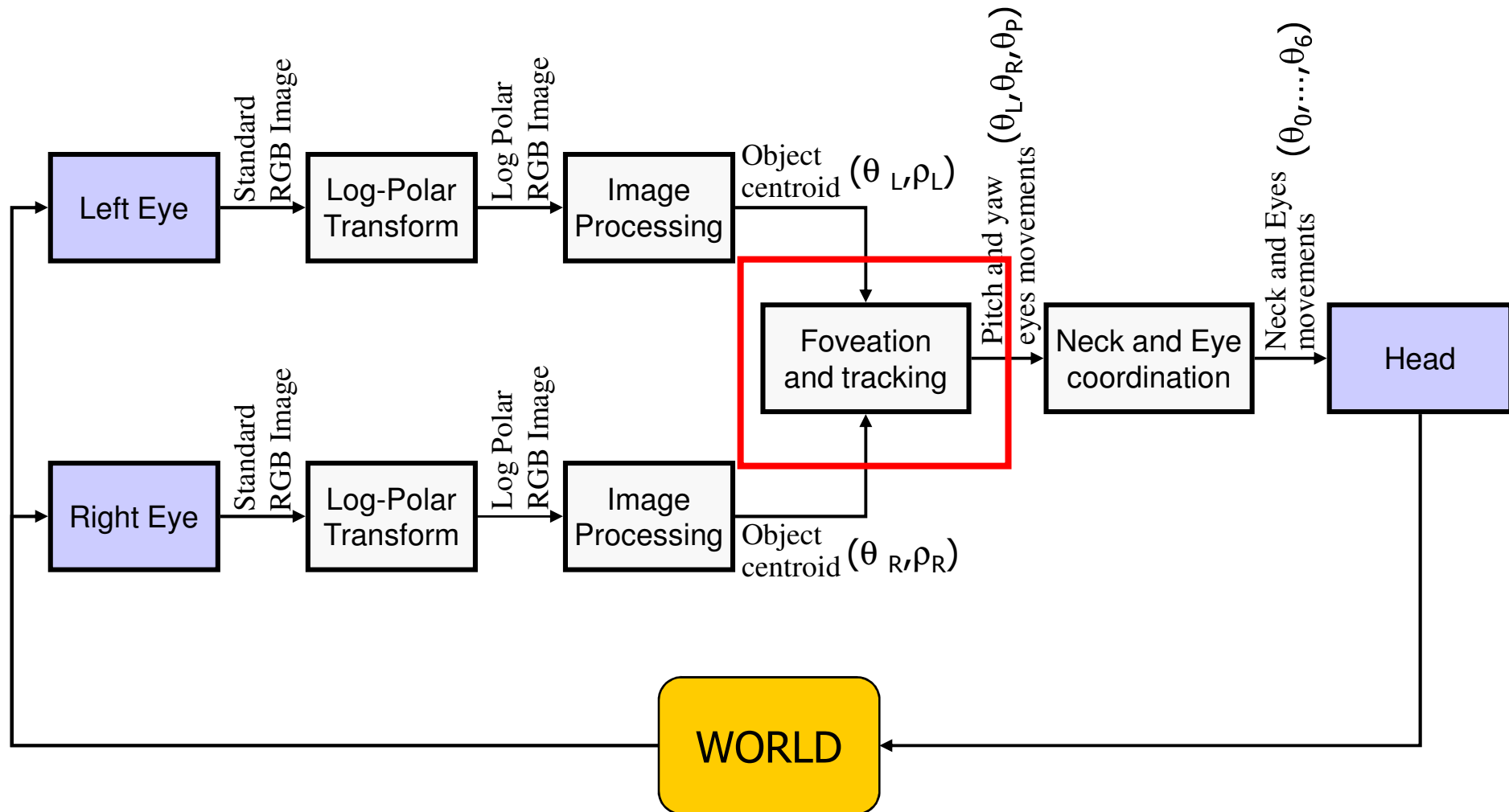
Detected lines (Boundaries)



Boundary reconstruction based on eye positions



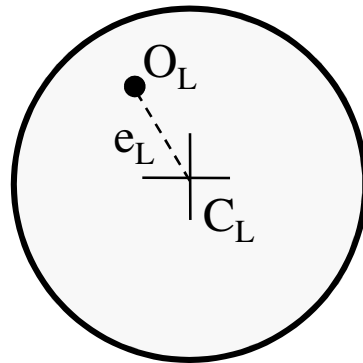
Overall sensory-motor scheme of the visual apparatus



Foveation of the object centroid

Proportional control based on the visual error

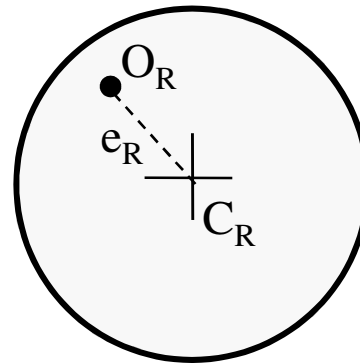
Left Image



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

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

Right Image

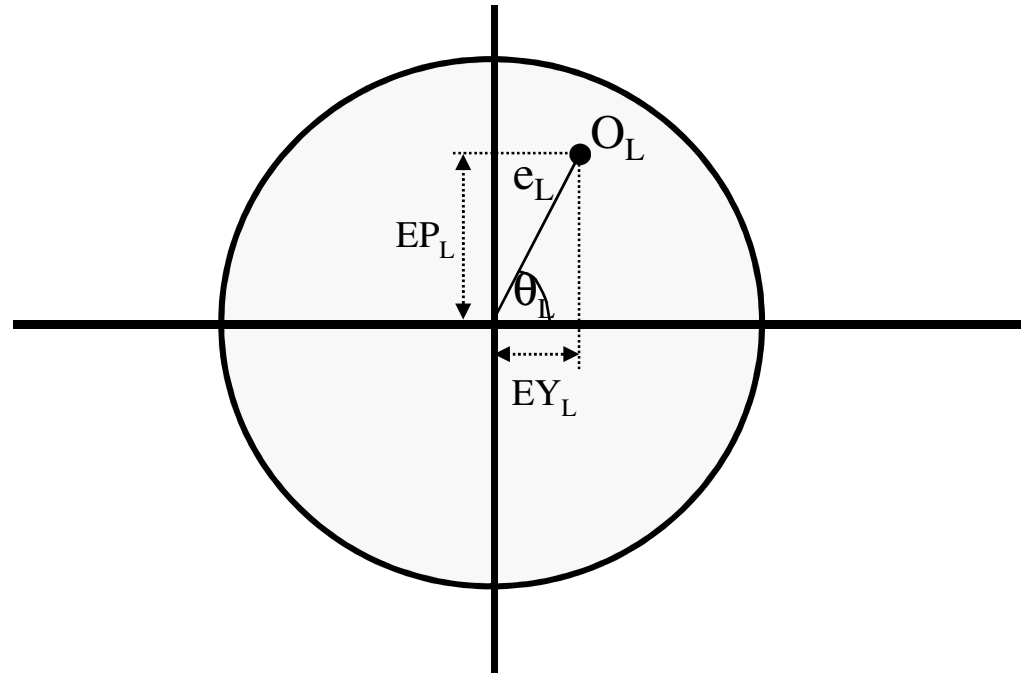


$$O_R = (\rho_R, \theta_R)$$

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

M_{r0} is the maximum ρ value (i.e. 152)

Computation of yaw and pitch eye movements



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

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

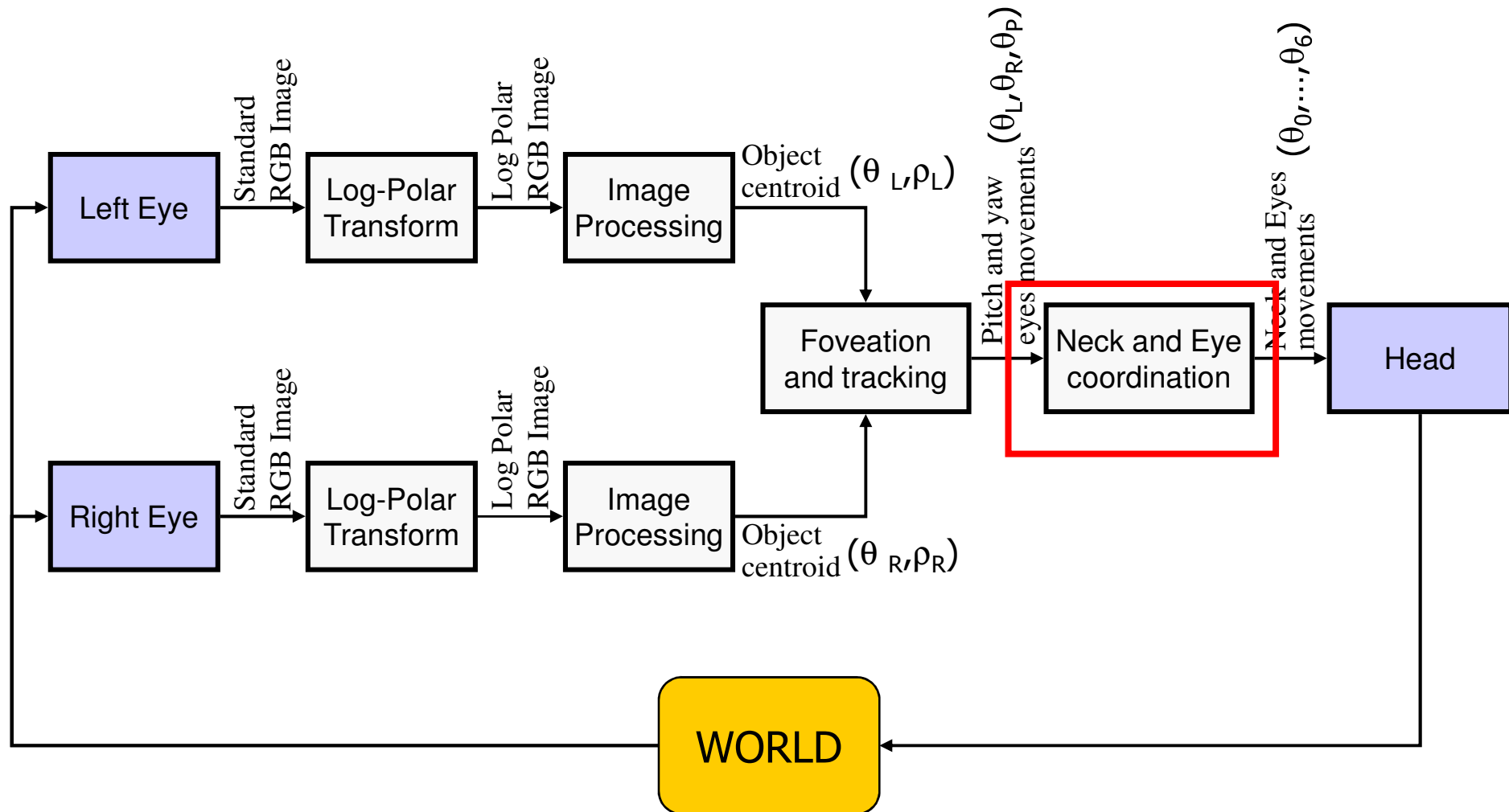
$$EY_R = e_R * \cos(\theta_R) * P_R$$

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

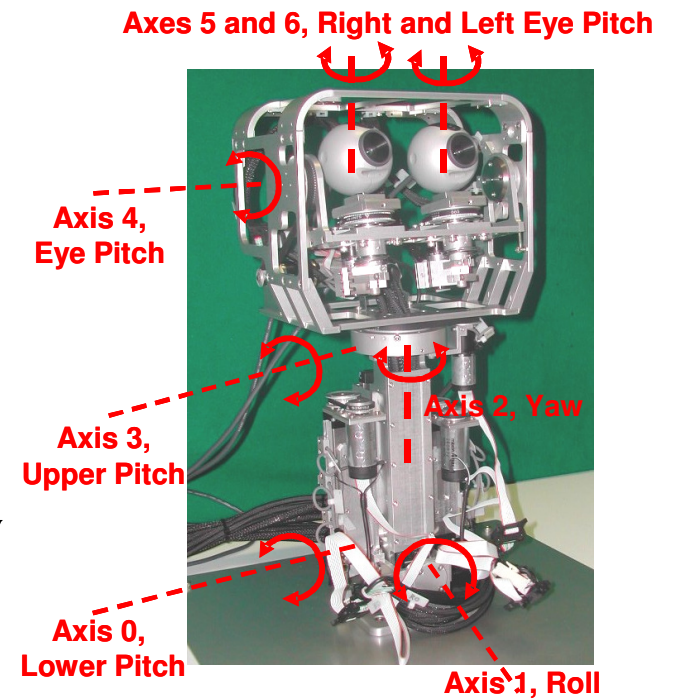
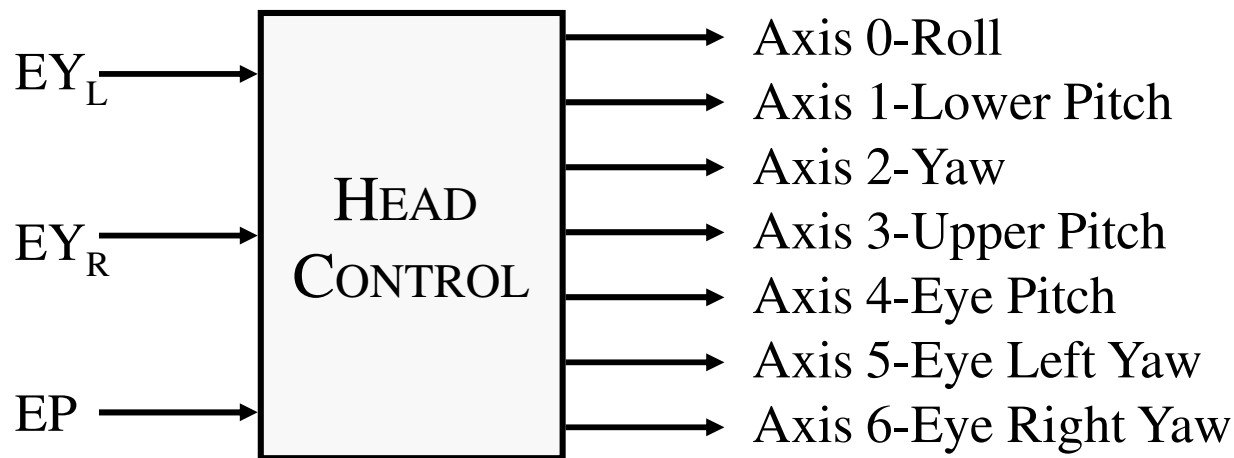
$$EP = (EP_L + EP_R) / 2$$

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



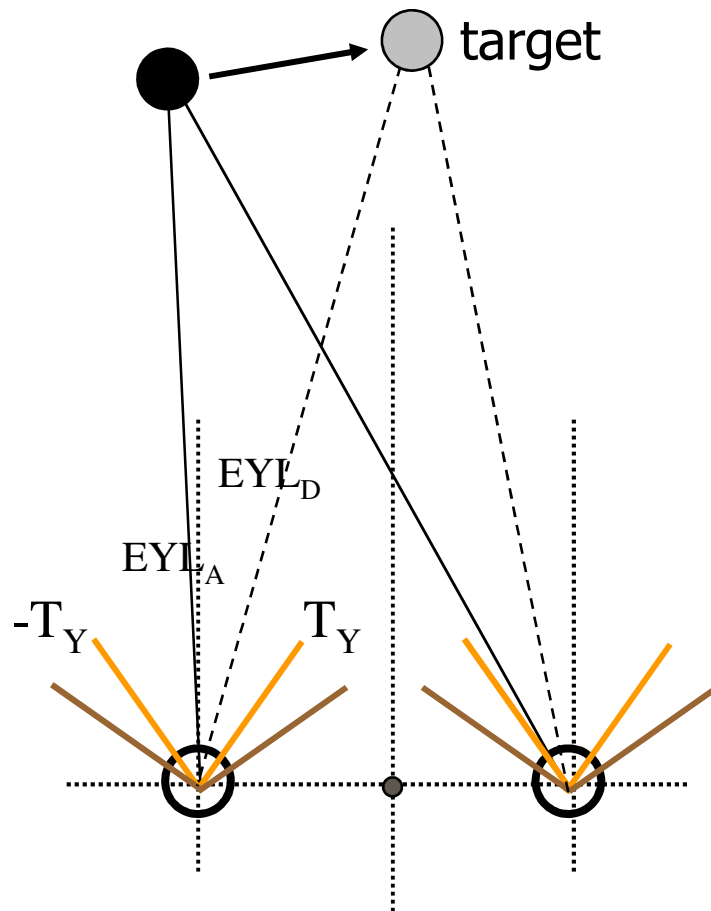
Solution 1



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

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

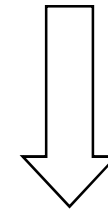
- If the movement is small, it is executed by the eyes, only



$$EYL_A + EY_L < T_Y$$

and

$$EYR_A + EY_R < T_Y$$



Left Eye Yaw

$$EYL_D = EYL_A + EY_L$$

Right Eye Yaw

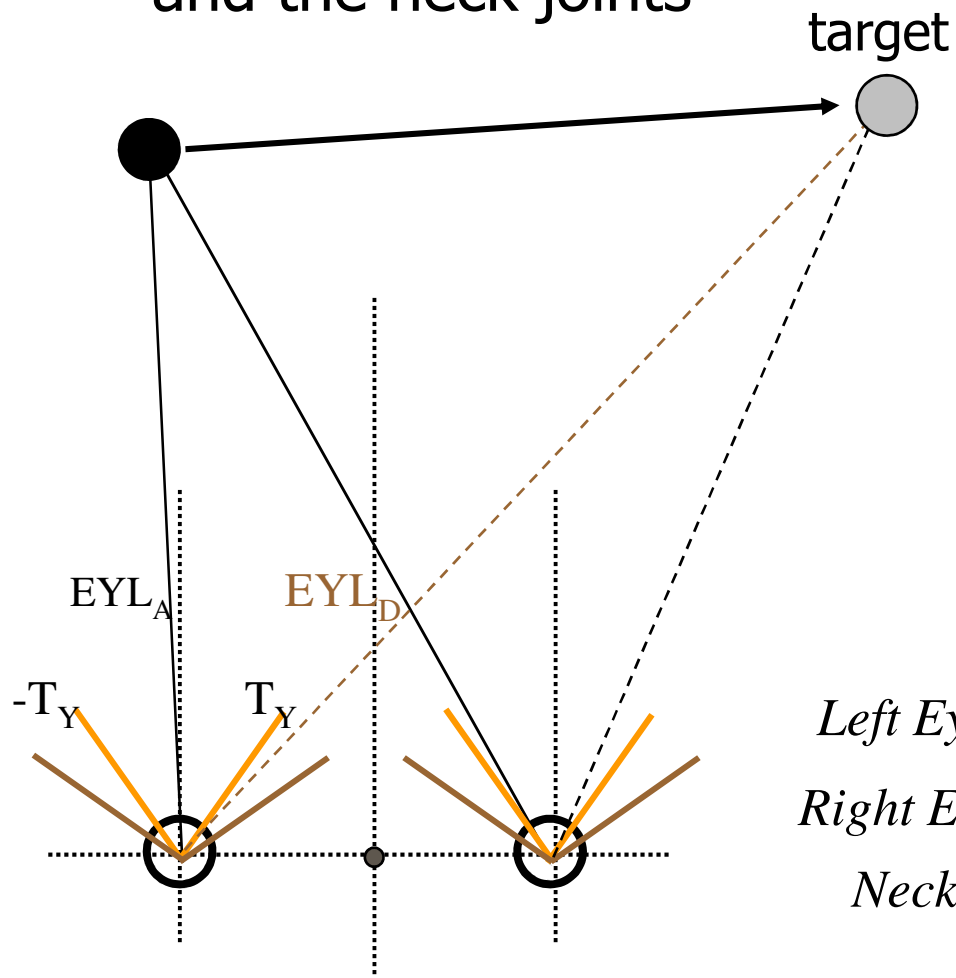
$$EYR_D = EYR_A + EY_R$$

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.

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

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

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

$$EYP_D = EYP_A + EP * K1$$

$$EUP_D = EUP_A + EP * K2$$

$$ELP_D = ELP_A + EP * K3$$

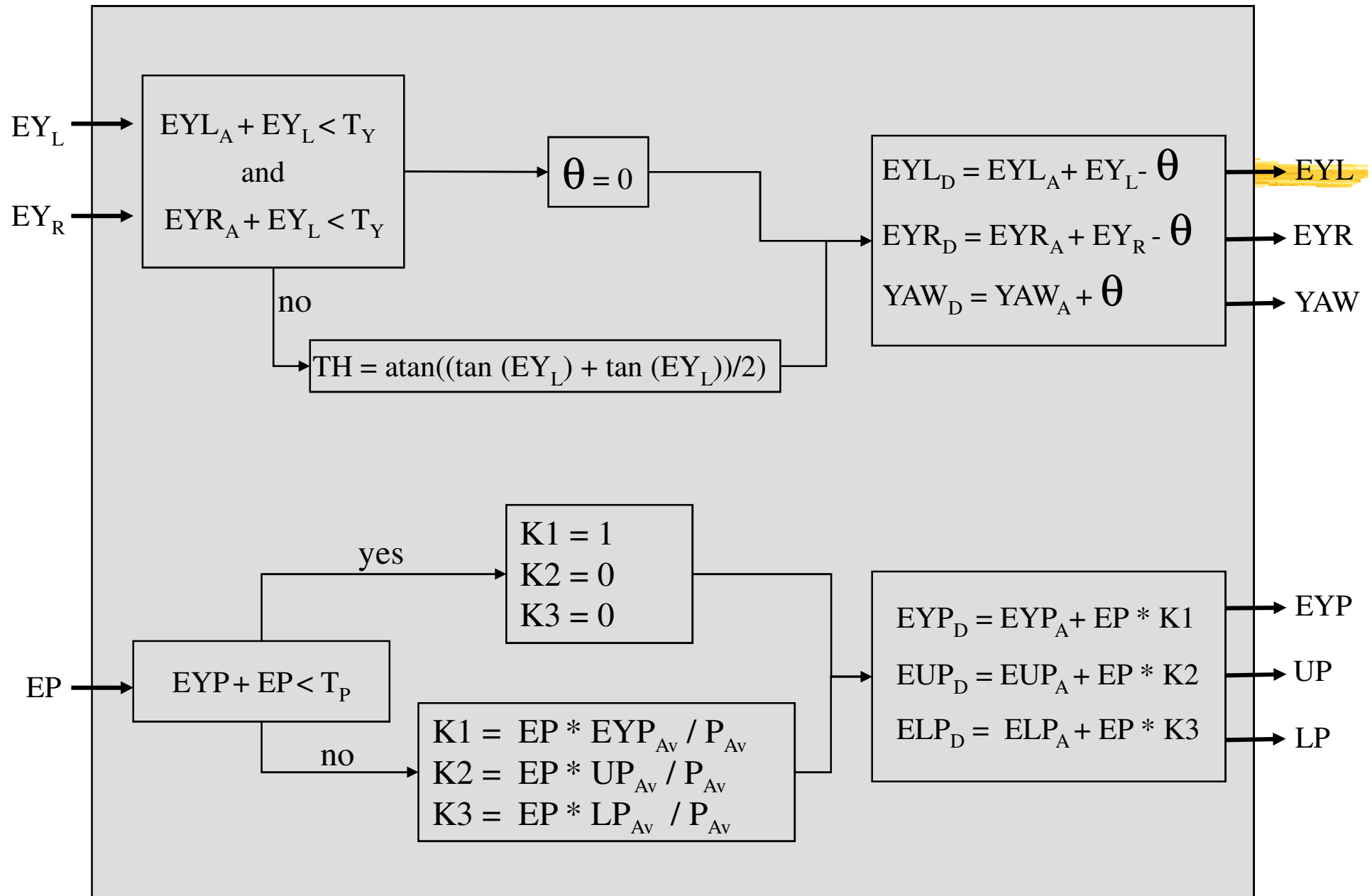
$$EYP_{AV} = EYP_M - EYP_A$$

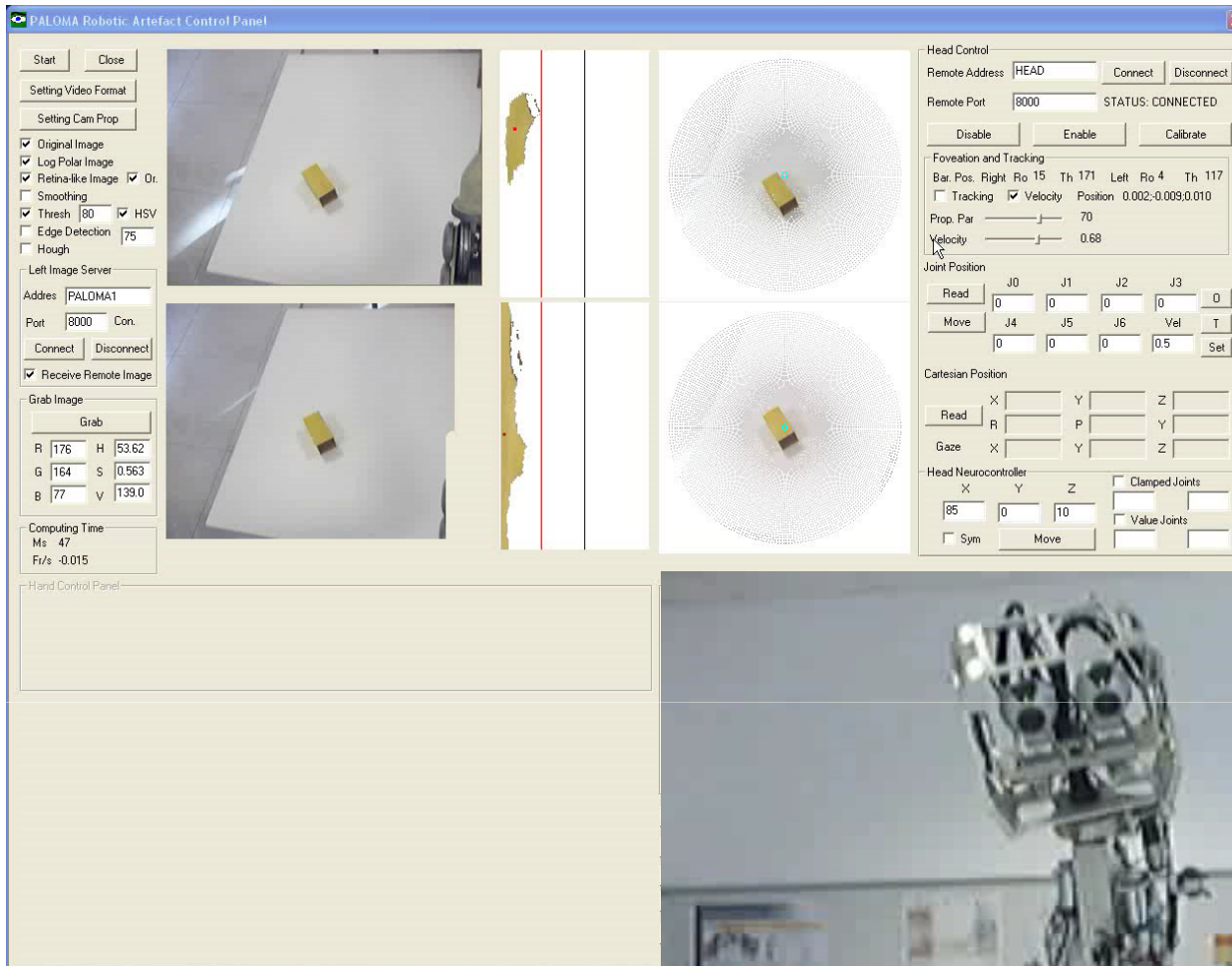
$$UP_{AV} = UP_M - UP_A$$

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

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

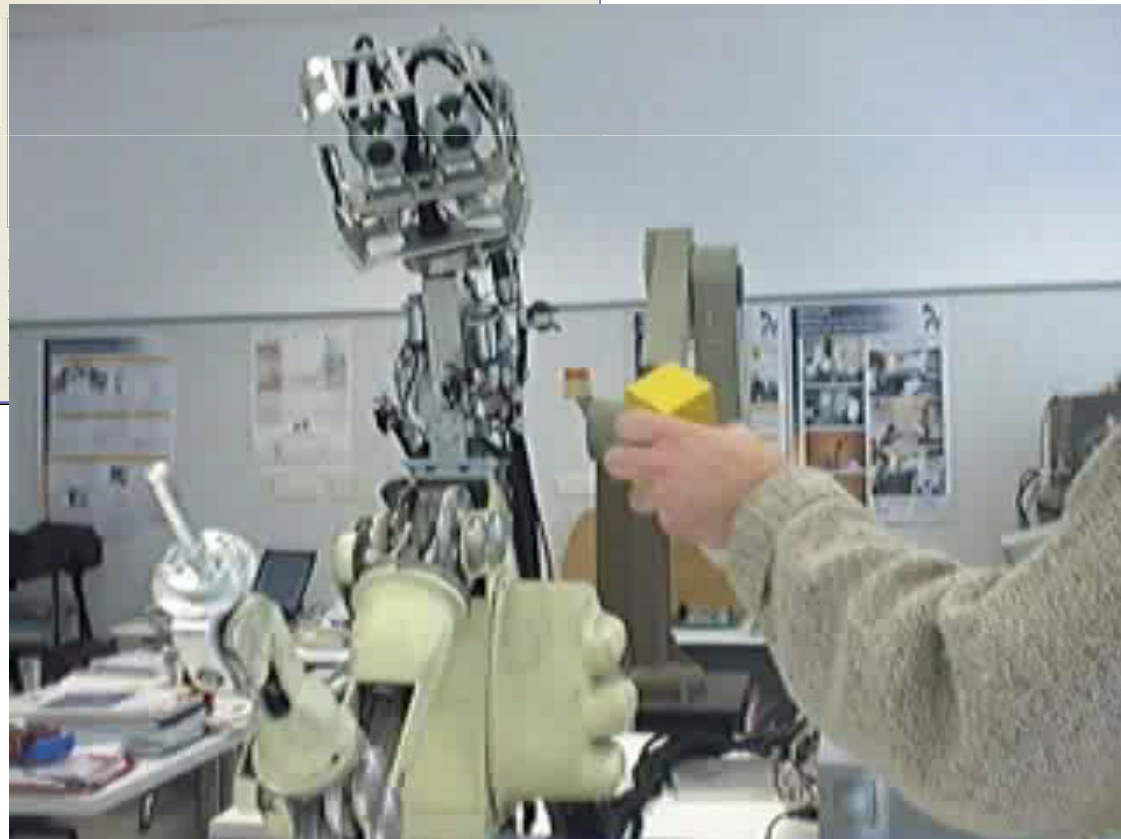
EYP_M , UP_M and LP_M are the range limits respectively for eye pitch, upper pitch and lower pitch axis

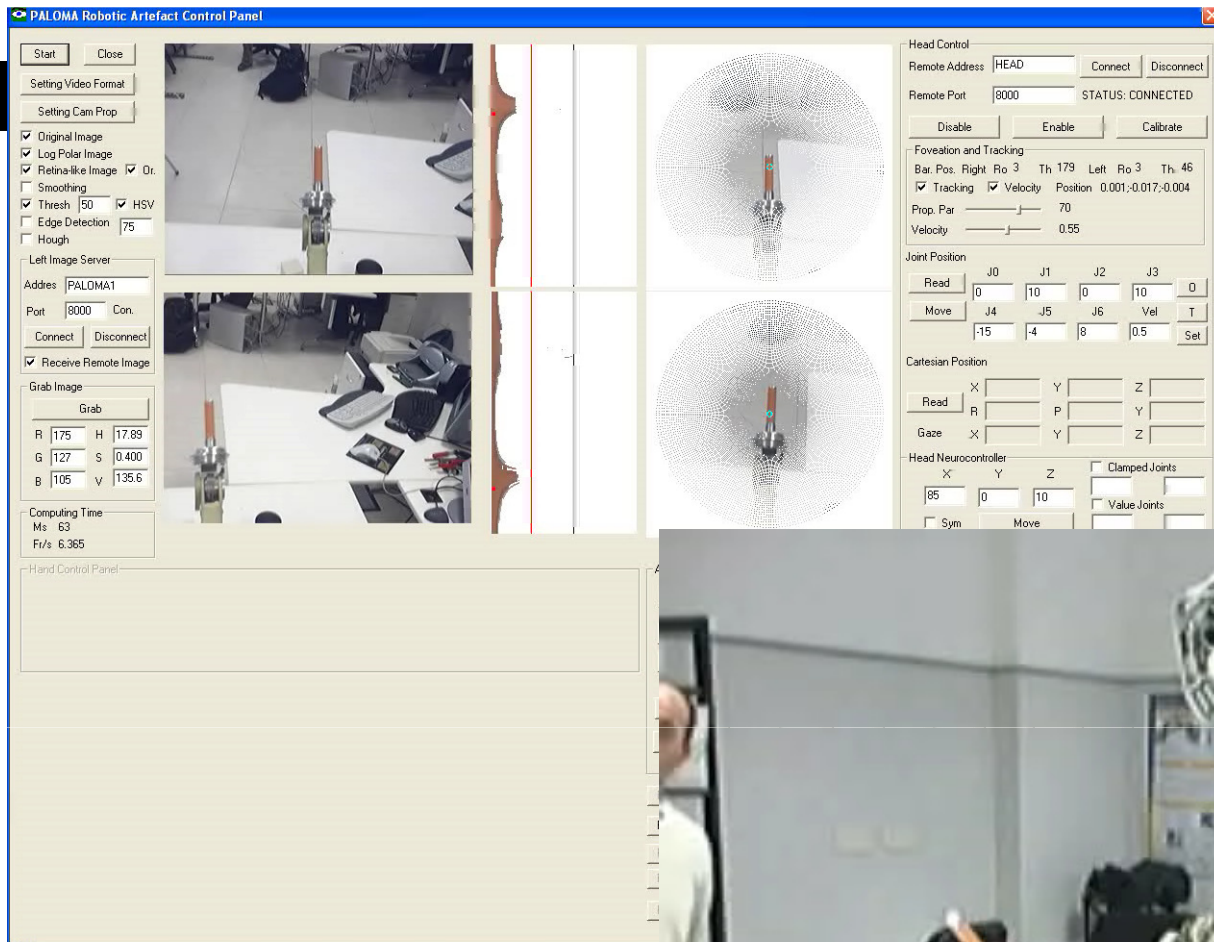




Pursuit Movement

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







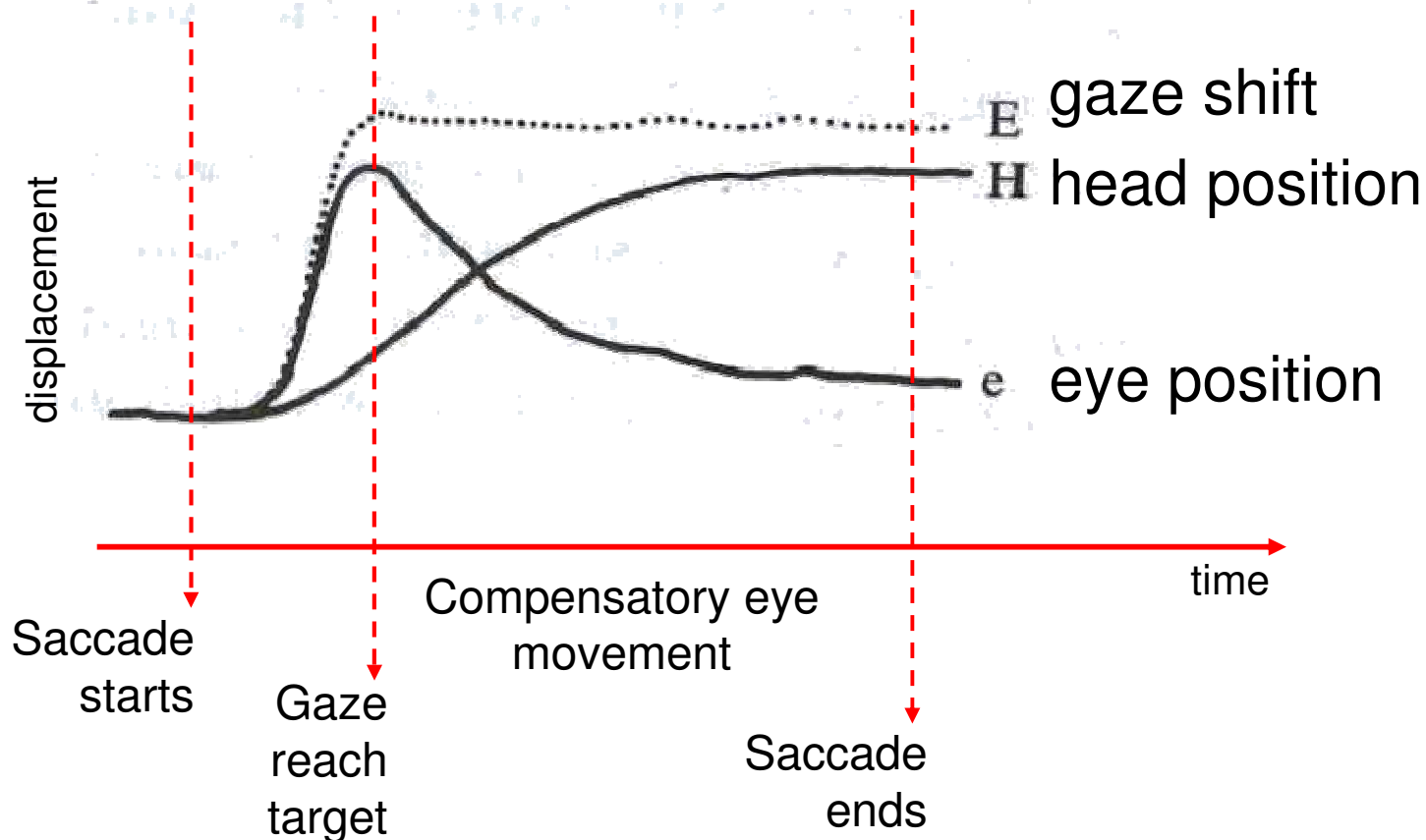
Solution 2

Implementation of a biological model of head-eye coordination

E.S. Maini, G. Teti, C. Laschi, M. Rubino, P. Dario, "Bio-inspired control of eye-head coordination in a robotic anthropomorphic head", *IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*, Pisa, Italy, February 20-22, 2006

What happens in Humans

to foveate a visual target in the periphery, gaze shifts are usually made with combined eye and head movements

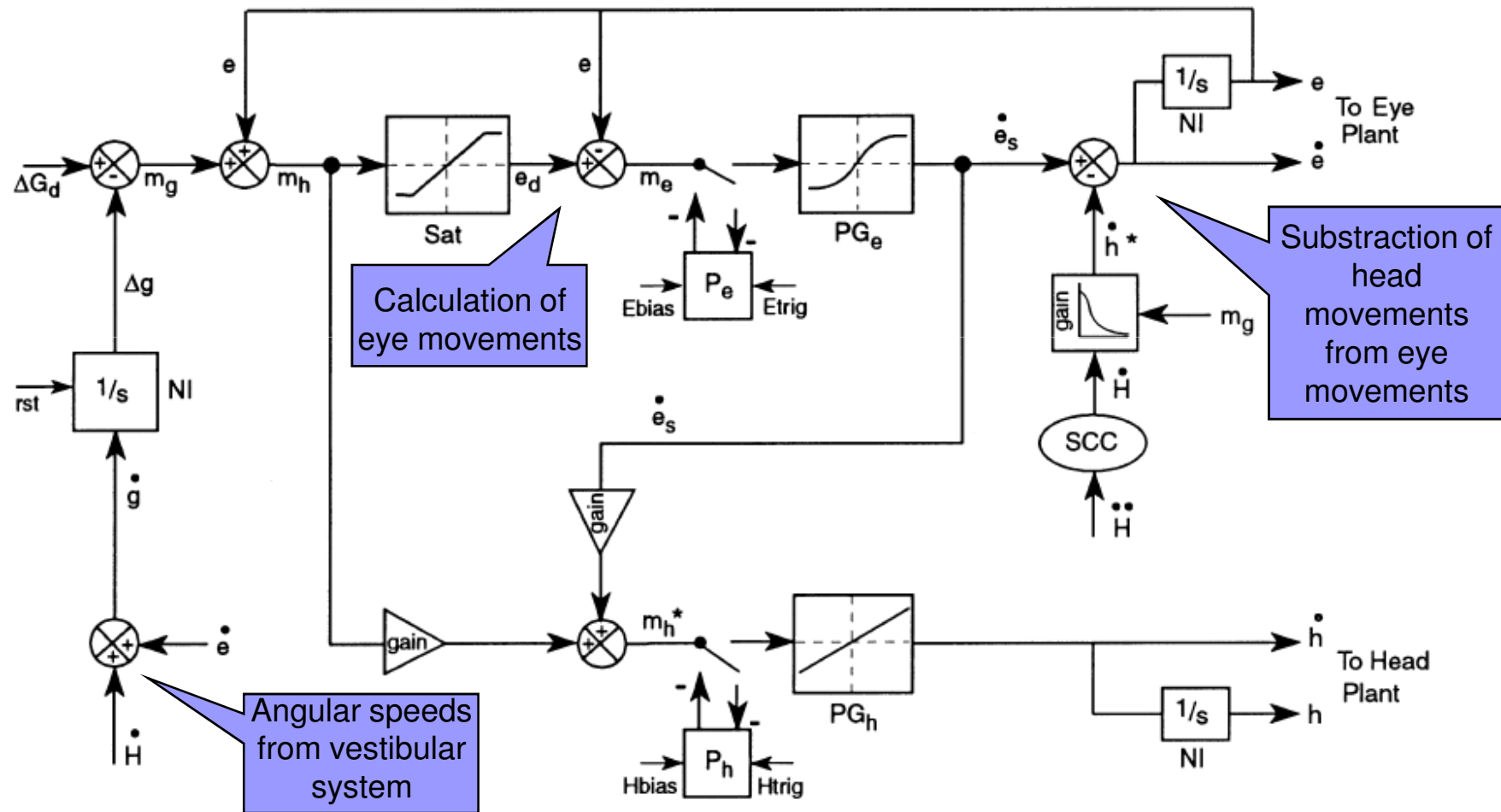




Biological Models

- *Laurutis V.P. and Robinson D.A., "The vestibulo-ocular reflex during human saccadic eye movements.", J Physiol (Lond) 1986, Vol. 373, pp. 209–233*
- *Lefèvre P. et al, "Experimental study and modeling of vestibulo-ocular reflex modulation during large shifts of gaze in humans.", Exp Brain Res 1992, Vol. 91, pp. 496–508*
- *Guitton D. and Volle M., "Gaze control in humans: eye-head coordination during orienting movements to targets within and beyond the oculomotor range.", J Neurophysiol 1987, Vol. 58, pp.427–459*
- *Guitton D., "Control of eye-head coordination during orienting gaze shifts.", Trends Neurosci 1992, Vol.15, pp.174–179*
- *Goossens H.H. and Van Opstal A.J., "Human eye-head coordination in two dimensions under different sensorimotor conditions", Exp. Brain Res.1997, Vol. 114, pp. 542–560*
- *Suzuki T. and Hirai N., "Reaction times of head movements occurring in association with express saccades during human gaze shifts", Neuroscience Letters 1998, Vol. 254, pp. 61–64*
- *Tweed D. et al, "Eye-head coordination during large gaze shifts.", J Neurophysiol 1995, Vol. 73, pp. 766–779*

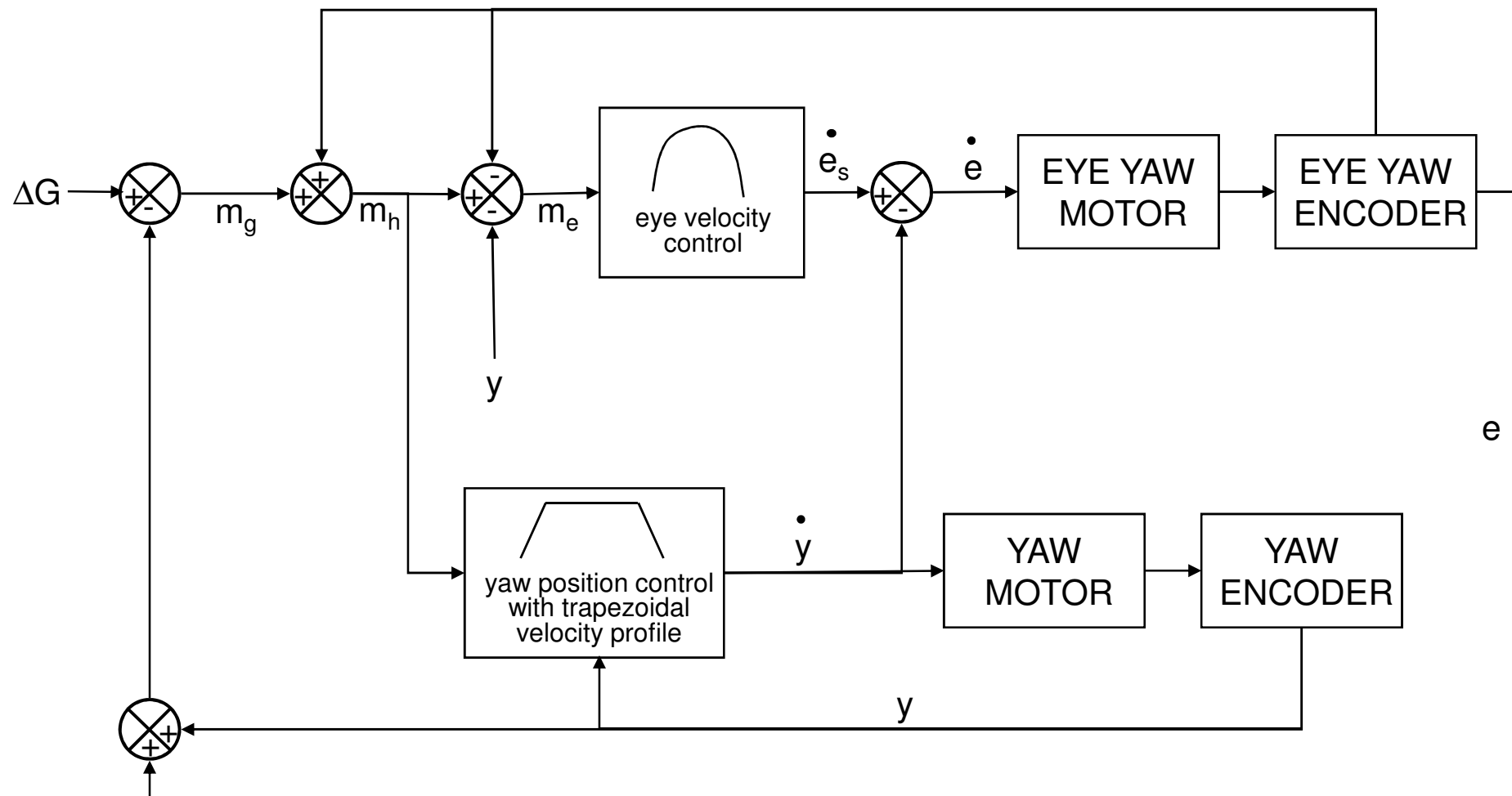
The Biological Model



Goossens H.H. and Van Opstal A.J., "Human eye-head coordination in two dimensions under different sensorimotor conditions", *Exp. Brain Res.* 1997, Vol. 114, pp. 542–560

The Artificial Model for horizontal motion

e





Experimental trials

Objectives:

1. To investigate the effectiveness of the bio-inspired paradigm to achieve an appropriate control of a multi-DOF robotic head
2. To verify if the proposed control paradigm is able to produce a motor output coherent with the reported patterns of eye-head coordination in humans



Experimental Methodology

In accordance to Goossens & Van Opstal two kinds of experiments were done:

- ***aligned experiments***: eyes and head of the robot were aligned at a straight-ahead position.
- ***not aligned experiments***: the eyes of the robot were randomly deviated from the straight-ahead position with an initial deviation in the range 20° - 55° .
- from the starting position the gaze shift was presented to the robot and the resulting displacements and timing of eye-head movements were recorded through the proprioceptive sensors of the head



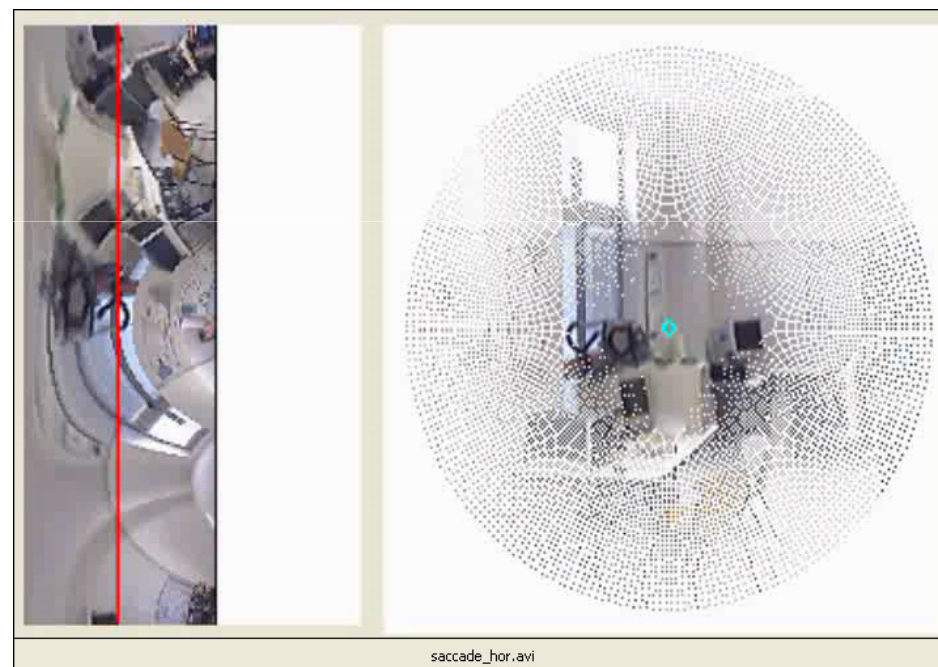
Experimental Trials

- **Aligned** Experiment:
horizontal saccadic movements of 45° on horizontal and vertical direction.
- **Not Aligned** experiment:
saccadic movement of 45° amplitude with an initial deviation of 25° in the same direction. Required movement for the eye was 20° whereas the head moved 45°
- 10 trials for each experiment
- Maximum velocities have been set to 400 deg/s and 300 deg/s respectively for eye yaw motion and eye tilt motion

Experimental Trials: horizontal saccades



Left eye only

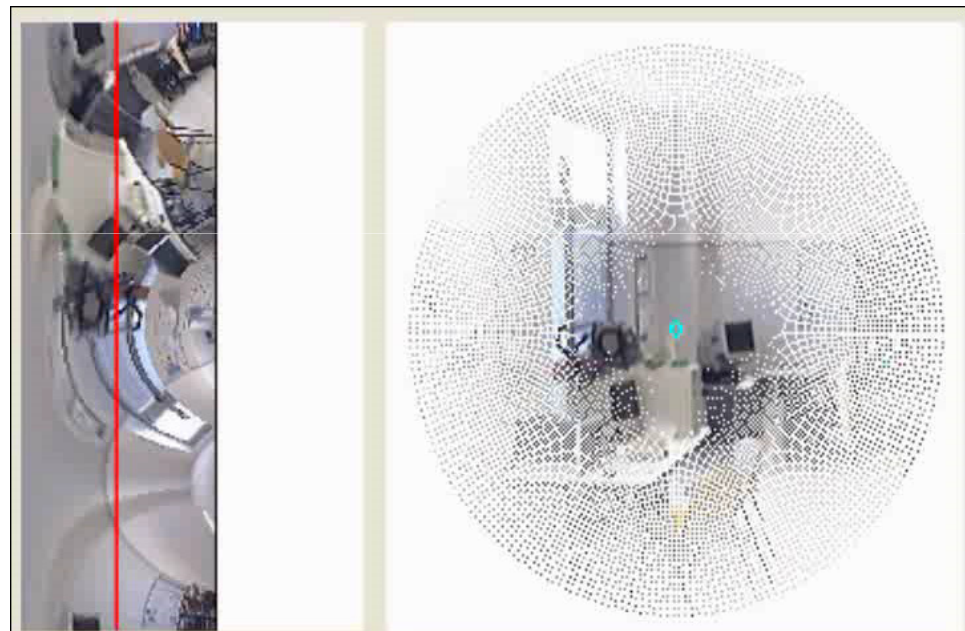


Camera View

Experimental Trials: vertical saccades



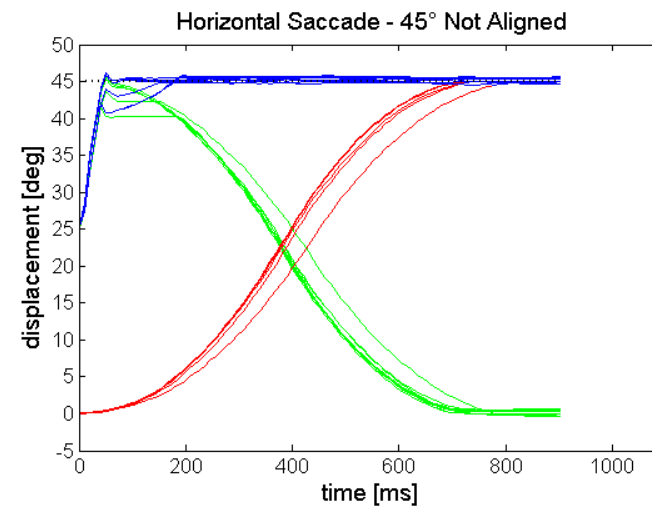
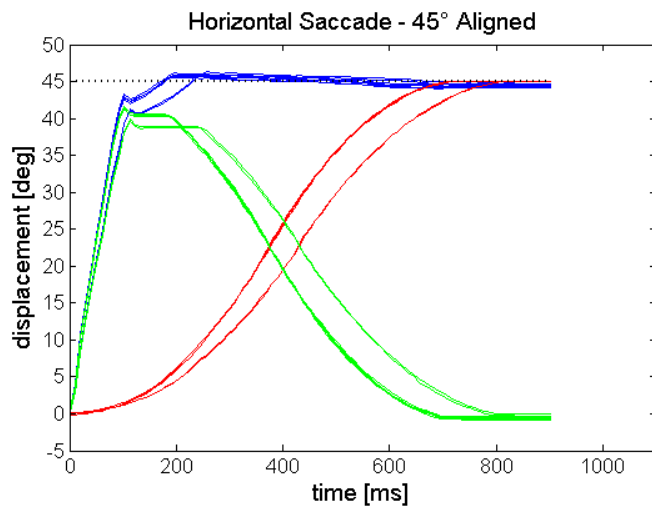
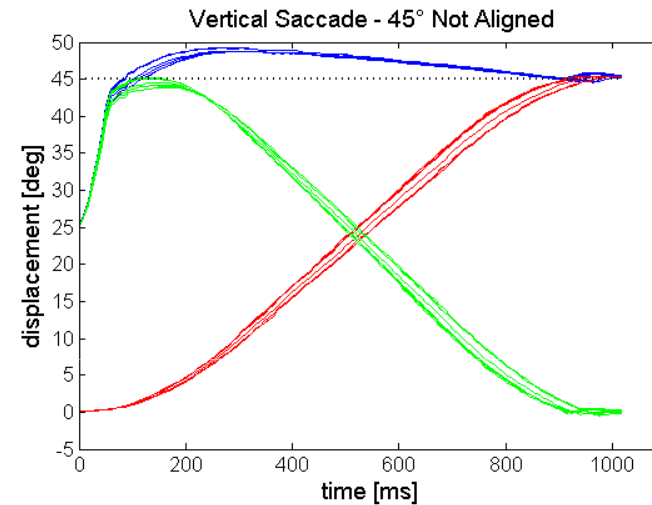
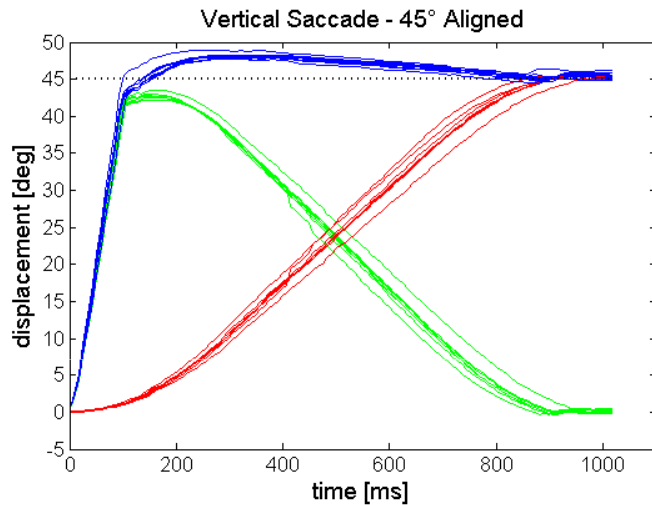
ext_saccade_ver.avi



saccade_ver.avi

Camera View

Experimental Results



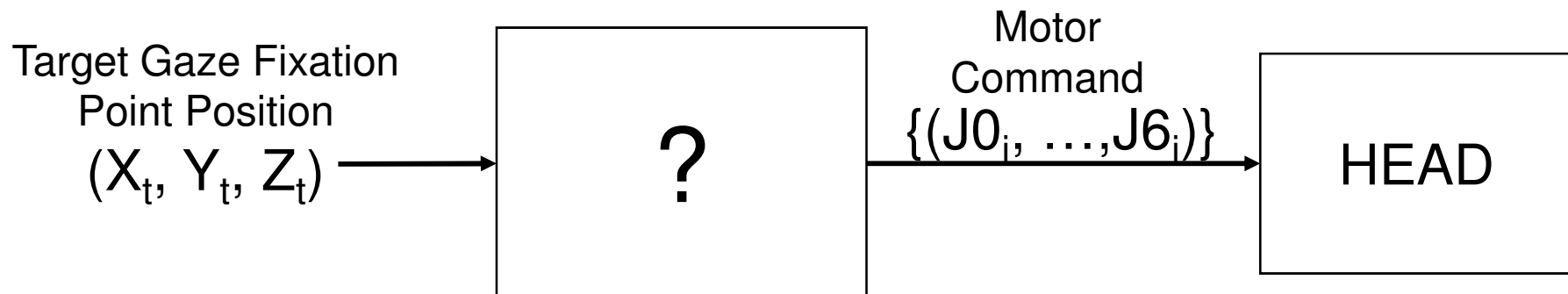
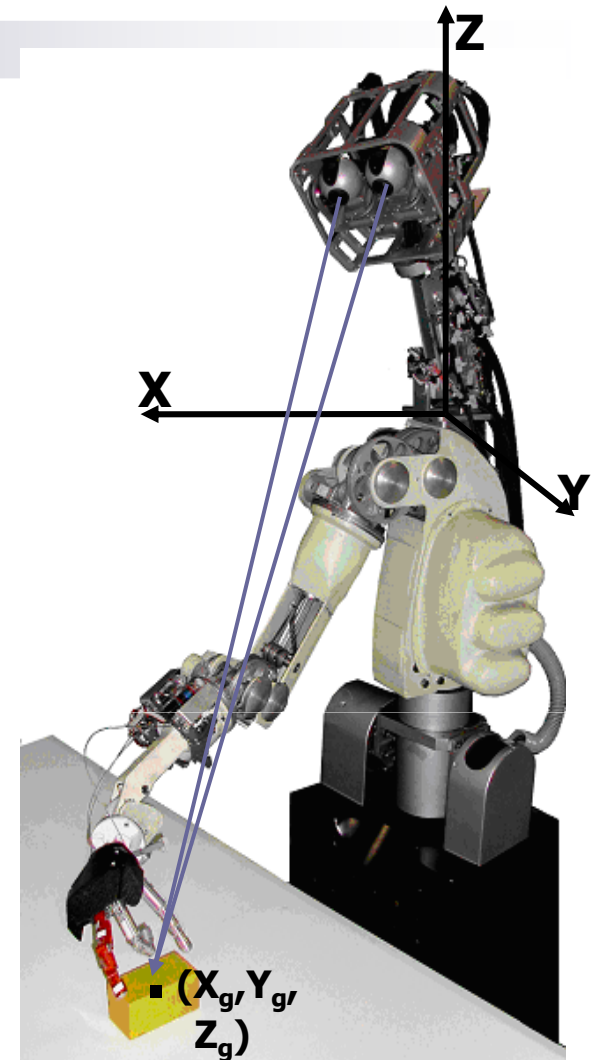
Solution 3

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

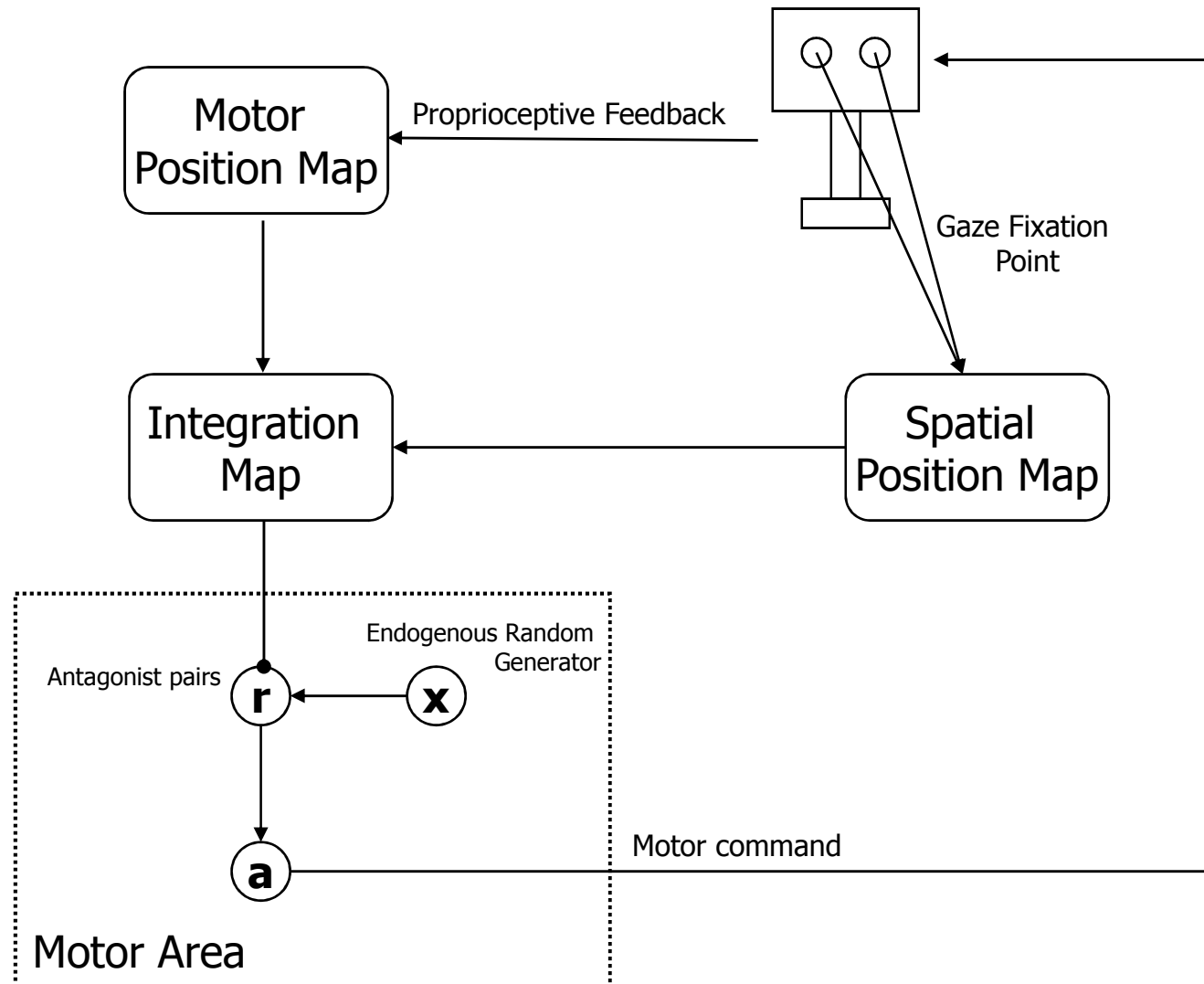
G. Asuni, G. Teti, C. Laschi, E. Guglielmelli, P. Dario, "A Robotic Head Neuro-controller Based on Biologically-Inspired Neural Models", *IEEE International Conference on Robotics and Automation – ICRA 2005*, Barcelona, Spain, April 18-22, 2005, pp.2373-2378.

Addressed Problem

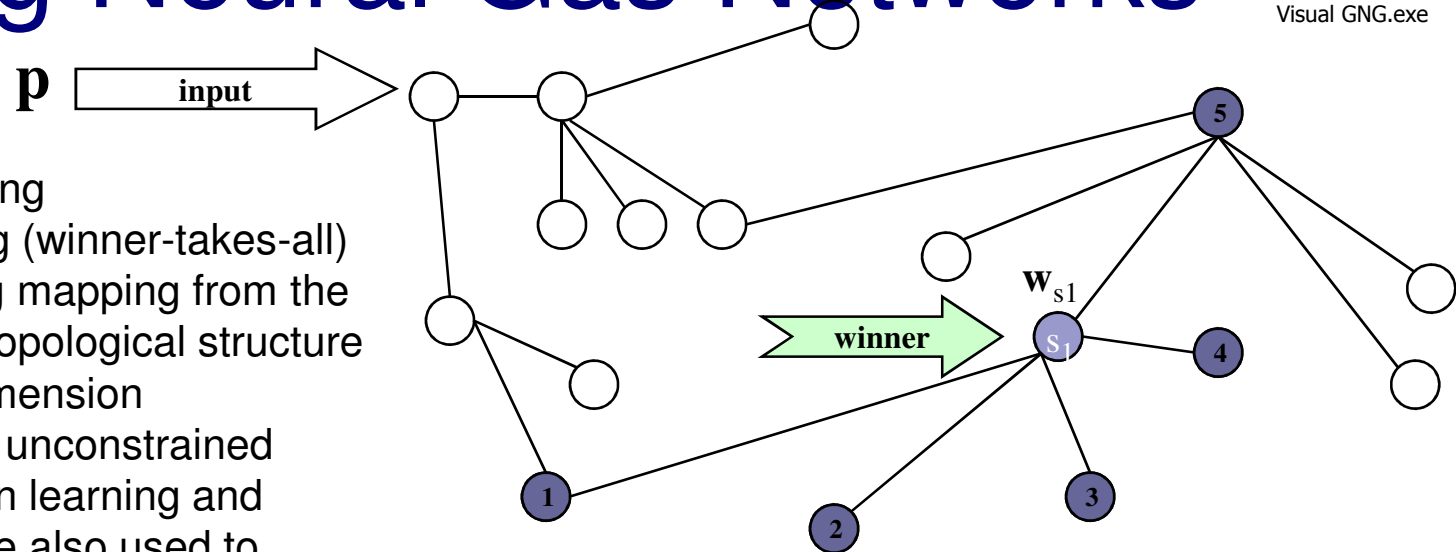
To develop a control module that receives in input a target gaze position and provides in output a command sequence able to reach it



The proposed neural model



Implementation tools: Growing Neural Gas Networks



- 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 interrupted 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)

$$N_{s_1} = \{ \textcircled{1} \textcircled{2} \textcircled{3} \textcircled{4} \textcircled{5} \}$$

Updating rules:

$$\begin{aligned} \mathbf{w}_{s_1} &= \mathbf{w}_{s_1} + \mathcal{E}_b (\mathbf{p} - \mathbf{w}_{s_1}) \\ \mathbf{w}_i &= \mathbf{w}_i + \mathcal{E}_n (\mathbf{p} - \mathbf{w}_i) \quad (\forall i \in N_{s_1}) \end{aligned}$$

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

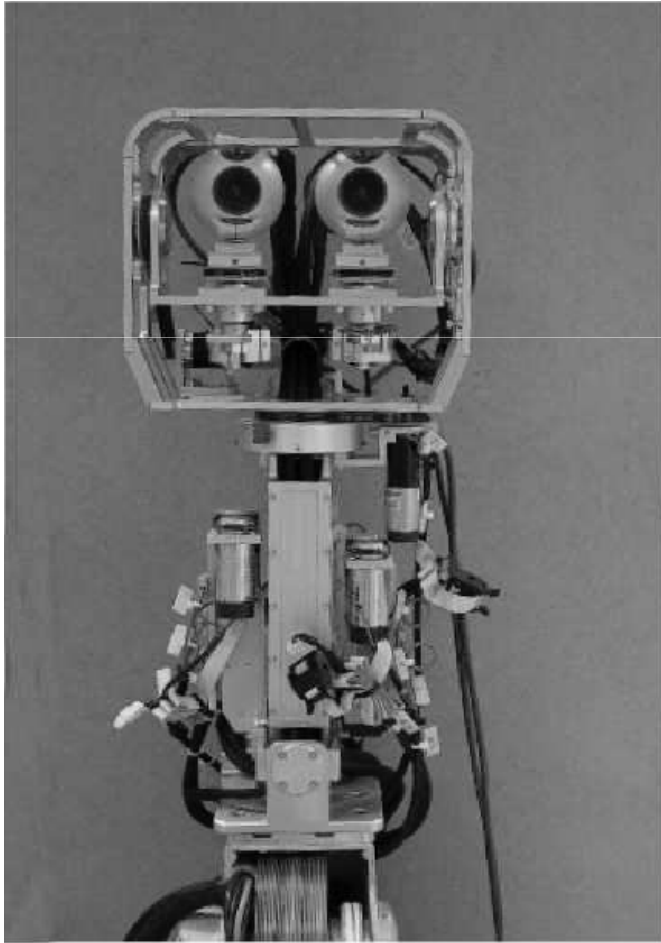


Testing phase

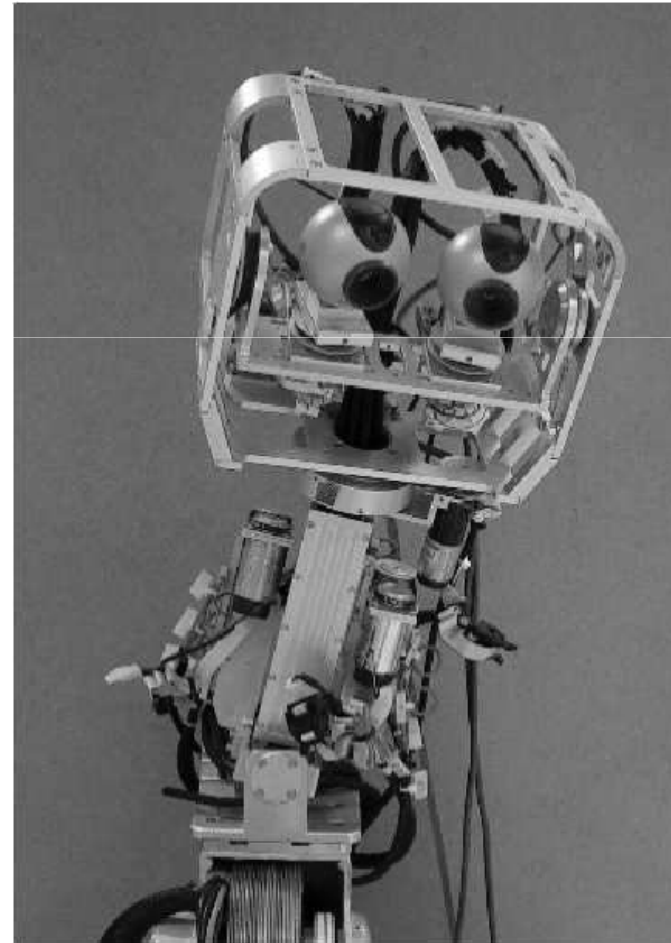
- After the training phase, given a target fixation point the system provides the joint rotations that drives the current gaze fixation point in the target point
- Three different modalities:
 1. Normal (without any constraint)
 2. With a clamped joint 0
 3. With symmetric angles for eye joints

All trials have been executed without additional learning

Experimental results: normal gazing



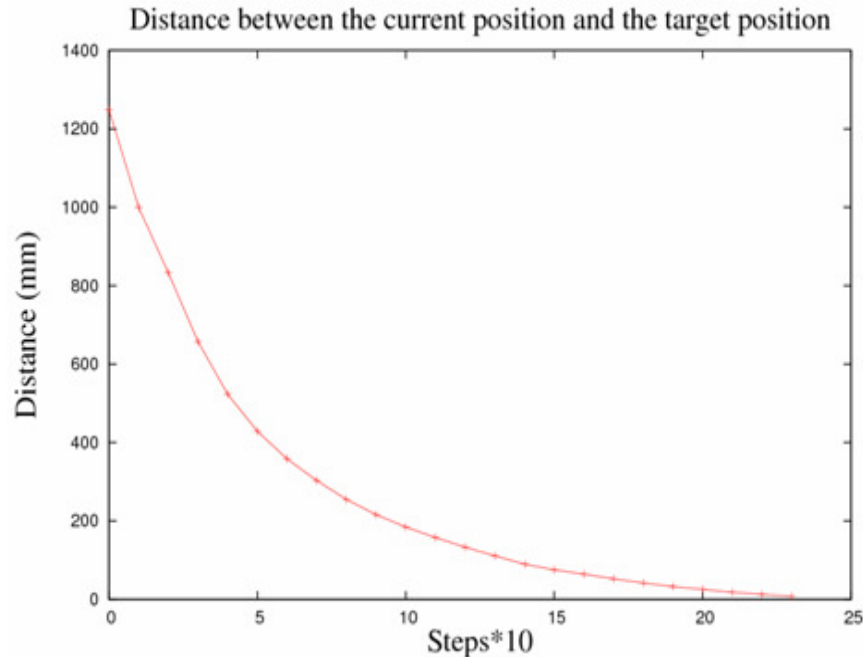
Initial posture



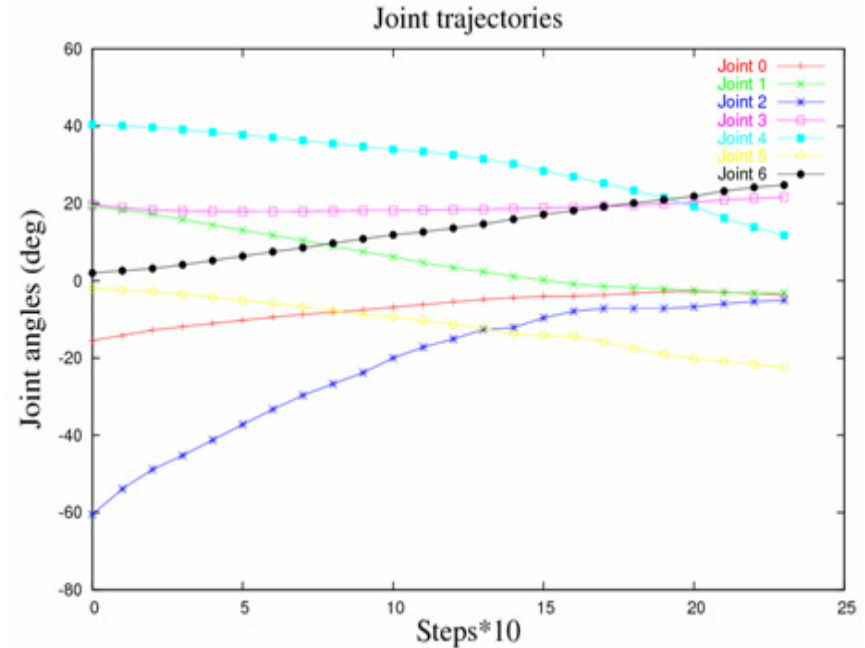
Final posture (normal)

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

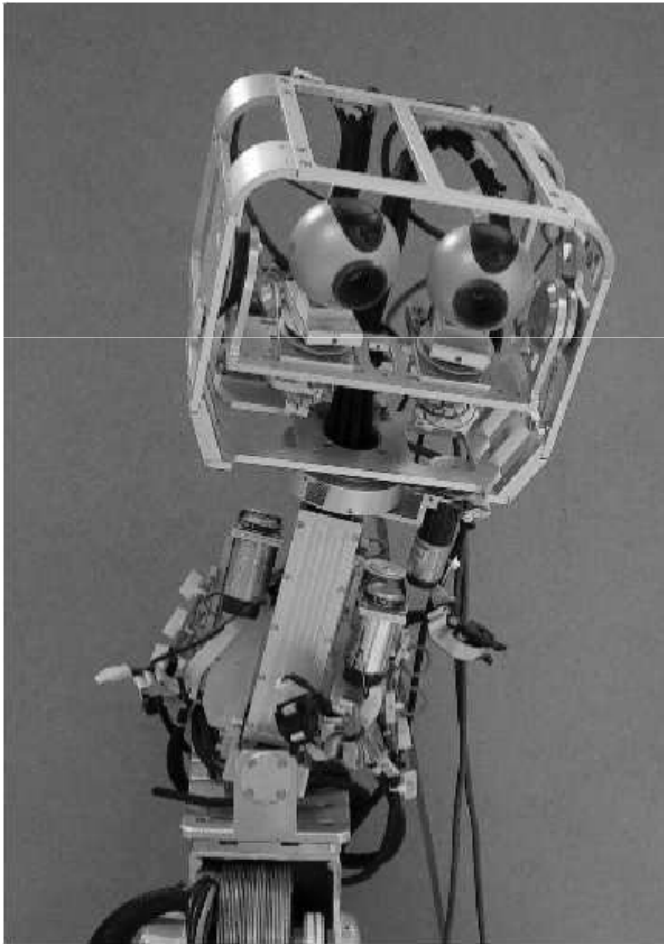
Distance between the current gaze
fixation point and the target:
monotonic trend



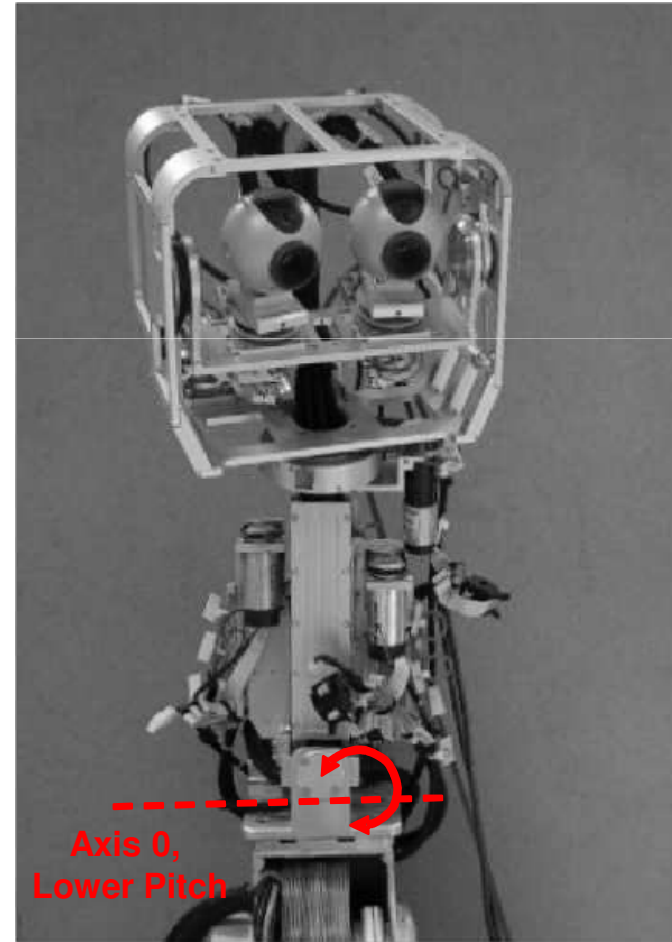
Joint trajectory



Experimental results: gazing with a clamped joint

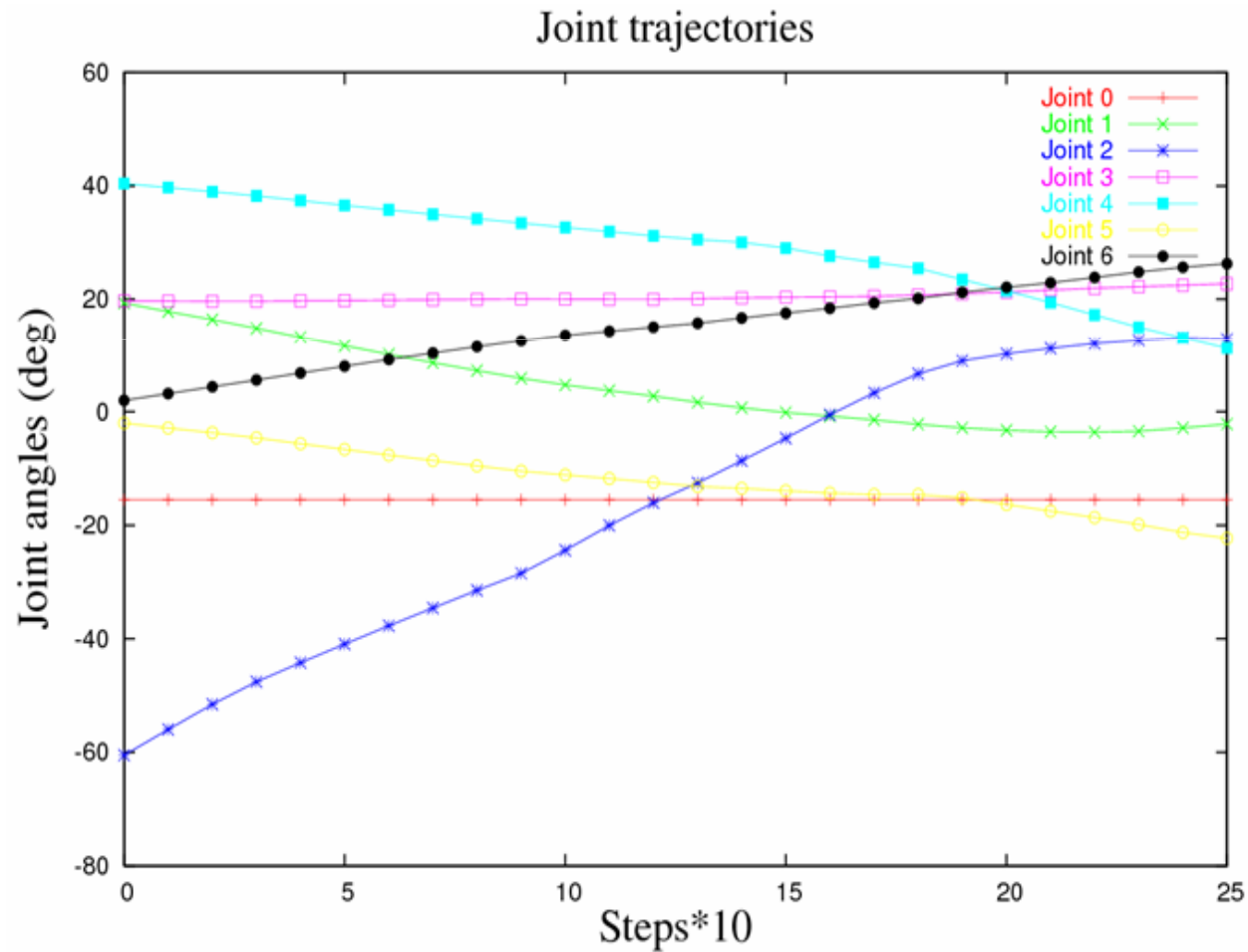


Final posture in normal mode



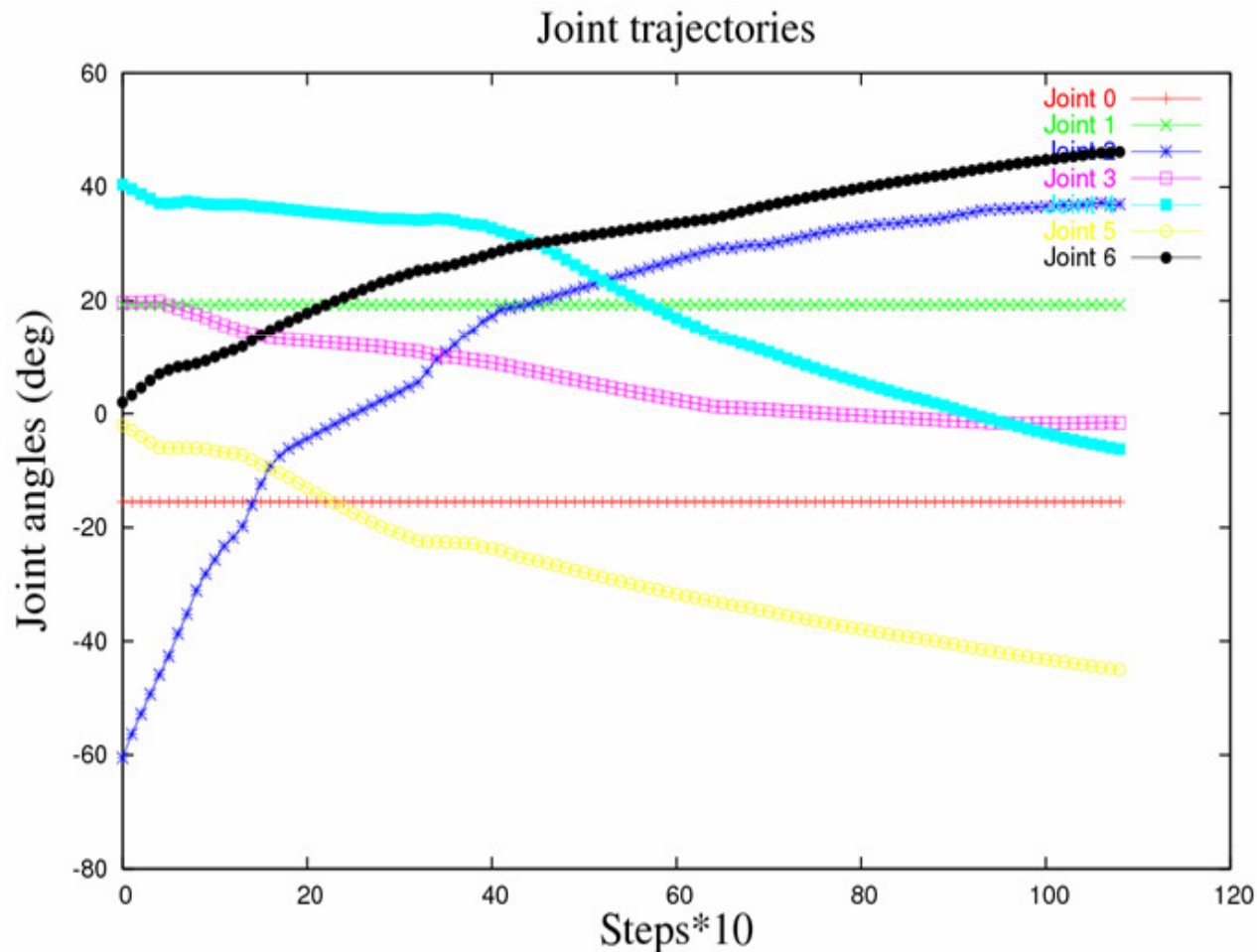
Final posture (clamped joint 0)

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



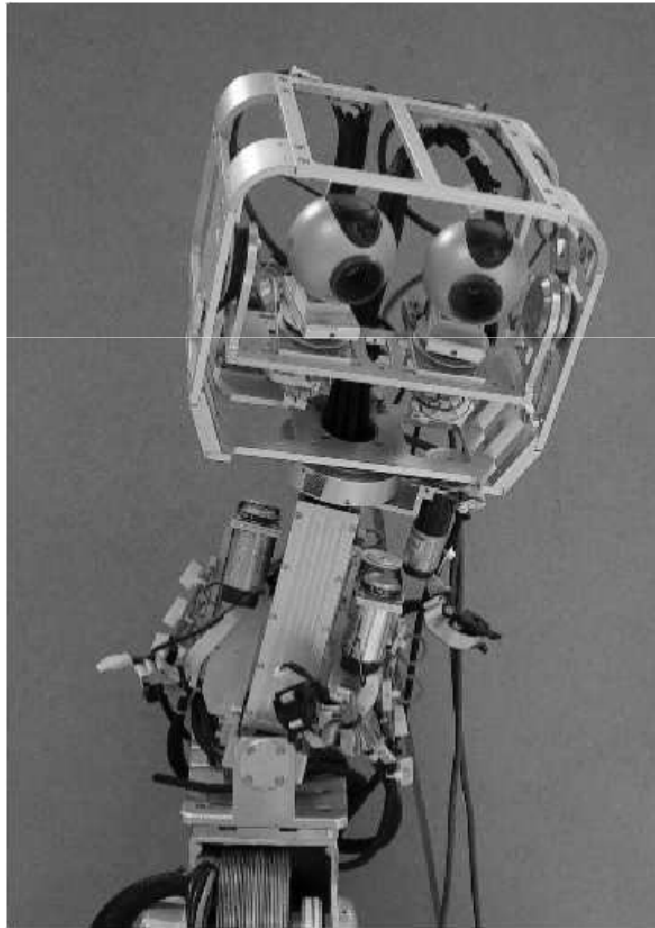
Joint trajectory: clamped joint 0

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

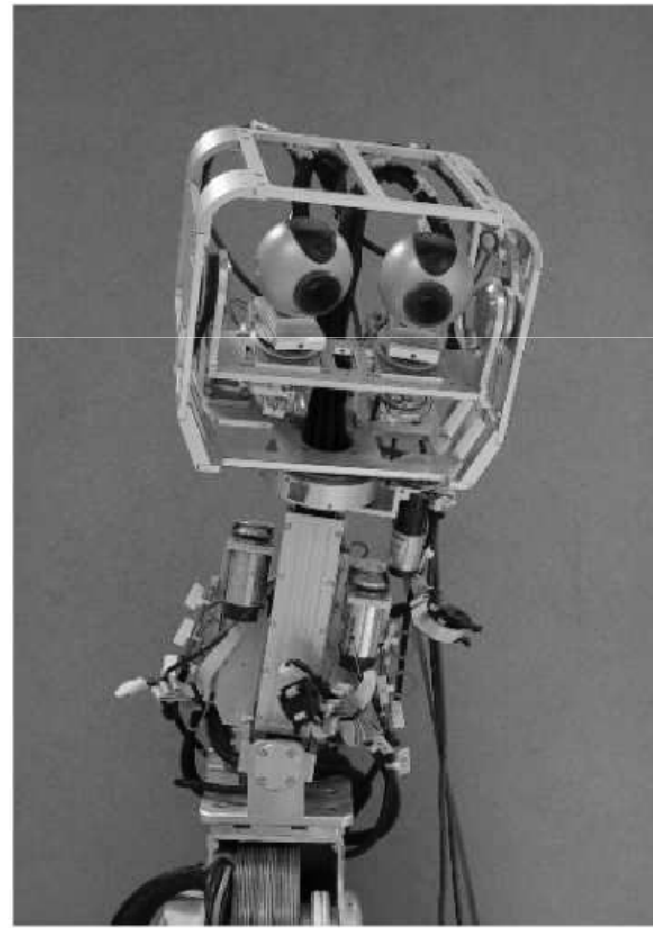


Joint trajectory: clamped joint 0 and joint 1

Experimental results: gazing with symmetric eye angles

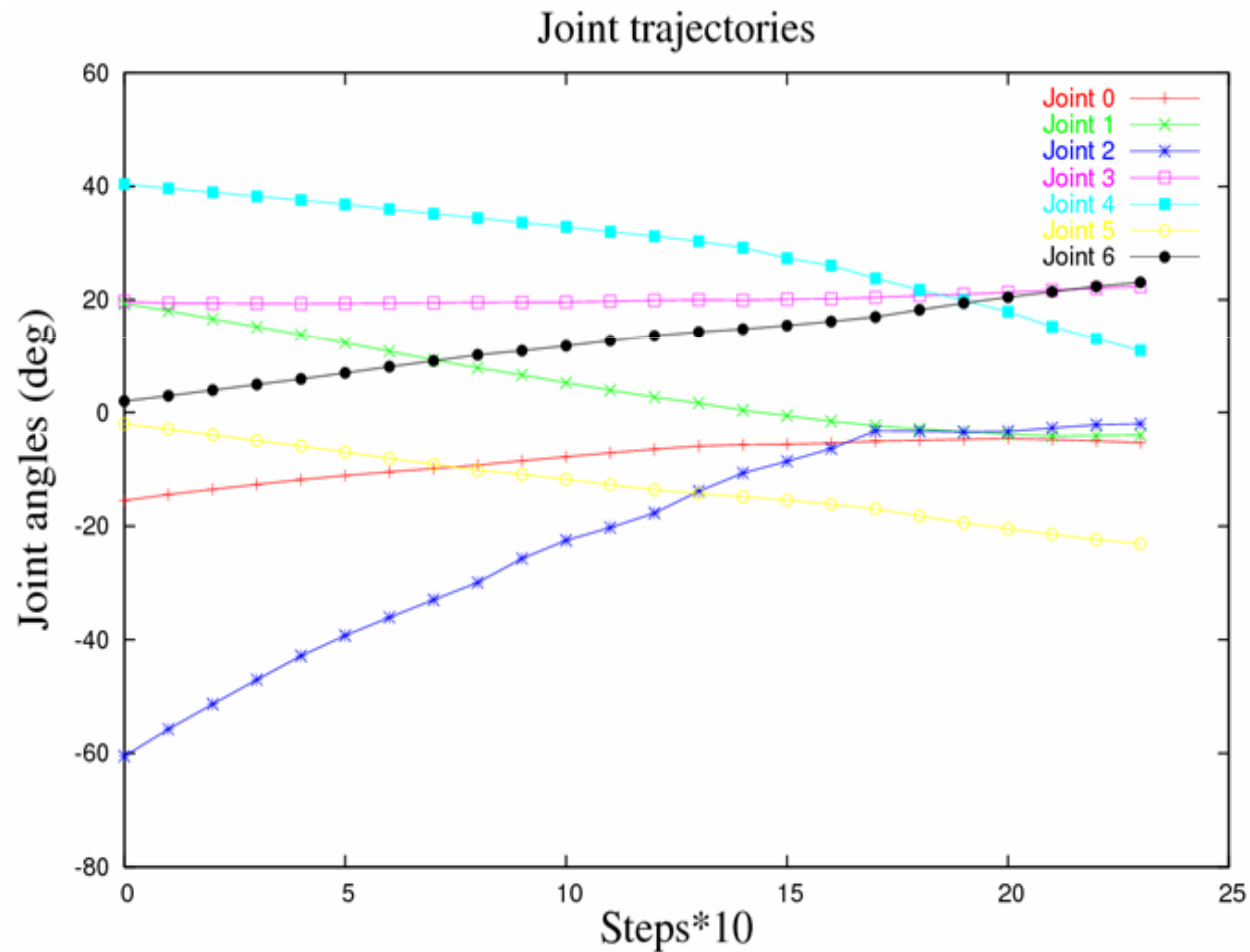


Final posture in normal mode



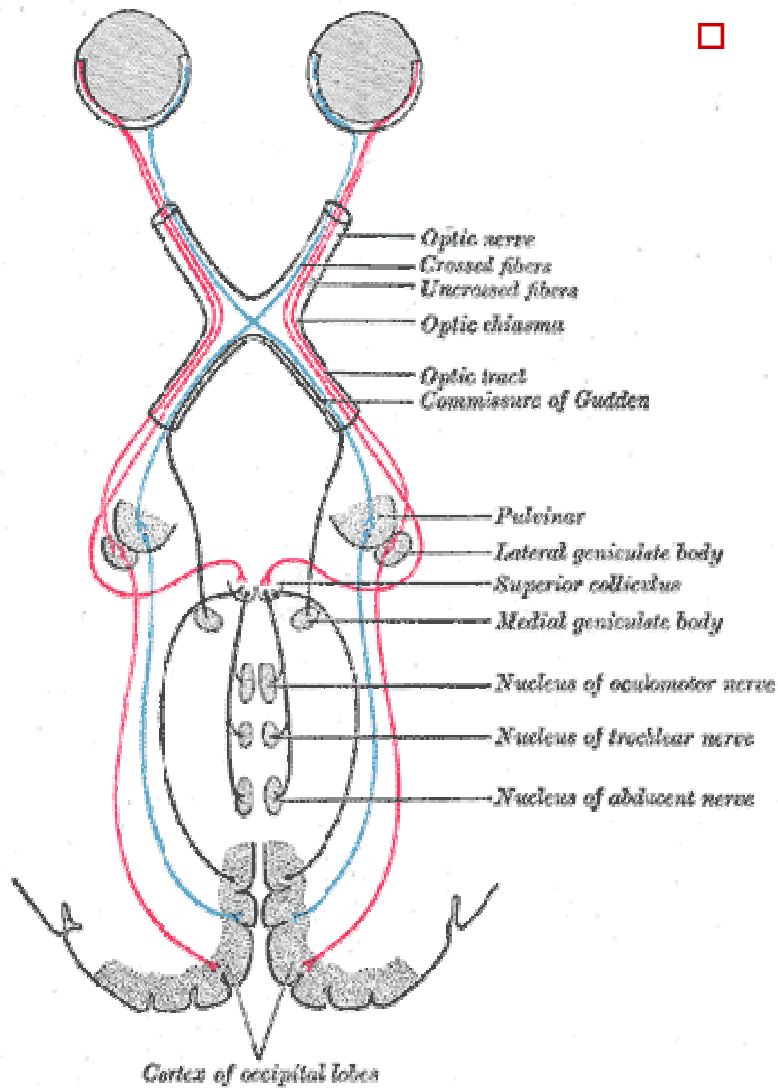
Final posture with symmetric
angles for eye joints

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

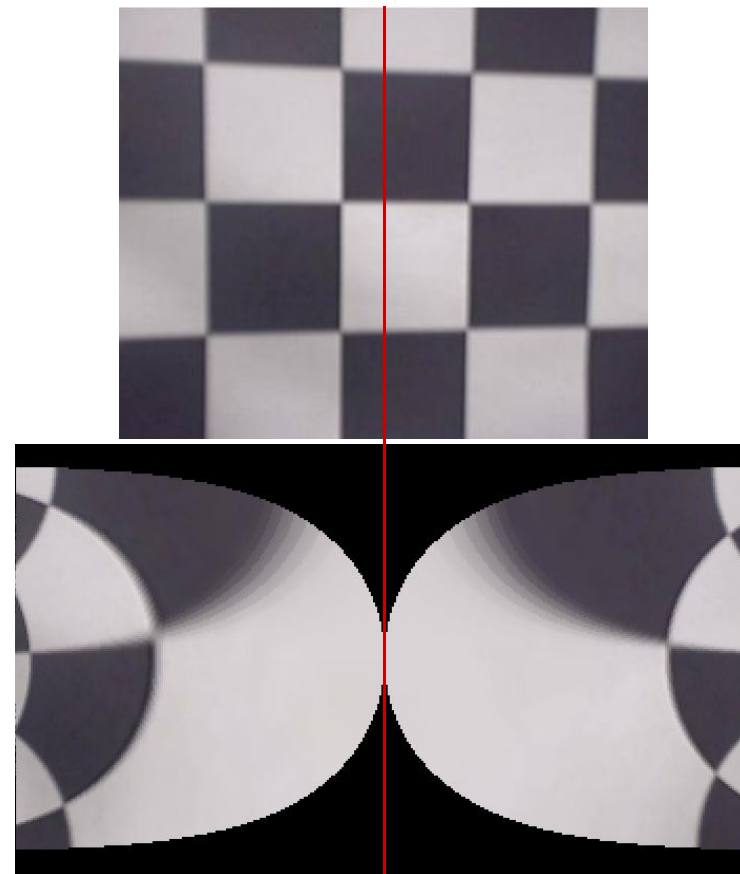


Joint trajectory: symmetric angles for eye joints (vergence)

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



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

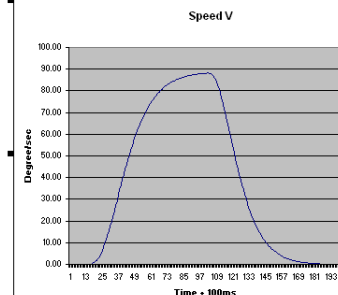
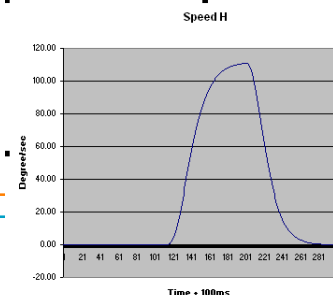
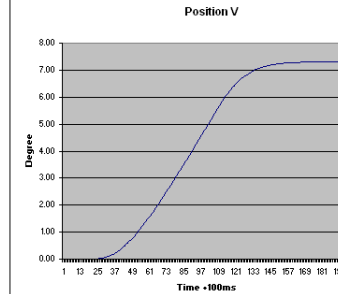
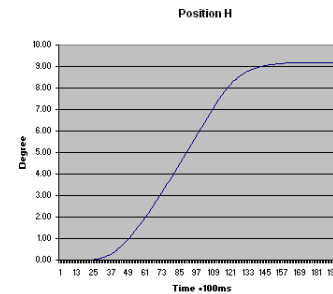
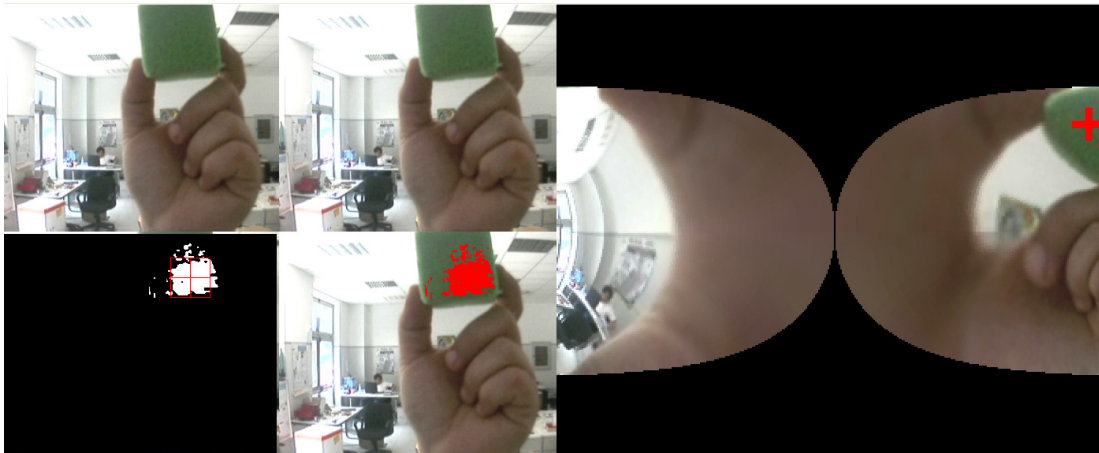
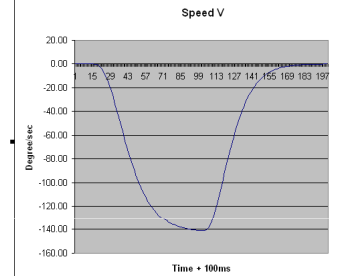
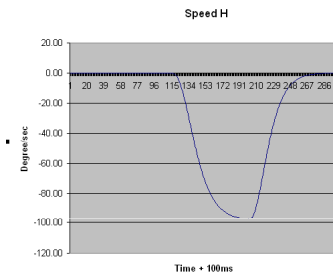
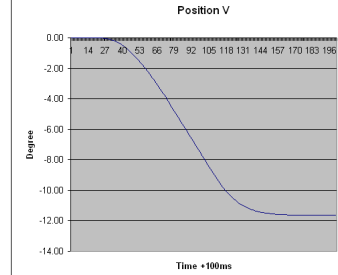
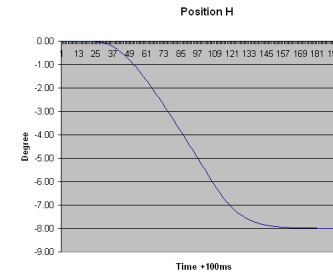
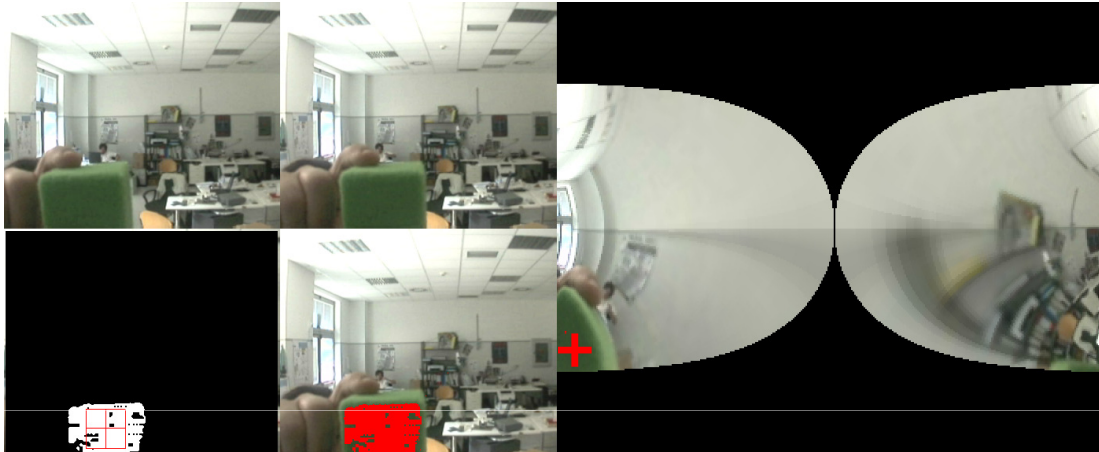
- Implementation of the mapping from the polar coordinates in visual space to the superior colliculus coordinates system, according to the model
- Generation of saccade movements:
 - A stimulus of a given colour can be detected in the map and the coordinates calculated in the superior colliculus, in 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)

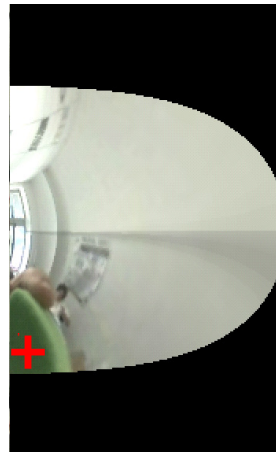
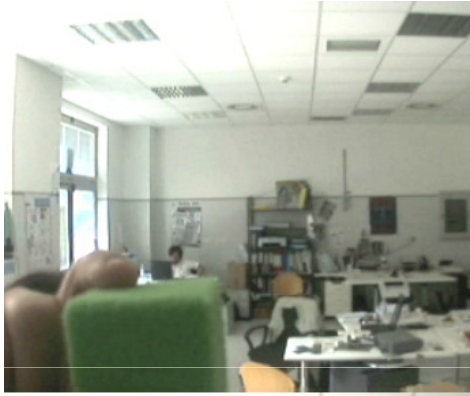


Generation of saccade movements



Generation of saccade movements

Stimulus #1



Stimulus #2



Saccades executed by the right eye