Gaze Control and Manipulation

Giulio Sandini
Giorgio Metta
Lorenzo Natale
Francesco Orabona

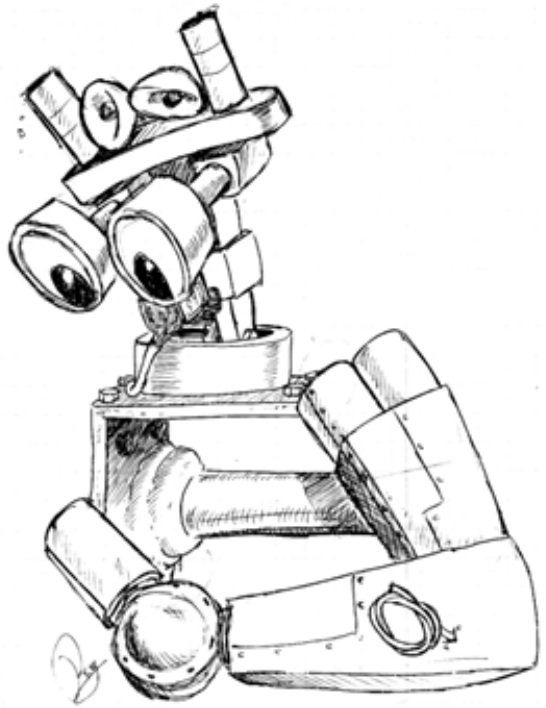
LIRA-Lab, DIST
University of Genova – Italy
http://www.liralab.it

Workshop on *Bulding Humanoid Heads*
Santa Monica,
November 10, 2004

The LIRA-Lab Team

Development and Robotics

Study the **development** of sensorimotor coordination and cognitive abilities in humans by building artificial systems (humanoids) that “grow”



It is the **goal** (understand cognition) but it is also the **means** (building complex systems)

Summary of talk

Development of oculomotor control

Role of gaze in the control of manipulation
(gaze control is more that just looking…)

Some new hardware

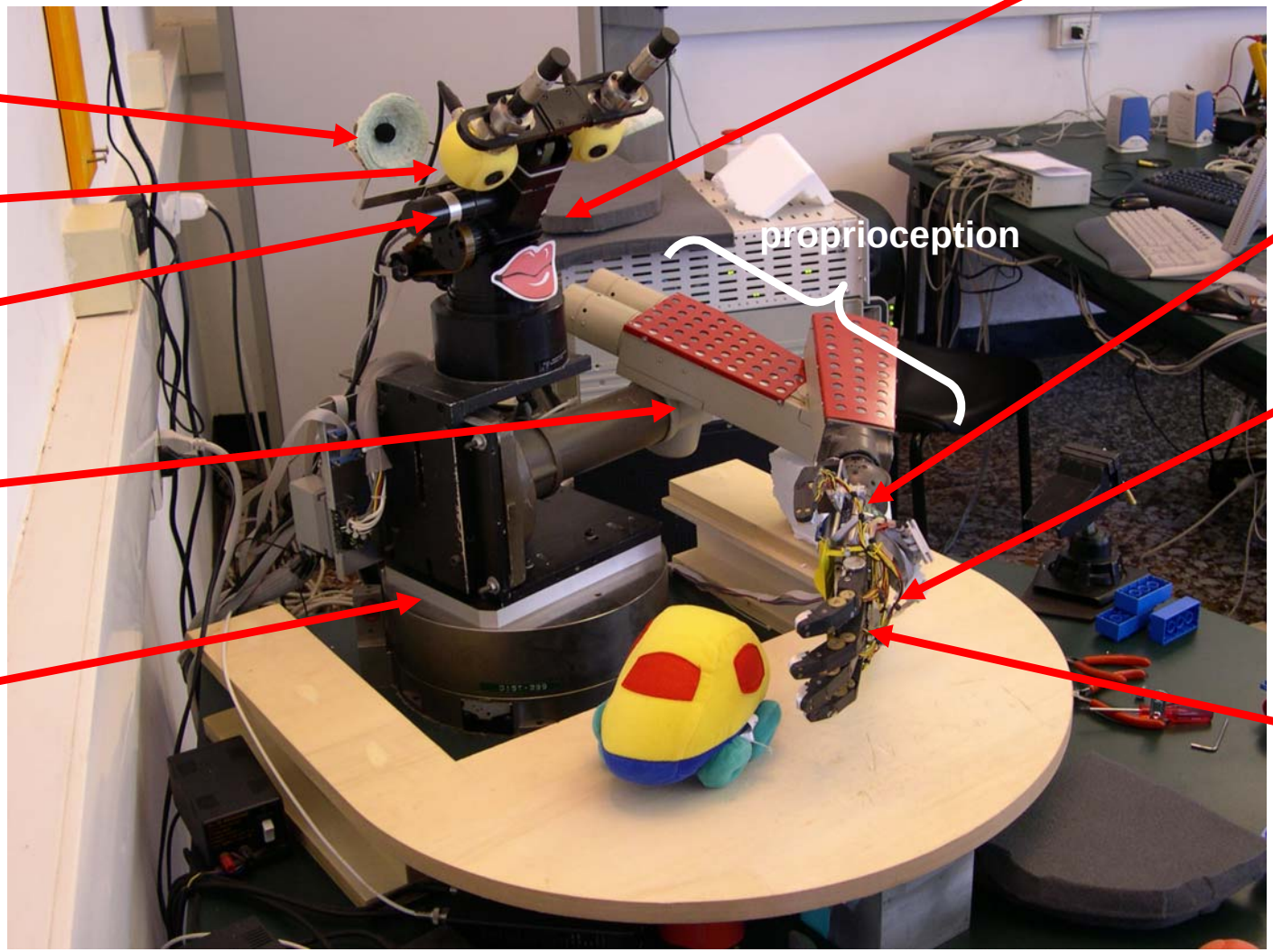
…heads…

4 d.o.f.

5 d.o.f.

8 d.o.f.

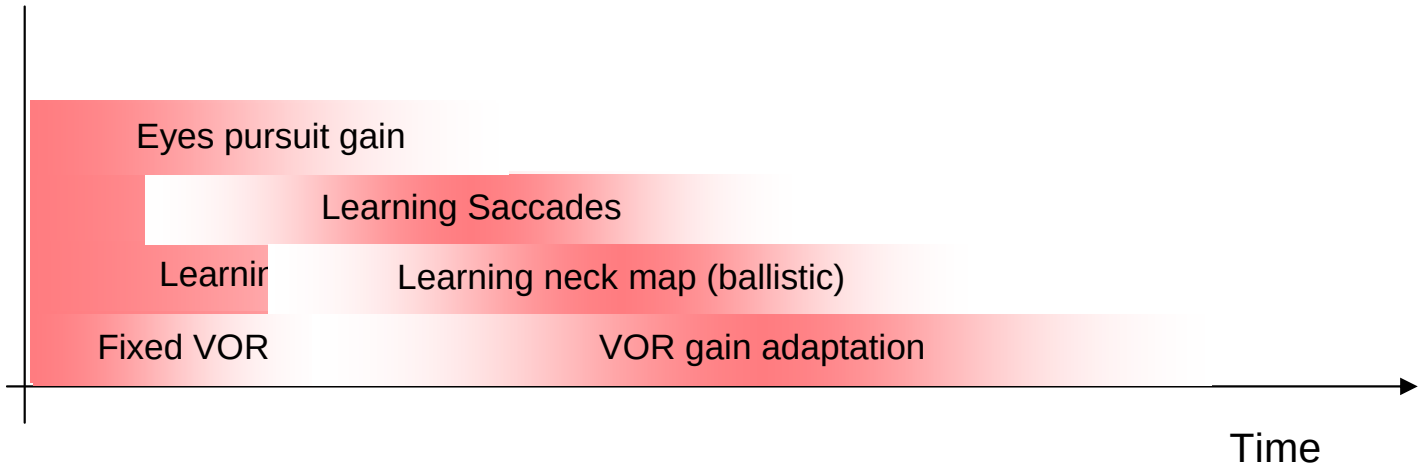
Babybot's body



Babybot

With Babybot we are studying the development of sensorimotor coordination and cognition (building an “adult” human-like being may be impossible).

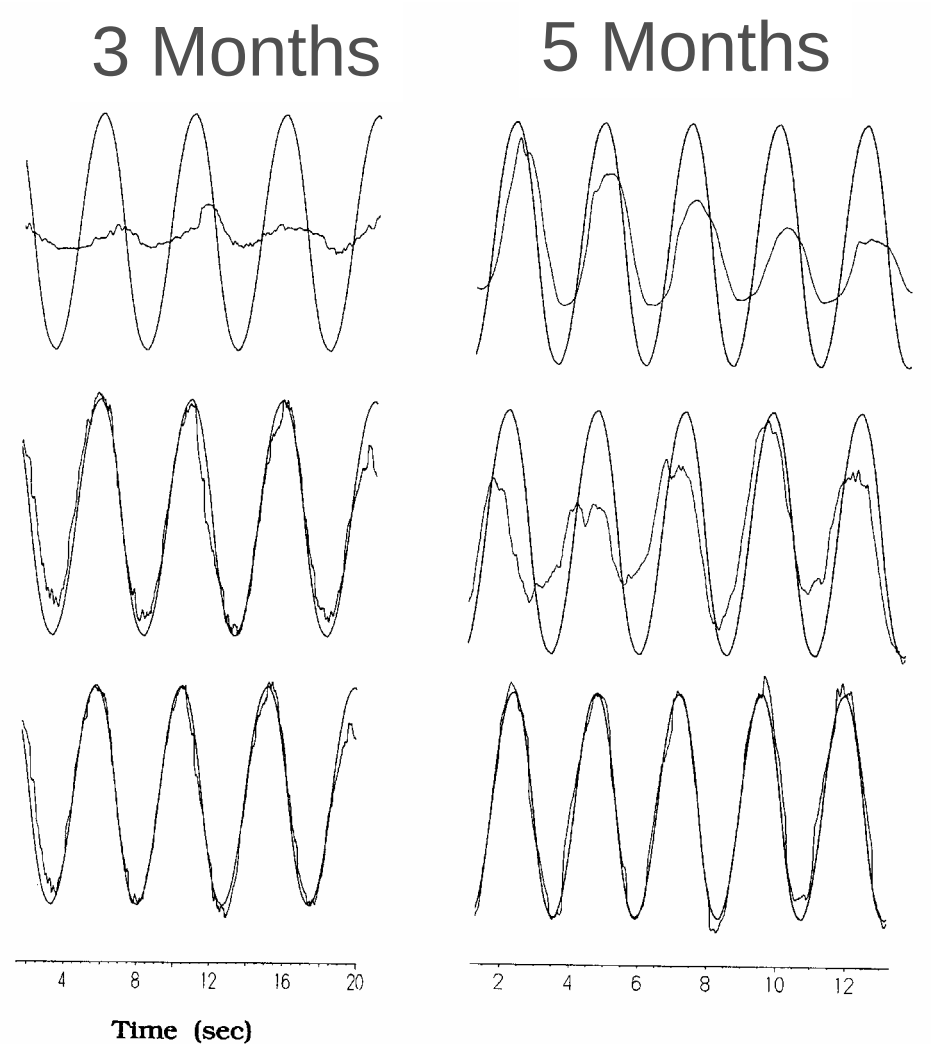
e.g.: Development of Eye-head Coordination



What is the sequence?
Why is a certain sequence better?
What moves the system from one stage to another?
Are there stages?

Babies Learning

(visual tracking)



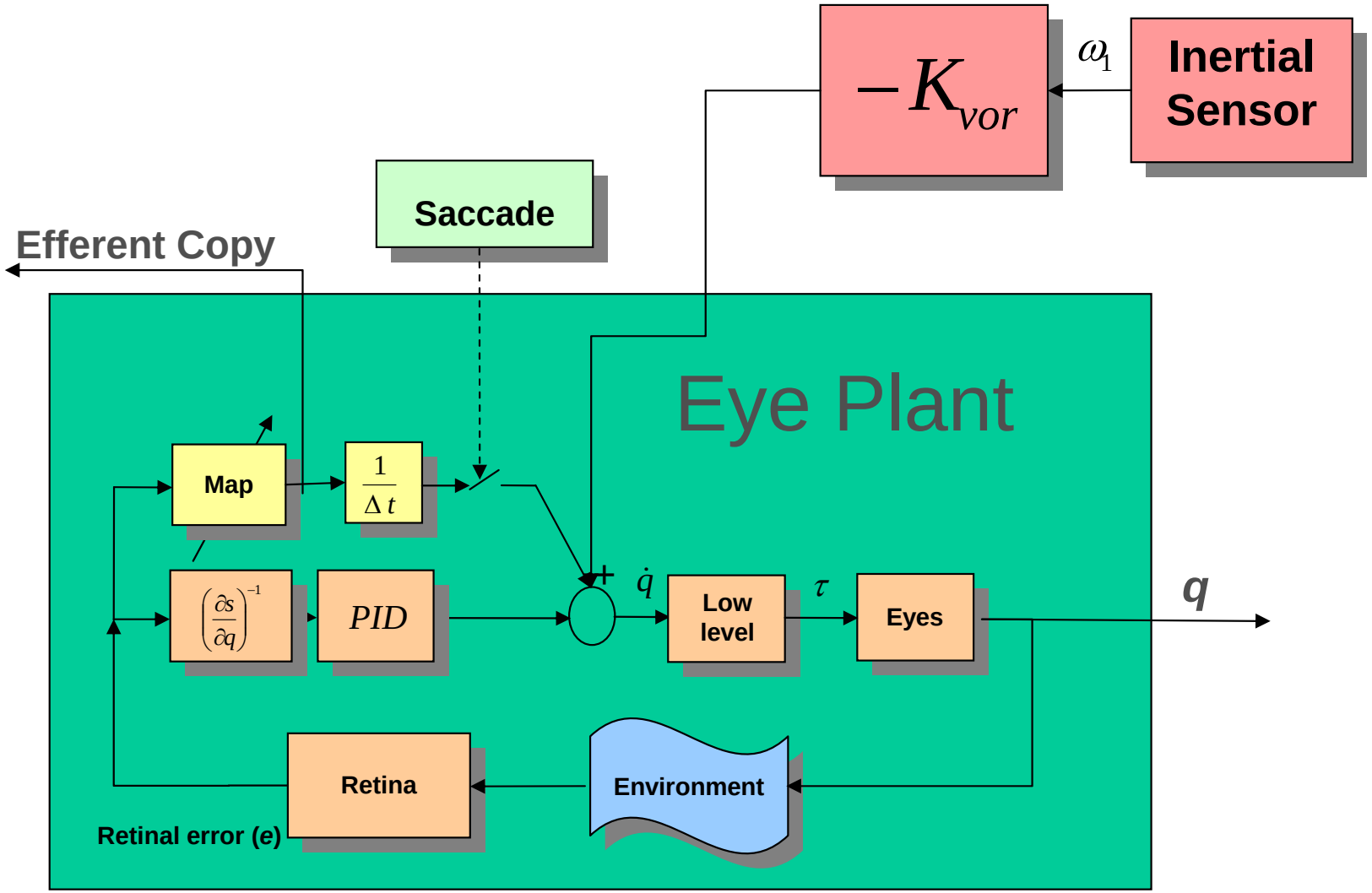
From: Von Hofsten, C. and K. Rosander, *Development of Smooth Pursuit Tracking in Young Infants.* Vision Research, 1997. **37**(13): p. 1799-1810.

Babybot's head

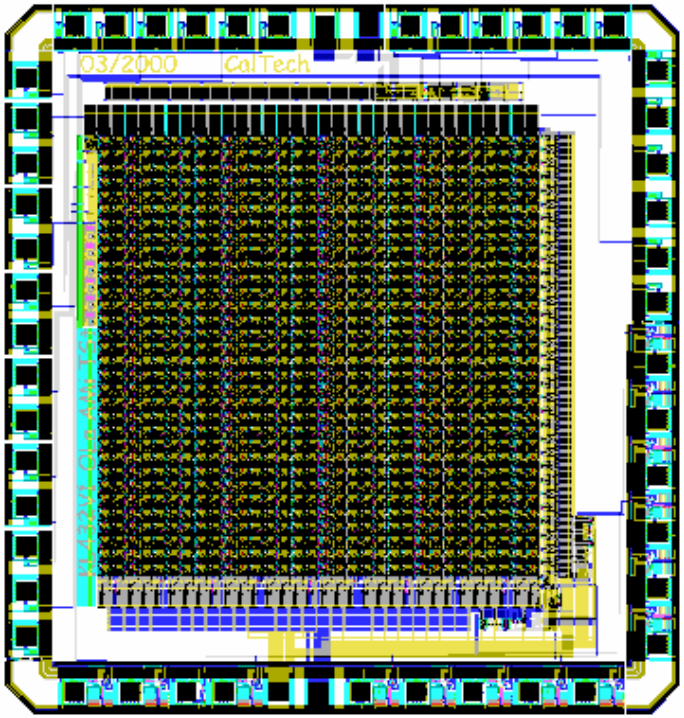
- 5 degrees of freedom

Visual Tracking
Stereo vision
Optical flow and image stabilization
Visuo-acustic integration
Visuo-inertial integration
Learning and development

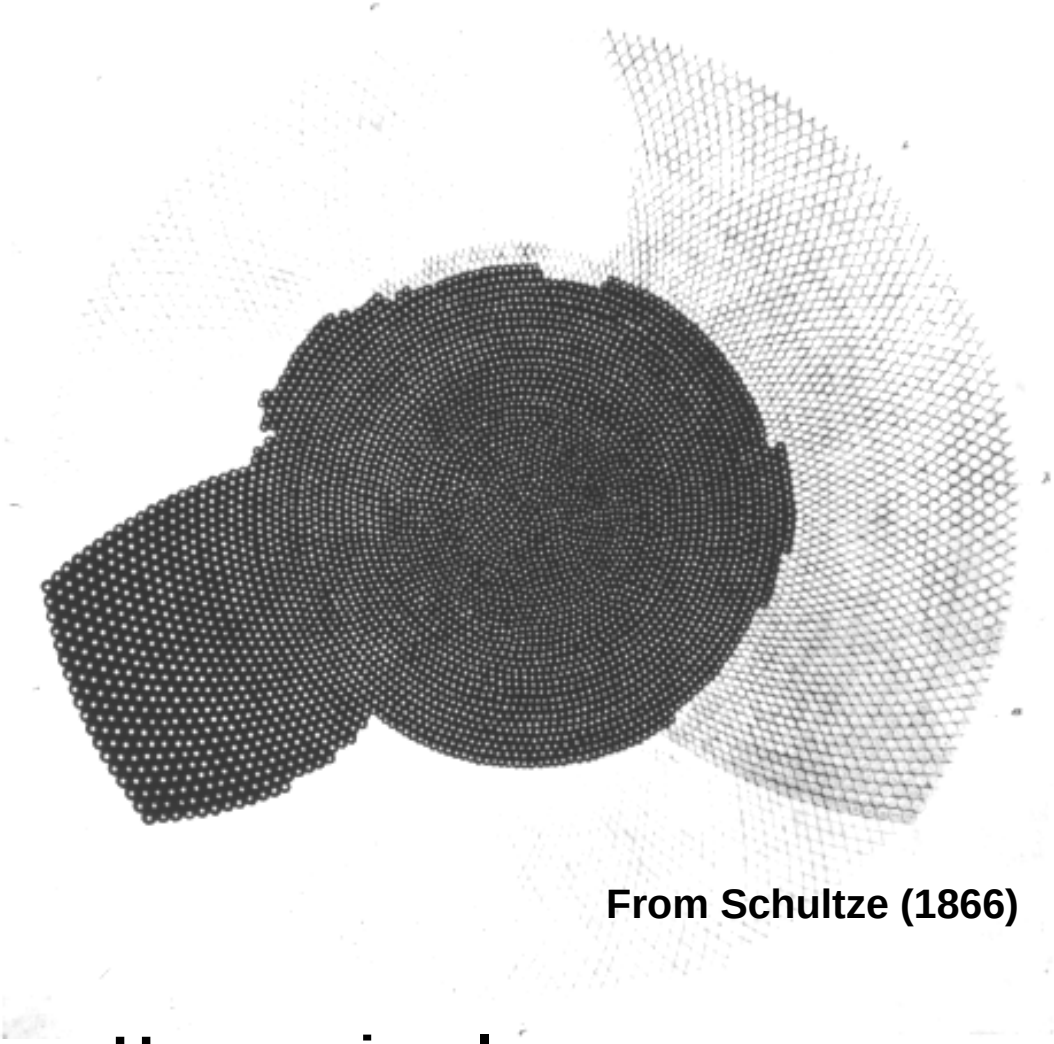
Eyes Plant


$-K_{vor}$
ω_1
Inertial Sensor
Saccade
Efferent Copy
Eye Plant
Map
$\frac{1}{\Delta t}$
$\left(\frac{\partial s}{\partial q}\right)^{-1}$
PID
+
$\dot{q}$
Low level
τ
Eyes
q
Retina
Environment
Retinal error (e)

Babybot Eyes….

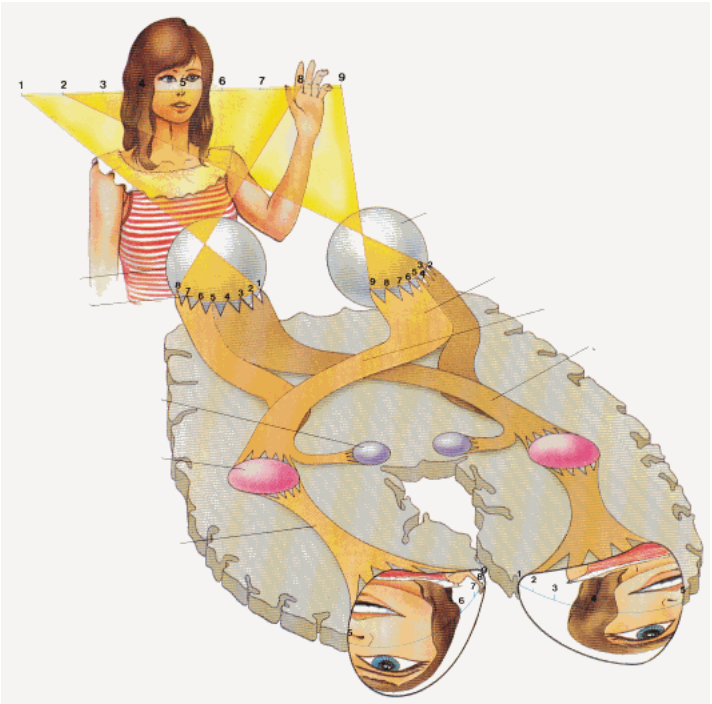


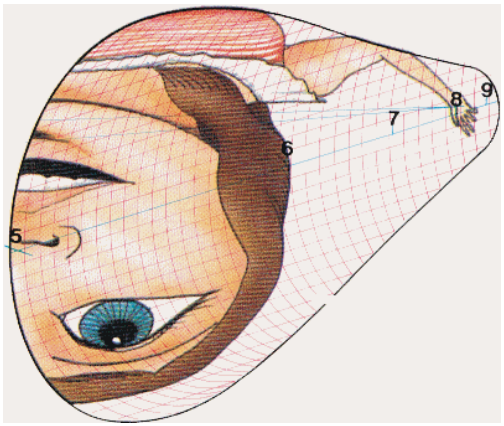
Conventional visual sensor

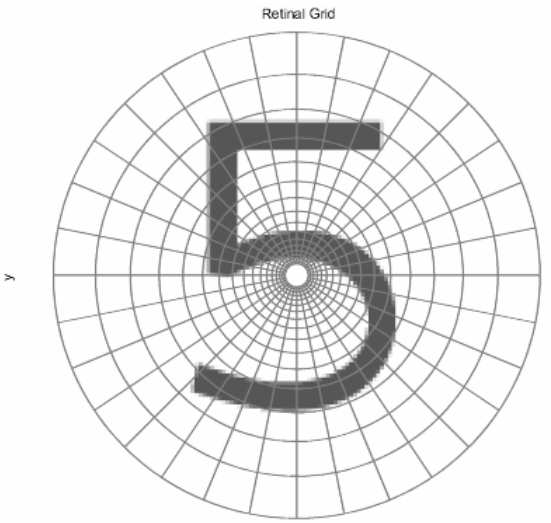


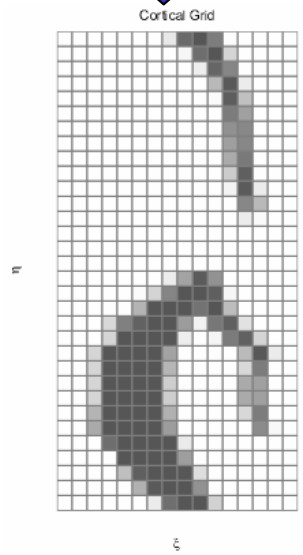
From Schultze (1866)

Human visual sensor (1-2 deg)



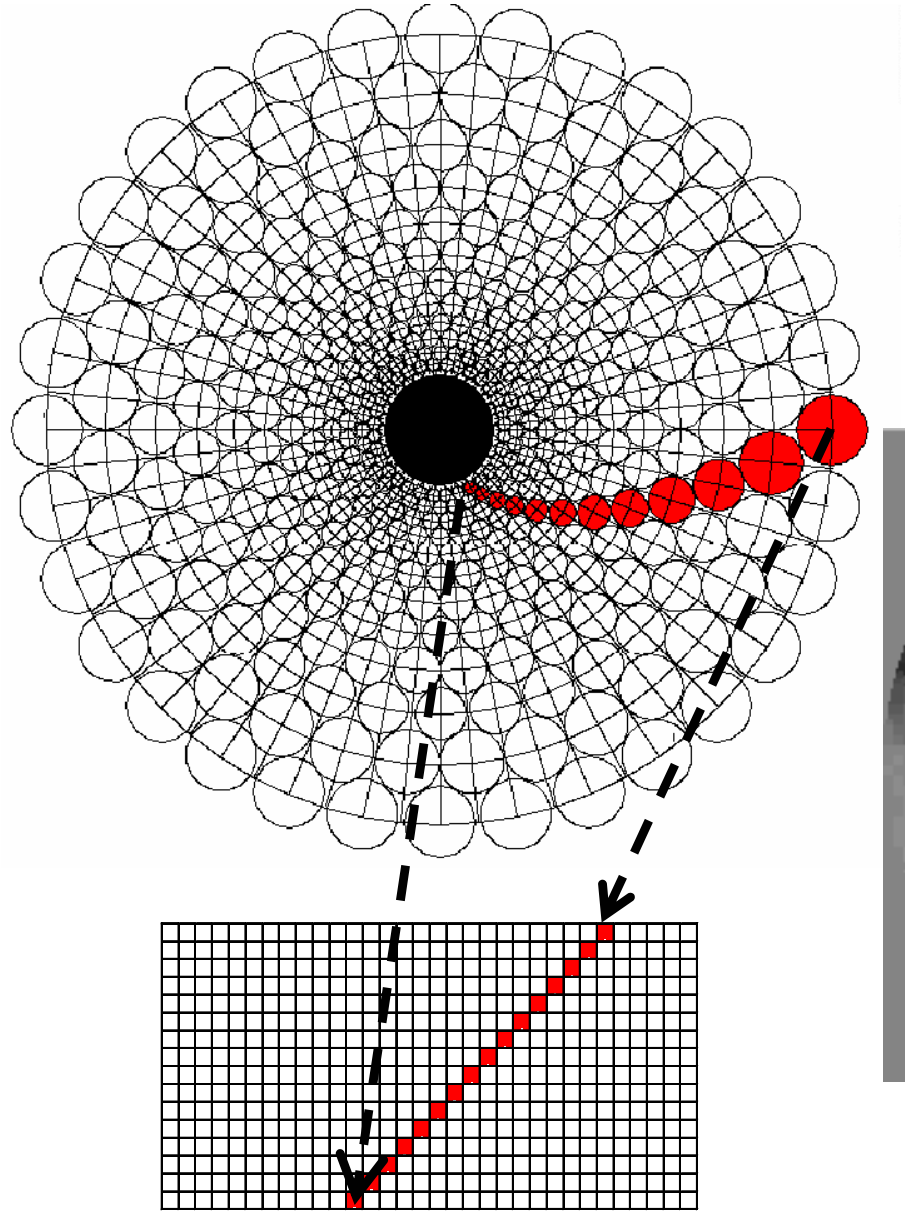




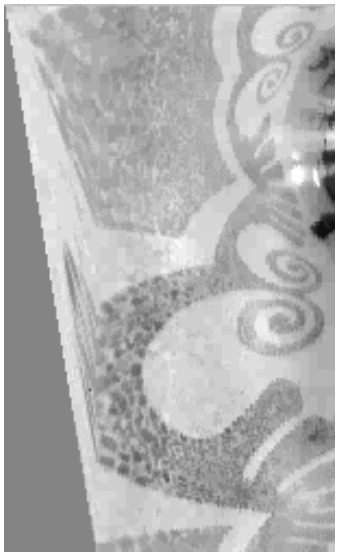


From: Sandini, G., Tagliasco V.: An anthropomorphic Retina-like Sensor for Scene Analysis, Computer Graphics and Image Processing. 1980

Log-polar (Cortical) Images



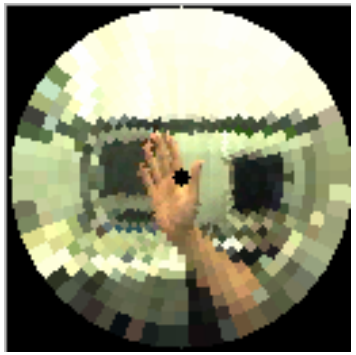
Images acquired by a retina-like sensor are rectangular



Some Questions asked

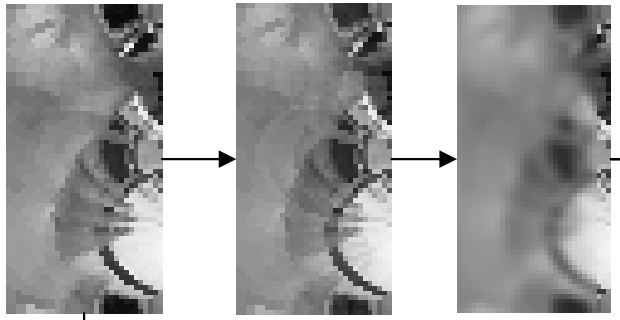
- Log-polar geometry and Time-to-impact
- Log-polar geometry and vergence control
- Integration of visual and vestibular data
- Role of vergence on tuning of VOR gain
- Motor-motor coordination in learning visually guided reaching
- Role of VOR in early development of oculomotor control
- Registration of visual and acoustic maps

Artificial Visual Cortex



Local feature detectors

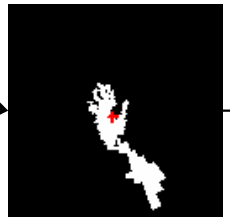
Left Eye



Cortical filtering

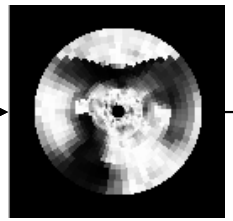
Global feature detectors

Optical flow



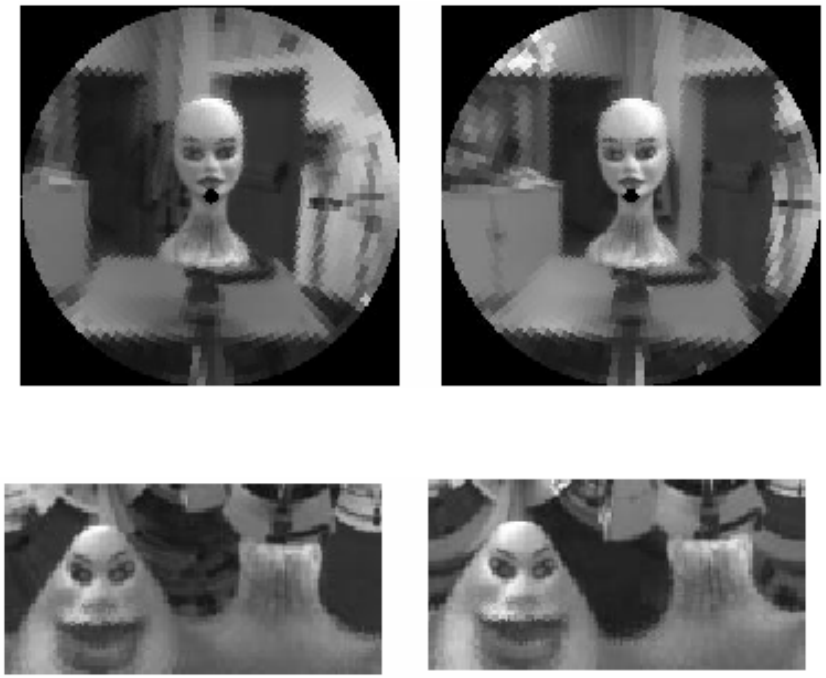
Color segmentation

From right eye

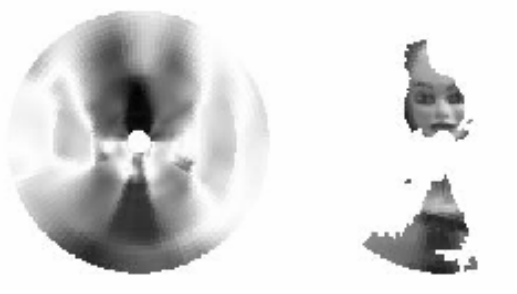


Disparity

Binocular Fusion

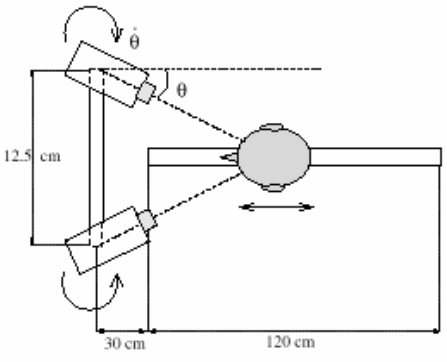


Zero Disparity Filters

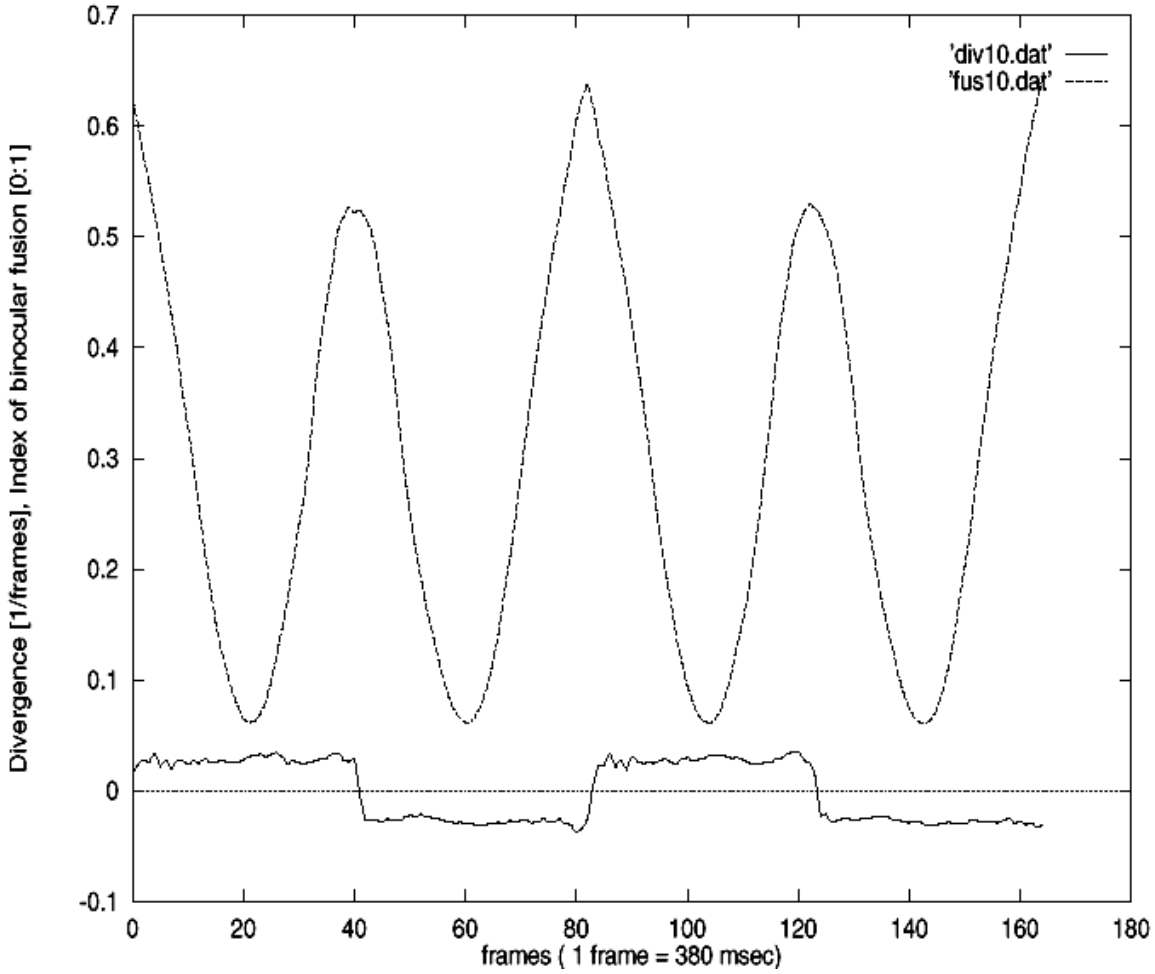


The fused part of the image in the log-polar plane is comparatively larger than in a constant resolution one

Dynamic Vergence



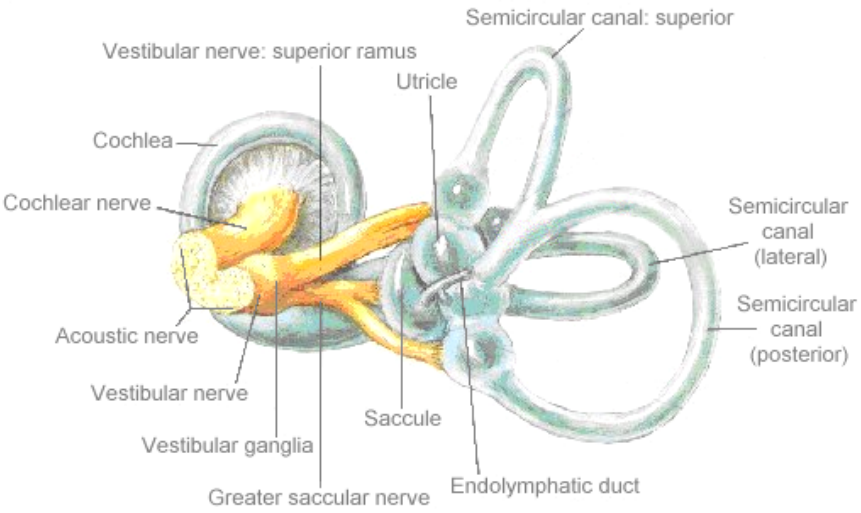
Radial Optic Flow as “fast loop” for vergence control

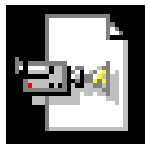


From: Capurro, C., F. Panerai, and G. Sandini, *Dynamic Vergence using Log-polar Images.* International Journal of Computer Vision, 1997. **24**(1): p. 79-94..

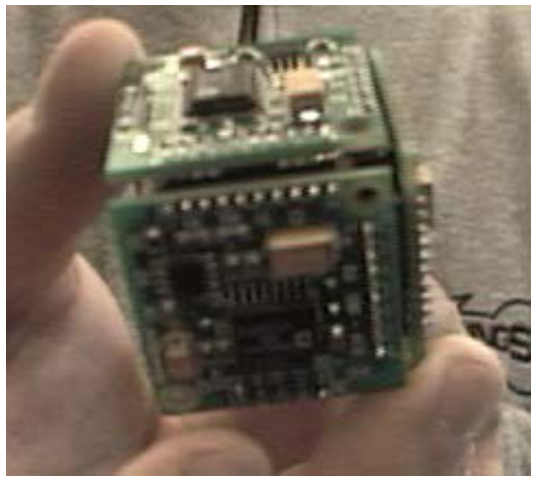
Busettini, C., G.S. Masson, and F.A. Miles, *Radial optic flow induces vergence eye movements at ultra-short latencies.* Nature,, 1997. **390**: p. 512--515

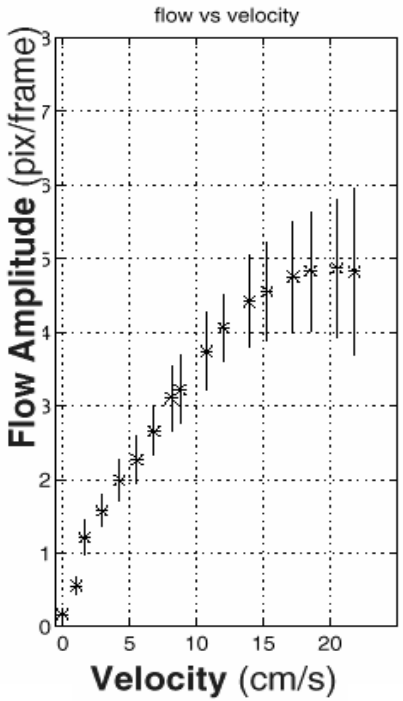
The artificial vestibular system



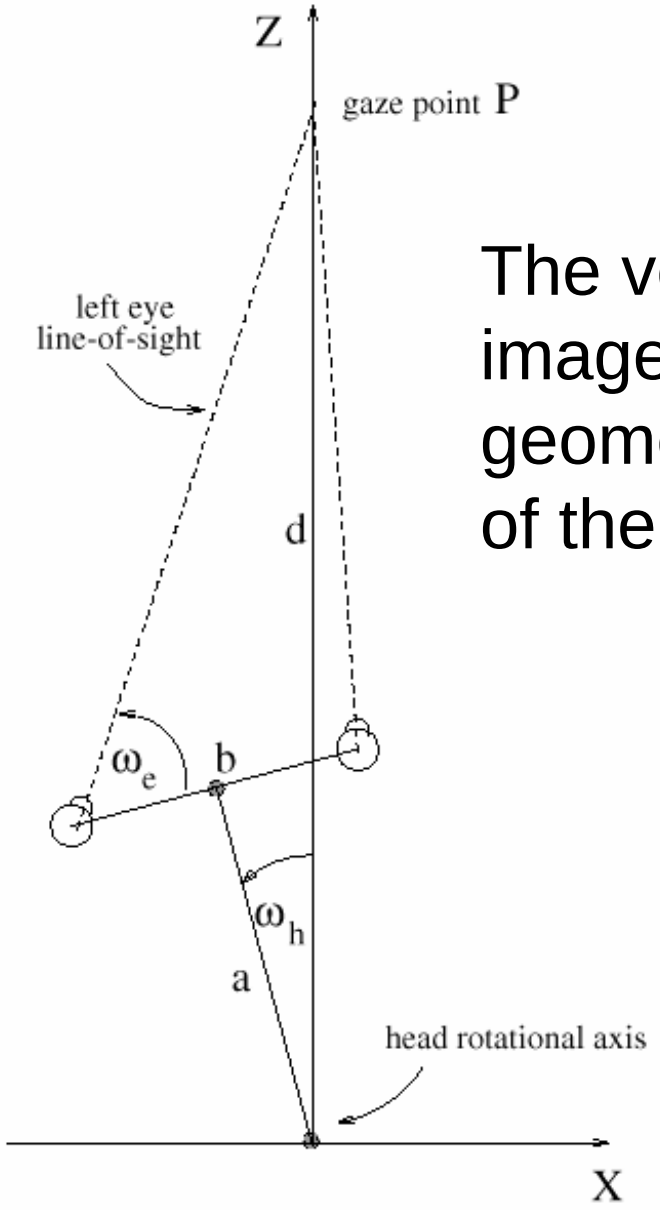


Video clip

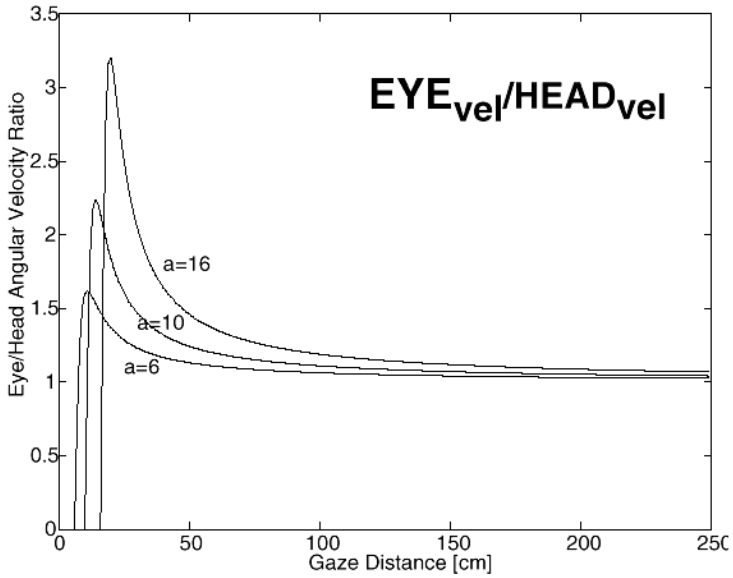




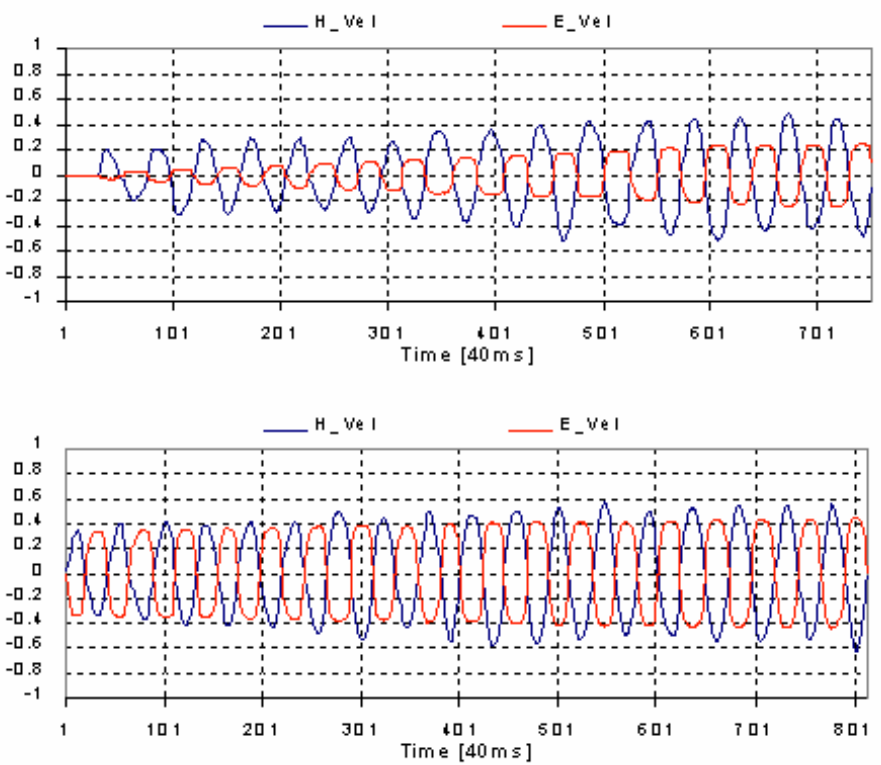
Geometric Relationship



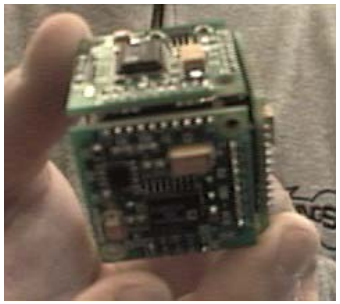
The velocity required to stabilize the image of an object depends on geometric parameters and the distance of the fixation point

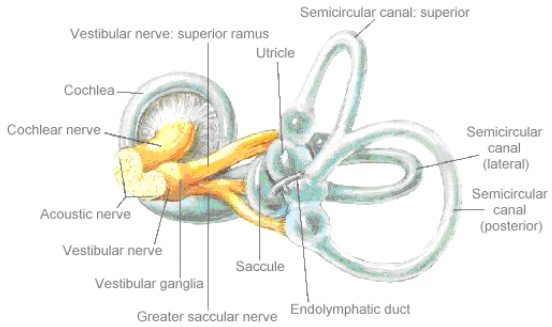


How does VOR Gain adapts?



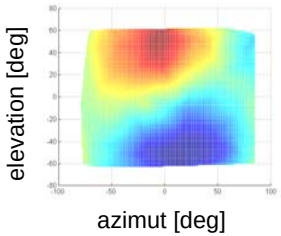
See: Finocchio, D.V., K.L. Preston, and A.F. Fuchs, *Infant Eye Movements: Quantification of the Vestibulo-Ocular Reflex and Visual-Vestibular Interactions.* Vision Research, 1991. **31**: p. 1717-1730.

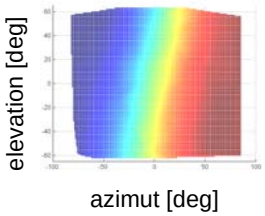


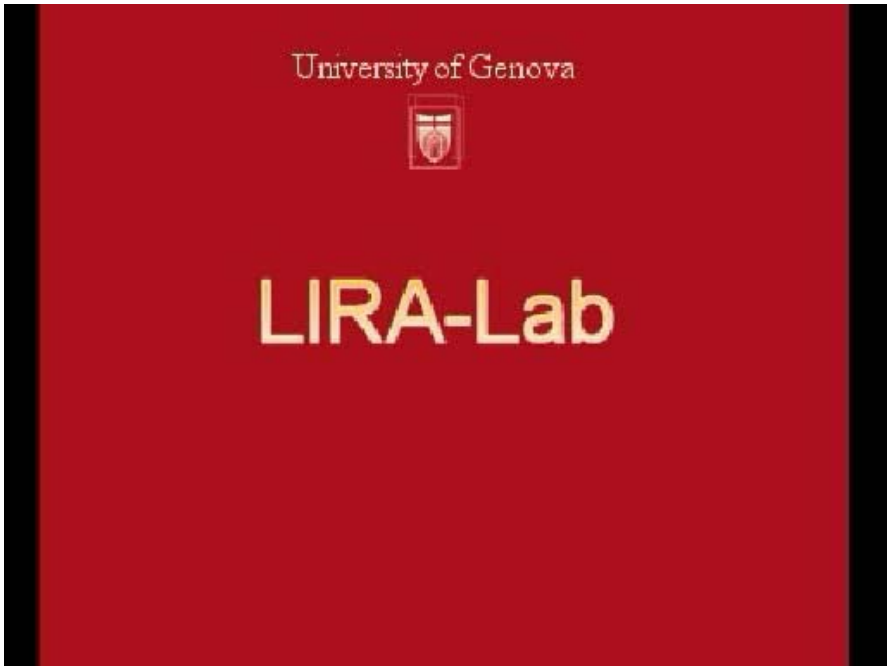


See: Sandini, G., F. Panerai, and F.A. Miles, *The Role of Inertial and Visual Mechanisms in the Stabilization of Gaze in Natural and Artificial Systems*, in *Motion Vision, Computational, Neural, and Ecological Constraints*, J.M. Zanker and J. Zeil, Editors. 2000, Springer. p. 189-218.

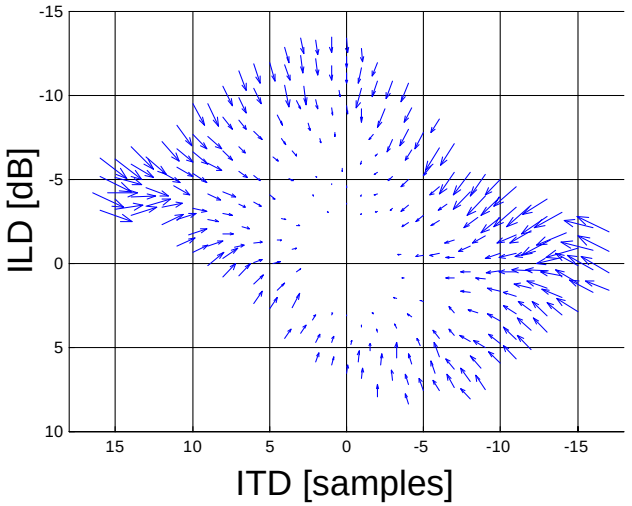
Learning Sound Directed Gaze



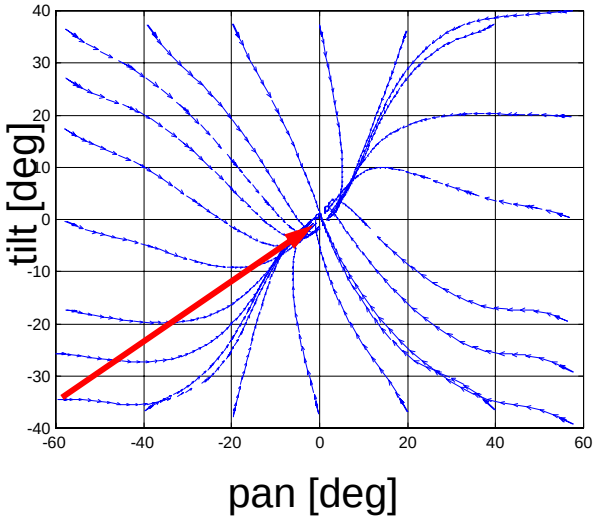




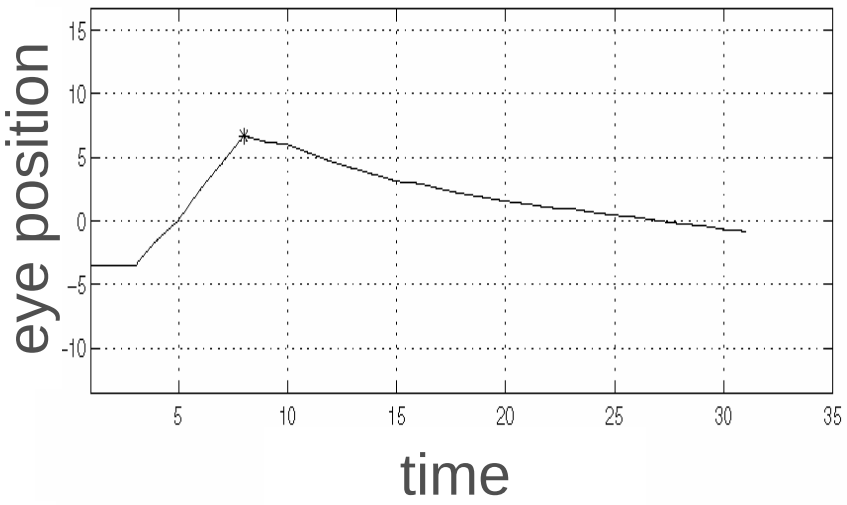
Auditory Saccades Map

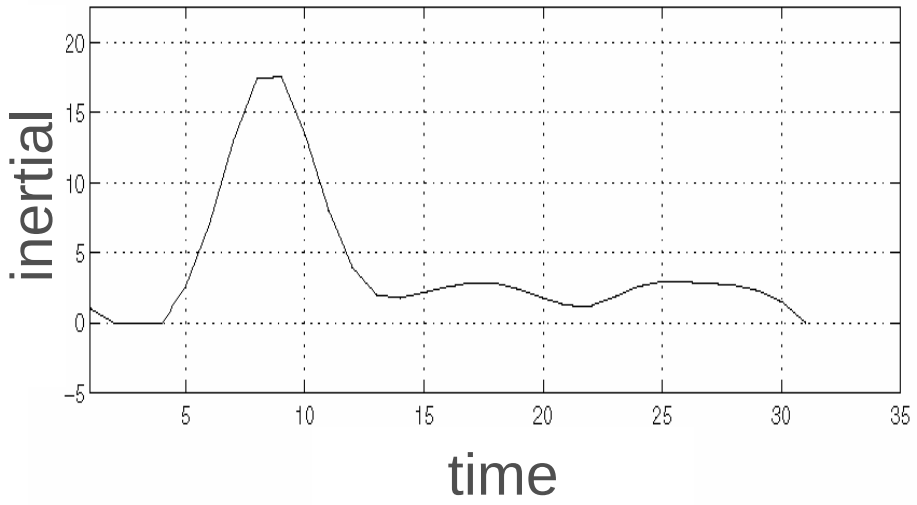


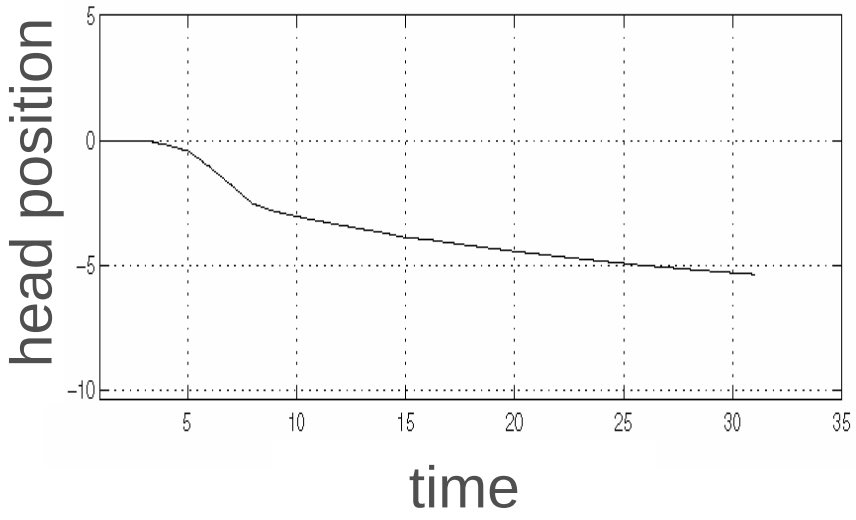
Closed Loop Trajectories

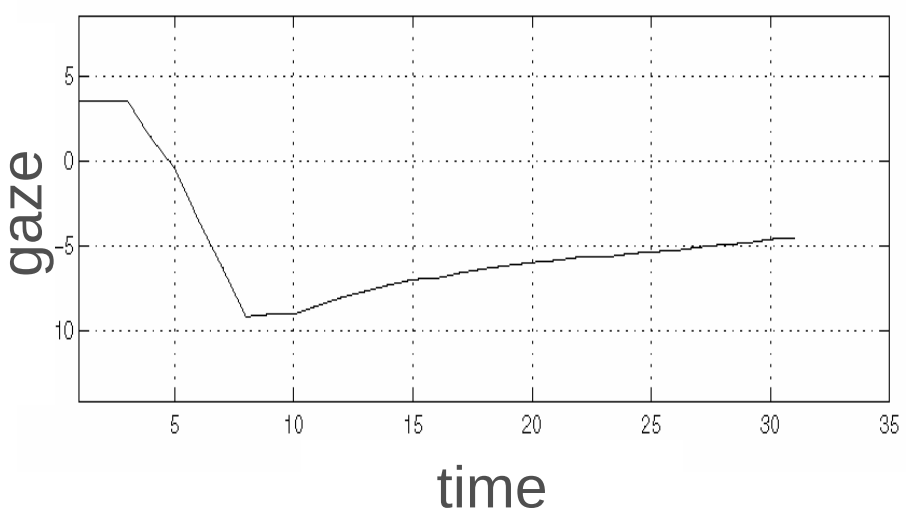


Saccades & VOR





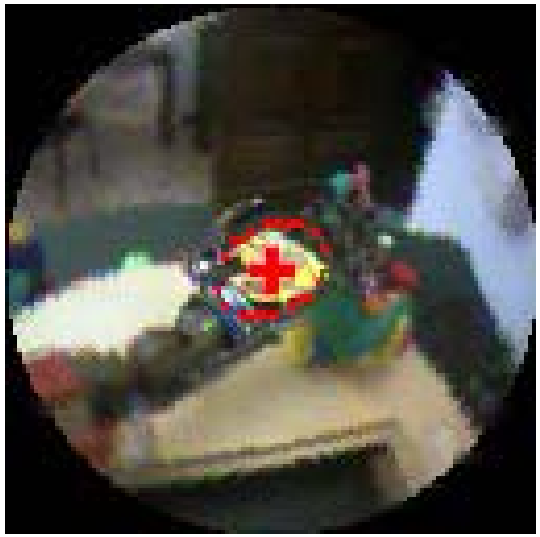


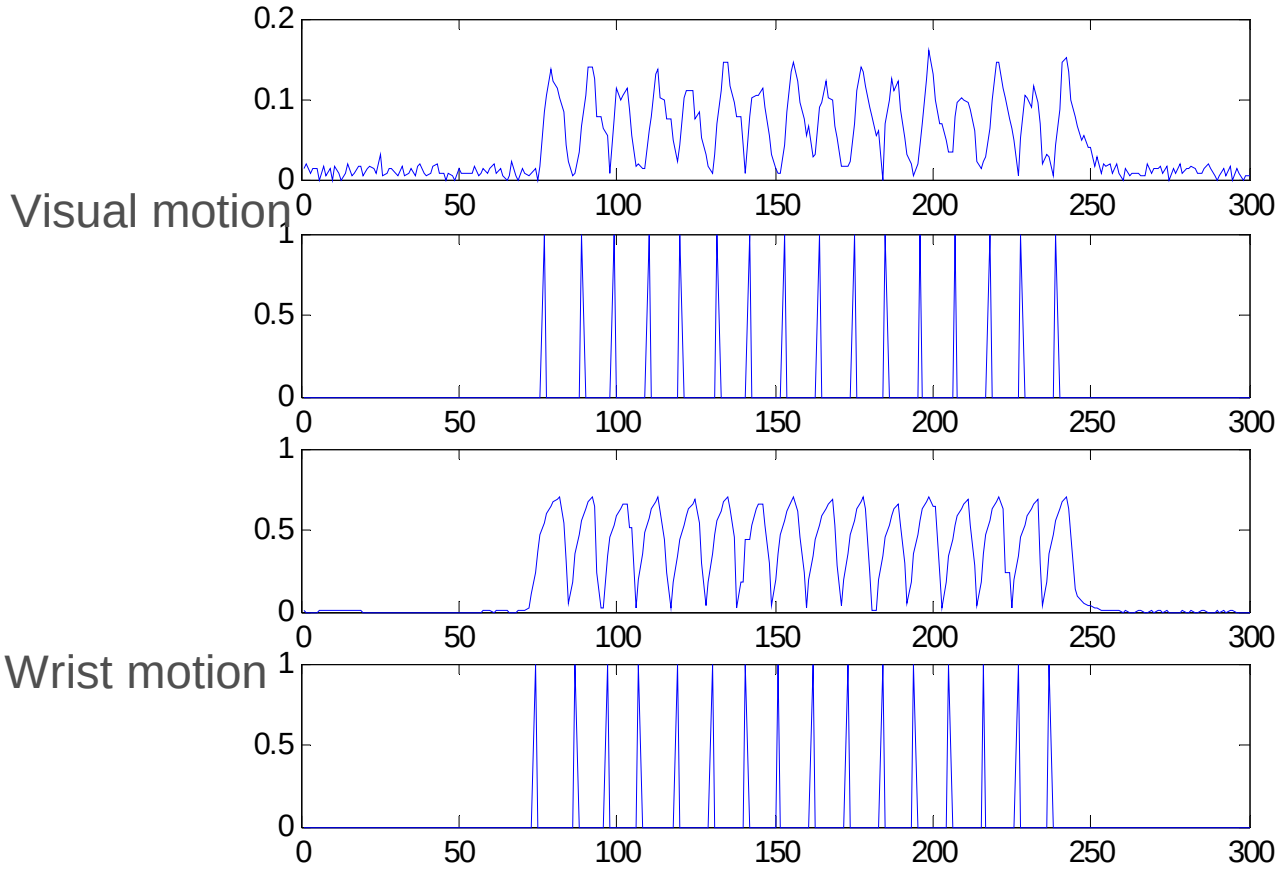


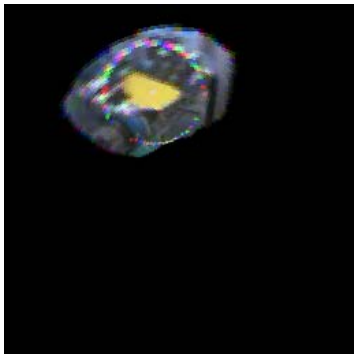
Gaze and manipulation (Body Image)

Learn about your own body (Where is my hand?)

- Exploit self-generated actions
- Exploit correlated motion







Hand segmentation: clips

Hand localization

(proprioception and color)

Tack Hand

Predict hand position and size

Visually Guided Reaching

It is possible to simplify the problem by reaching for the fixation point.

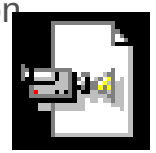
This suggests that head-eye-hand coordination plays an important role in the organization of these movements and leads to the hypothesis that a representation of current gaze direction may serve as a reference signal for arm motor control.

From: Flanders, M., L. Daghestani, and A. Berthoz, *Reaching beyond reach.* Experimental Brain Research, 1999. **126**(1): p. 19-30.

Motor-motor coordination.
Coordination is obtained by mapping motor commands from the eye-head space into motor commands to the arm-hand.

F. Gandolfo, G. Sandini, E. Bizzi, “A Field-based Approach to Visuo-motor Coordination,” Workshop on sensorimotor coordination: amphibians, models, and comparative studies, Sedona, USA, 1996.

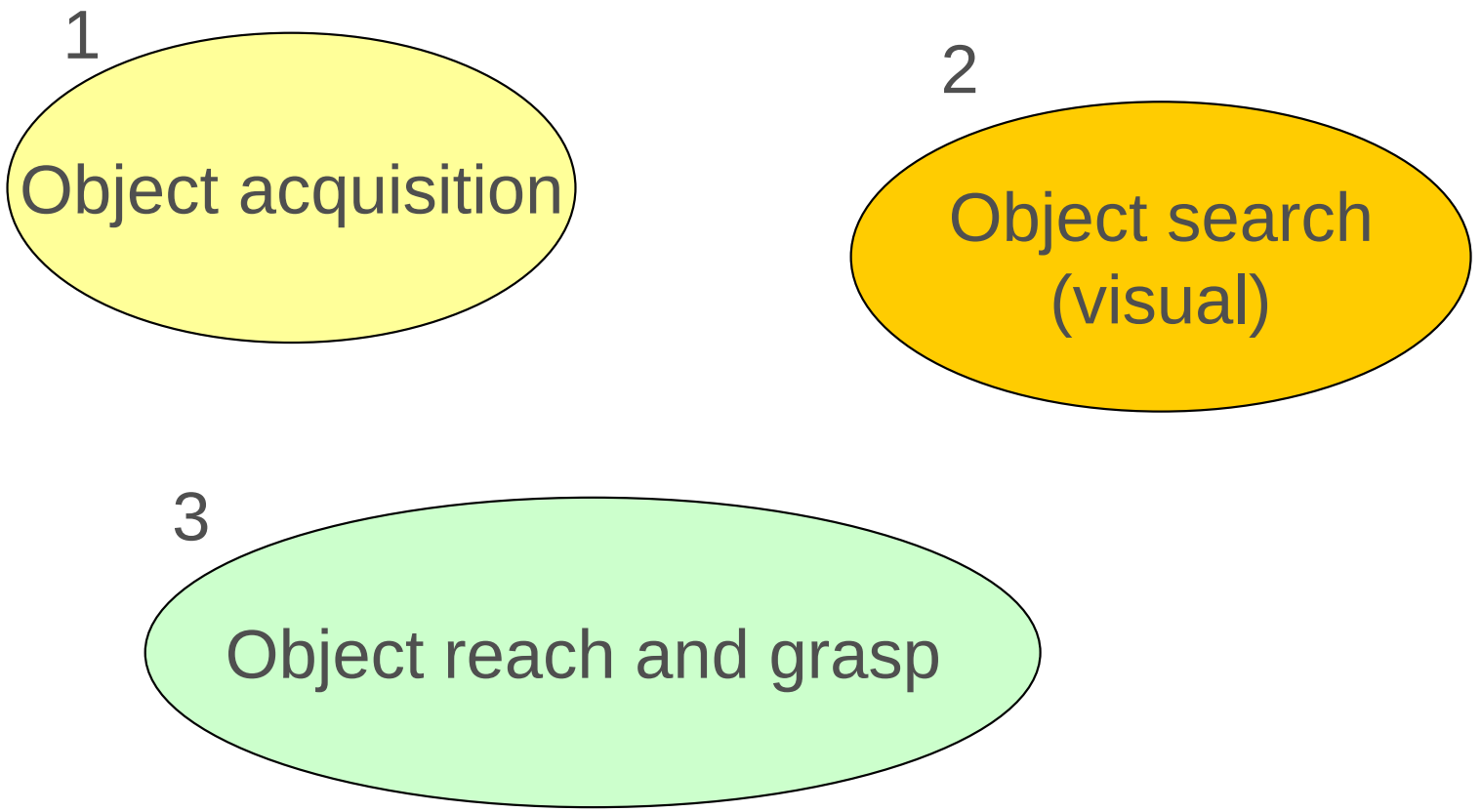
Metta, G., G. Sandini, and J. Konczak. *A Developmental Approach to Sensori-motor Coordination in Artificial Systems*. in *IEEE Conference on System, Man and Cybernetics*. 1998. San Diego (USA).



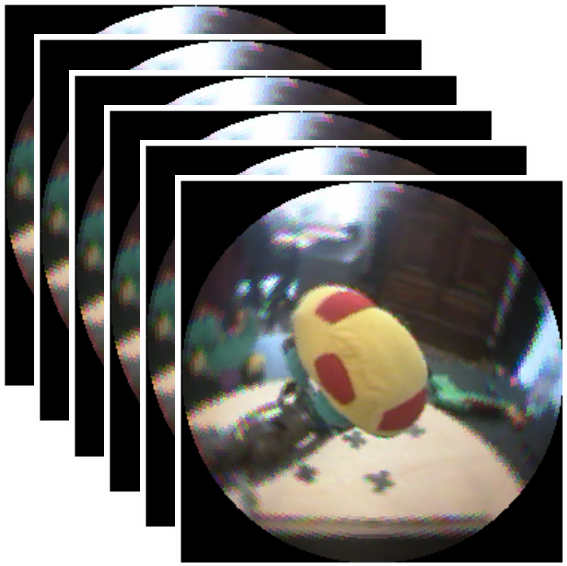
Video Clip

…a complex behavior…

A robot grasping behavior based on three phases



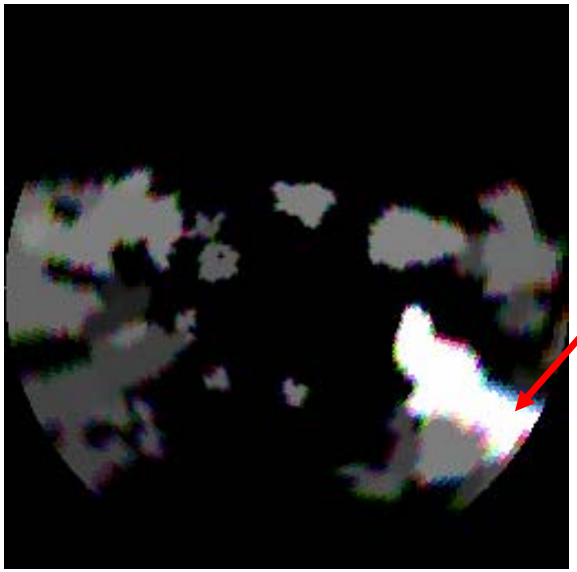
Object Acquisition



Blob processing

Estimating probabilities

- Watching the hand holding the object
- Hypothesis: central blob $\in$ object
- Estimate: $P(blob_i \in object \mid \text{fixating object \&\& !hand})$



Most salient blob

Object search
(visual)

Looking for airplane

Fixating

Segmentation
(watershed)

Full segmentation
(recovered from model)

Object reach and grasp

1. Robot moves the hand over the object
2. Robot moves down the hand
3. Robot grasps the object
4. If object is grasped, drop it outside table
5. If object is not grasped go back to visual search

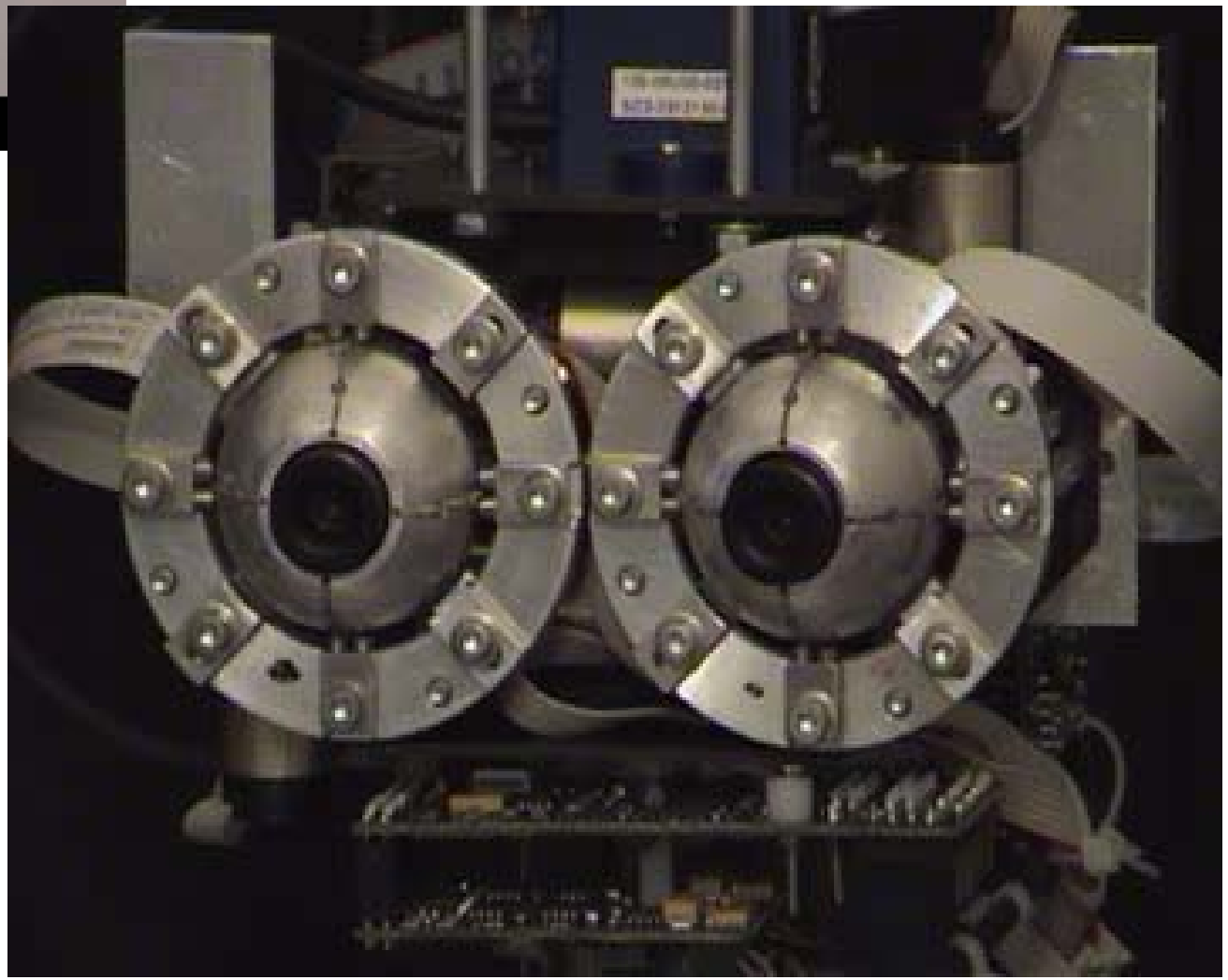
Two examples

Future directions…

RobotCub
(see poster…)

- Humanoid: as much as possible
- Child size
- Reasonable weight
- Robust: it might “fall”
- Degrees of freedom: 55?

New Head



The end