From natural
to
artificial cognitive systems

Giorgio Metta

LIRA-Lab
Dept. of Computer and Systems Sciences
University of Genoa, Italy

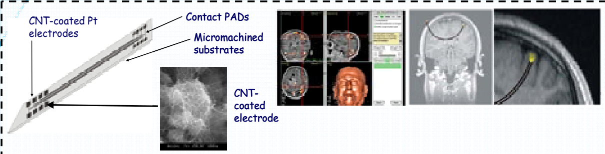
Dept. of Robotics, Brain and Cognitive Sciences
Italian Institute of Technology
Genoa

RobotCub

The Italian Institute of Technology

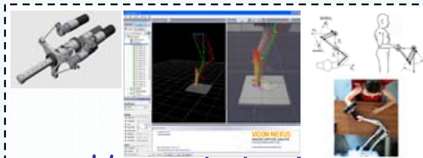
Cogsys Cognitive Systems

Robotics, Brain and Cognitive Sciences Department



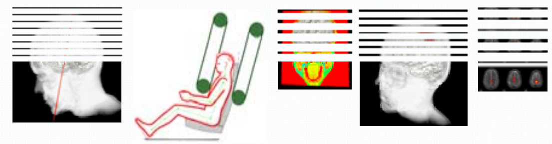
Brain machine interface

Design and fabrication of neural probes for in-vivo applications bases on conductive and biocompatible nano-materials for electrodes



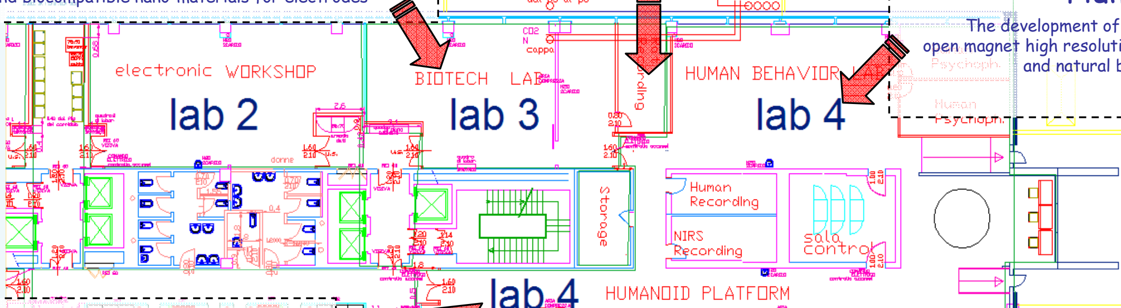
Human behavior

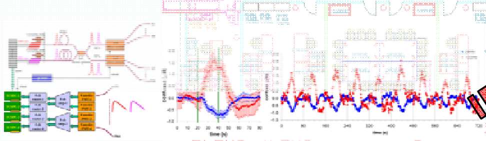
Exploring the brain mechanisms at the basis of motor behavior and motor learning



Human behavior

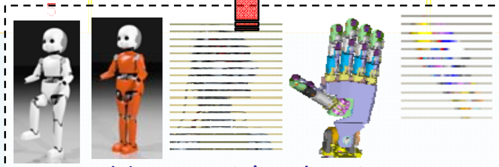
The development of new generation brain imaging systems: open magnet high resolution FMRI system targeted to brain research and natural behavior analysis (manipulation).





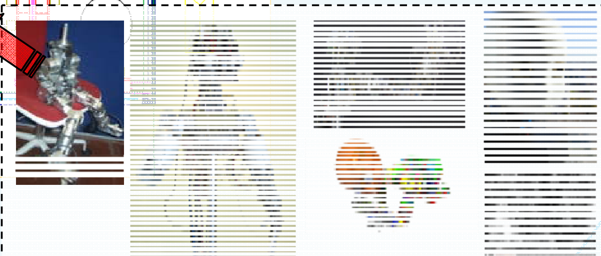
Brain machine interface

The near-infrared spectroscopy: a new non-invasive technique to investigate brain functions in human subjects



Humanoid robotics

Designing new generation, intelligent and interacting platforms



Humanoid robotics

Toward next generation hybrid systems realized with soft materials and with cognitive abilities allowing operating in human environment

The tale of the Wright brothers (and three messages)

1. Reverse engineering:
   - Looking and copying bird flight → aircraft design
   - Reverse of reverse engineering → aerodynamics led to better understanding of bird flight (forward engineering)
2. Models:
   - Wright started from a previous model of Lilienthal (which was wrong) → but then they had (after 2 years) to produce their own models (and test them), they built a wind tunnel (very modern)!
3. Stability and control:
   - Separate models didn't work well (either stability or control)
   - Discovered that the key to stability and control is by rolling → turn by rolling! Separate models don't work, holistic approach is required and this was done by looking also at birds.
   - Understanding at the systems level

* From The Computational Neurobiology of Reaching and Pointing, Shadmehr and Wise, MIT press

…and the story goes

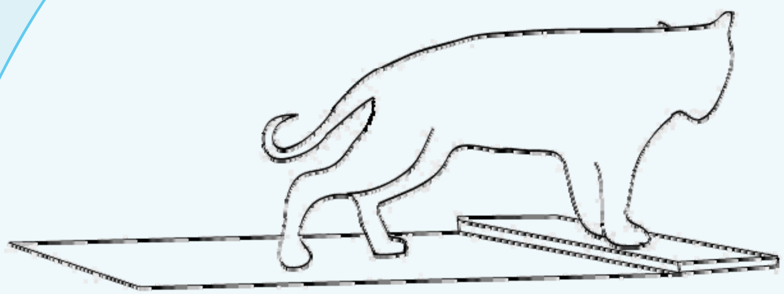
- Late 1903, first powered flight (35m, 10km/hour)
- 5 yrs later, 2 hours flight
- 8 yrs later, across North America
- 24 yrs later, New York to Paris
- 65 yrs later, three people to the moon
- Now, small seats and screaming infants

The three messages again!

1. Reverse and forward engineering
2. Mathematical modeling and empirical testing
3. Systems-level approach

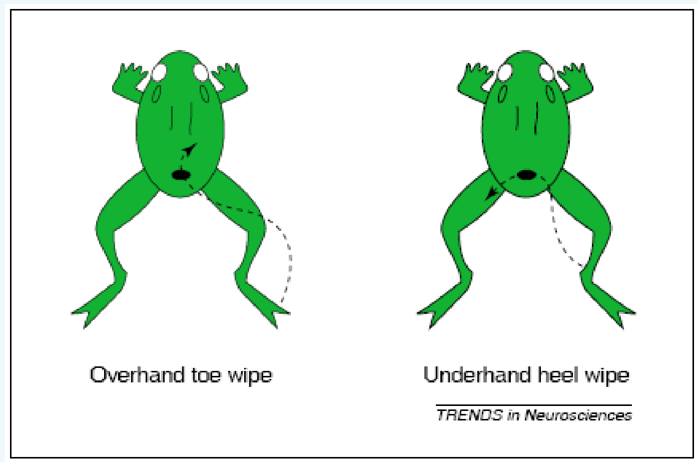
Reverse engineering

Spinal behaviors



Walking behavior: cat rehabilitated to walk after complete spinal cord transection

Wiping reflex: an irritating stimulus elicits a wiping movement precisely directed at the stimulus location

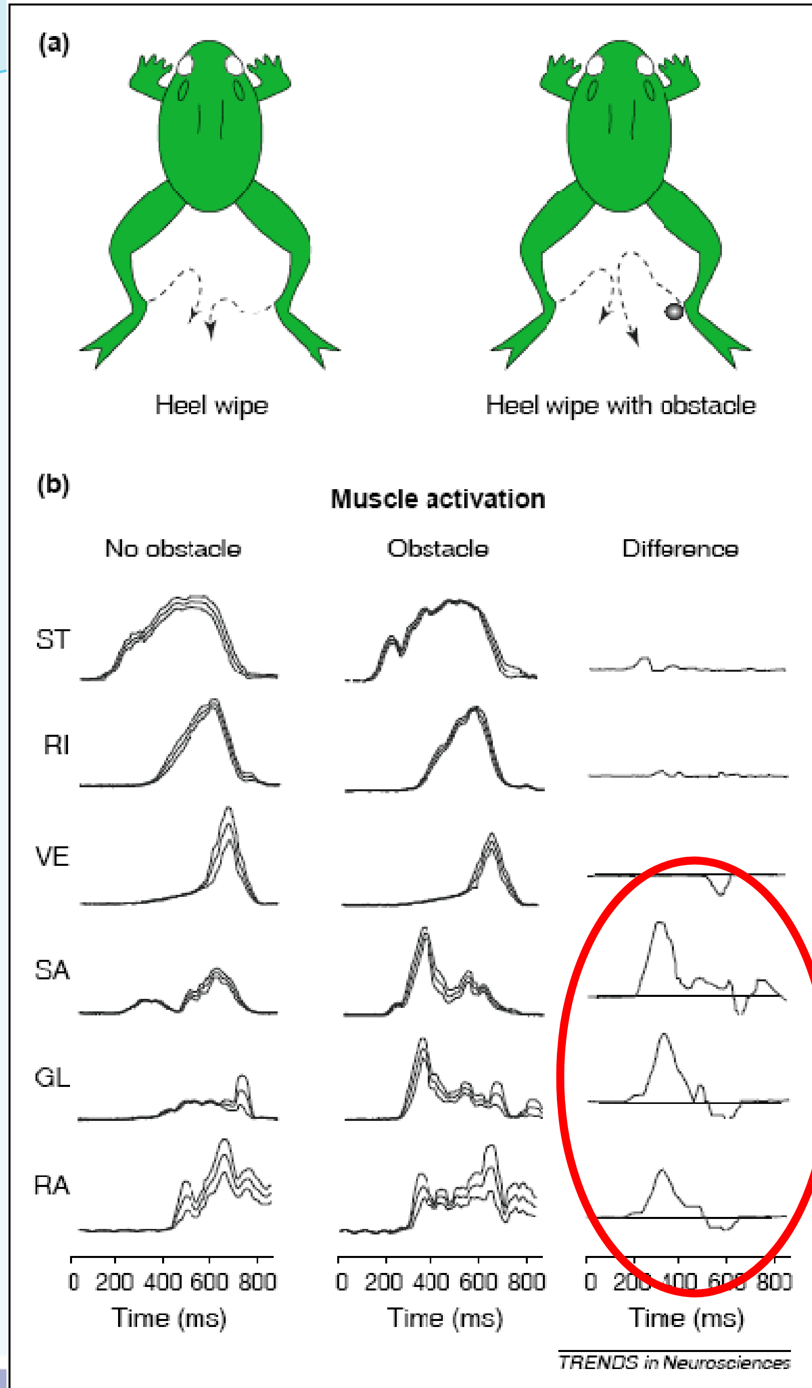


Poppele, R., & Bosco, G. (2003). Sophisticated spinal contributions to motor control. *Trends in Neurosciences, 26*(5), 269-276.

Spinal sensorimotor coordination

(not simple reflexes after all)

- Stimulus-response coordinate transformations
  - Evidence for:
    - Combination of basic behaviors
    - Reach body parts that move respect to each other
    - Adapt and avoid obstacles
    - Use different sets of muscles
  - Inverse dynamics (less compelling to me):
    - Move → dynamics
    - Descending "kinematics" pathways (e.g. optic tectum, orienting behavior)
    - Walking, CPG's, etc.
  - Degrees of freedom problem:
    - Evidence of synergies
    - Muscles activate together
    - Multi-joint muscles

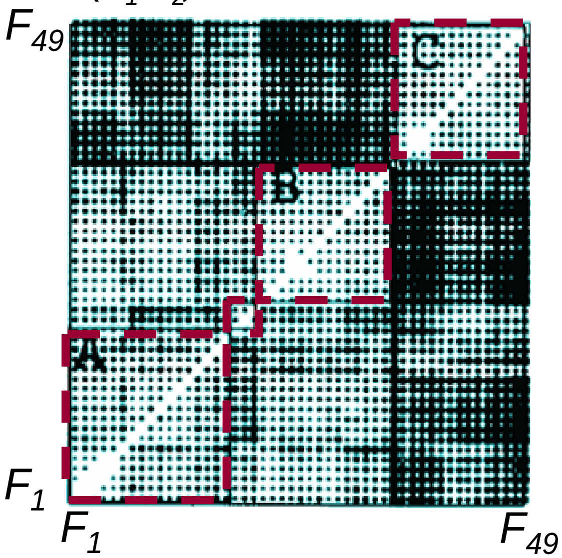


Differences in the muscle activation

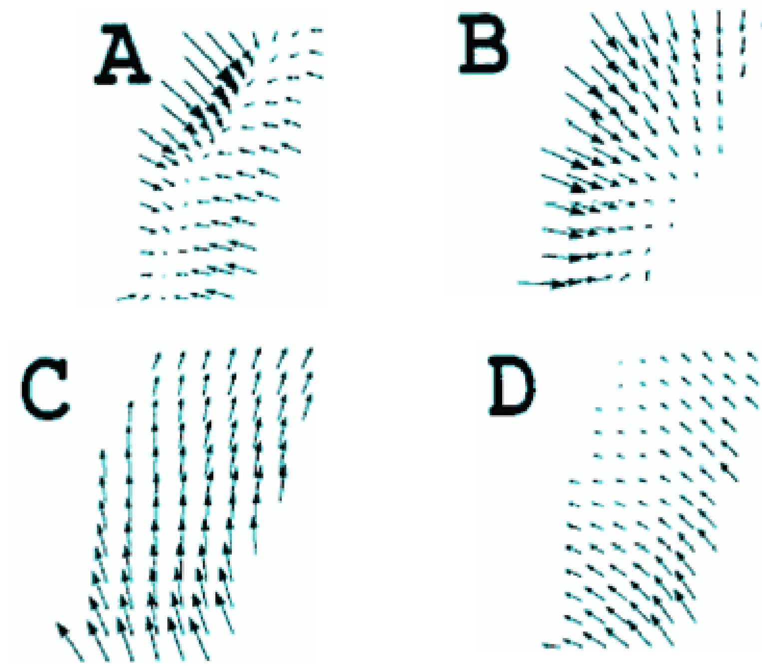
Modeling the spinal controller: Francesco Nori

- Systematic stimulation of different regions of the spinal cord produced only a **few types force fields (at least four)**.

Dissimilarity matrix of 49 force fields. Darker circles represent *d(F_1,F_2)<<1*



4 clusters can be identified



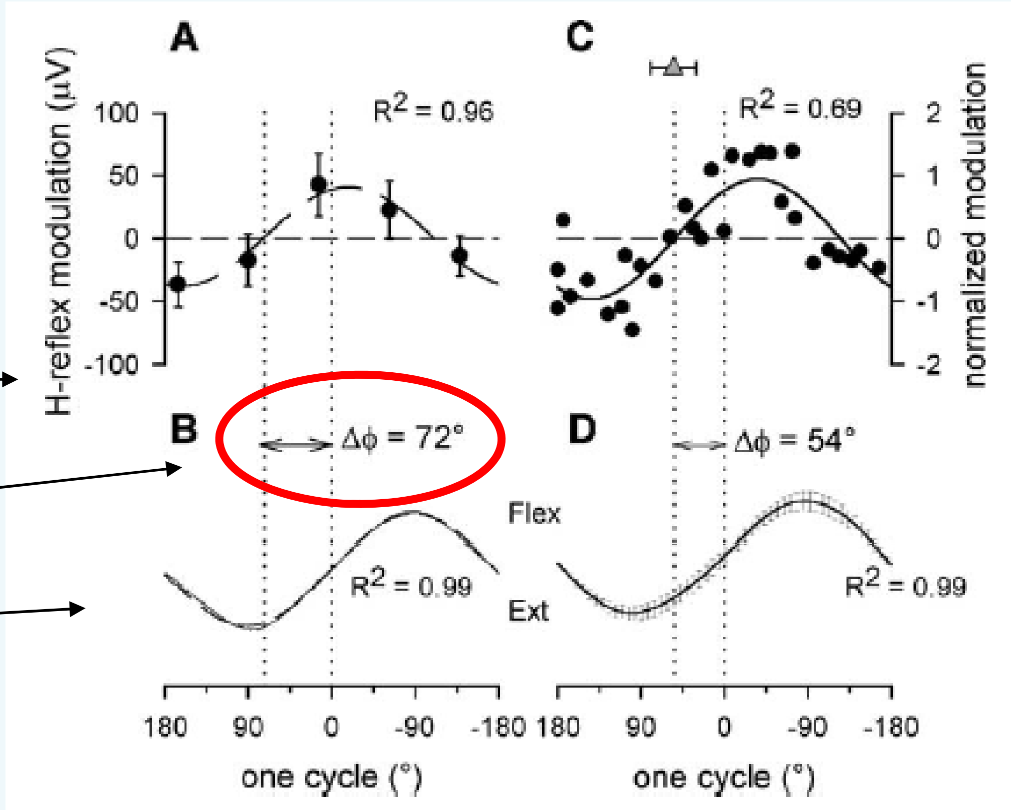
The presence of only few units of motor output within the spinal cord is difficult to reconcile with the obvious ability of the nervous system to produce a wide range of movements.

Looking at muscles

H-reflex deviation from the mean
(conductance of the spinal-muscle nerves)

Muscle anticipate the kinematics

kinematics

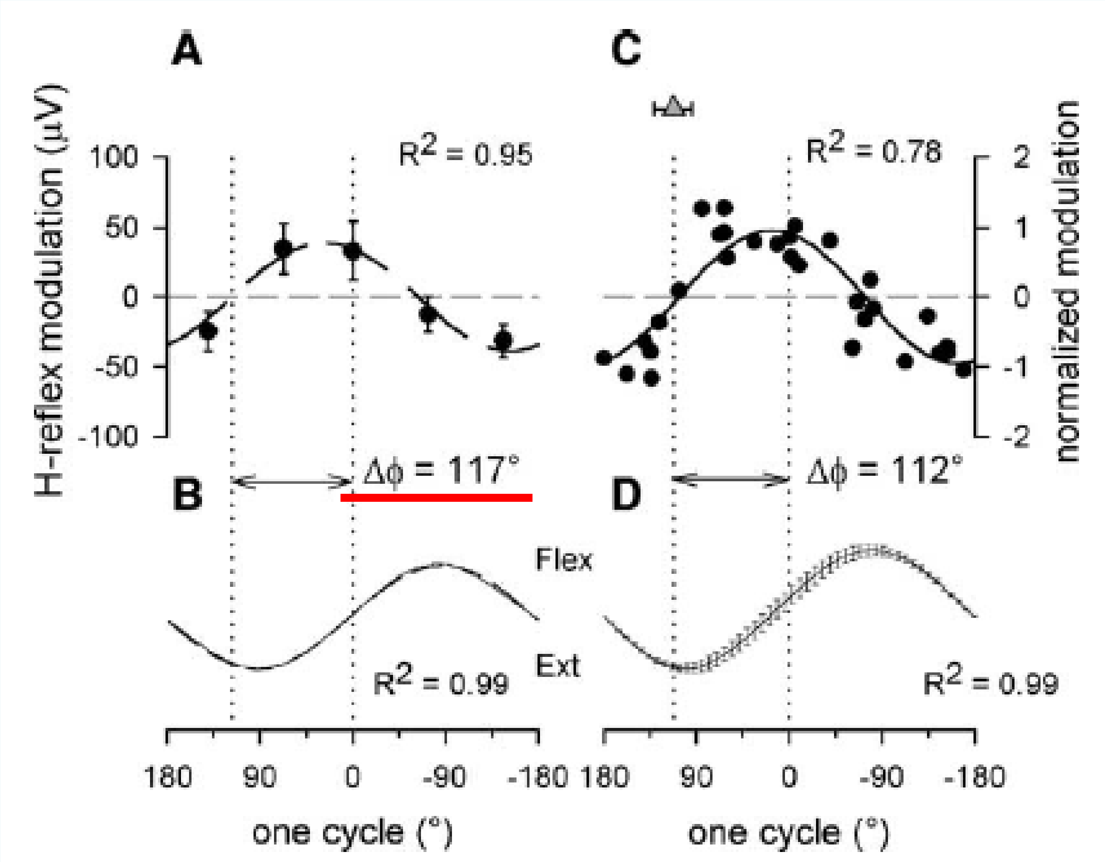


Prone position, wrist ext/flexion (ECR, FCR muscles)

Borroni, P., Montagna, M., Cerri, G., & Baldissera, F. (2005). Cyclic time course of motor excitability modulation during the observation of a cyclic hand movement. *Brain Research, 1065*, 115-124.

Supine position

As before, but different phase difference btw the kinematics and the muscular activation

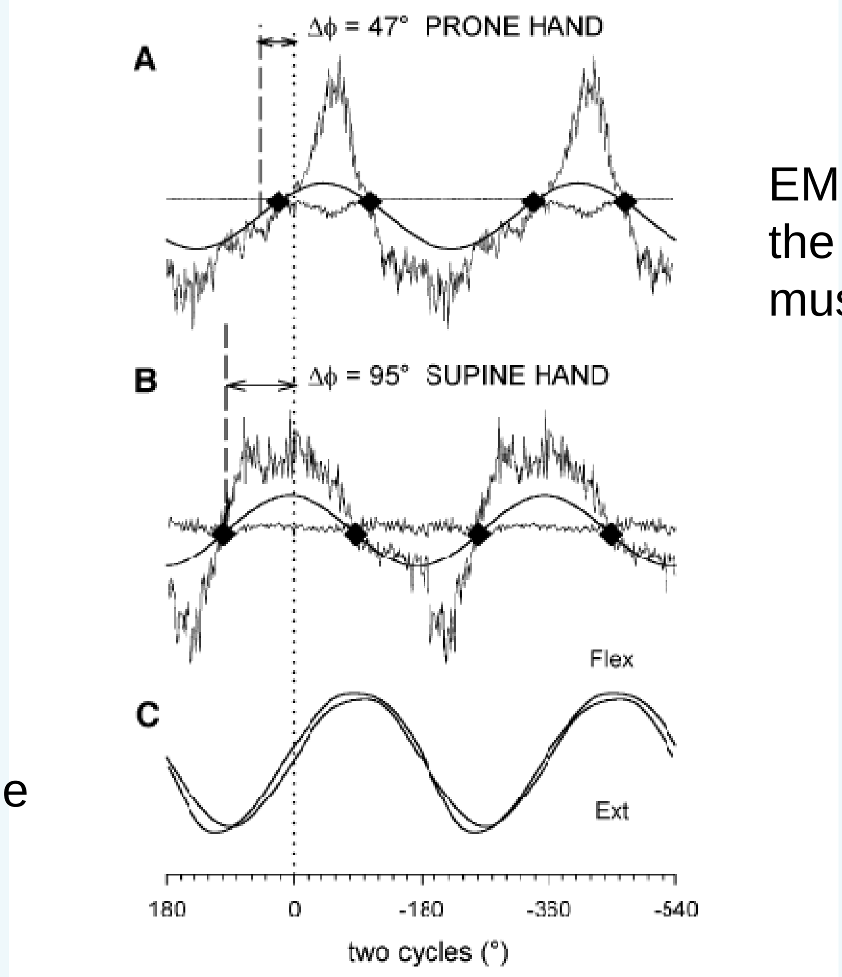


Then compare with actual action

Which was 54° on average

Which was 112° on average

The movement doesn't change



EMG signals from the FCR and ECR muscles

RobotCub

…but let me take a little leap forward

European Journal of Neuroscience, Vol. 15, pp. 399–402, 2002

SHORT COMMUNICATION
Speech listening specifically modulates the excitability of tongue muscles: a TMS study

Luciano Fadiga,[1] Laila Craighero,[1,2] Giovanni Buccino[2] and Giacomo Rizzolatti[2]

[1]Dipartimento di Scienze Biomediche e Terapie Avanzate, Sezione di Fisiologia Umana, Università di Ferrara, via Fossato di Mortara 17/19, 44100 Ferrara, Italy

[2]Istituto di Fisiologia Umana, Università di Parma, via Volturno 39, 43100 Parma, Italy

Keywords: mirror neurons, motor-evoked potentials, motor system, motor theory of speech perception

Cogsys Cognitive Systems

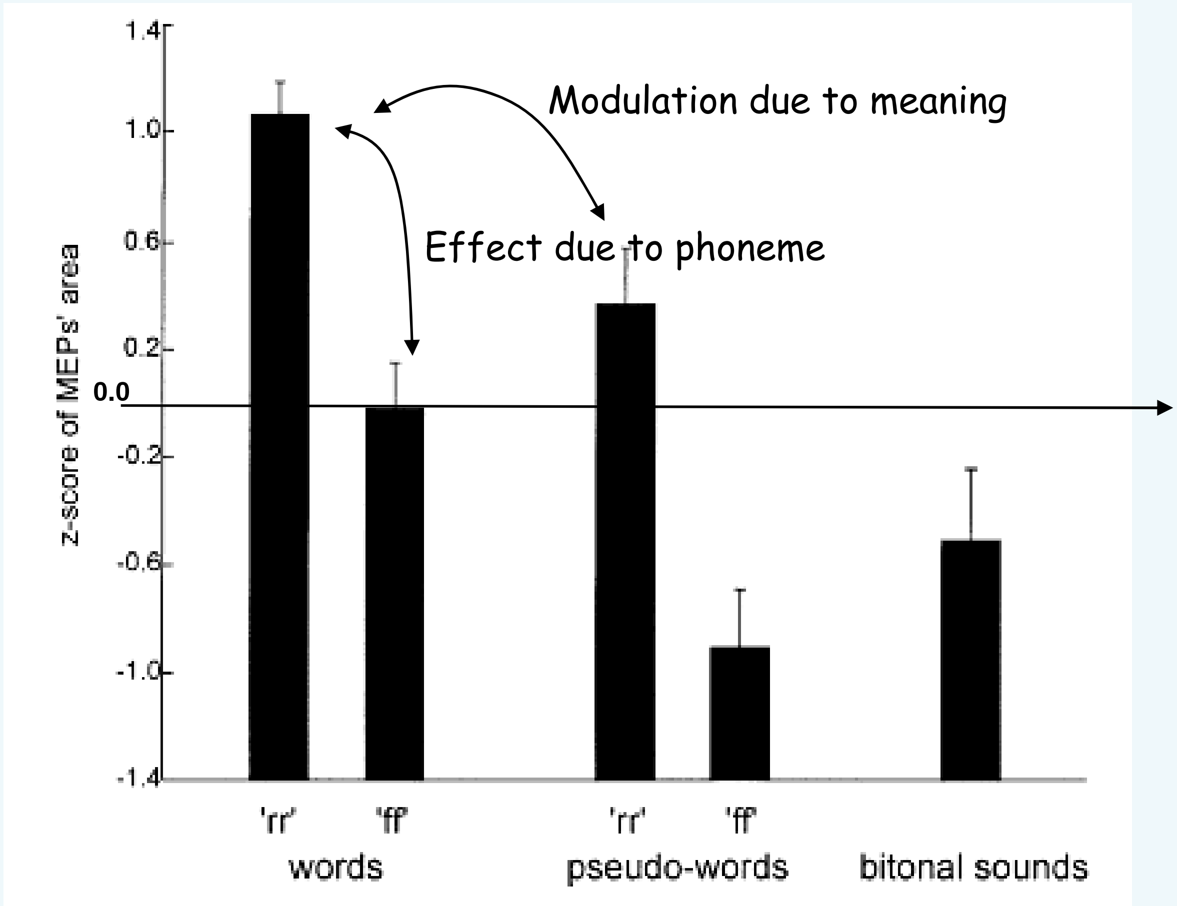
The experiment

- Listening: three categories of stimuli (words, pseudo-words, bi-tonal sounds).
- Two phonemes 'rr' requires strong tongue tip movement, 'ff' requires slight tongue tip movement.
- TMS of the under-threshold motor cortex.
- Recording of the MEP (motor-evoked potential) from the tongue muscles.

Examples of word/pseudo-words

Labiodental fricative consonant, 'rr'		Lingua-palatal fricative consonant, 'ff'	
Words	Pseudo-words	Words	Pseudo-words
birra (bier)	berro	baffo (moustache)	biffo
carro (cart)	firra	beffa (hoax)	ciffo
cirro (cirrus)	forro	buffo (funny)	leffa
farro (spelt)	furra	ceffo (snout)	meffa
ferro (iron)	marro	coffa (crow's nest)	paffo
mirra (myrrh)	merro	goffo (clumsy)	peffa
morra (morra)	parro	muffa (mold)	poffa
porro (leek)	perro	puffo (smurf)	seffa
serra (greenhouse)	vorro	tuffo (dive)	viffa
terra (ground)	vurro	zaffo (plug)	voffo

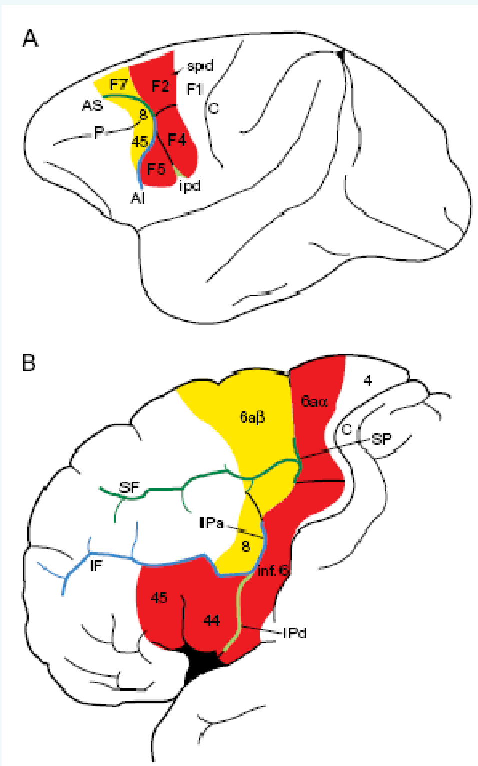
Results (in short)



Language within our grasp

Giacomo Rizzolatti and Michael A. Arbib

In monkeys, the rostral part of ventral premotor cortex (area F5) contains neur
both when the monkey grasps or manipulates objects and when it observes
making similar actions. These neurons (mirror neurons) appear to represent a sy



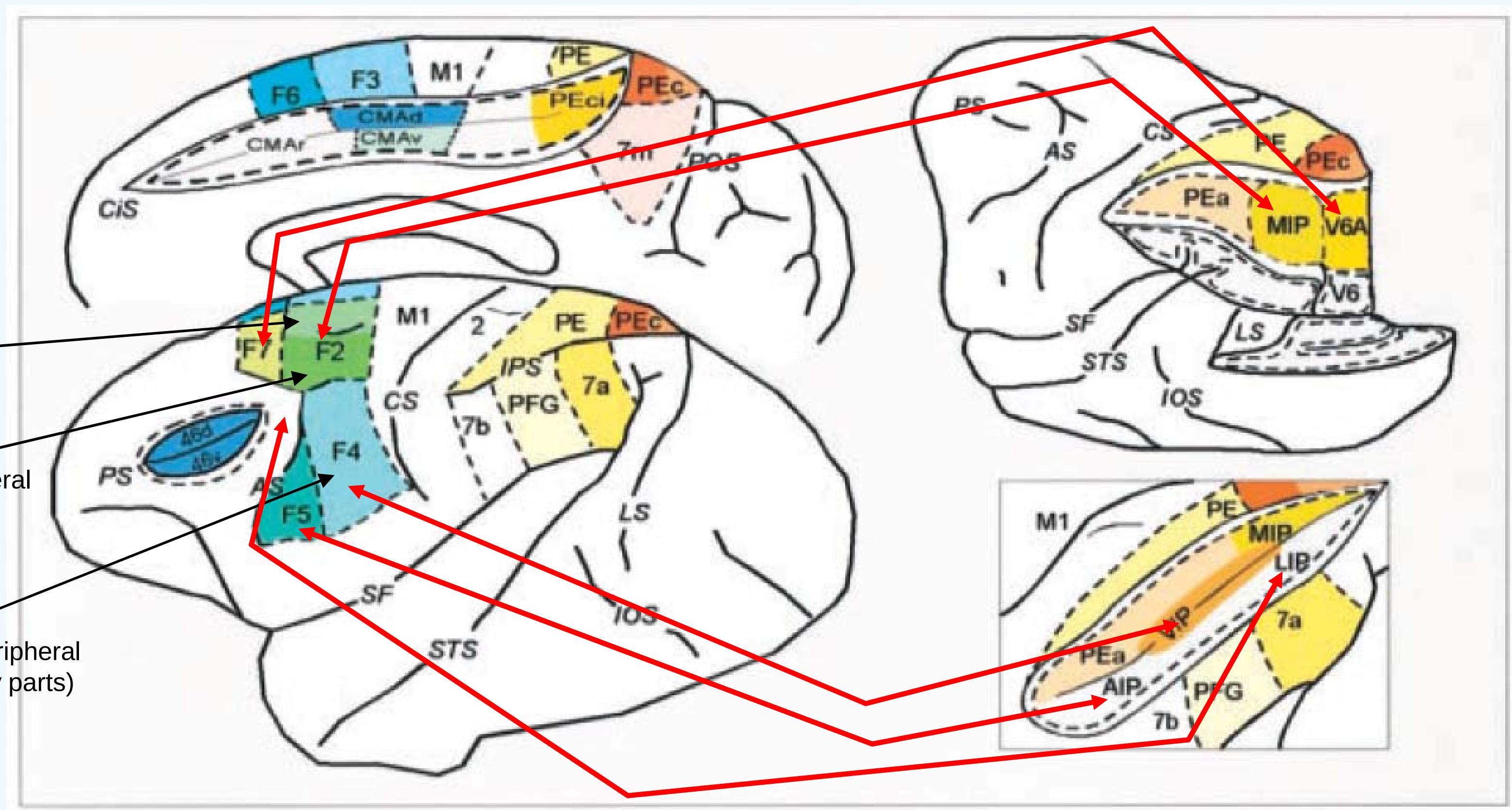
"In all communication, sender and receiver must be bound by a common understanding about what counts; what counts for the sender must count for the receiver, else communication does not occur. Moreover the processes of production and perception must somehow be linked; their representation must, at some point, be the same."

[Alvin Liberman, 1993]

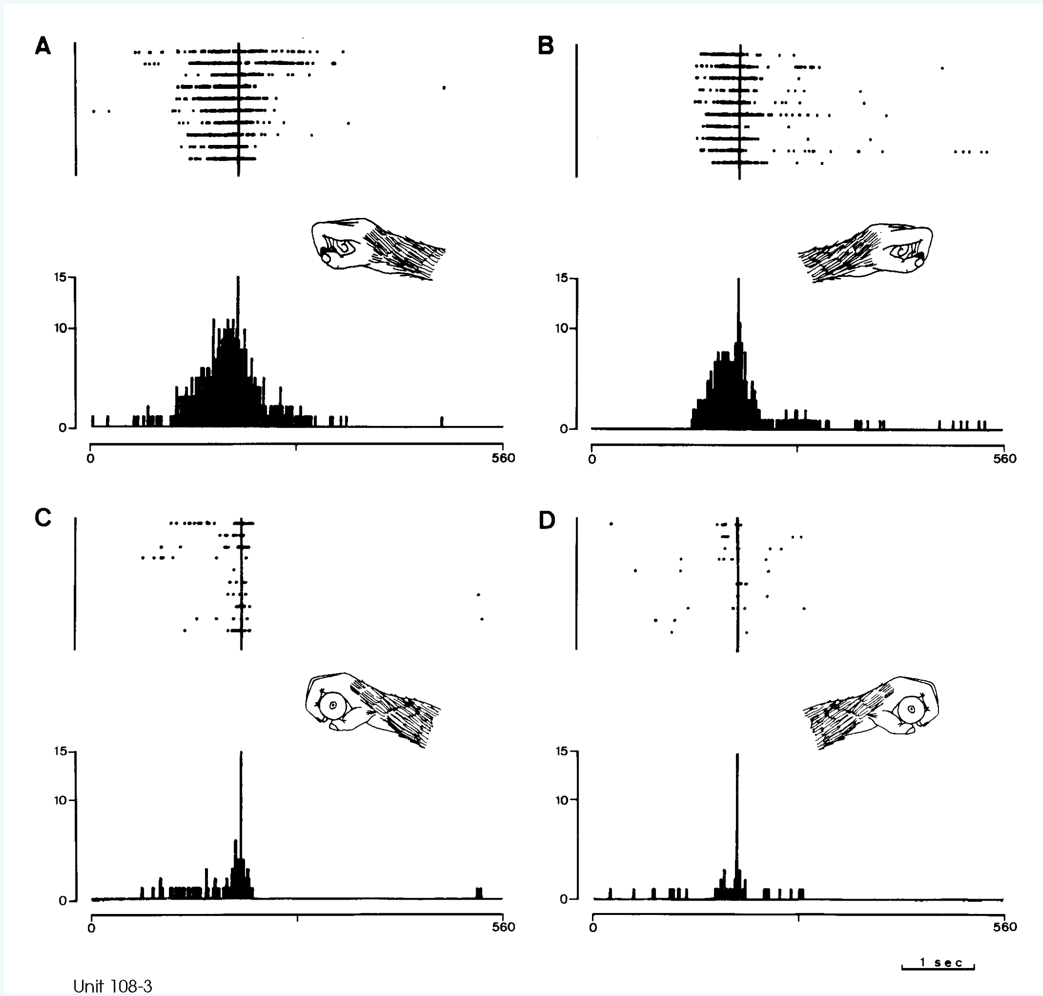
More central

More peripheral

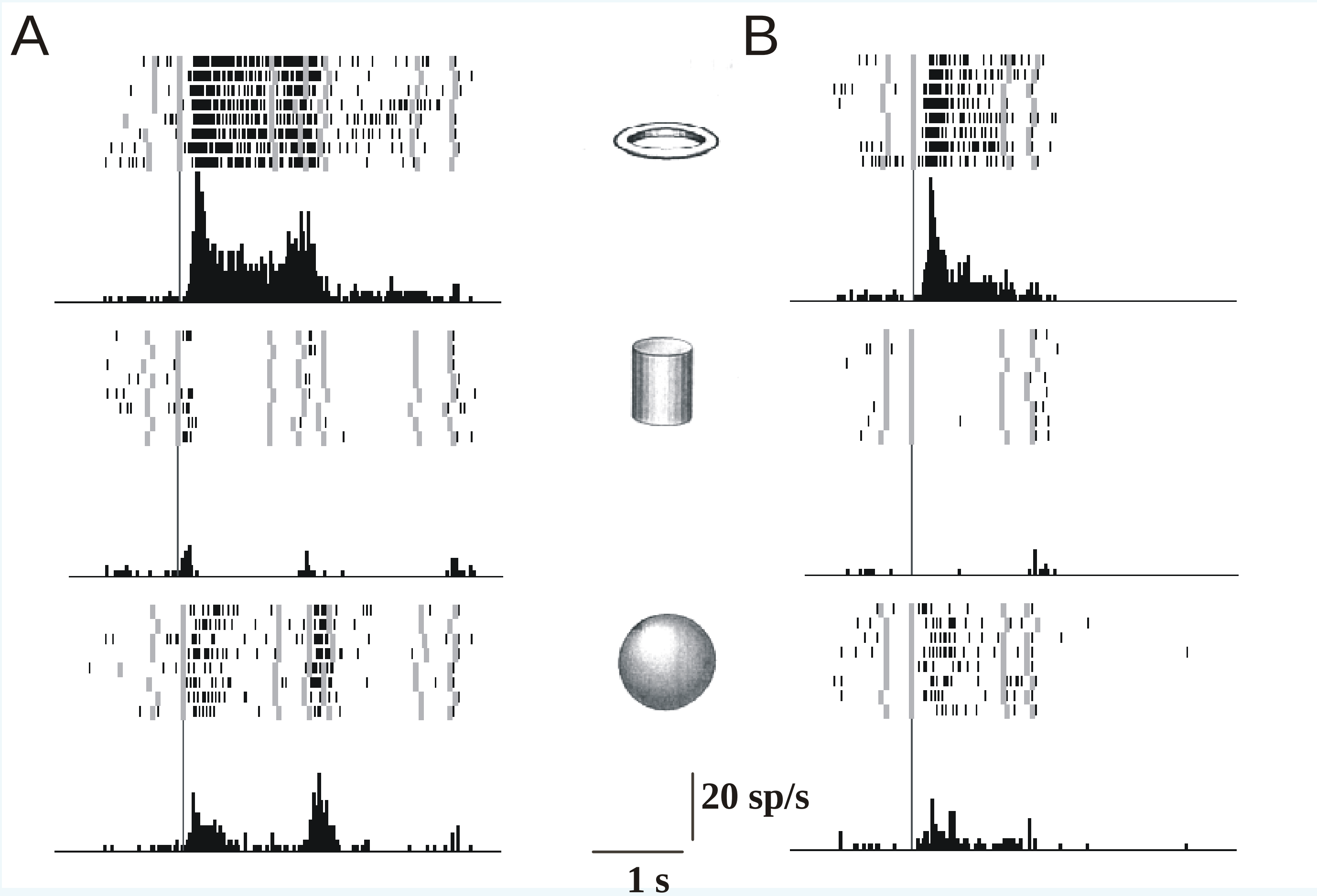
Even more peripheral
(multiple body parts)



Grasping neurons



F5 canonical neurons

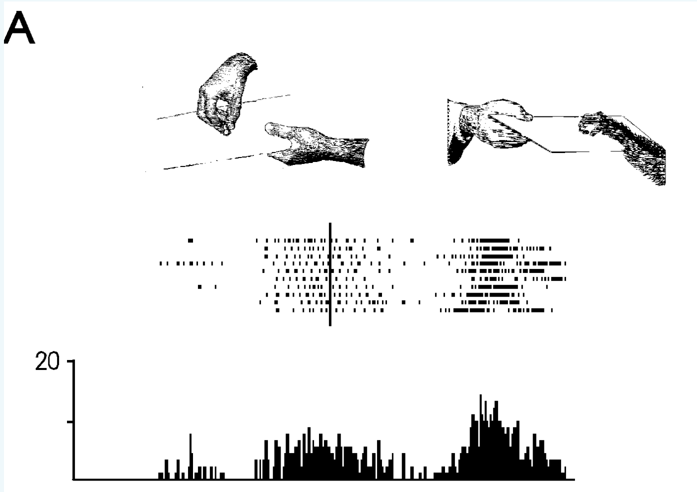


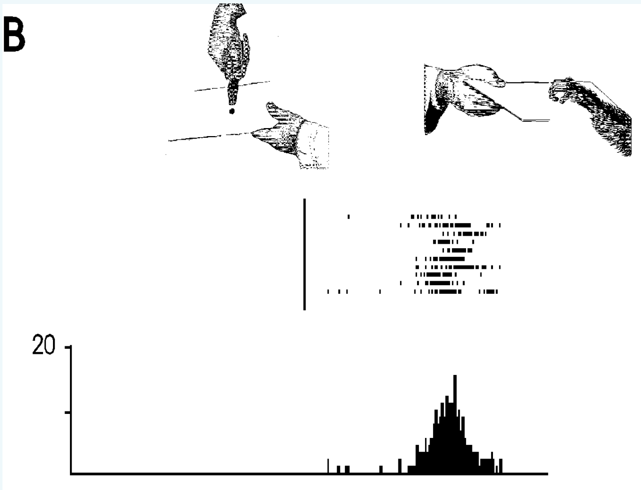
Observation + action

Observation only

Mirror Neurons

The neuron is activated by “seeing” someone else’s hand performing a manipulative action **and** while the monkey is performing the same action

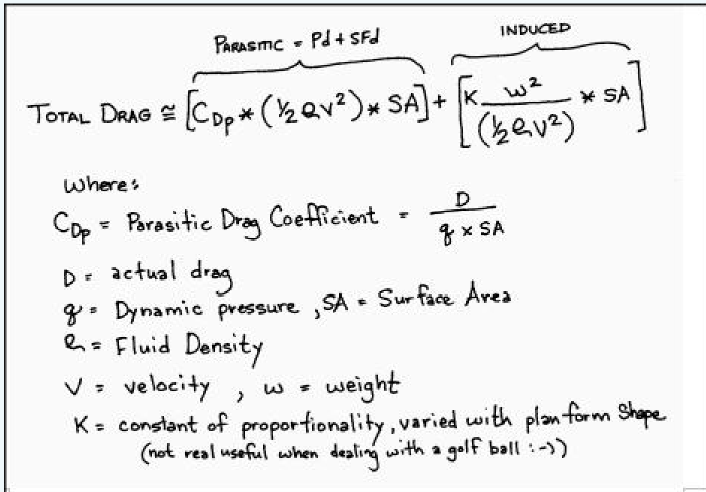


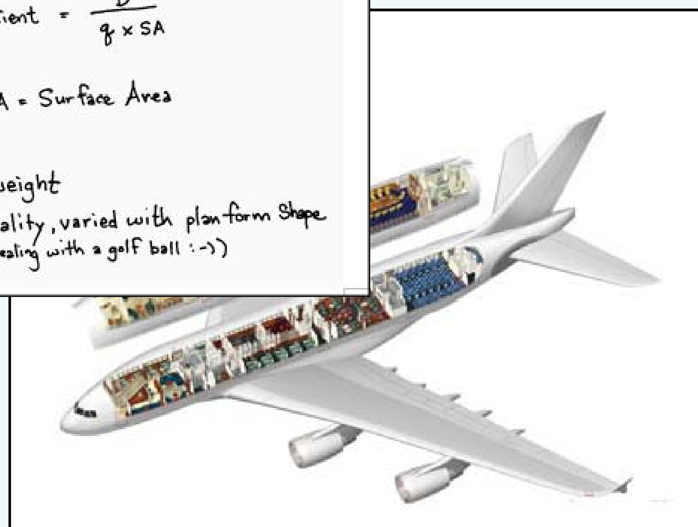


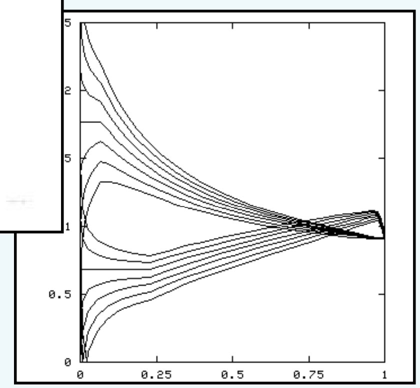
The type of action seen is relevant

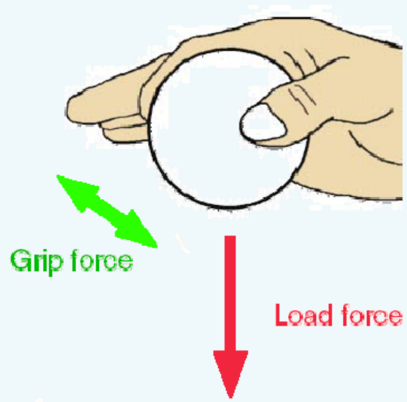
From: Fadiga, L., L. Fogassi, V. Gallese, and G. Rizzolatti, *Visuomotor Neurons: ambiguity of the discharge or "motor“ Perception?* Internation Journal of Psychophysiology, 2000. **35**: p. 165-177.

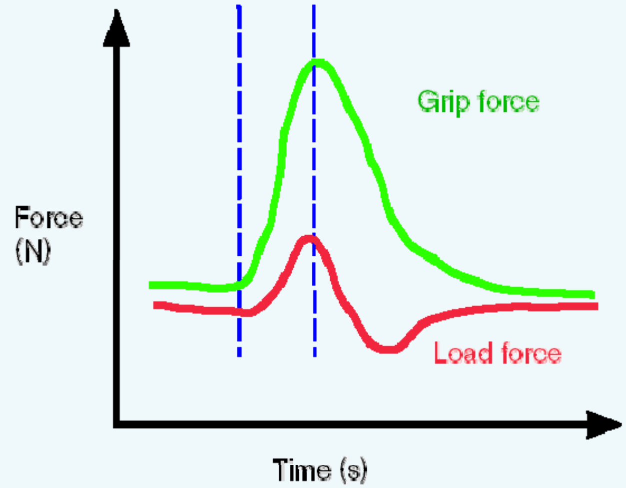
Models

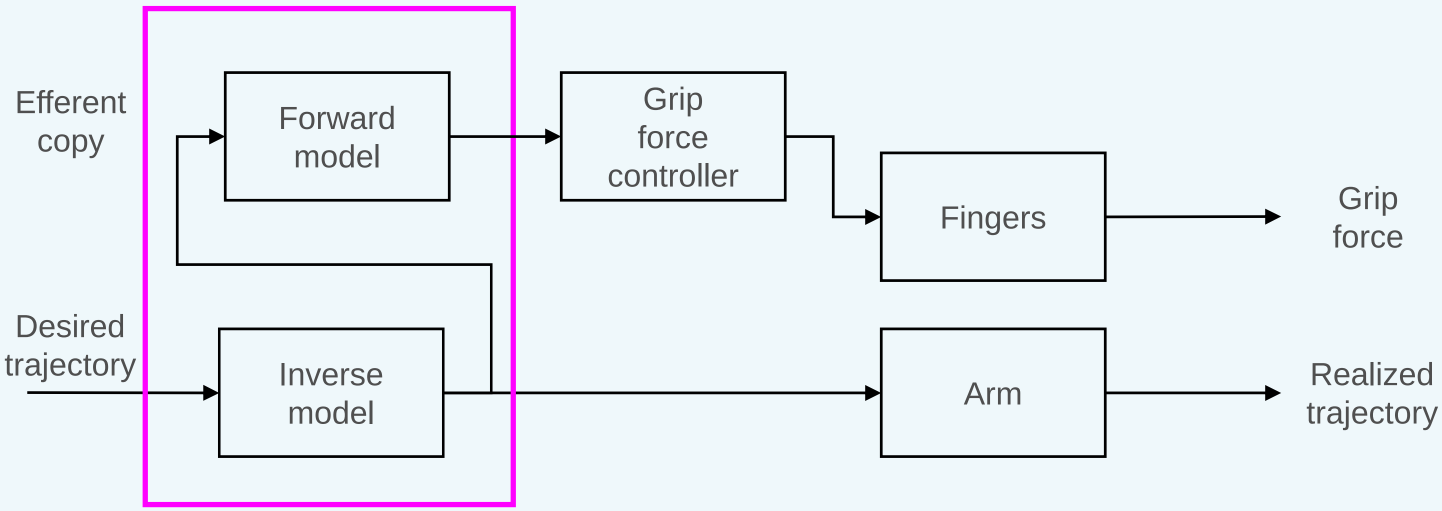

PARASITIC = Pd + SFd
INDUCED
TOTAL DRAG ≅ [CDp * (½ ϱ V²) * SA] + [K w² / (½ ϱ V²) * SA]
where:
CDp = Parasitic Drag Coefficient = D / q × SA
D = actual drag
q = Dynamic pressure , SA = Surface Area
ϱ = Fluid Density
V = velocity , w = weight
K = constant of proportionality, varied with planform Shape
(not real useful when dealing with a golf ball :-))



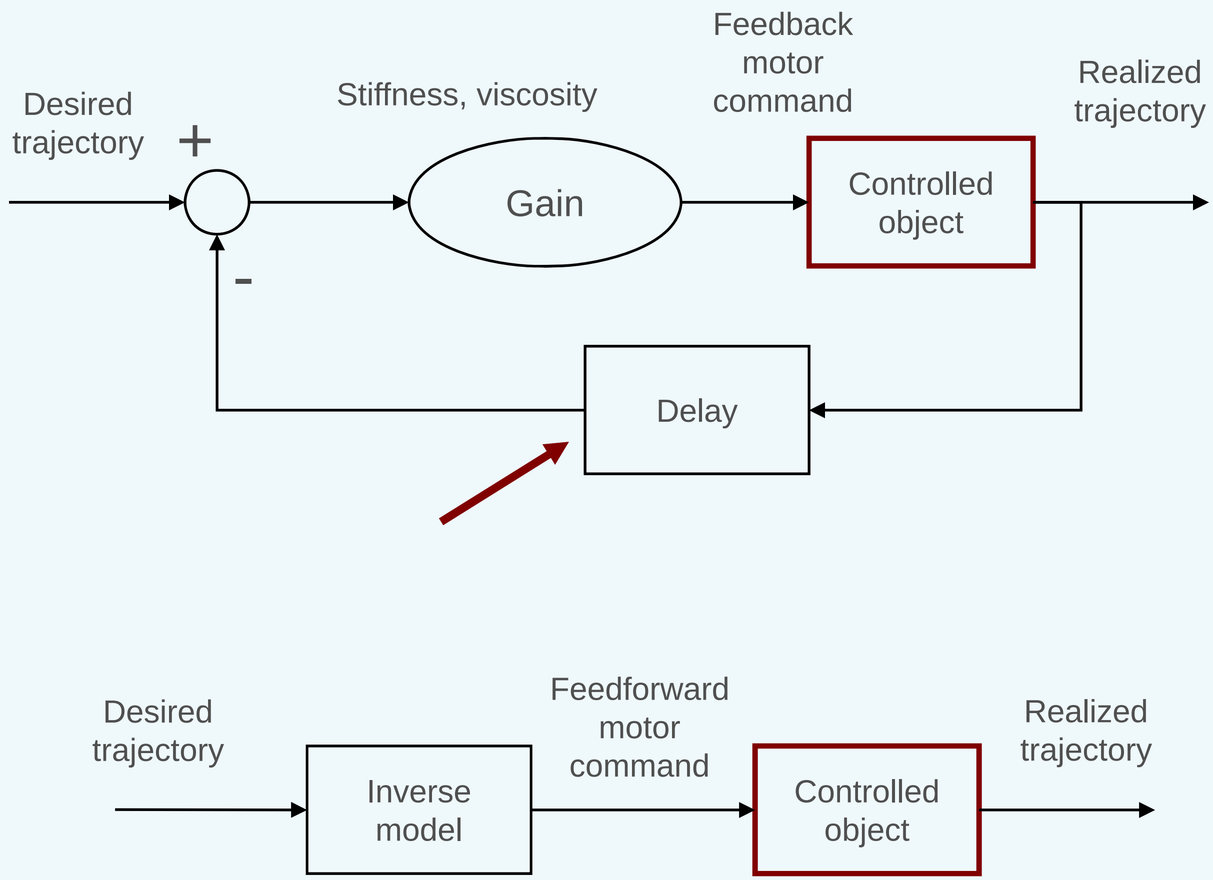

0
0.5
1
0.25
0.75







<u>From:</u> Flanagan JR, Wing AM. ***The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads.*** **Journal of Neuroscience 1997, 17:1519-1528.**



From: M. Kawato. *Internal models for motor control and trajectory planning*. Current Opinion in Neurobiology 1999, 9:718–727

Effect of delays

Make feedback control either poor or unstable altogether

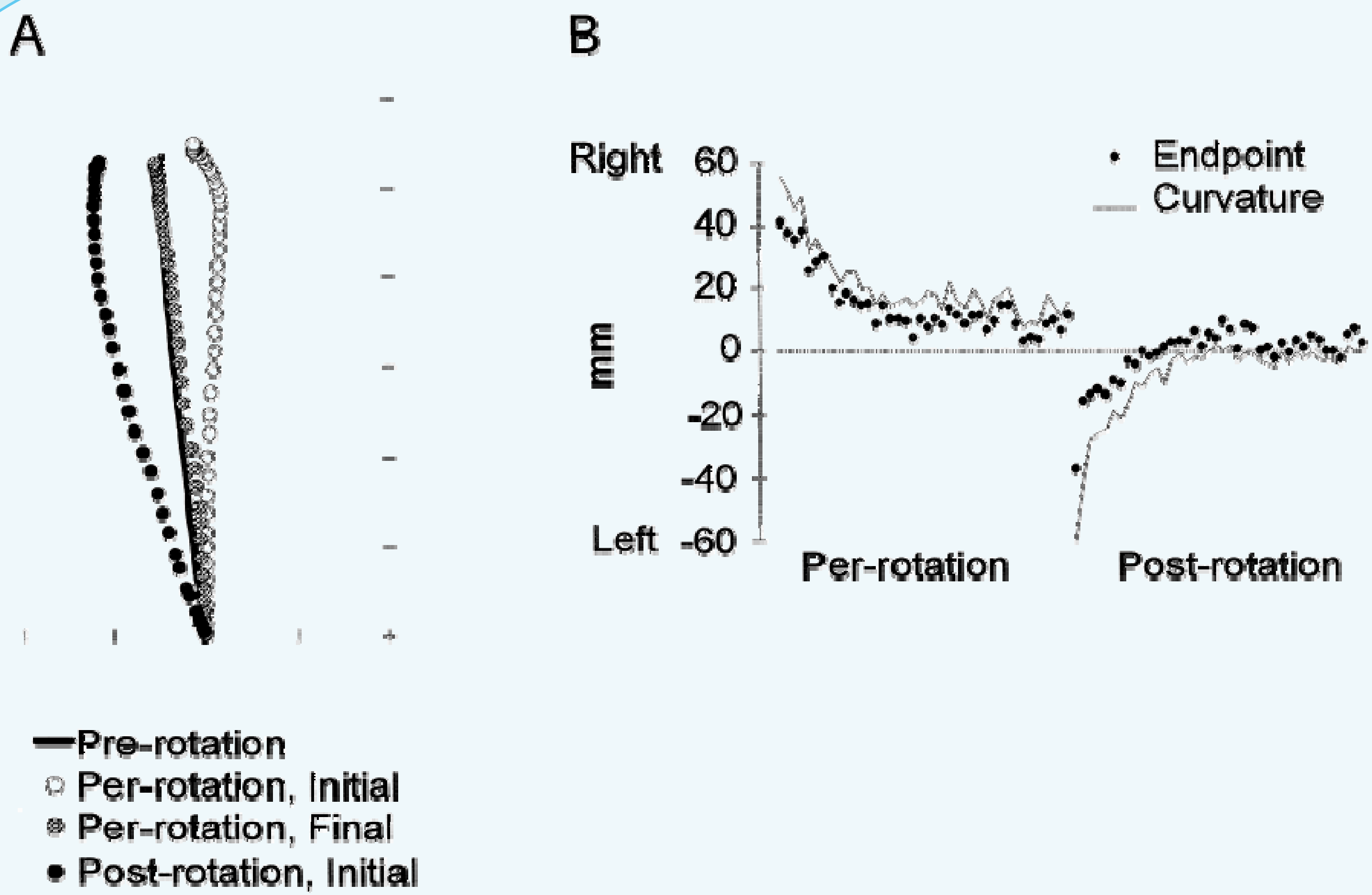
- Engineering control systems
  - Delays: 500μs
  - Movement duration: seconds
  - Gain of the controller: can be made high
- BTW: maintaining fast control loops is not an easy feat

- Humans
  - Delays:
    - 20-50ms (spinal)
    - 150-250ms (vision)
  - Movement duration: 150-500ms
  - Gain of the controller: stiffness and muscle viscoelastic properties (comparatively low)

Building the internal models

- Rote-learning would be unpractical:
  - Too many possible actions (dof) for the available number of neurons (although they're quite a lot!)
- Generalizing past experience:
  - Past experience is bound (unfortunately) to "represent" only a portion of the whole state space
- Developing and extending the control structure to new behaviors
  - Sequencing and combining primitive behaviors appropriately
- Predicting the future course of action
  - It might turn out to be useful!

RobotCub



From: Lackner JR, Dizio P. *Gravitoinertial force background level affects adaptation to Coriolis force perturbations of reaching movements.* Journal of Neurophysiology 1998, 80:546-553.

Internal models for interception

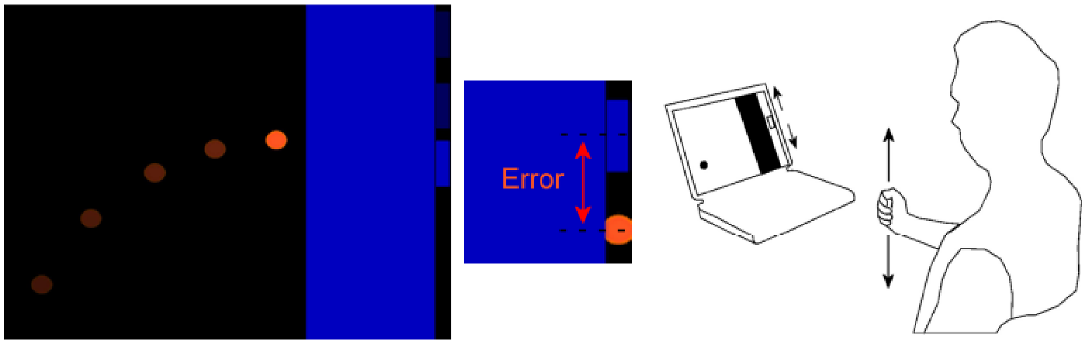


Fig. 1: Experimental setup: (right) a schema of the game played by subjects; (center) error measure; (left) a sketch of an experiment. **28 subjects** participated in the experiment.

Experimental setup

Stimuli (3 conditions)

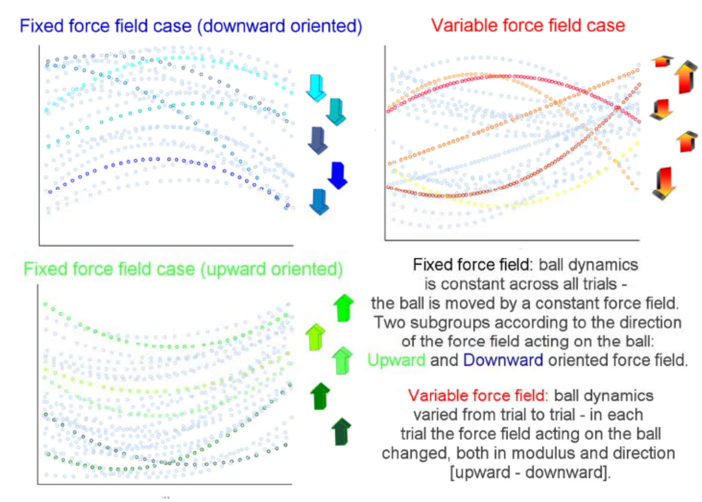
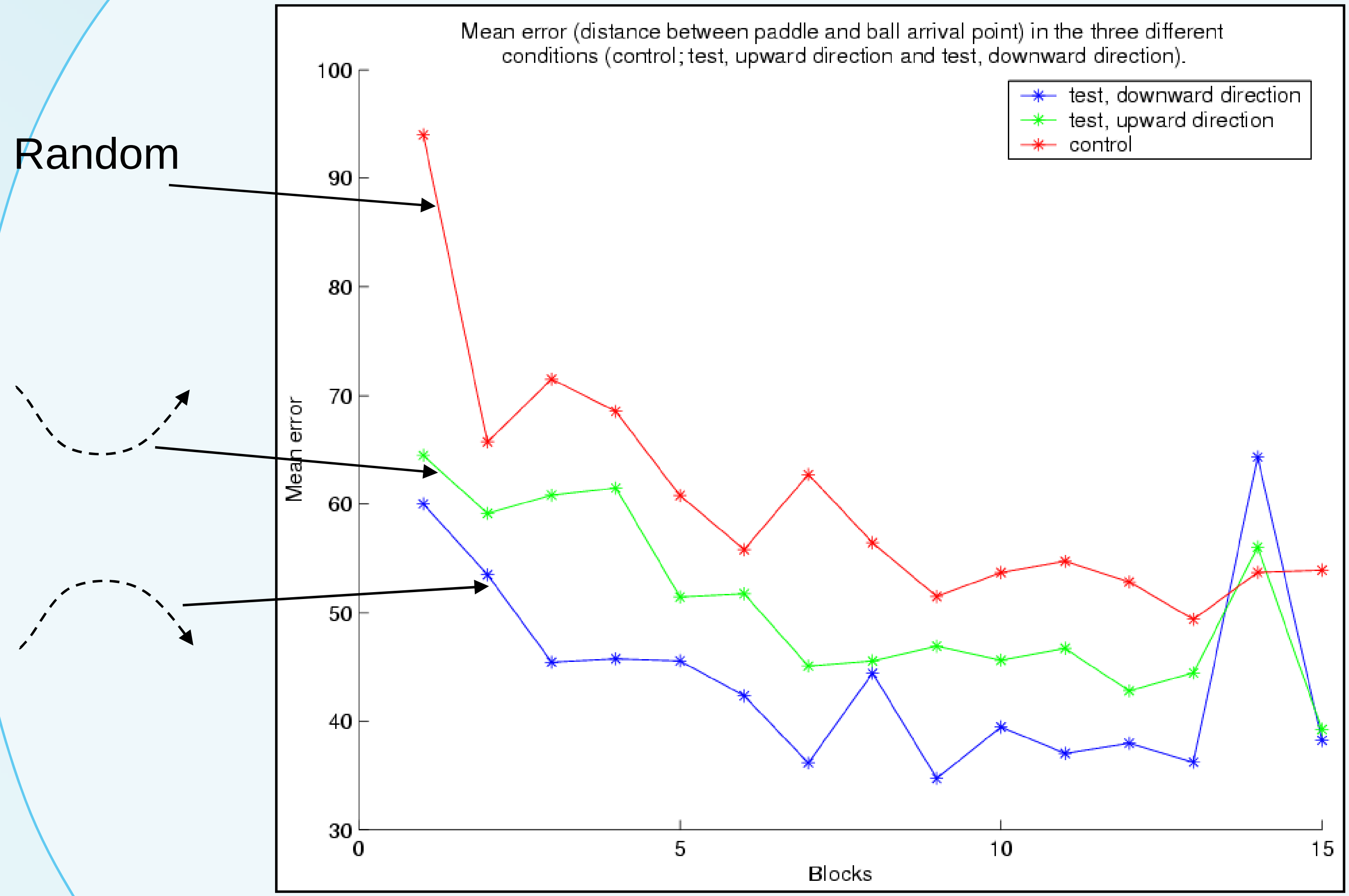


Fig. 2: Examples of ball trajectories and representation of the force fields applied to the ball in the three experimental cases.

With Alessandra Sciutti, Francesco Nori, Thierry Pozzo

Some results



Experiment with motor information

The wind tunnel

The iCub: quick summary

The **iCub** is the humanoid baby-robot being designed as part of the **RobotCub** project

- The iCub is a full humanoid robot sized as a three and half year-old child.
- The total height is 104cm.
- It has 53 degrees of freedom, including articulated hands to be used for manipulation and gesturing.
- The robot will be (once the software is done) able to crawl and sit and autonomously transition from crawling to sitting and vice-versa.
- The robot is GPL/FDL: software, hardware, drawings, documentation, etc.

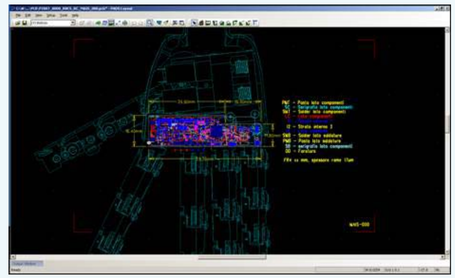
Degrees of freedom

- 53: 9 in each hand
- Sensors: position, torque, temperature
- And also: cameras, microphones, gyroscopes, linear accelerometers
- For the future: tactile sensors, skin…
  - Low-resolution:
    - Distributed many sensing points
  - Fingertips:
    - Localized, high-resolution

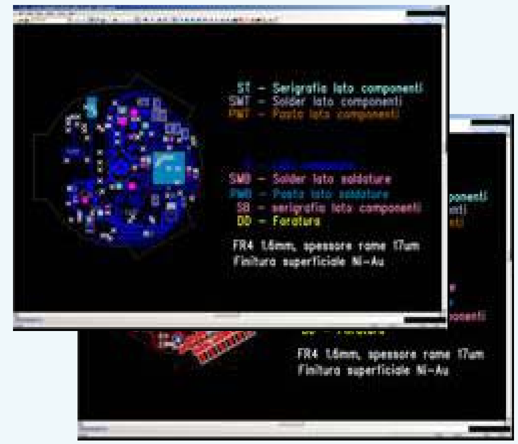
* Old version

A few examples

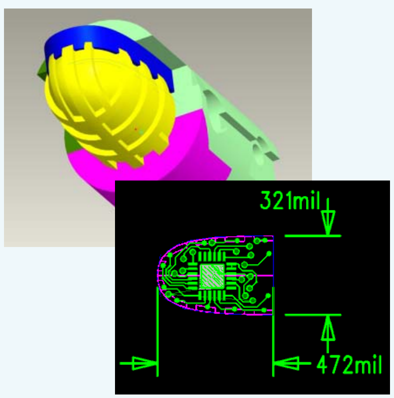
Custom design



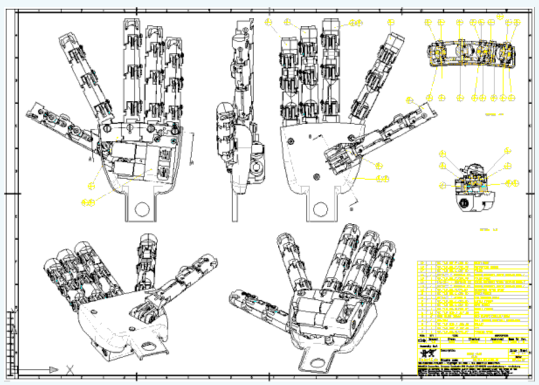
Hand sensor sampling PIC-based card



Force/torque sensor fitted into the sensor



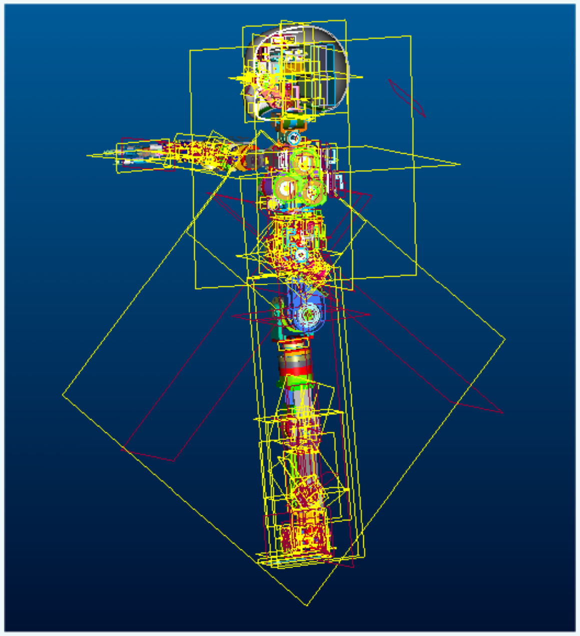
Fingertip sensorization



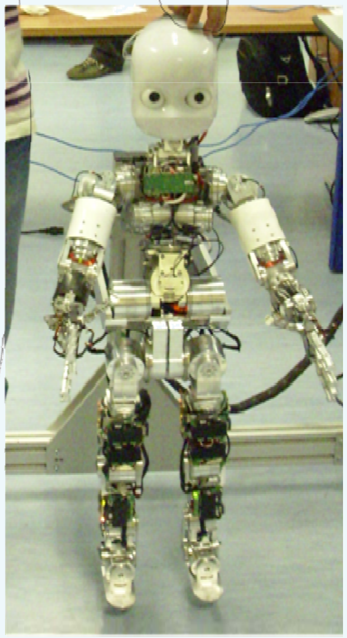
Design and documentation

Wired with 25micron coated wires

Current status

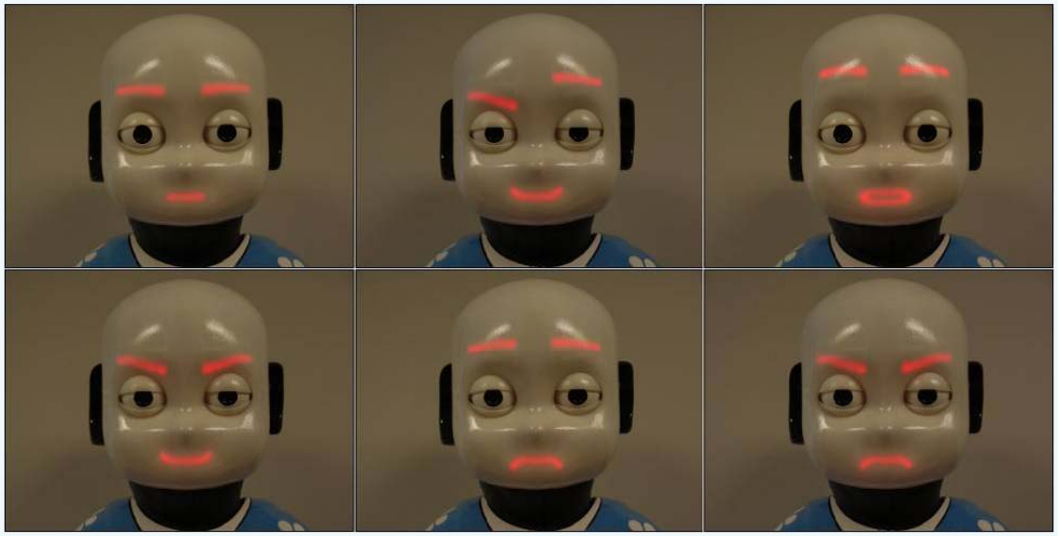


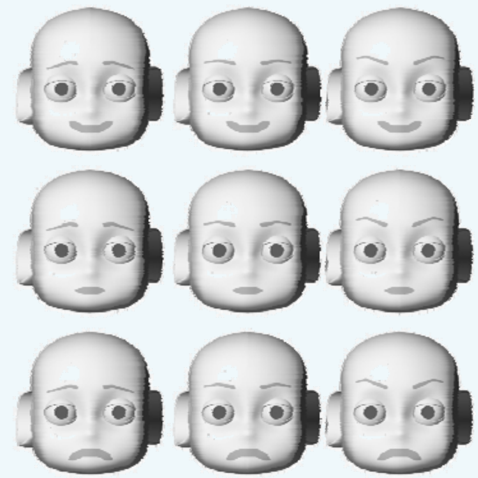
Design

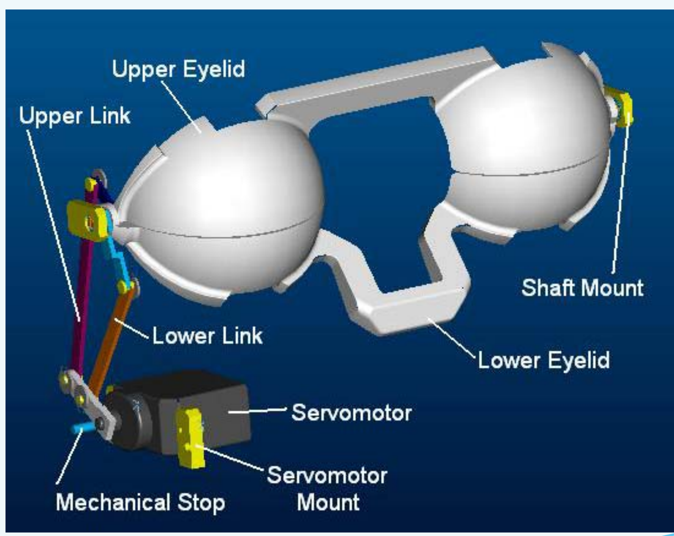


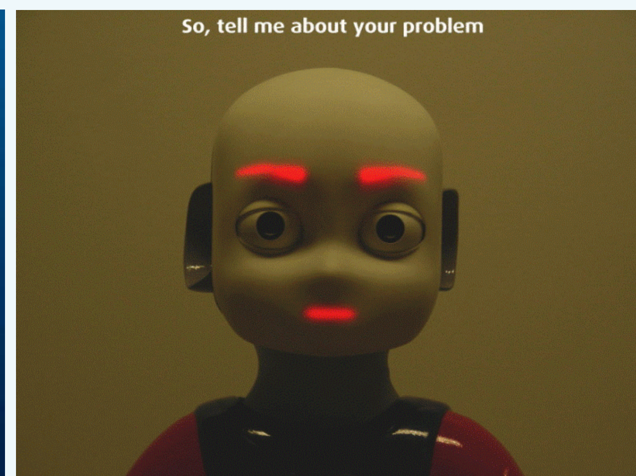
Reality

Facial expressions



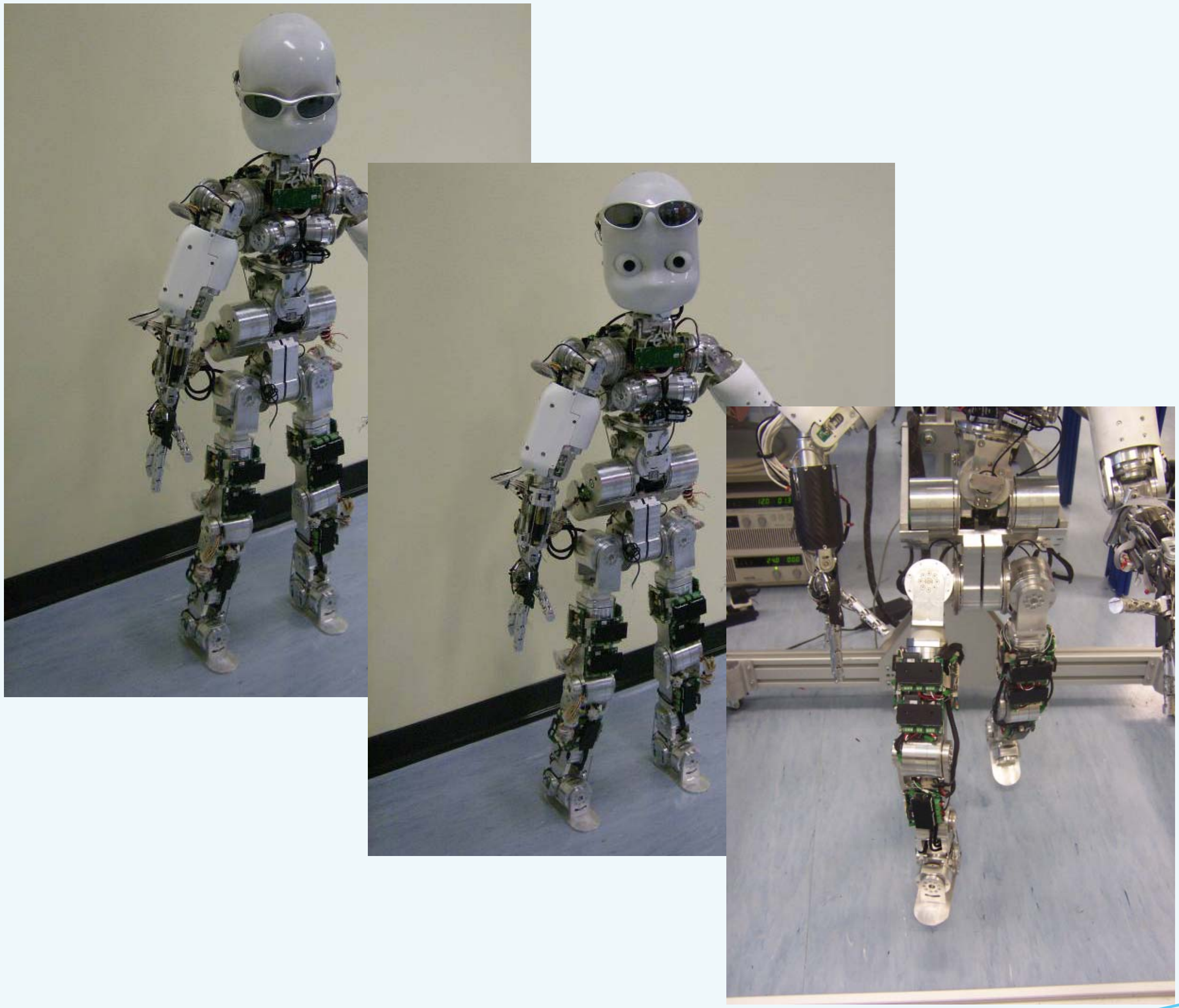






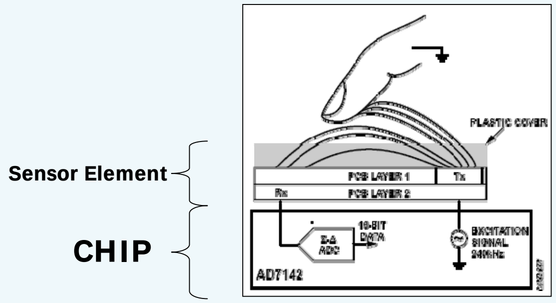
The soul of a new machine

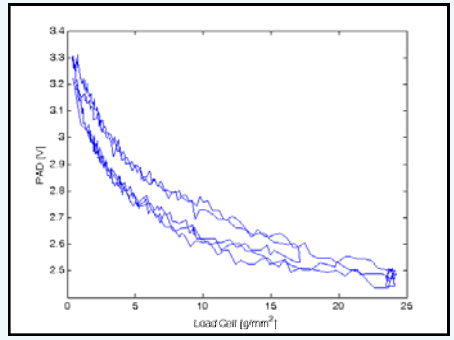
More integration

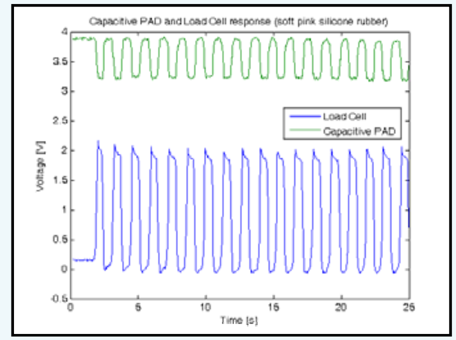


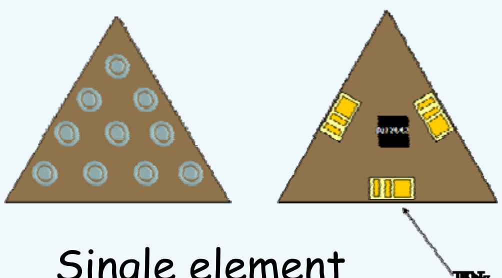
… and robot Yoga

Skin, architecture by Marco Maggiali

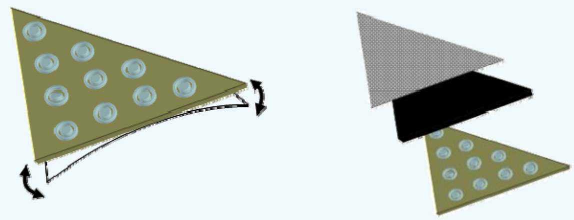




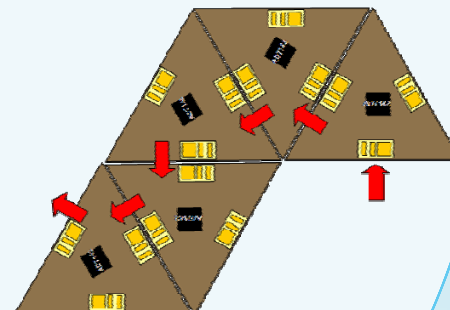




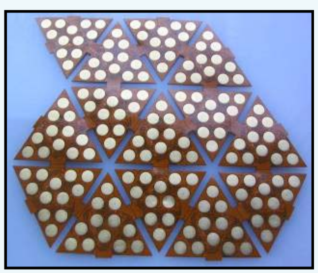
Single element

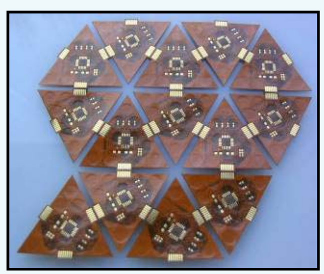


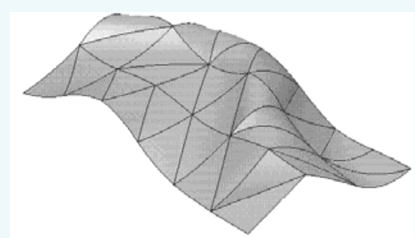
Flexible PCB



Architecture

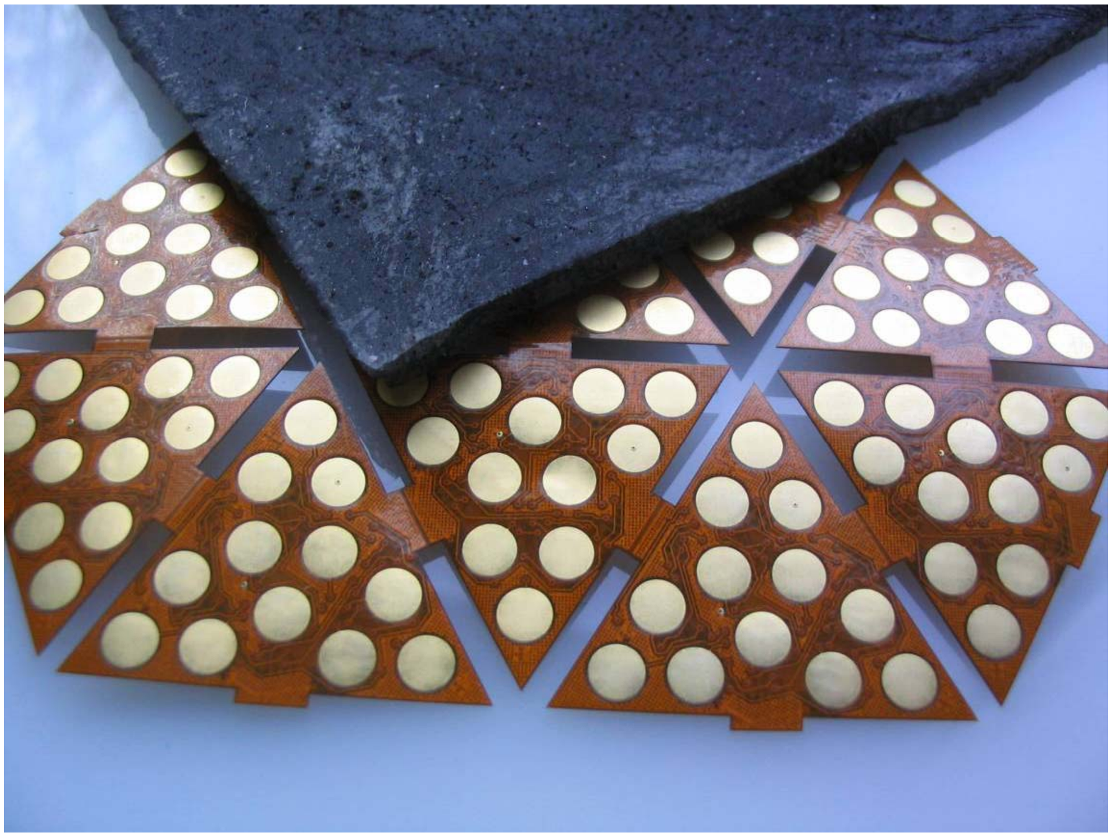


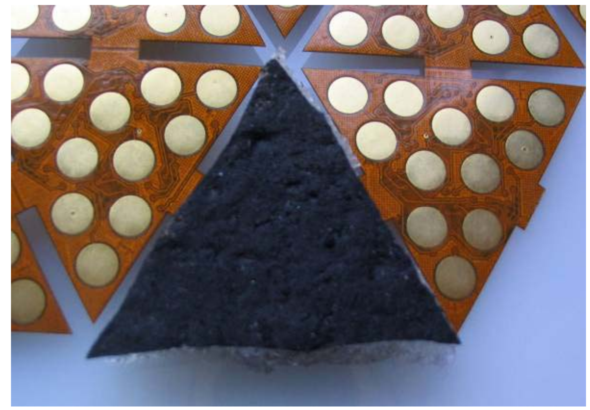


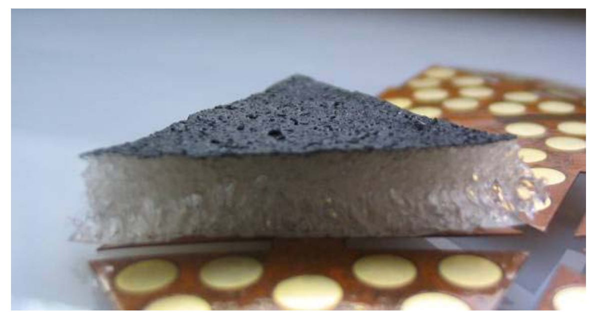


With Giorgio Cannata (Univ. of Genoa)

Prototype covered by silicon rubber

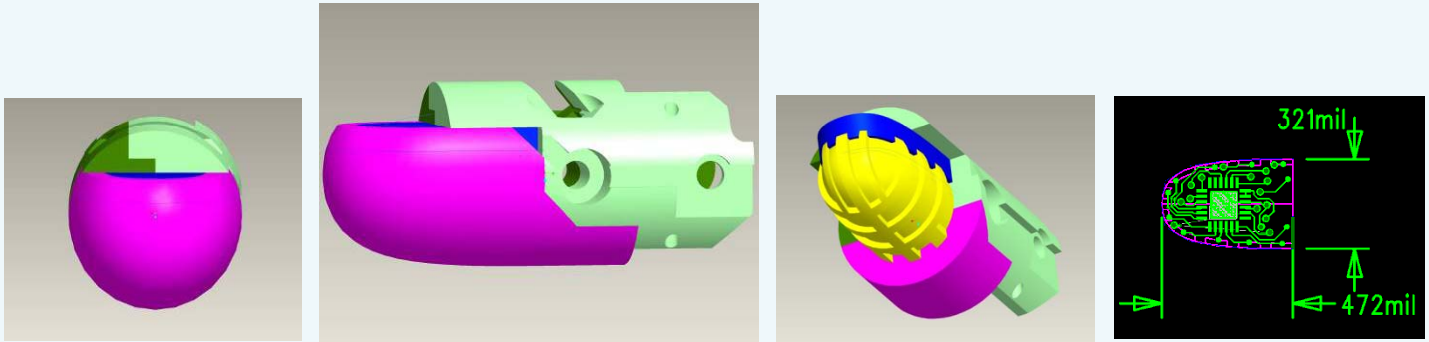


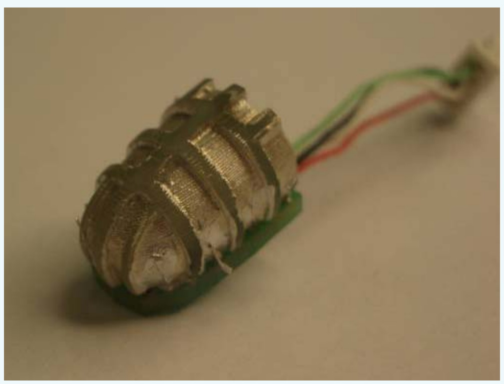




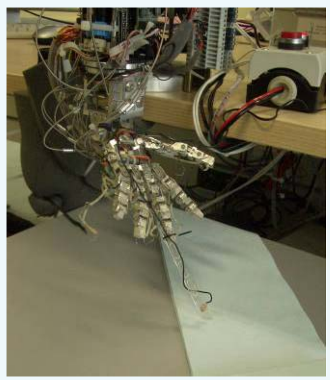
Fingertip sensors

CAD/concept

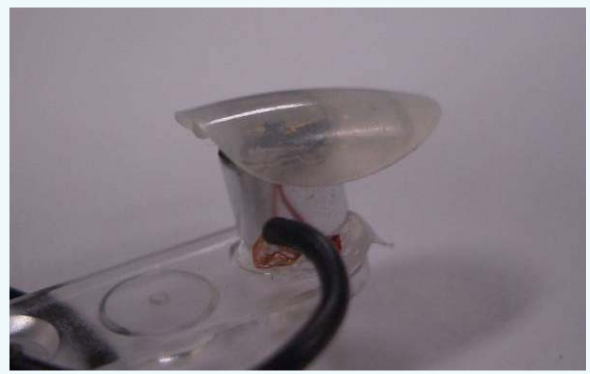




Prototype



Some testing



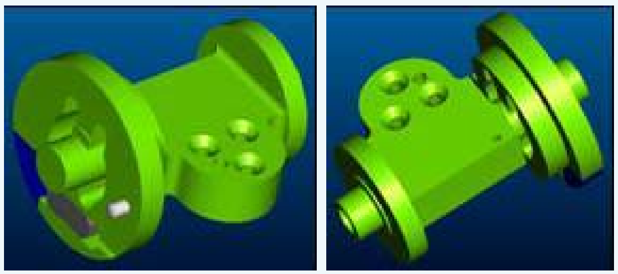
Fingernail + microphone

By Alexander Schmitz, Marco Randazzo,
Marco Maggiali and Lorenzo Natale

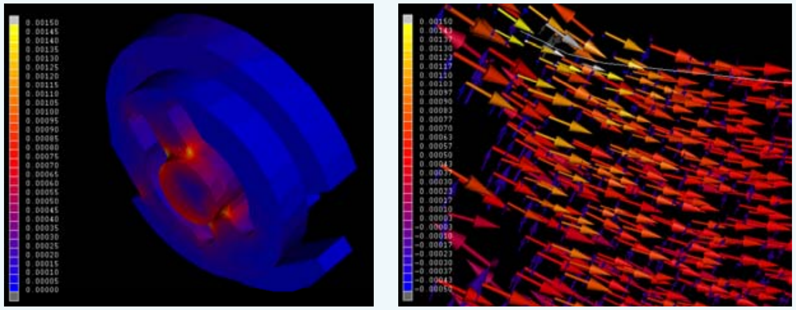
RobotCub

The hands

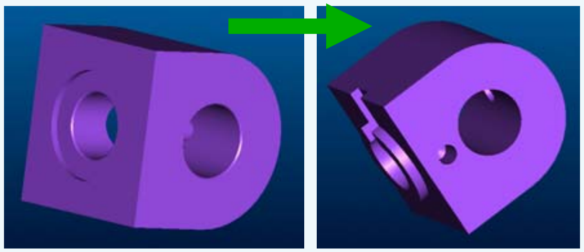
Joint-level torque sensing



Existing parts



FEM analysis of deformation

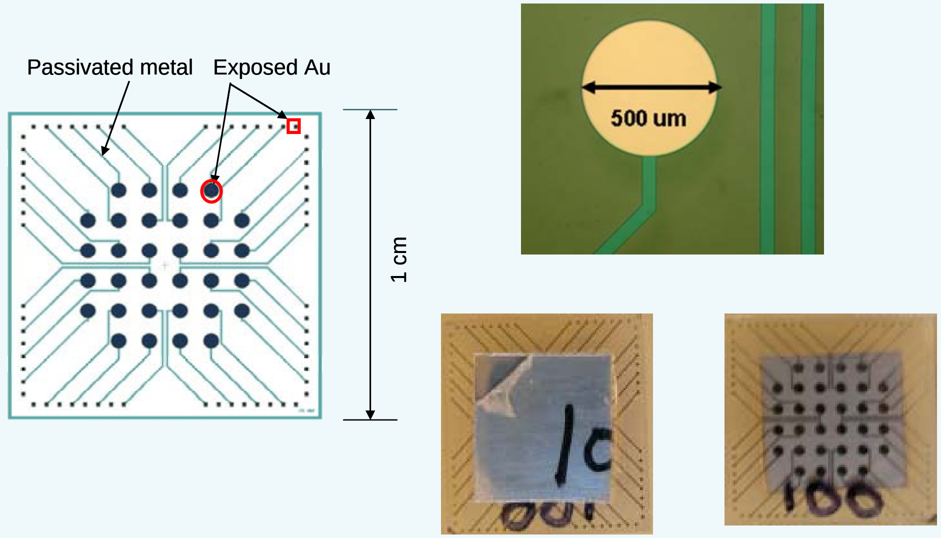


Changes (under implementation)

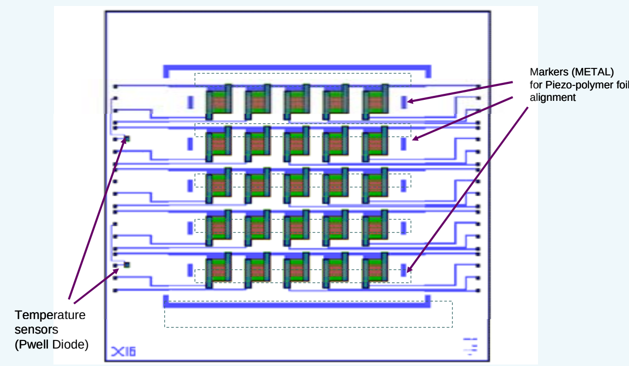
Changes (under implementation)

By Alberto Parmiggiani and collaborating with Nikos Tsagarakis

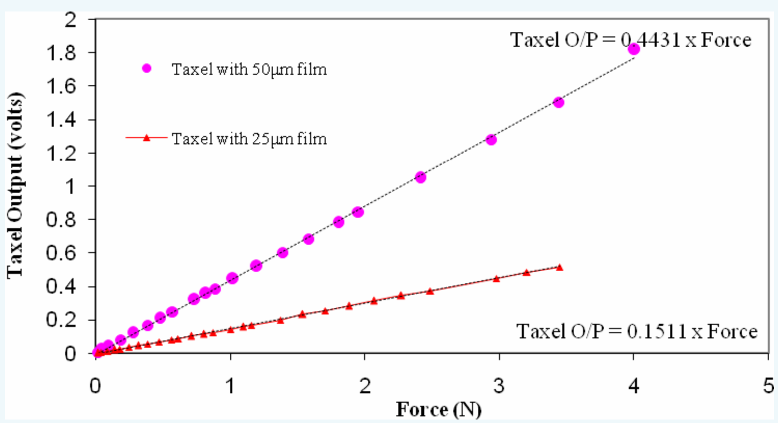
High-res tactile sensors



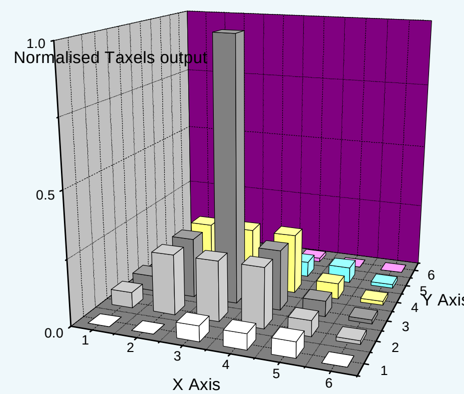
Contact (transistor gates) + PVDF



FETs

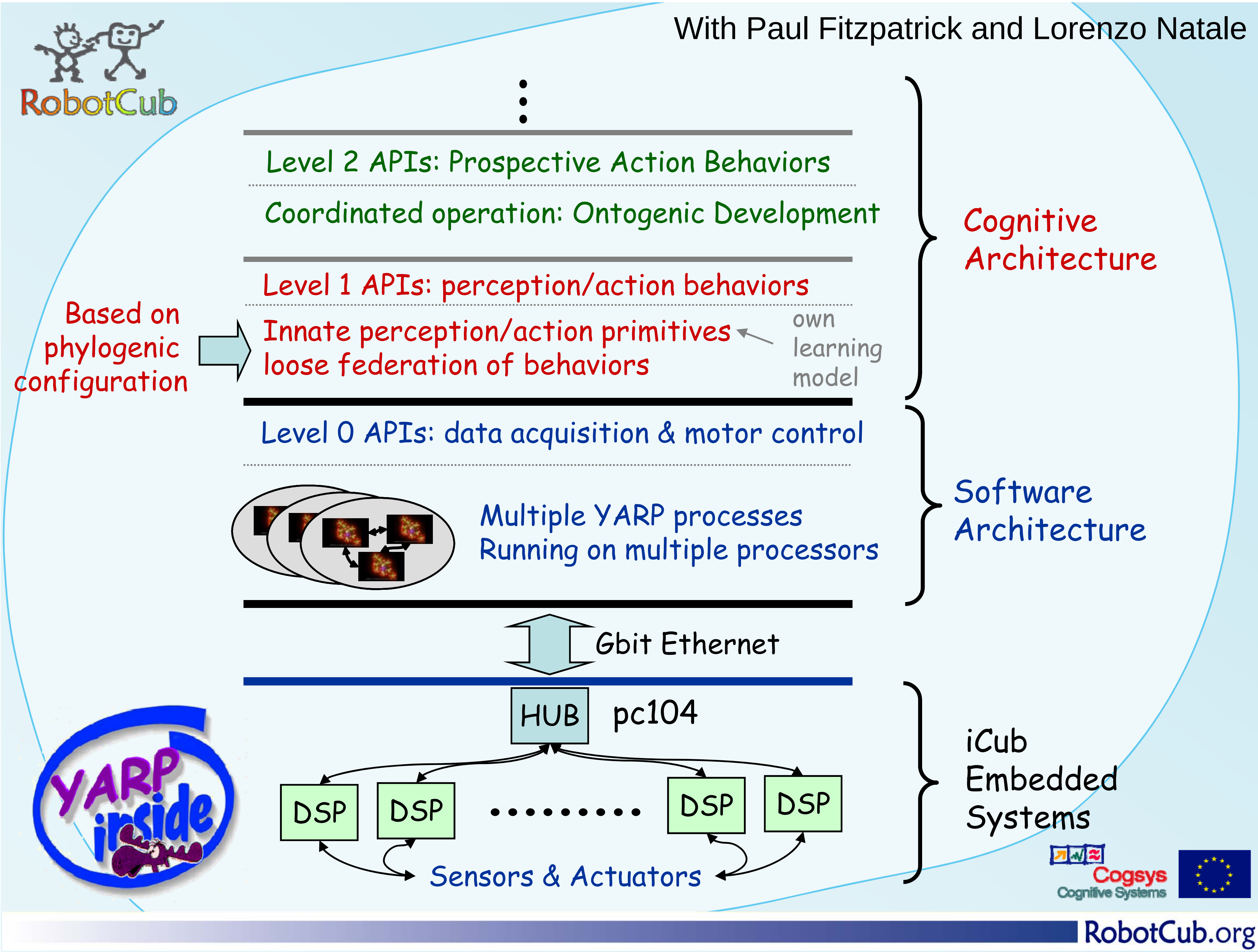


Testing the piezo- film deposition



Cross-talk btw. taxels

By Ravinder Dahiya, Maurizio Valle (Unige) & Leandro Lorenzelli (Trento)


With Paul Fitzpatrick and Lorenzo Natale
RobotCub
Level 2 APIs: Prospective Action Behaviors
Coordinated operation: Ontogenic Development
Level 1 APIs: perception/action behaviors
Innate perception/action primitives
loose federation of behaviors
own learning model
Based on phylogenic configuration
Cognitive Architecture
Level 0 APIs: data acquisition & motor control
Multiple YARP processes
Running on multiple processors
Software Architecture
Gbit Ethernet
HUB
pc104
DSP
DSP
DSP
DSP
Sensors & Actuators
iCub Embedded Systems
YARP inside
Cogsys
Cognitive Systems
RobotCub.org

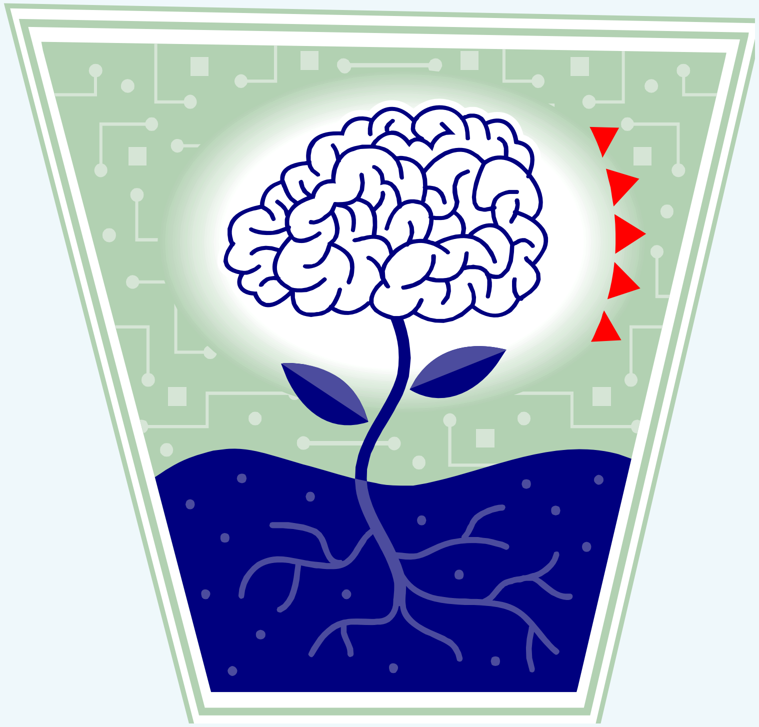
…more work in progress

- Ryo Saegusa: sensory prediction
- Andrew Dankers: models of vision
- Matteo Fumagalli: force control
- Boris Duran: dynamical systems control
- Lorenzo Jamone: grasping
- Serena Ivaldi: optimal control
- Massimiliano Izzo: internal models

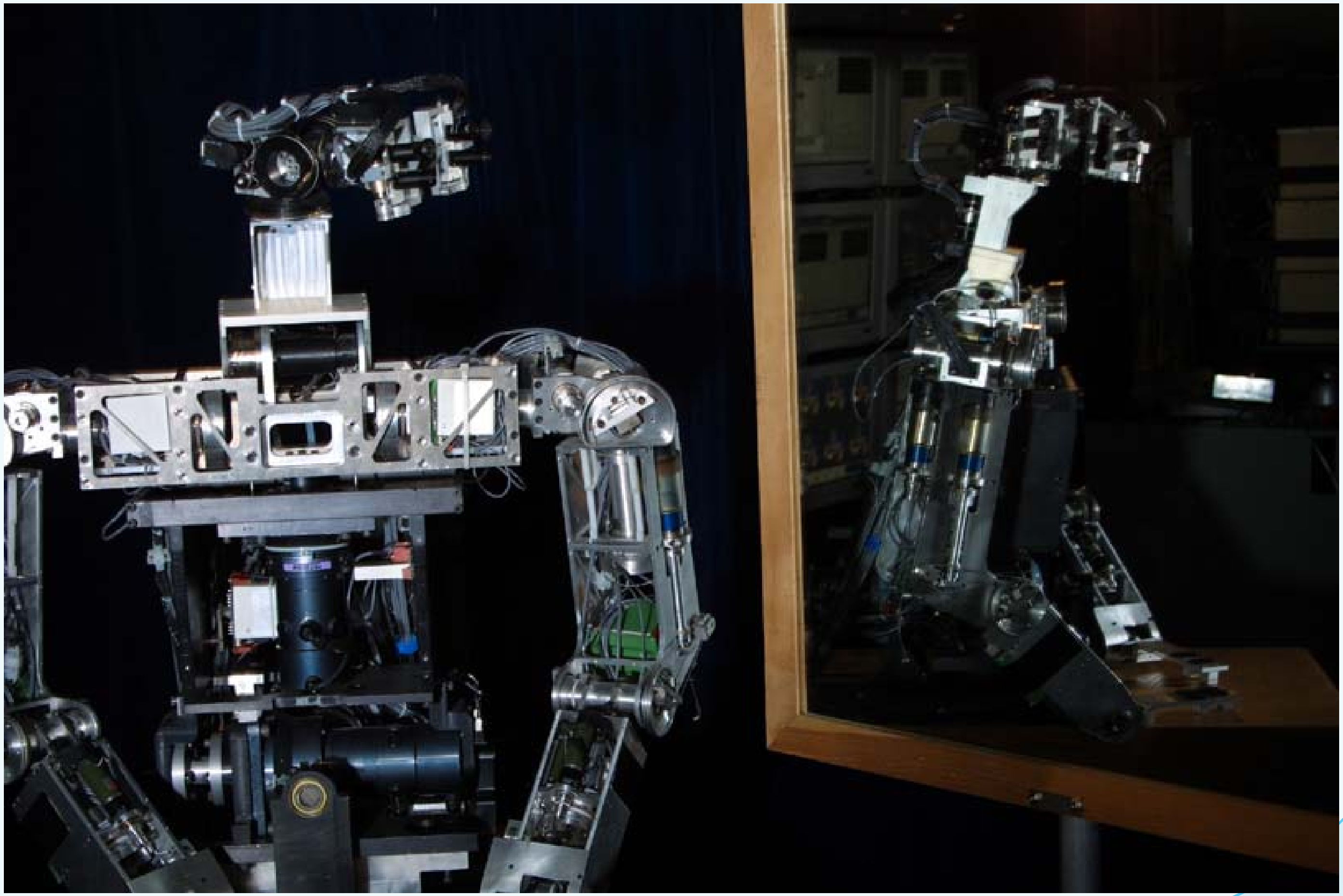
Lots of people

- Lorenzo Natale, Francesco Nori: Software, testing, calibration
- Marco Maggiali, Marco Randazzo: firmware, DSP libraries, tactile sensing
- Francesco Becchi, Paolo Pino, Giulio Maggiolo, Gabriele Careddu: design and integration
- Gabriele Tabbita, Walter Fancellu: assembly
- Nikos Tsagarakis, William Hinojosa: legs and spine, force/torque sensors
- Bruno Bonino, Fabrizio Larosa, Claudio Lorini: electronics and wiring
- Luciano Pittera: wiring
- Mattia Salvi: CAD maintenance
- Alberto Zolezzi: managing quotes, orders and spare parts
- Giovanni Stellin: hand
- Ricardo Beira, Luis Vargas, Miguel Praca: design of the head and face
- Paul Fitzpatrick & Alessandro Scalzo: software middleware
- Alberto Parmiggiani: joint level sensing
- Alexander Schmitz: fingertips
- Nestor Nava: small Harmonic Drive integration
- Ravinder Dahiya: FET-PVDF tactile senors
- Lorenzo Jamone: fingertips
- Daniel Roussy: construction
- Ludovic Righetti: simulation and initial torque specification

Other tunnels



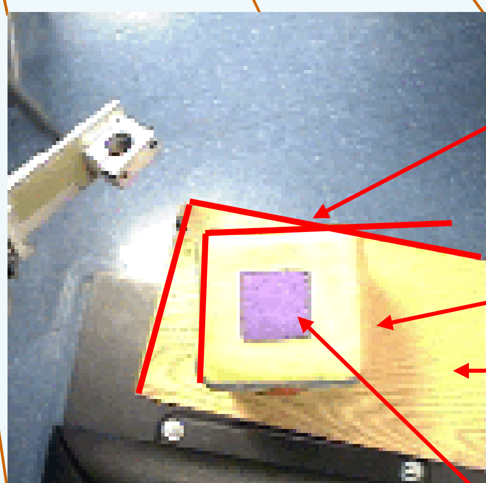
Experimental setup…



The initial idea…

Objects come to existence because they are manipulated

Fixate target | *Track visual motion…* | *(…including cast shadows)* | *Detect moment of impact* | *Separate arm, object motion* | *Segment object*



Which edge should be considered?

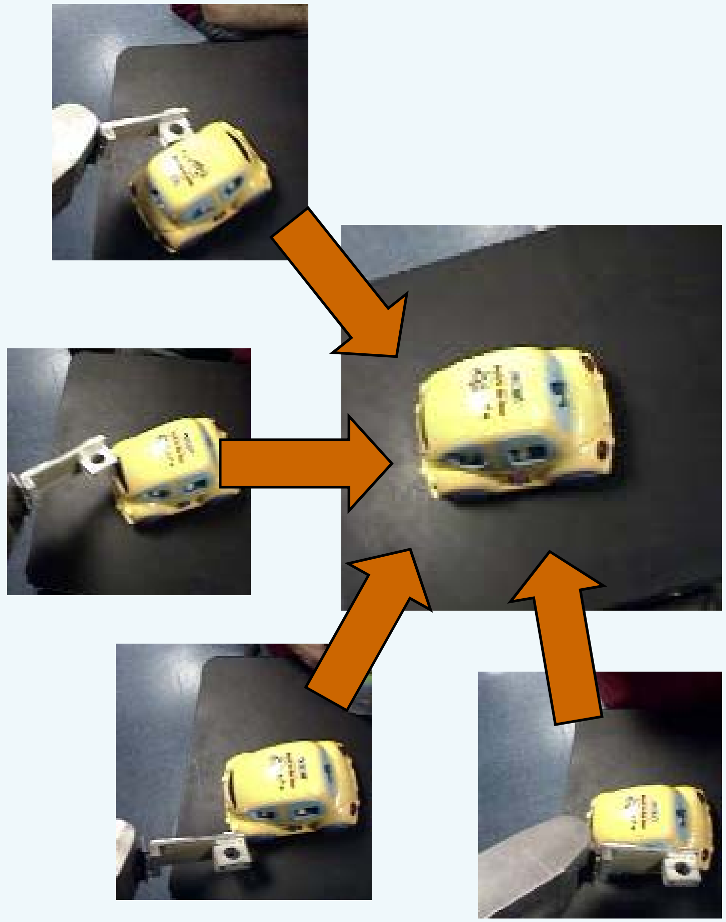
Maybe some cruel grad-student glued the cube to the table

Color of cube and table are poorly separated

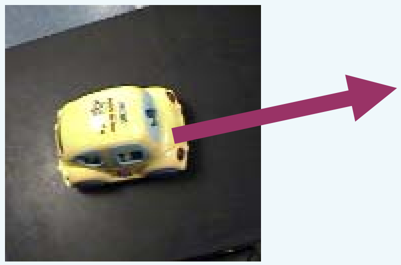
Cube has misleading surface pattern

by Paul Fitzpatrick

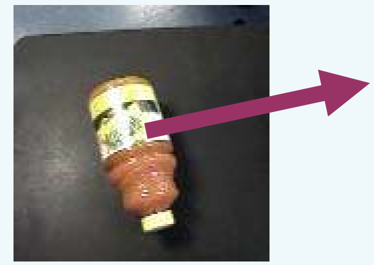
Gesture “vocabulary”



Exploring an affordance: rolling

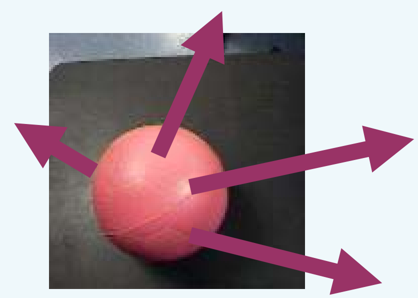


A toy car: it rolls in the direction of its principal axis



A bottle: it rolls orthogonal to the direction of its principal axis

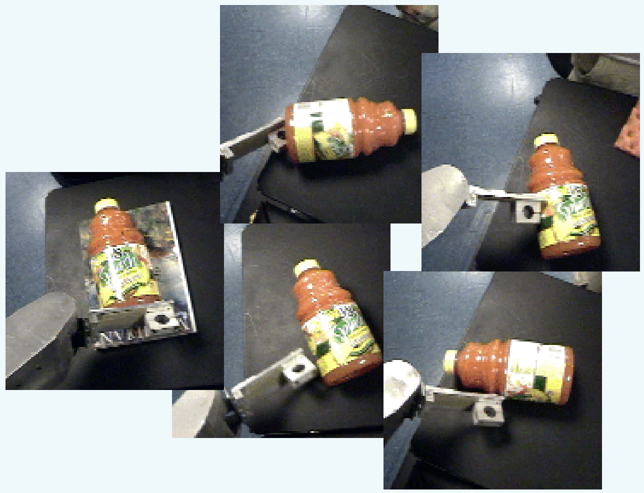
A toy cube: it doesn't roll, it doesn't have a principal axis

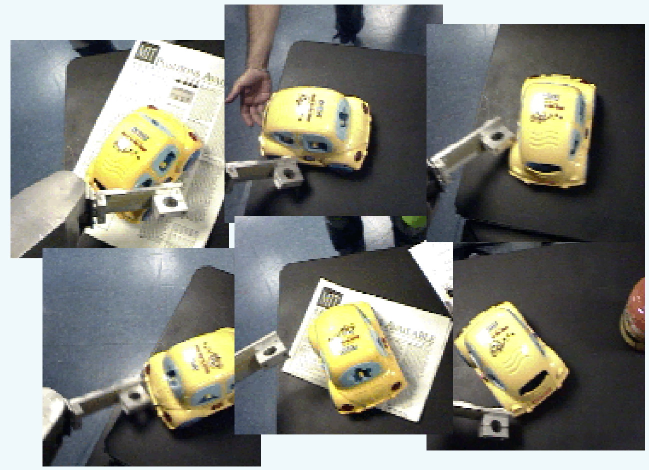


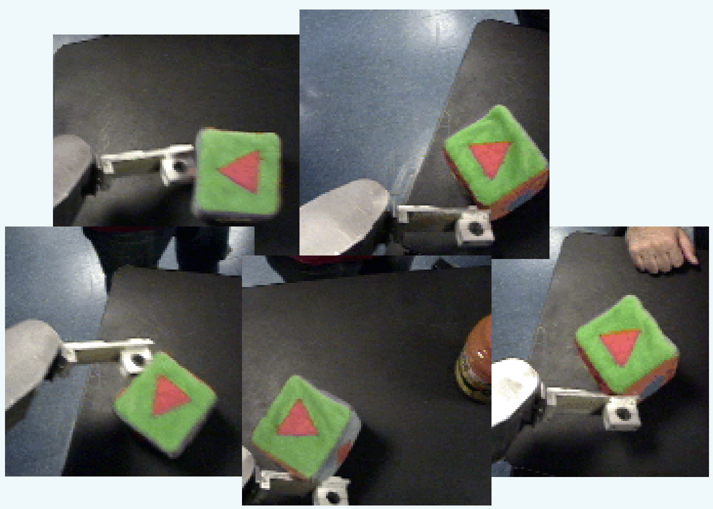
A ball: it rolls, it doesn't have a principal axis

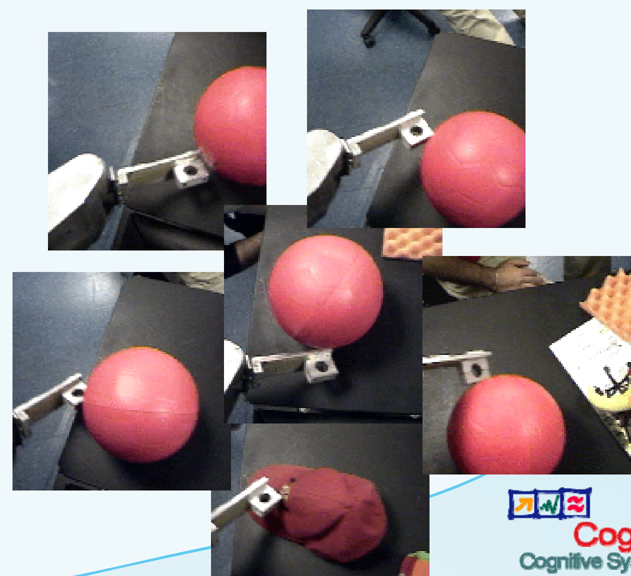
RobotCub

Forming object clusters

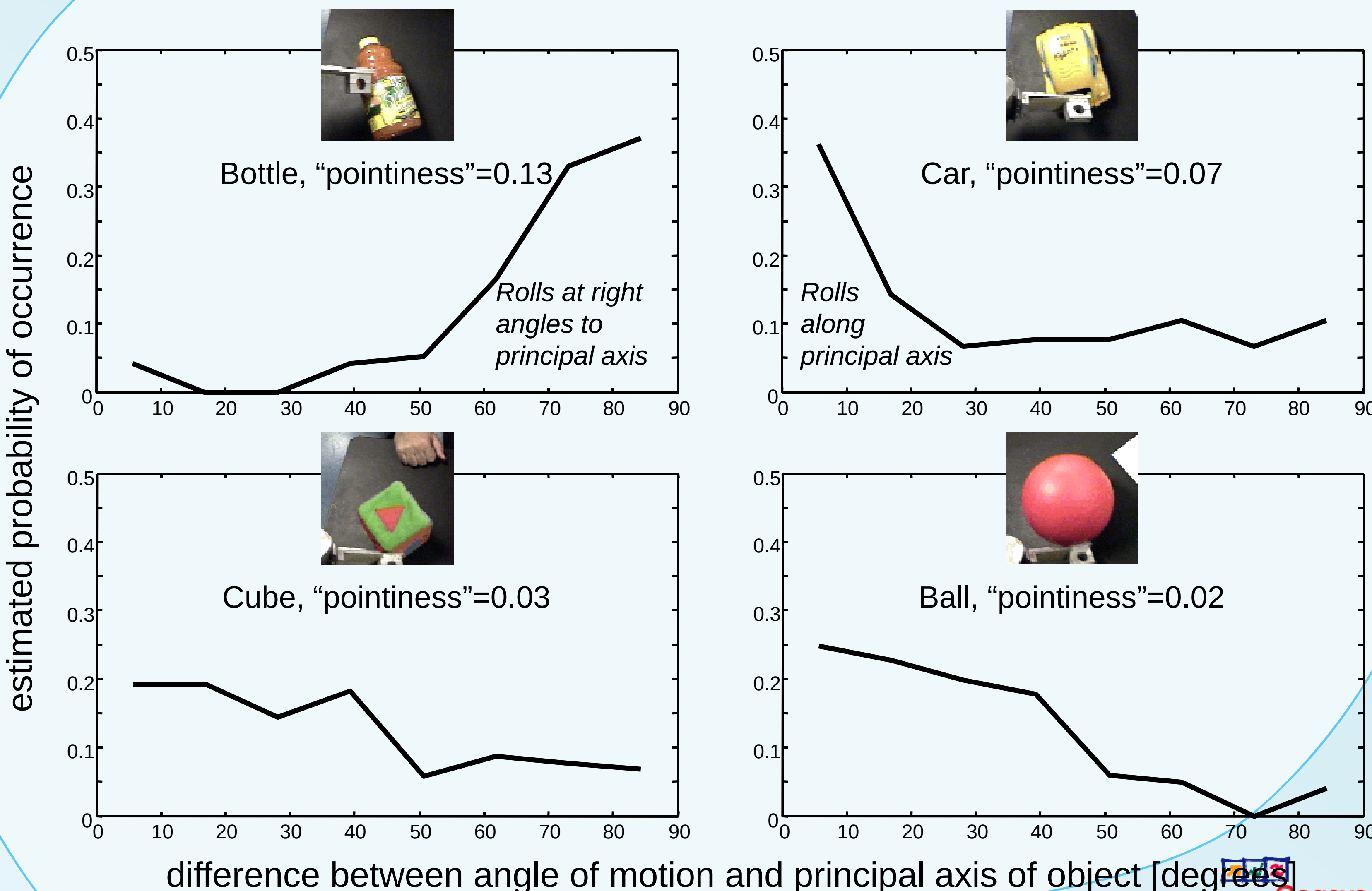




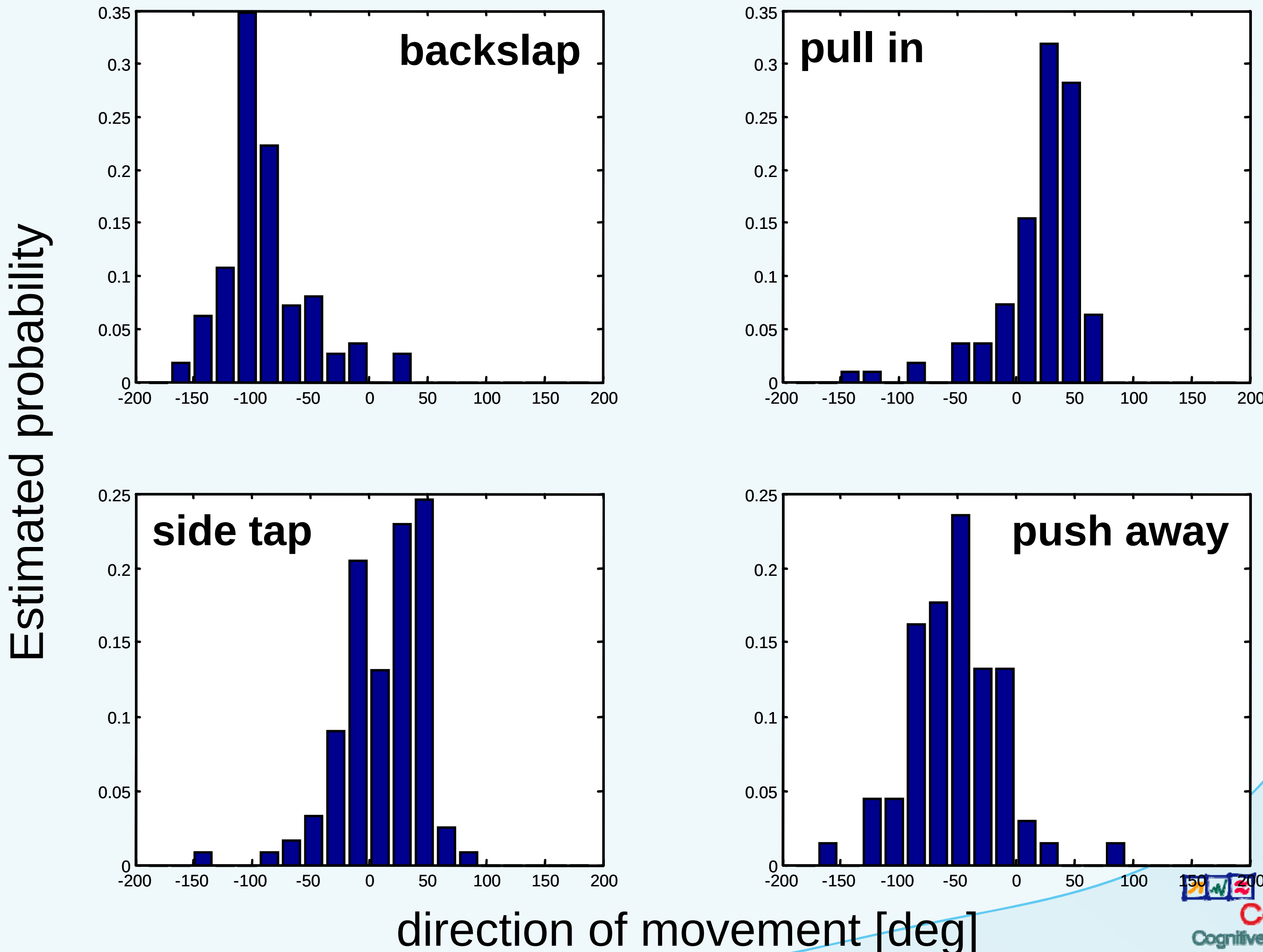




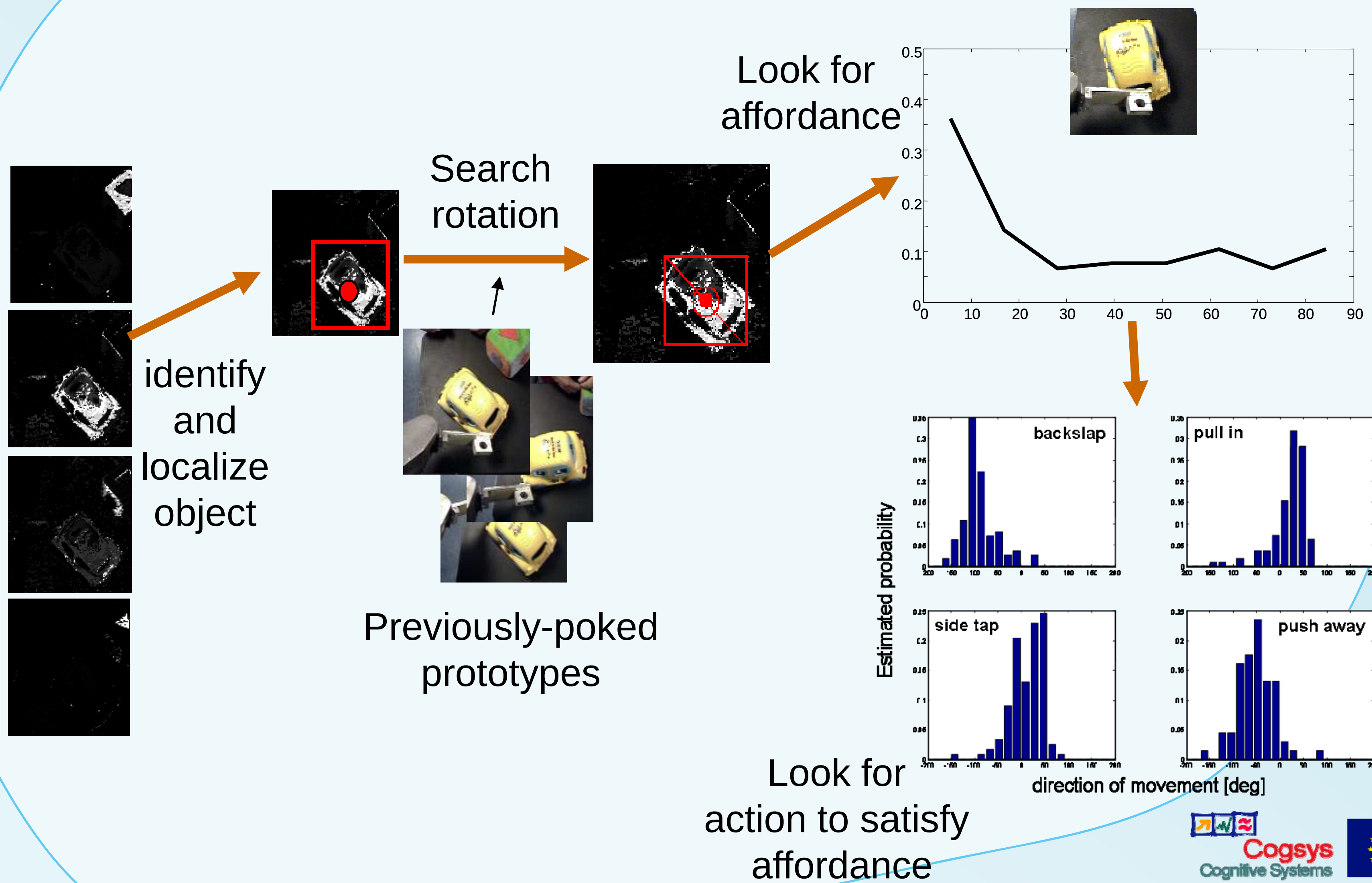
Into object affordances…



The geometry of poking

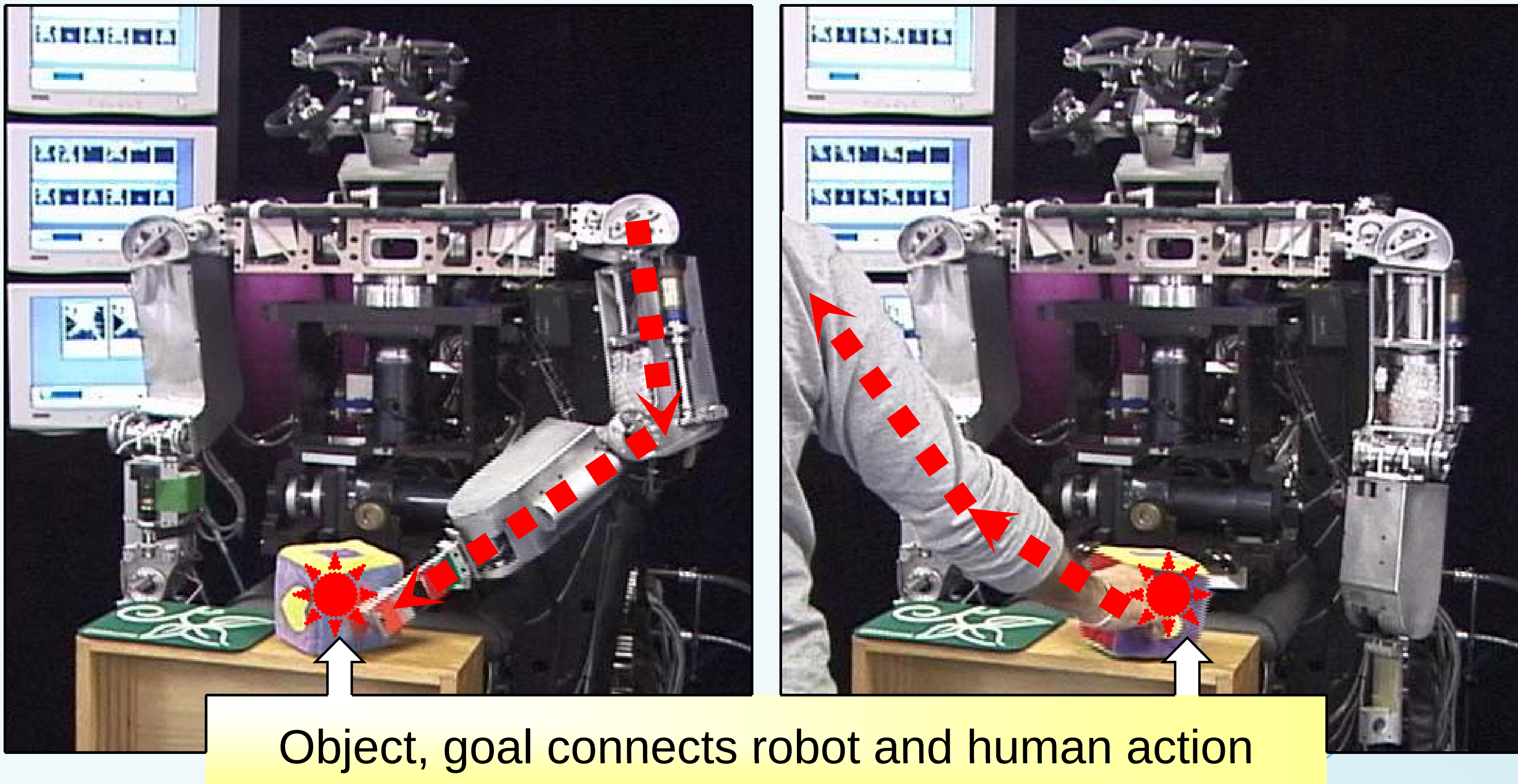


Behavior: poking according to affordance



Behavior: poking according to affordance

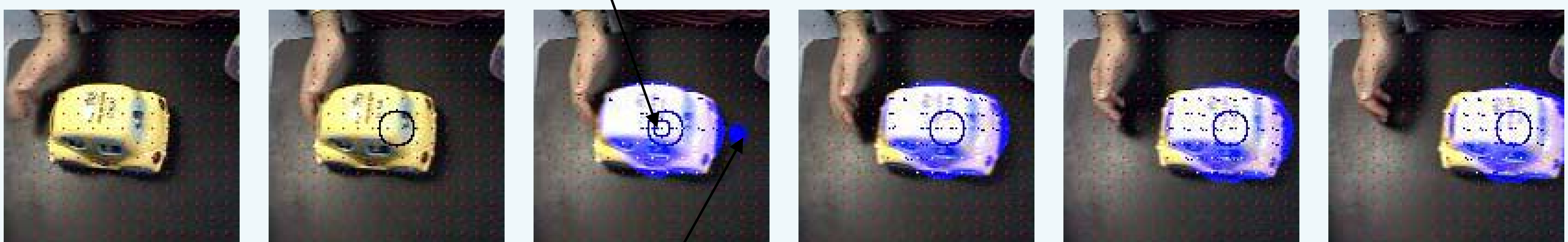
Understanding a foreign manipulator



Interpreting observations

"The robot could actually tell this was a side tap"

Initial position



Final position

A foreign manipulator (human) pokes an object
The direction of movement is compared with the object affordance

Interpreting observations

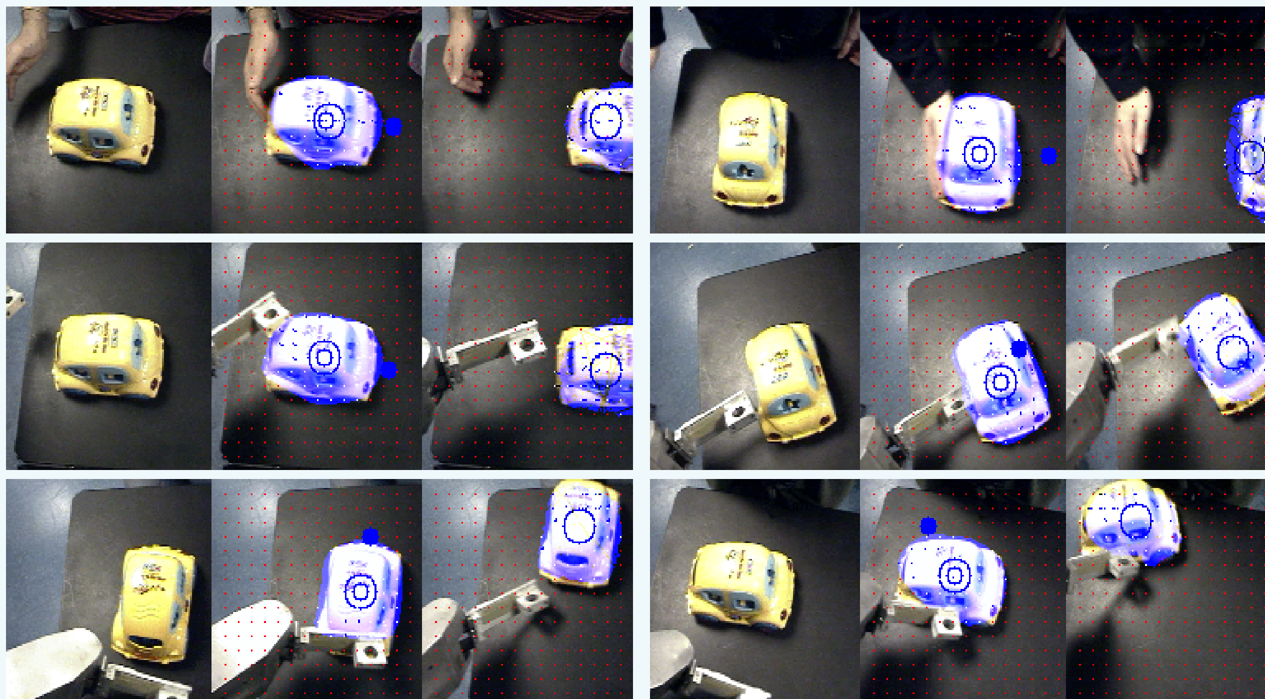
Invoking the object's natural rolling affordance

Going against the object's natural rolling affordance

Demonstration by human

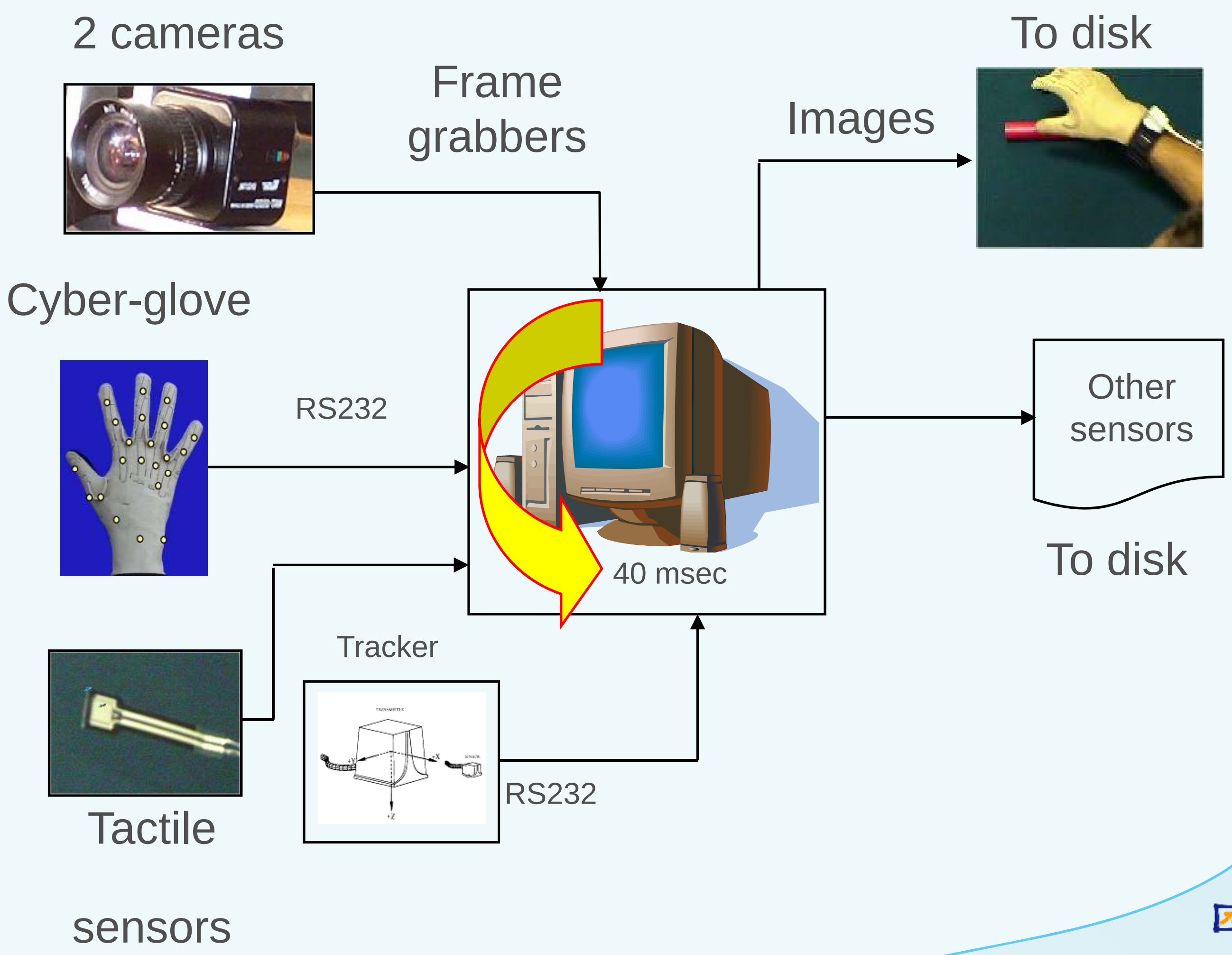
Mimicry in similar situation

Mimicry when object is rotated



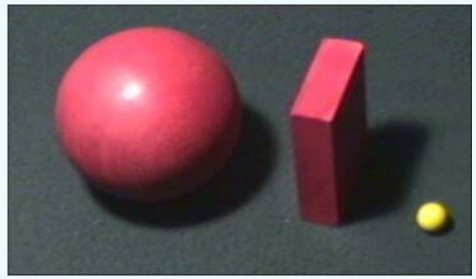
Mimicry

Data from human grasping

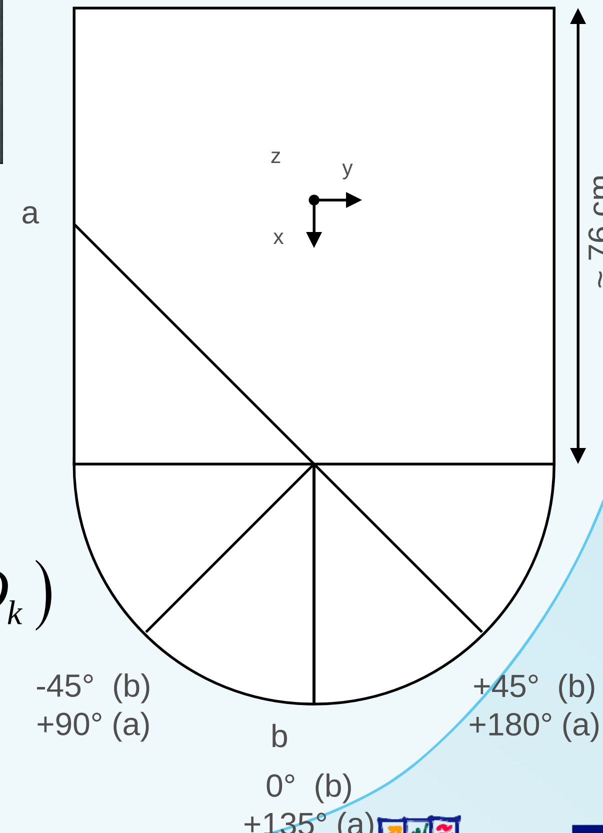


Bayesian classifier

{*Gi*}: set of gestures
F: observed features
{*Ok*}: set of objects



168 sequences per subject
10 subjects
6 complete sets

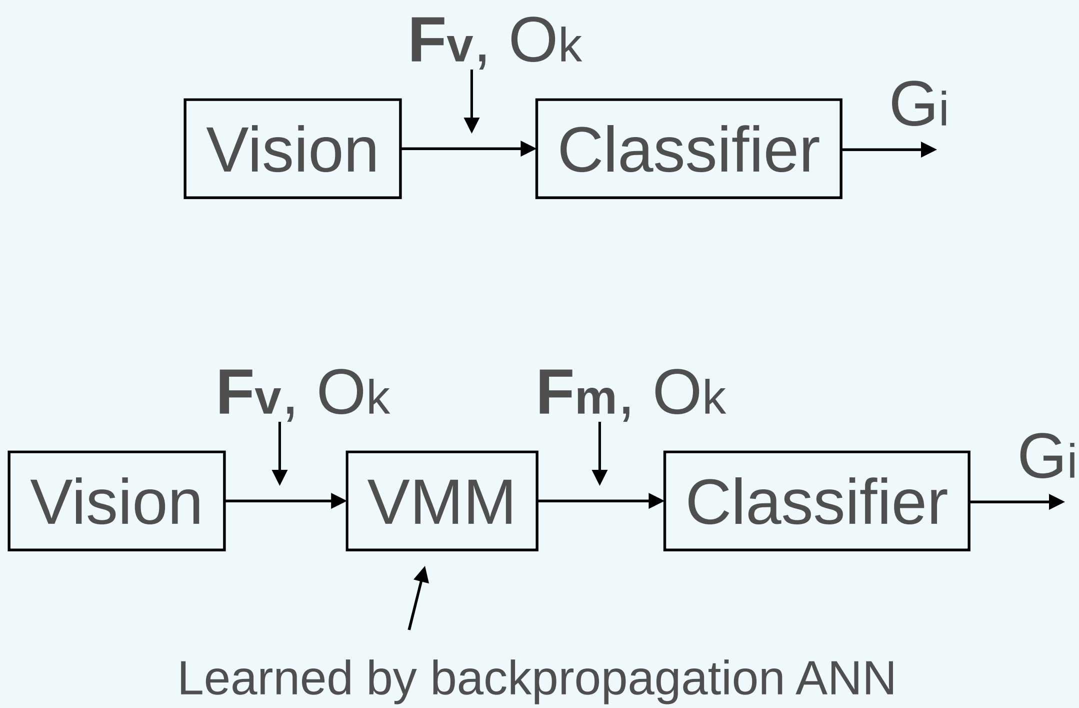


p(Gi|Ok): priors (affordances)
p(**F**|*Gi,Ok*): likelihood to observe **F**

$$p(G_i \mid \mathbf{F}, O_k) = p(\mathbf{F} \mid G_i, O_k)\, p(G_i \mid O_k) / p(\mathbf{F} \mid O_k)$$

$$\hat{G}_{MAP} = \arg\max_{G_i} (G_i \mid \mathbf{F}, O_k)$$

Two types of experiments



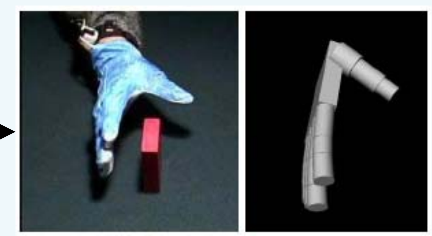
Estimation

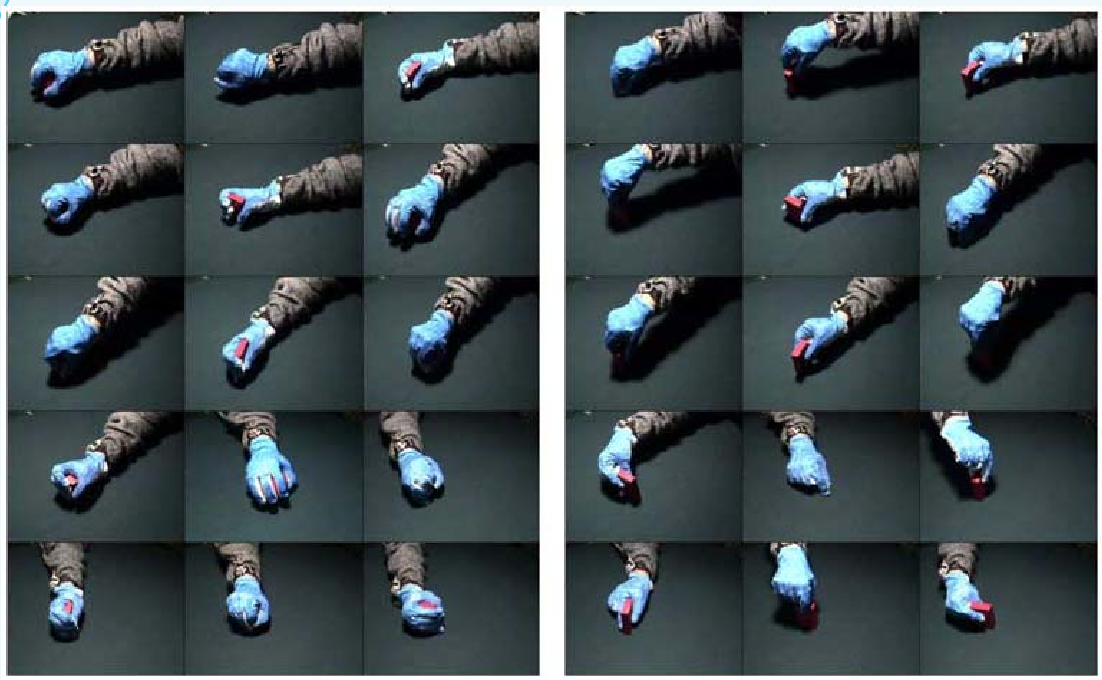
- p(Gi|Ok): affordances, by counting, estimated on the whole database
- p(**F**|Gi,Ok): EM algorithm on the parameters of a mixture of Gaussians (from Matlab implementation)
- VMM: Neural network, sigmoidal activation units, linear output, trained on the whole database

Role of motor information in action understanding

Object affordances (priors)

Visual space Motor space





Grasping actions

Classification (recognition)

Understanding mirror neurons: a bio-robotic approach. *G. Metta, G. Sandini, L. Natale, L. Craighero, L. Fadiga.* Interaction Studies. Volume 7 Issue 2. 2006

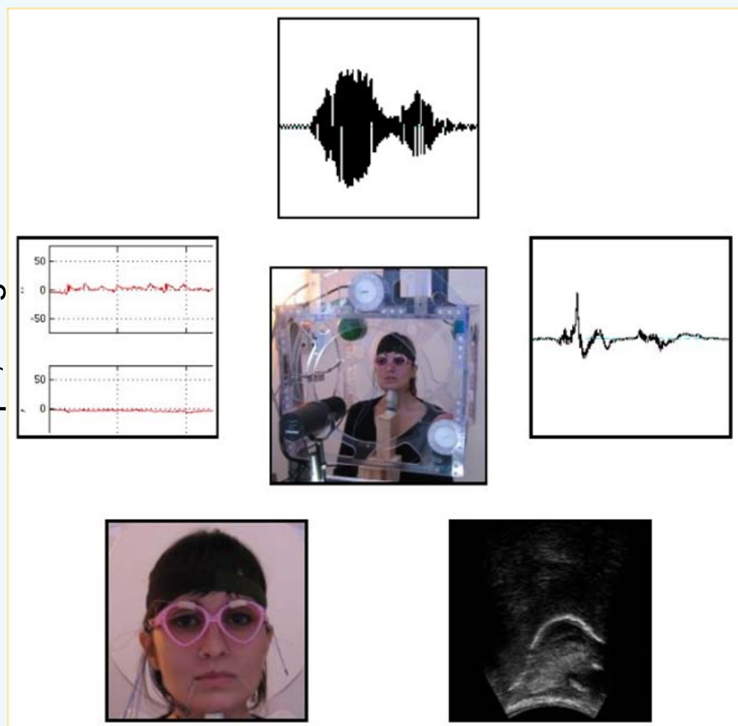
Some results…

	Exp. I (visual)	Exp. II (visual)	Exp. III (visual)	Exp. IV (motor)
	Training			
# Sequences	16	24	64	24
# of view points	1	1	4	1
Classification rate	100%	100%	97%	98%
# Features	5	5	5	15
# Modes	5-7	5-7	5-7	1-2
	Test			
# Sequences	8	96	32	96
# of view points	1	4	4	4
Classification rate	100%	30%	80%	97%

Speech again

Sound

Movement of the head lips, tongue

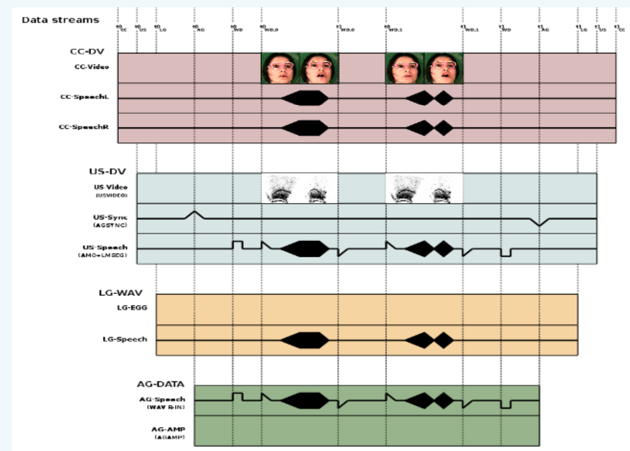


Larynx

Facial expression

Shape of the tongue

Ultrasound



Sound, video, etc.

With Michele Tavella & the CONTACT project team

System-level approach

Cognition

Cognition: a process by which a system achieves behaviour that is

- robust
- adaptive
- anticipatory
- autonomous

Entails embodied perception, action, and interaction

Our approach

Guiding Philosophy

- Cognition cannot be hand-coded
- Is necessarily the product of a process of embodied development
- Initially dealing with immediate events
- Increasingly acquiring a predictive capability

Cognition and perception are functionally-dependent on the richness of the action interface

Our Approach (contd.)

Emergent embodied cognitive systems:

- Given a rich set of innate action and perception capabilities
- Develop over time an increasing range of cognitive abilities
- Recruiting ever more complex actions
- Achieving an increasing degree of prospection (and, hence, adaptability and robustness)

RobotCub Cognitive Architecture

Grounded in neuroscience and psychology

Rooted in action-dependent perception

Focused on adaptive & prospective capabilities

Designed to facilitate development

Cognitive architecture ≡ (RobotCub) Phylogeny

In spite of the wiring problems…