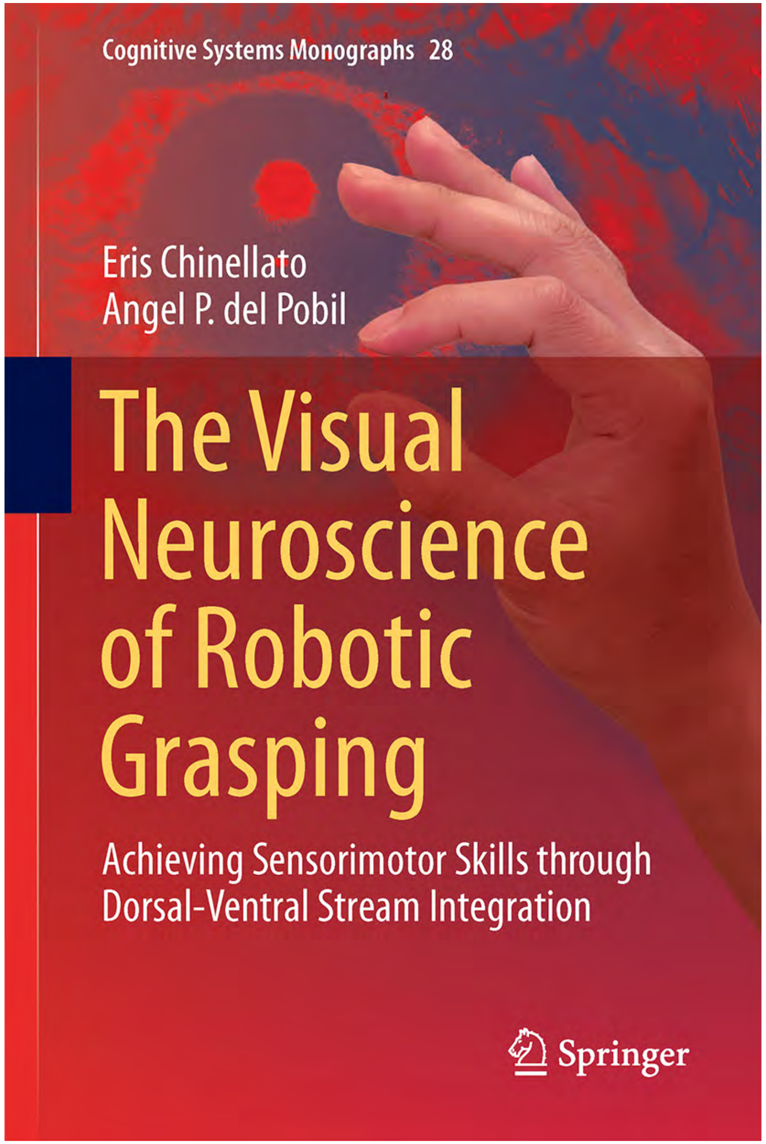
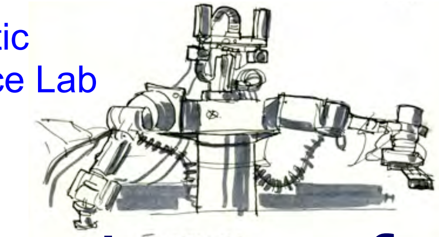


UJI Robotic
Intelligence Lab



The visual neuroscience of robotic grasping:

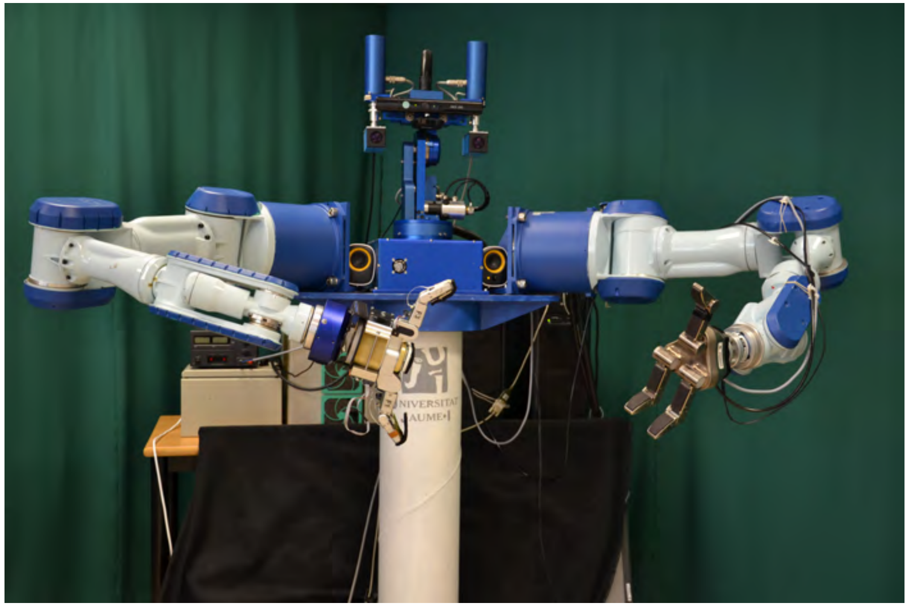
Ideas for functional computational models

Angel P. del Pobil

Robotic Intelligence Lab, Universitat Jaume I
Castellón, Spain,
The ShanghAI Lecturers 2015
17 December 2015

UNIVERSITAT JAUME·I

Vision based grasping: who does it better?



Tombatossals, the UJI Humanoid Torso

Thanks to UMass Lab for Perceptual Robotics

Artificial intelligence, neuroscience and robotics

Artificial intelligence *emulates* biological intelligence but:

- Humans (animals in general) work much better than robots in **unstructured** (real) environments
- Humans (and all other animals) perform a lot better than robots in non-cognitive (i.e. **sensorimotor**) tasks
- Robotic grasping is still very far from the versatility, dependability and autonomy of human and other primate grasping

cognition as symbol-processing vs. cognition emergent from sensory-motor and interaction processes

Free hands fostered intelligence

.

- Bipedalism freed hands for manipulating tools
- Multifinger manipulation was the cornerstone for human evolutionary development towards intelligent behavior
- Most of our physical interaction with the world is mediated by the use of our hands.
- An autonomous robot will be severely impaired if it is not able to reliably perform manipulation tasks in everyday environments, that is, it should have ***manual intelligence***

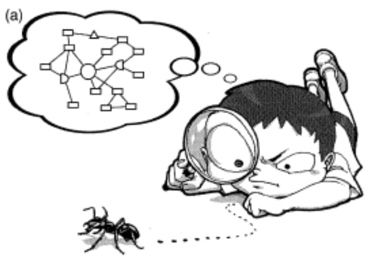
Byrne, R.W. 2004, "The manual skills and cognition that lie behind hominid tool use" in *The Evolution of Thought. Evolutionary Origins of Great Ape Intelligence*, edited by A.E. Russon and D.R. Begun, pp. 31-44, Cambridge University Press.

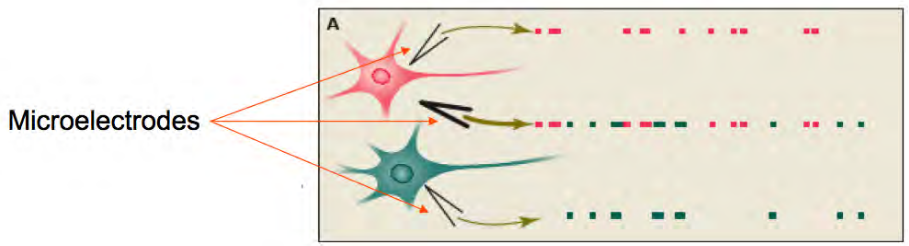
Research goals

Deal with the task of vision-based robotic grasping following a truly interdisciplinary approach, promoting a mutual enrichment between the fields of robotics and neuroscience (**synthetic methodology**):

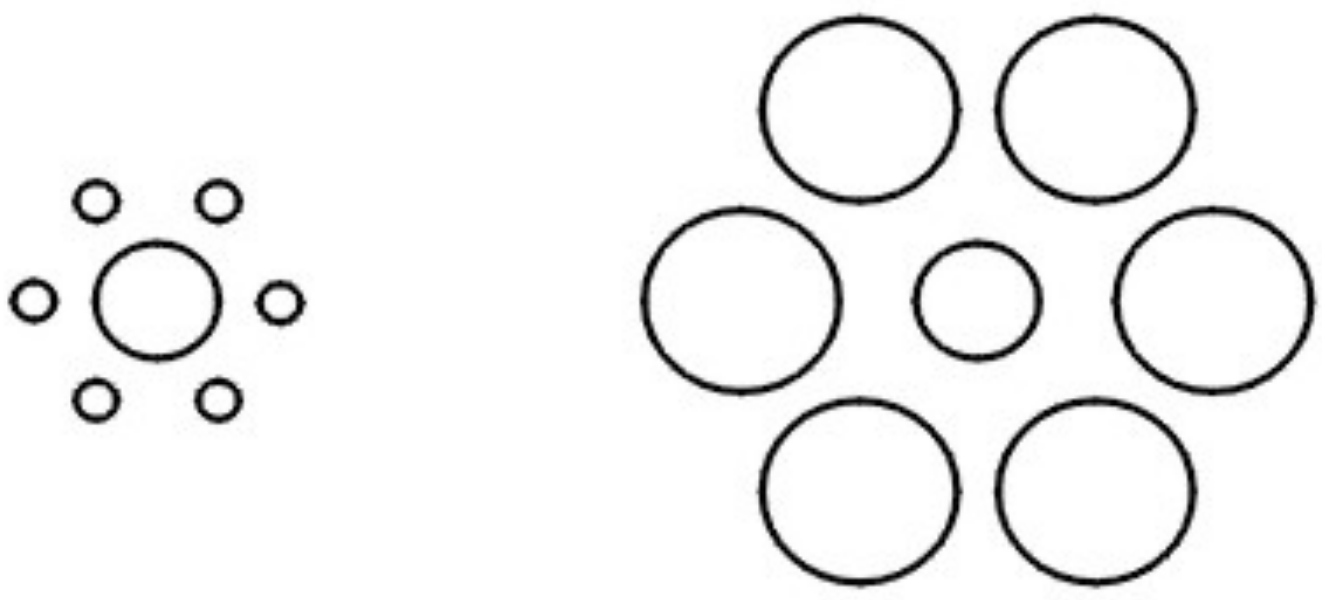
- Robots can achieve a better interaction with objects in their peripersonal space, mimicking the solutions evolved by natural systems
- Neuroscience hypotheses can be advanced, analyzed, compared and validated through computational modeling and robotic experiments

Where to start?



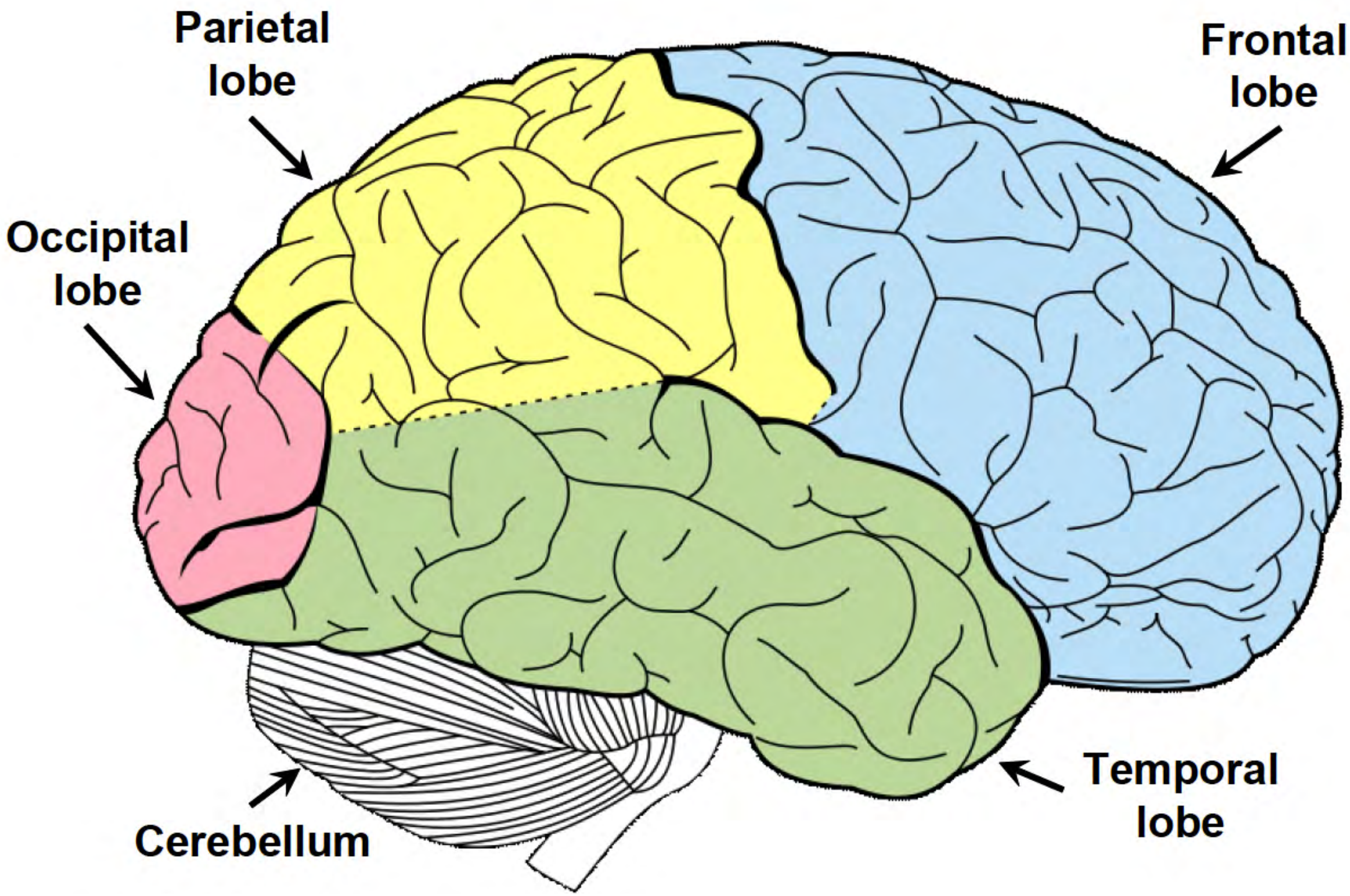


Fooled or not?

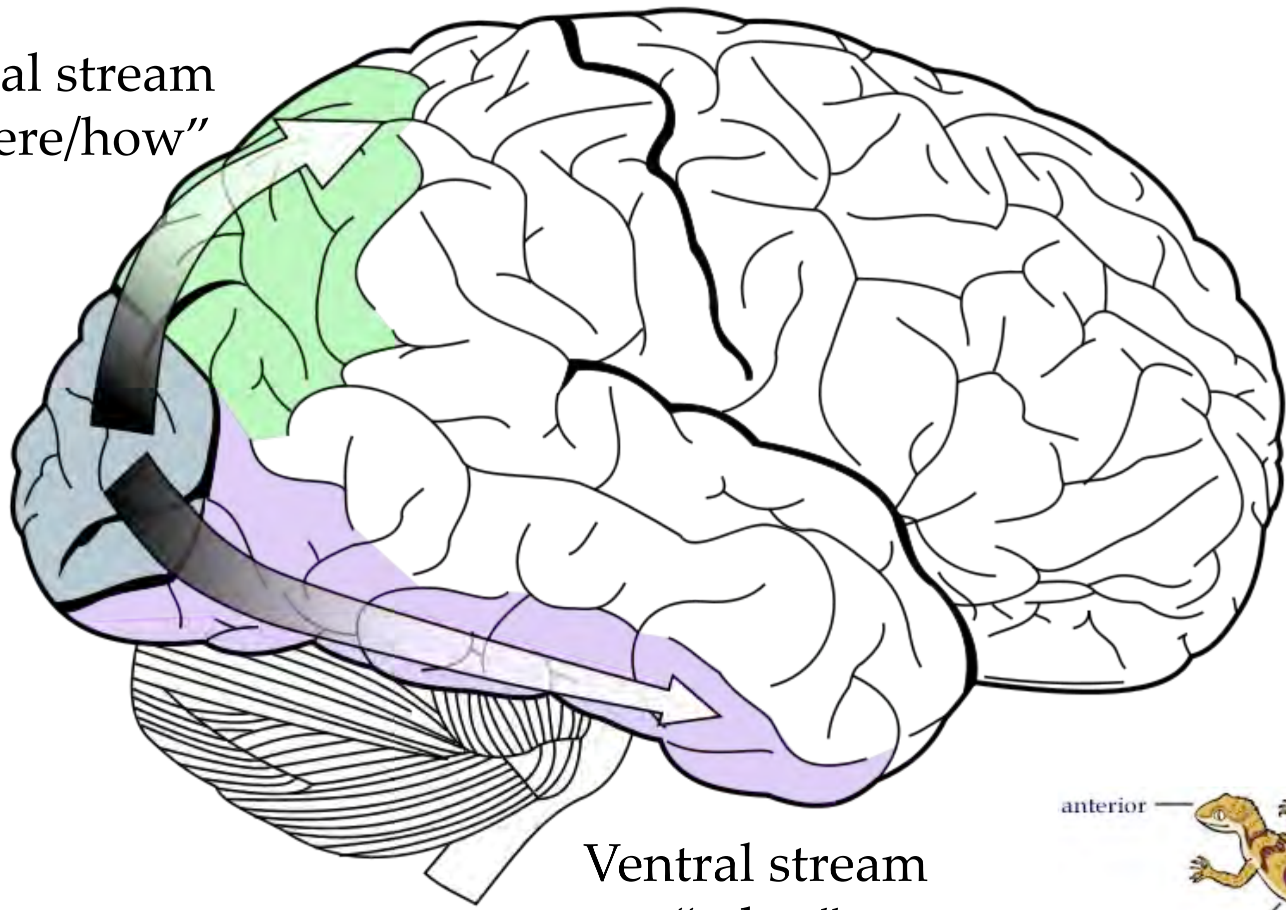


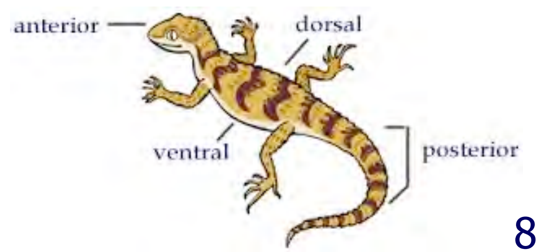
Cortical lobes and cerebellum

.



Dorsal and ventral streams

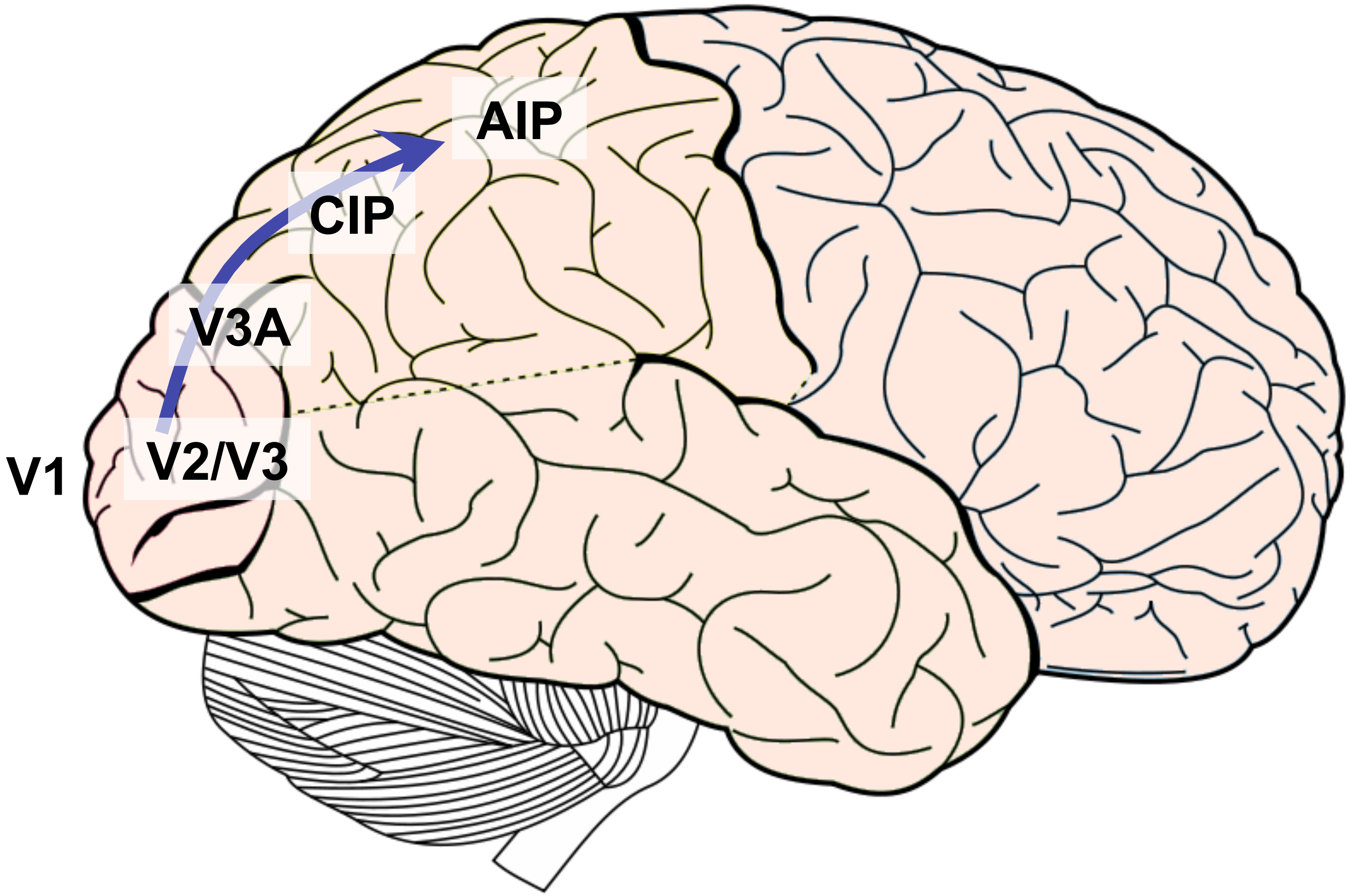




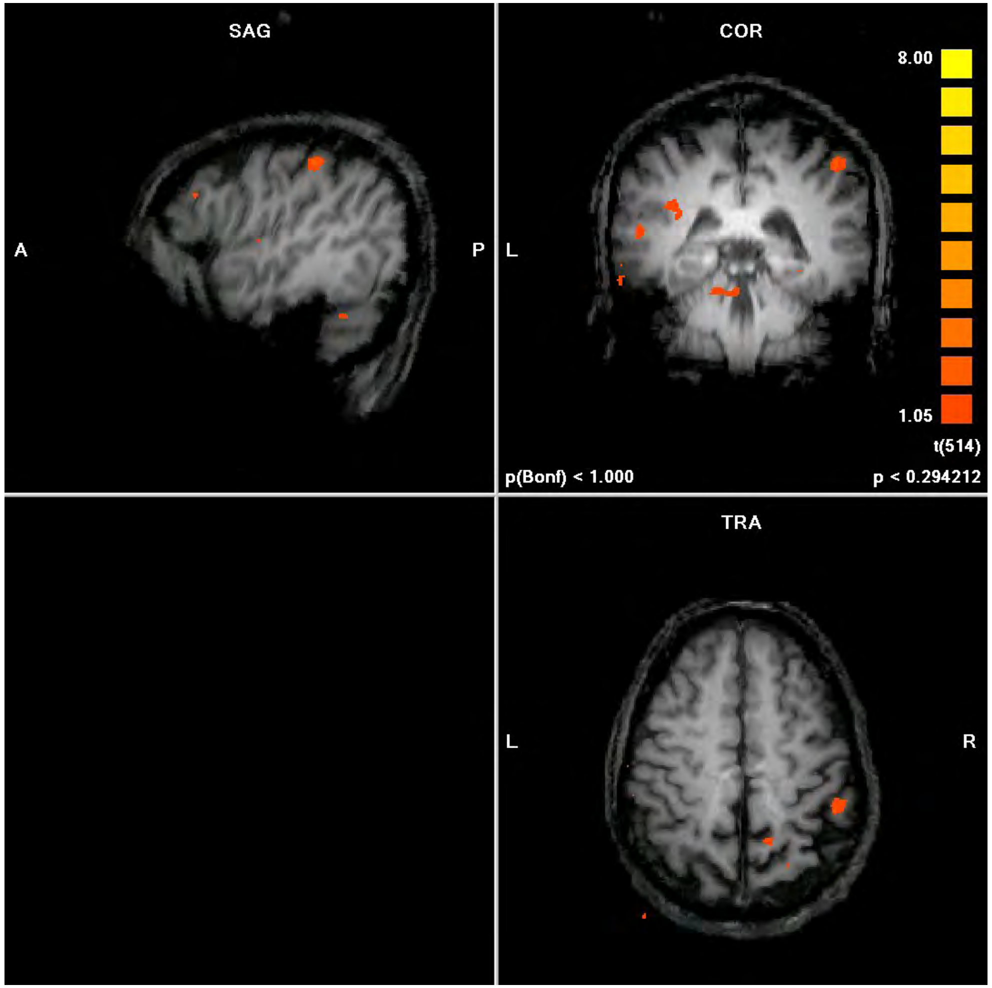
Complementary role of the two streams

Ventral Stream	Dorsal Stream
Object recognition	Visuomotor control
Global, invariant analysis	Local, feature analysis
Object weight, roughness, compliance	Object local shape, size
Object meaning	Object location
Previous experiences	Actual working conditions
Scene-based frame of reference	Effector-based frame of reference
Long-term representation	On-line computation

Dorsal and ventral brain areas



AIP
Anterior intraparietal sulcus



AIP activates:

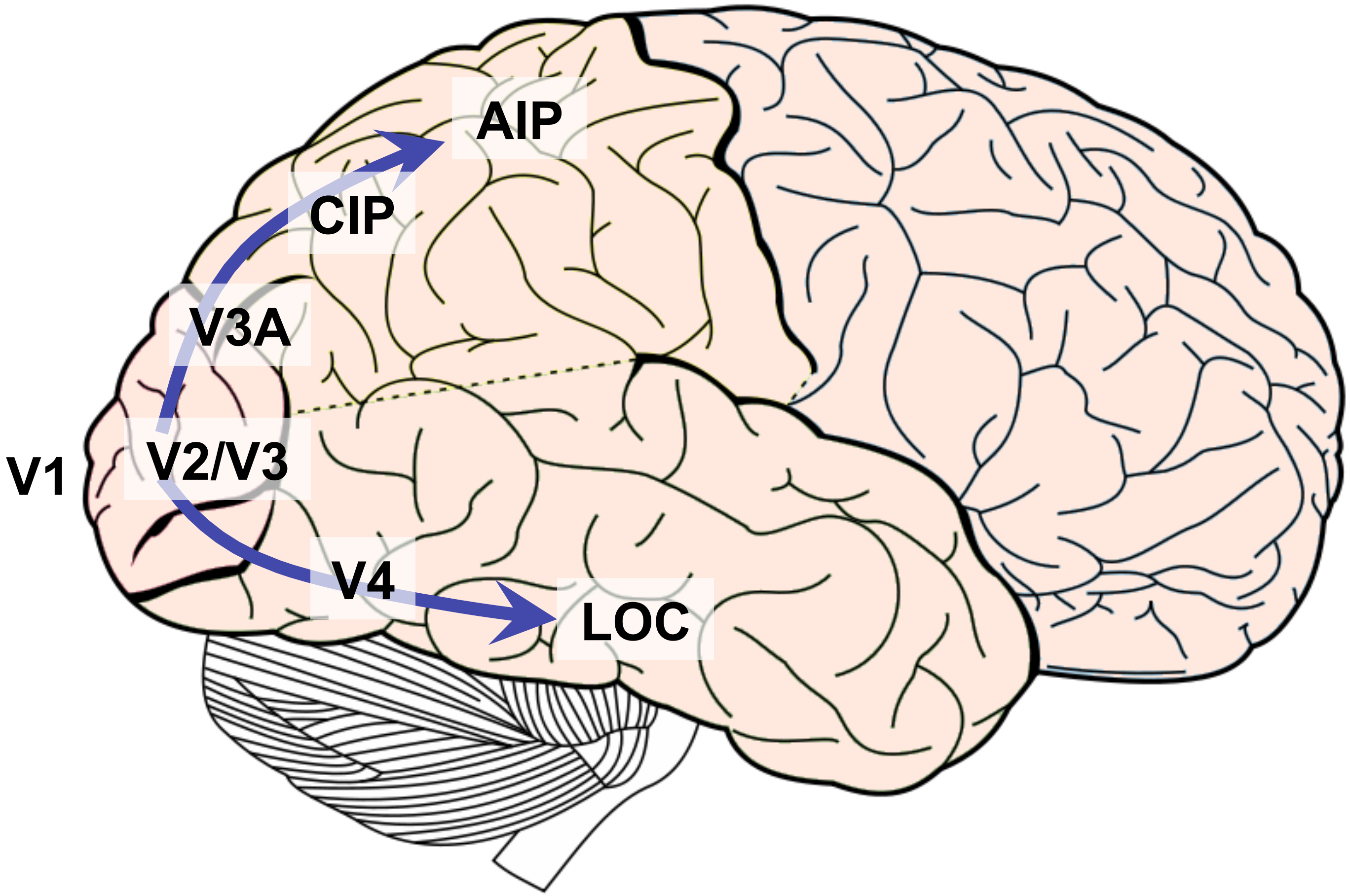
- when the subject grasps an object, especially if vision is involved
- when the subject looks at an object with grasping purposes
- differentially between grasping and reaching

Thanks to Jody Culham and Mel Goodale (UWO) for all fMRI images

AIP
The grasping area

- AIP receives visual data about object features that may be suitable for grasping (action-oriented vision), mainly from CIP
- Some neurons describe a **shape-based representation** of objects, whilst others code for **the hand configuration** suitable for grasping
- Integrating information from several other brain areas, including ventral stream areas, AIP filters and processes the visual features, **generating and selecting suitable grasp configurations**
- AIP projects to the **ventral premotor cortex (PMv**) where the final grasping action is "composed"
- AIP collaborates with PMv and probably the Cerebellum during grasp execution for monitoring the action progress

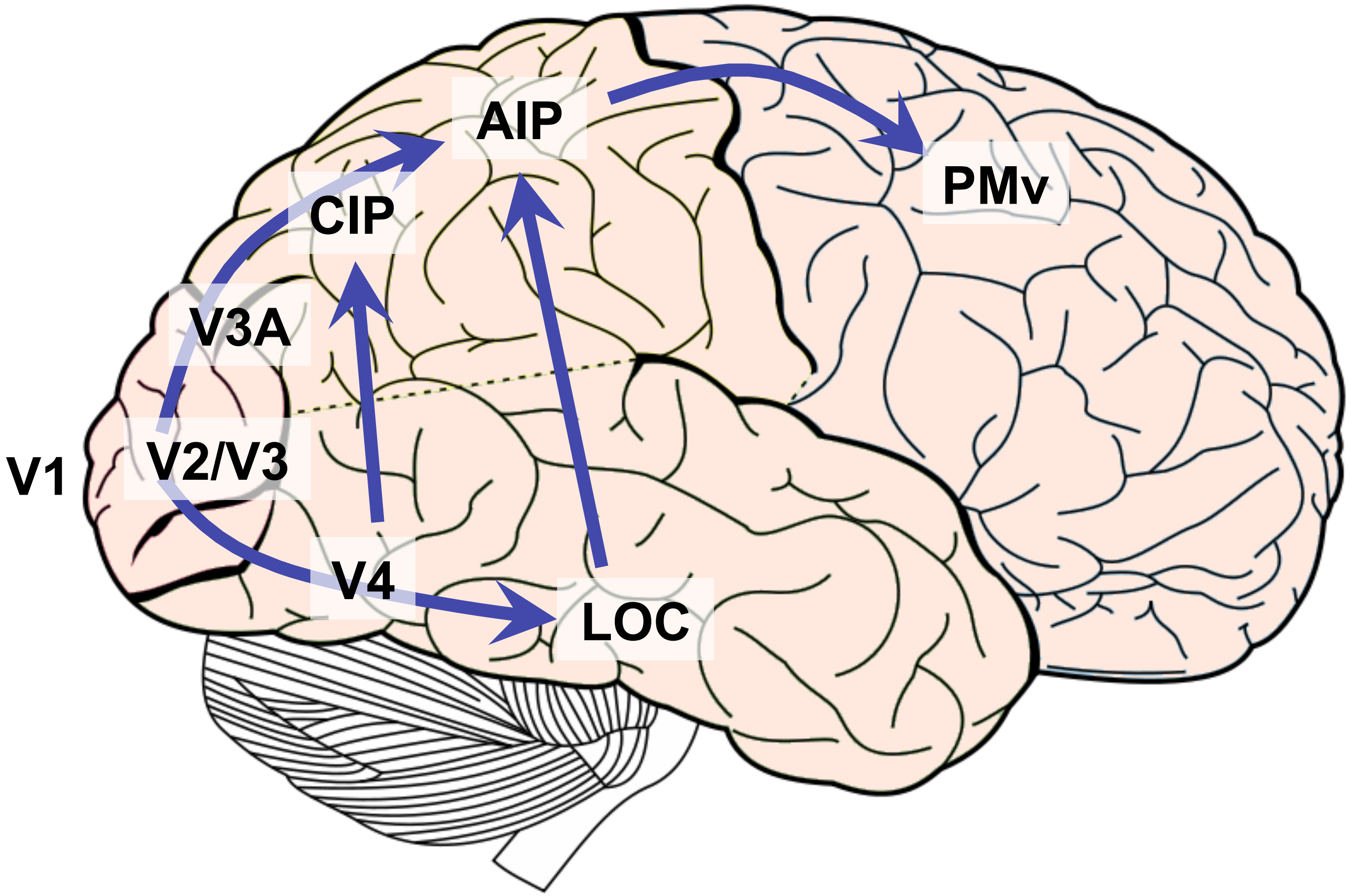
Dorsal and ventral brain areas



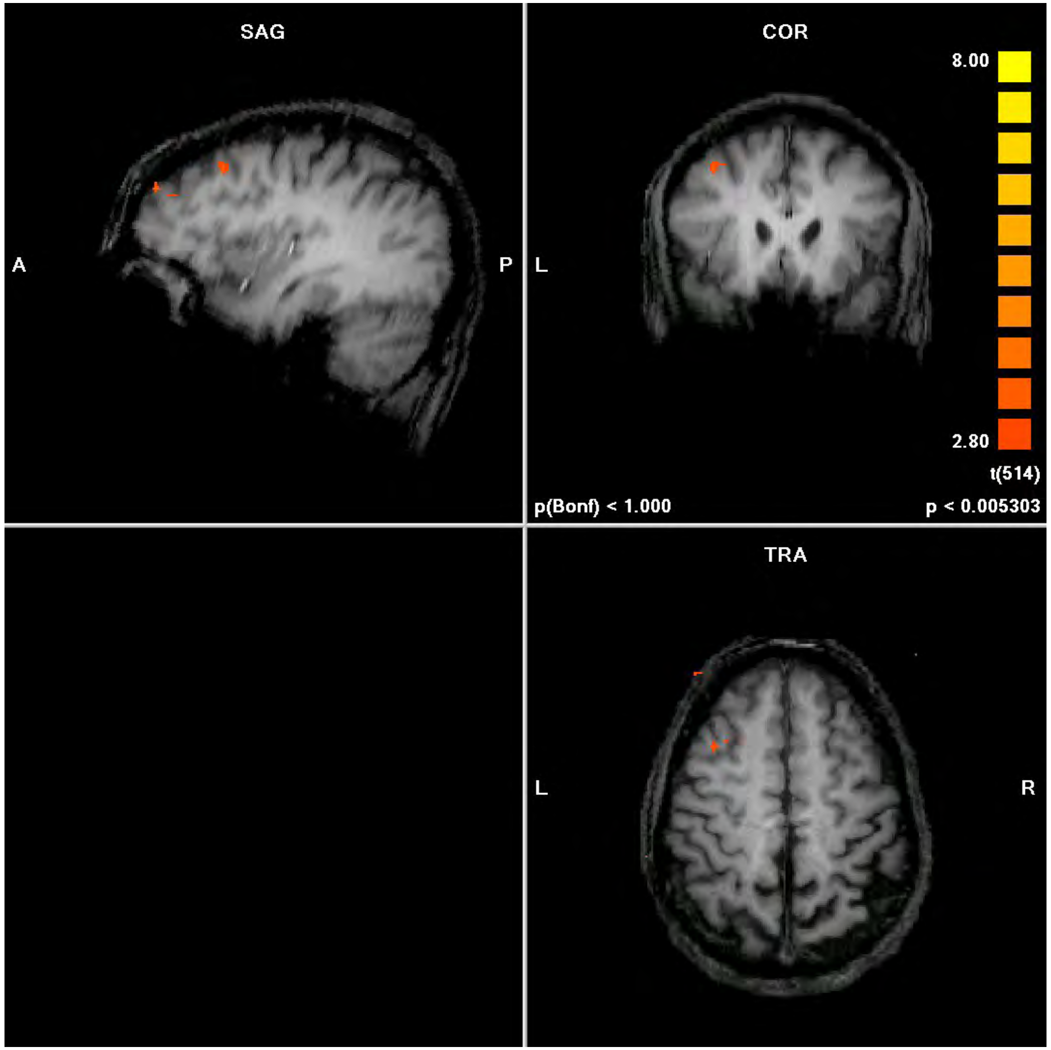
LOC
Lateral occipital complex

- The LOC is the part of the brain cortex dedicated to **object recognition**
- It receives high-level visual input from V4 and **integrates** elements that share similar attributes (orientation, color, depth)
- Object representation in LOC is **invariant** with stimulus type (3D or silhouettes, color maps, etc.): ***conceptual*** representation
- LOC is very likely involved in reaching and grasping actions (especially in some conditions, e.g. delayed grasping, familiar objects)
- Information from the ventral stream probably meets the dorsal stream at different stages, likely through both V4 and LOC
- LOC contributes to grasping actions providing access to **memories** of object data (weight, material, …) and past grasping **experiences**

Dorsal and ventral brain areas



PMv
Ventral premotor cortex



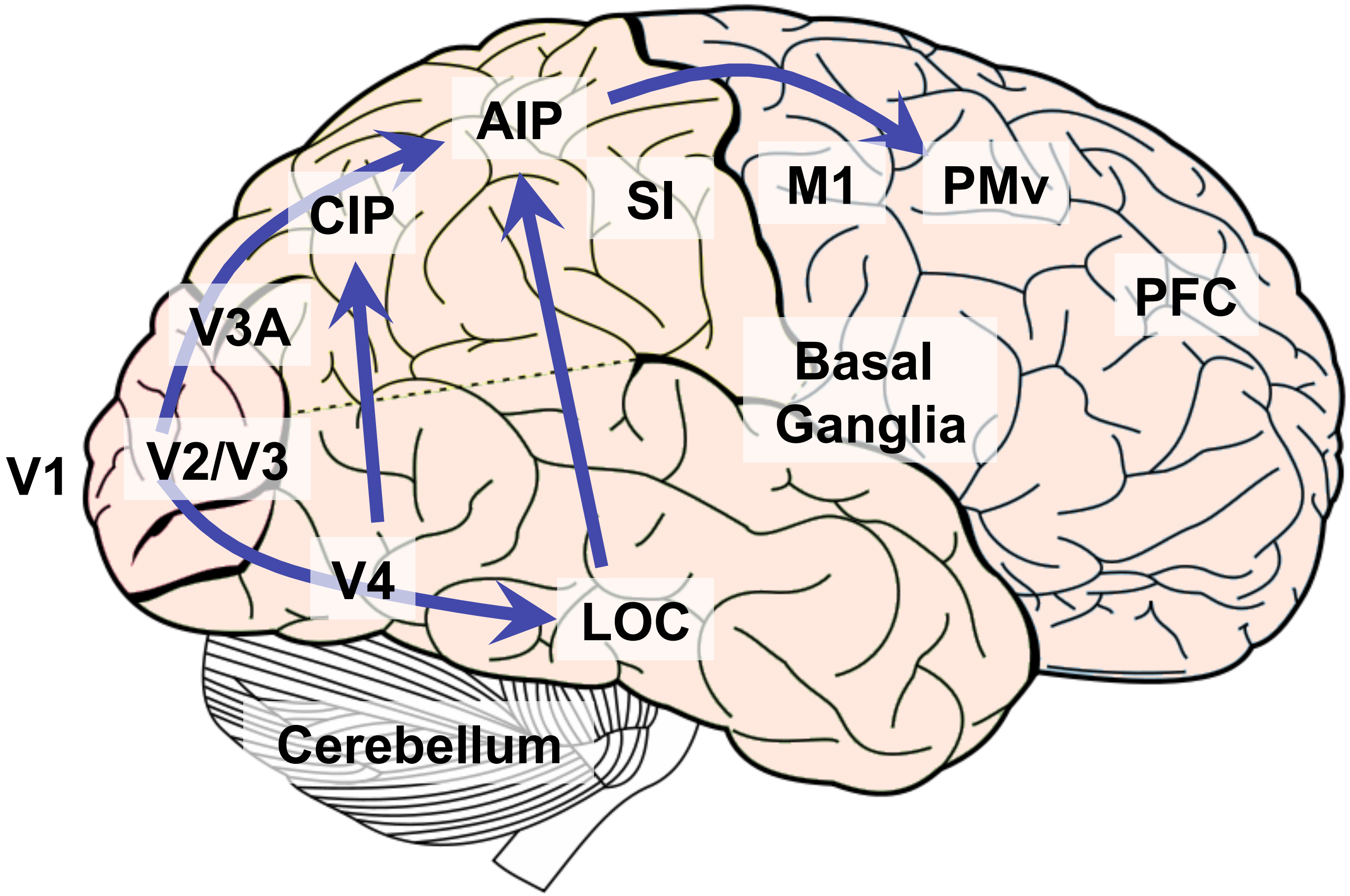
PMv (also known as F5 from monkey studies) activates:

- when the subject **performs** a given hand action
- when the subject **imagines** to be performing the same action
- when the subject **sees** someone else performing that action (mirror neurons)

PMv (F5)
The action composition area

- PMv contains a **vocabulary of motor actions** of different complexity, duration and significance
- Basic actions are **selected and combined** according to task, object shape and size, timing of action…
- Selection is very likely performed in collaboration with the posterior parietal cortex (especially with AIP)
- PMv projects to **primary motor areas (M1)**, which sends to the muscles the signals for action execution

Brain areas involved in vision-based grasping actions



Evidence for the two visual streams hypothesis

- Neurological patients (*optic ataxia / visual agnosia*)
- Data from primate studies (especially single cell experiments)
- Psychophysical experiments (e.g. optical illusions)
- Brain imaging experiments (fMRI, PET, TMS)

Optic Ataxia

Visual Agnosia

Bridging neuroscience findings with robotic applications

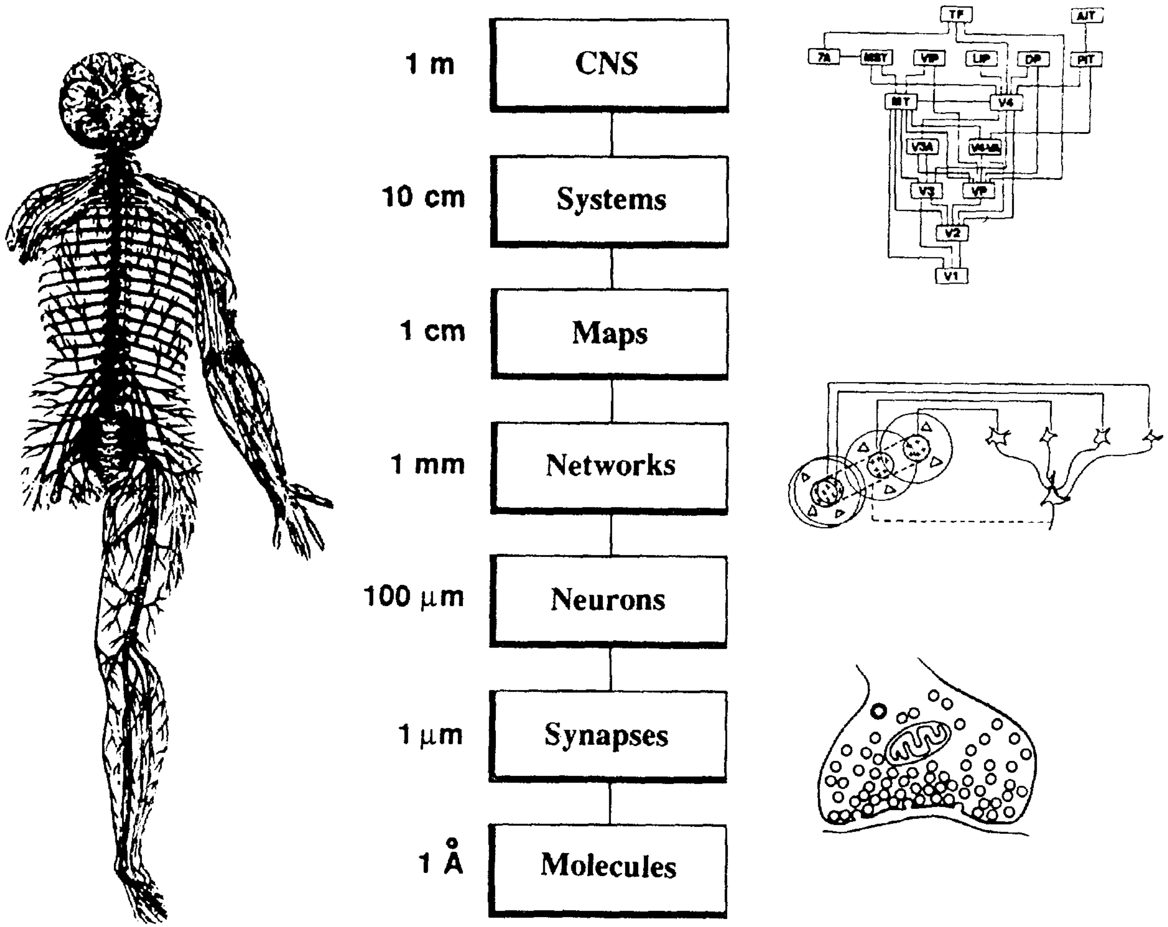
- The ventral and dorsal streams perform **complementary** tasks
- There is evidence for both control switching and modulated collaboration between the streams
- The ventral stream does intervene in grasping actions

The task - **develop a functional model of the brain areas most important for vision-based grasping**:

- special attention to interaction between dorsal and ventral streams
- faithfulness to **biological reality** and suitability to **robotic implementation**

Modeling the nervous system
Levels of abstraction

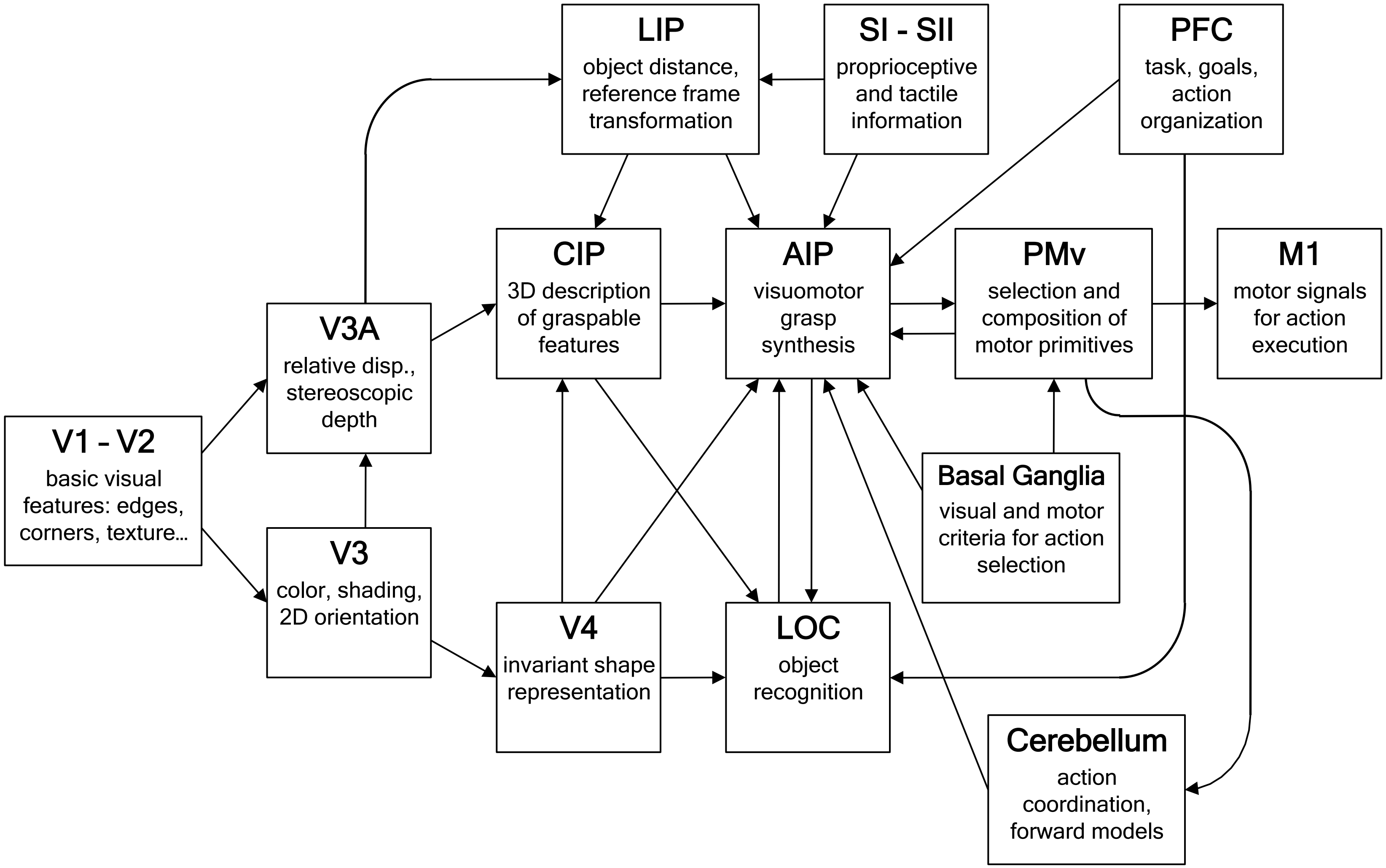
UNIVERSITAT JAUME•I



From *The Computational Brain*, Churchland & Sejnowski 1992

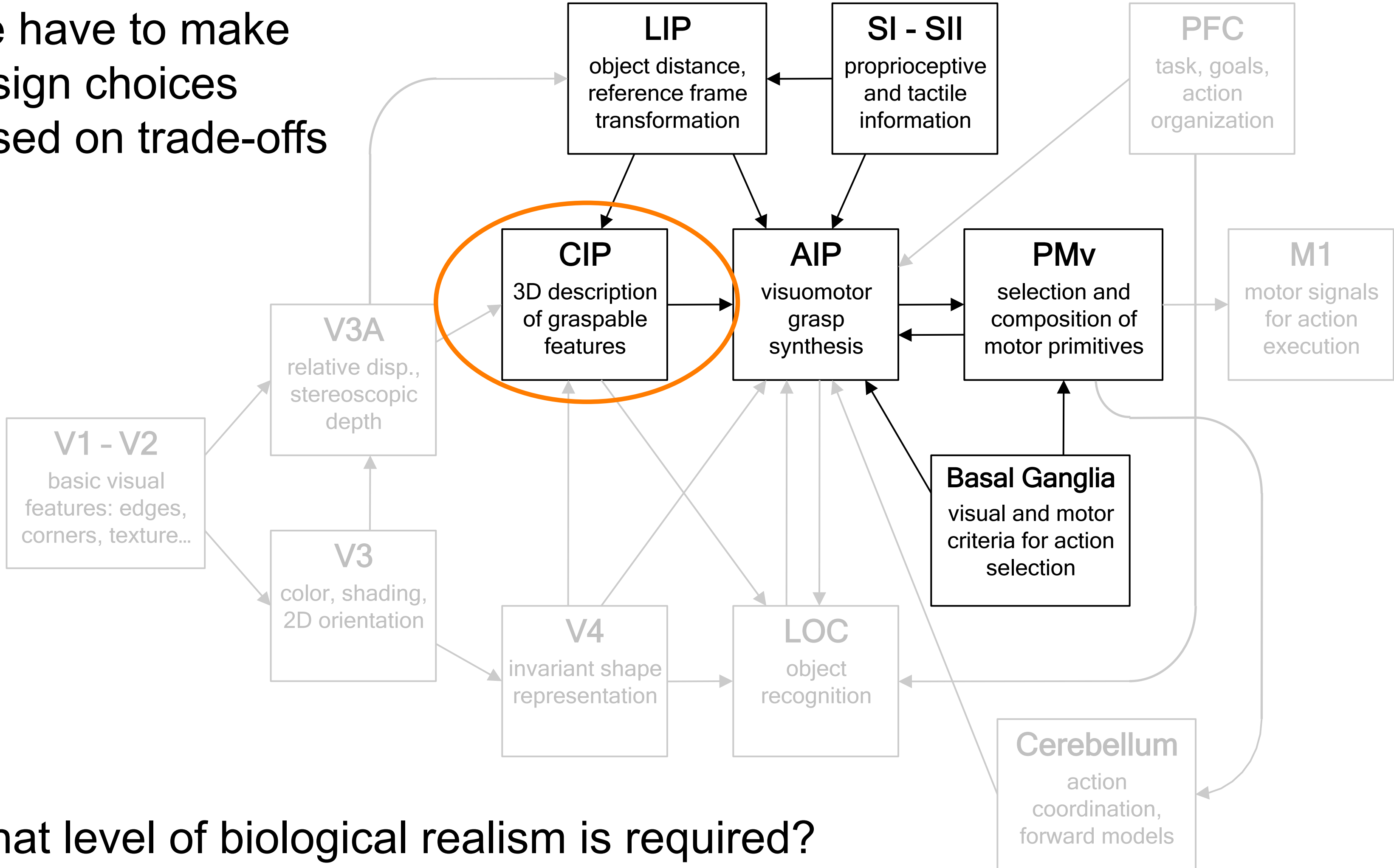
A functional model of vision-based grasping

UNIVERSITAT JAUME•I



From object shape to hand configuration

We have to make design choices based on trade-offs

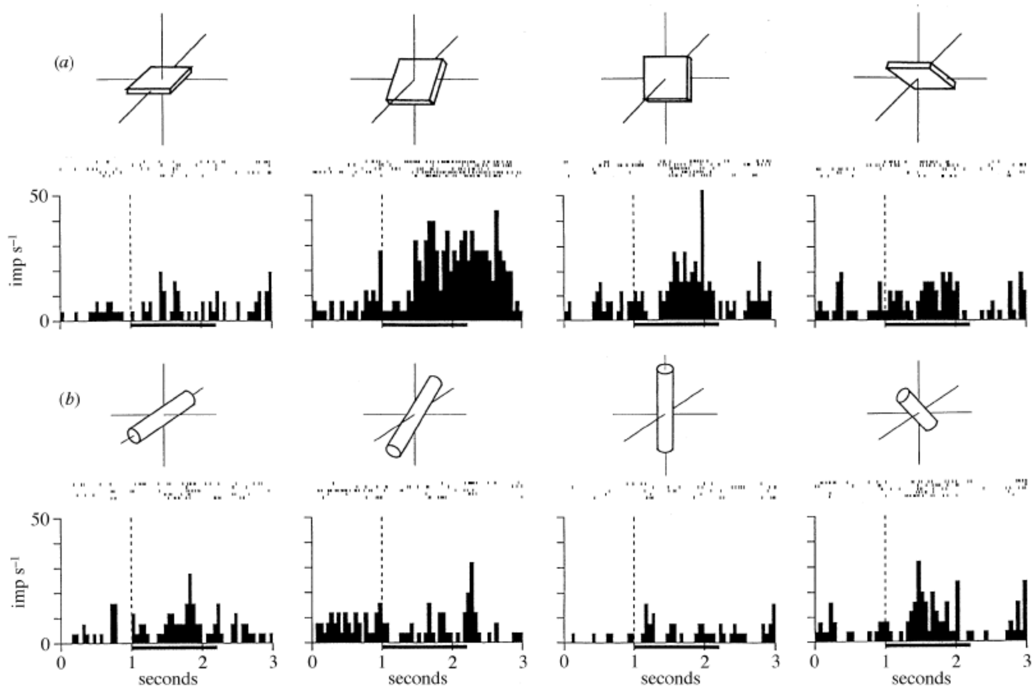


What level of biological realism is required?

Surface and axis orientation selective neurons: SOS & AOS

Visual neurons in the lateral bank of the caudal intraparietal sulcus (CIP or c-IPS area) in macaques are selective for orientation of visual features suitable for grasping, and for their proportions (Sakata et al., 1998);

Humans fMRI studies indicate a similar responsiveness (Shikata 2003)

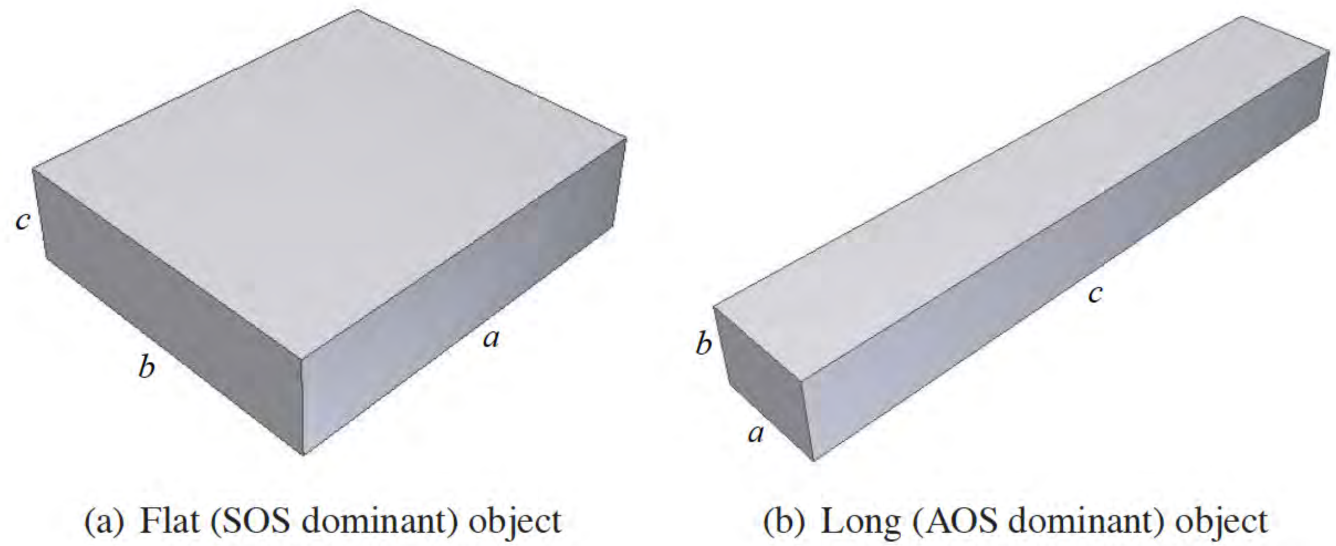


Responses of an SOS neuron that preferred a square plate tilted 45º backward. No response was obtained when a cylinder was presented at the same orientation.

From Sakata et al., *Phil.Trans. R. Soc. Lond. B* (1998)

Surface and axis orientation selective neurons: SOS & AOS

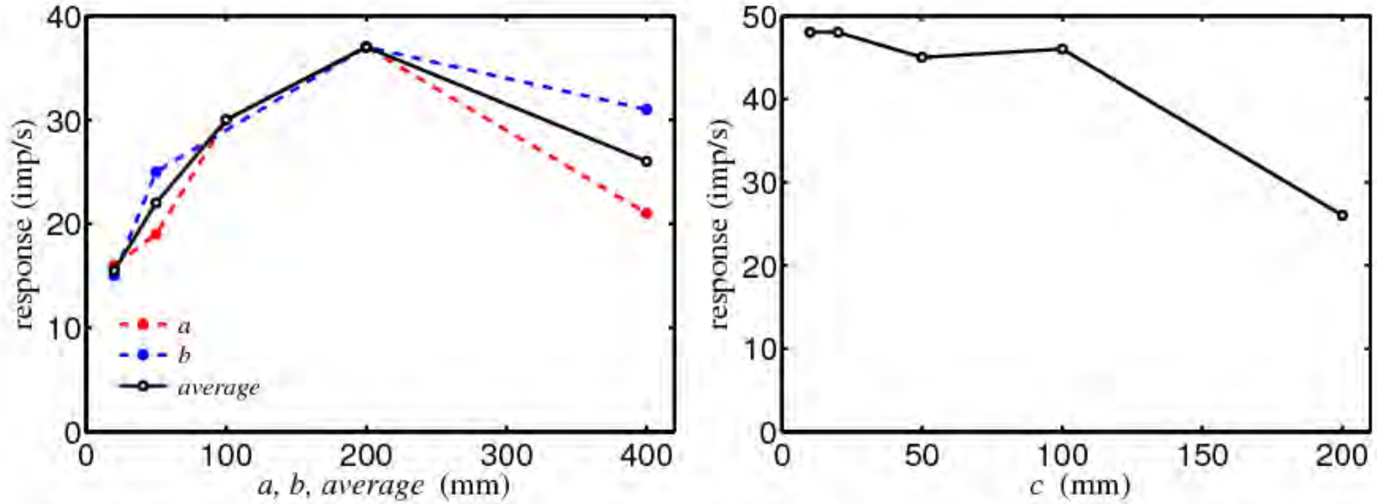
Examples of SOS and AOS dominant objects and size naming convention



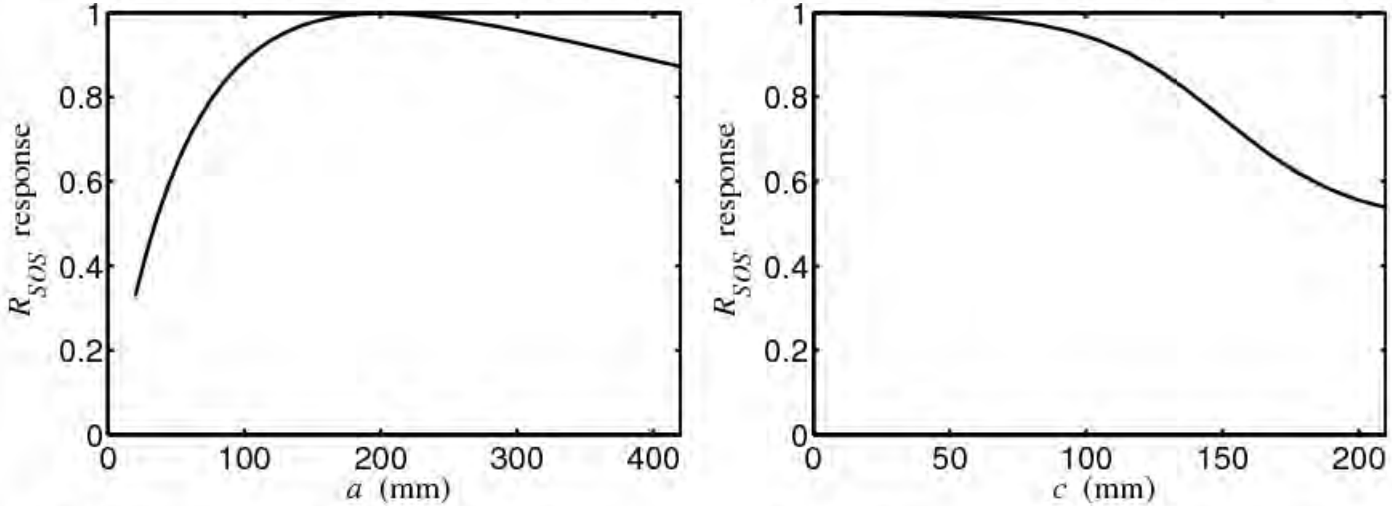
SOS neurons transfer function

.

Experimental data adapted from Shikata et al (1996).

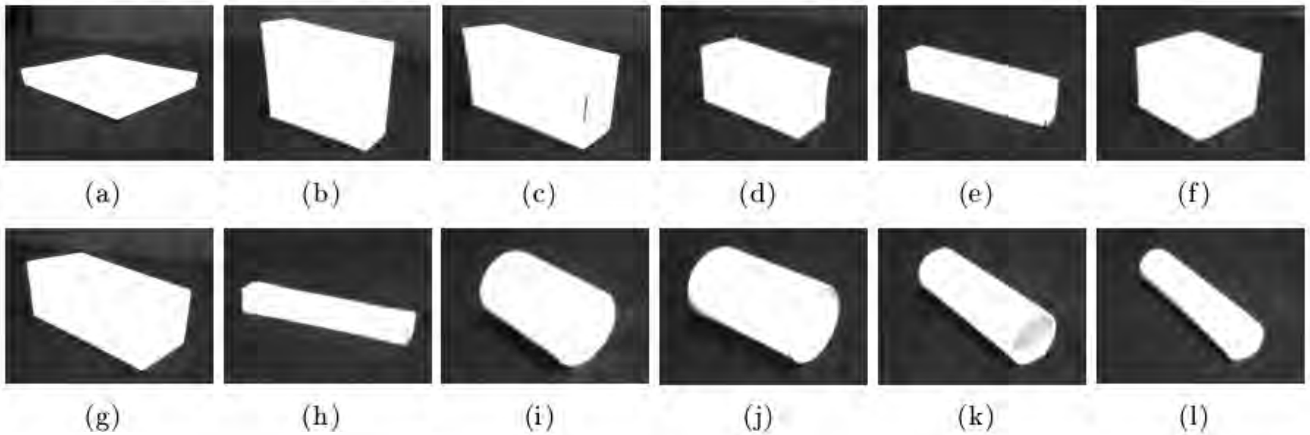


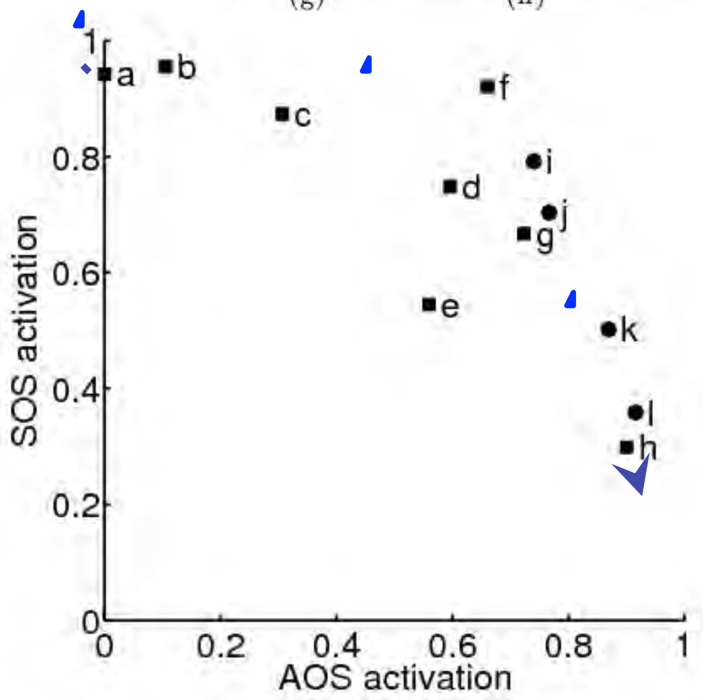
(a) Width response (thickness $c = 20$mm) (b) Thickness response (width $a = b = 200$mm)



(a) Width response ($b = 200$mm, $c = 20$mm) (b) Thickness response ($a = b = 200$mm)

Robotic SOS and AOS response





- Three clusters, with clear distinction between elongated and flat shapes
- A main distribution arc with few exceptions (graspability?)

But: more experiments are required to improve the SOS/AOS modeling

Grasp ingredients

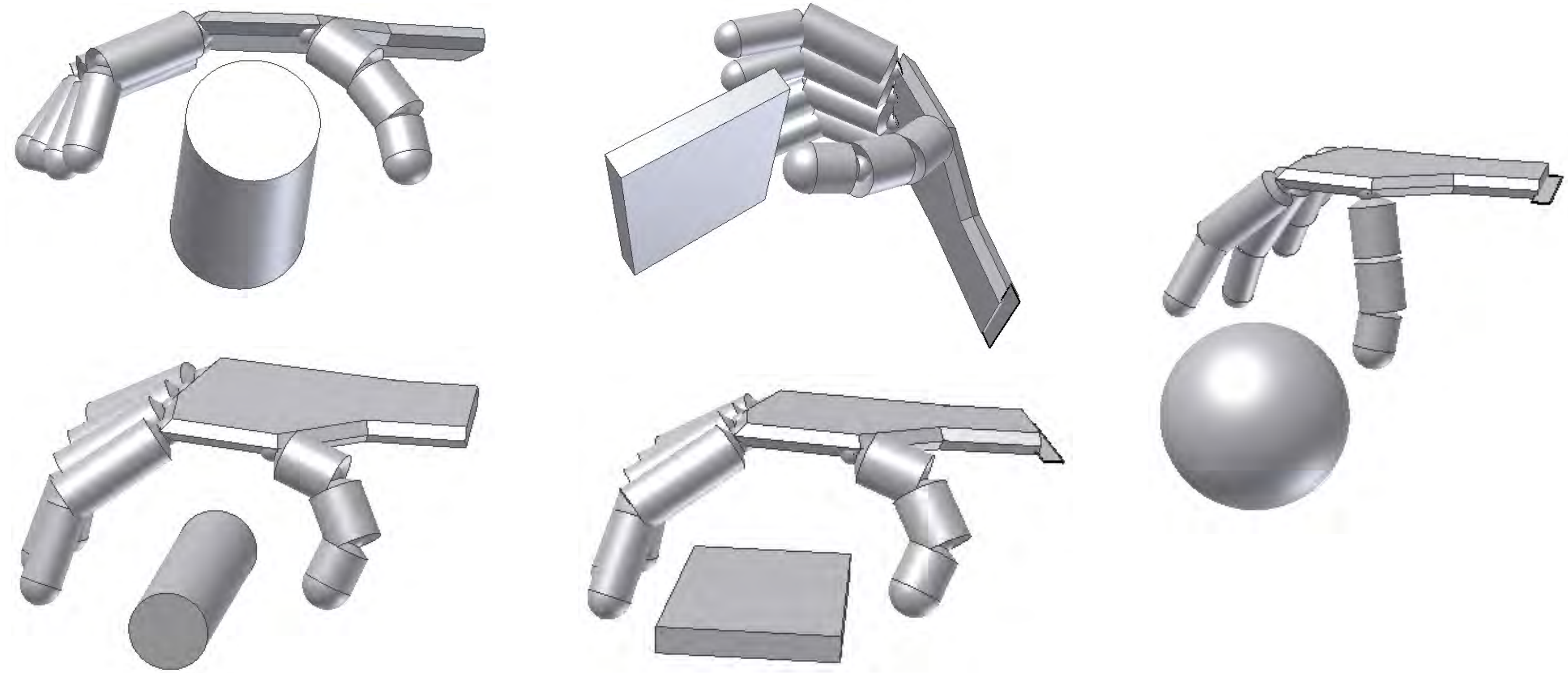
.

- SOS and AOS activations (from CIP) are complemented with absolute **object distance** (LIP) and **size** (CIP, AIP)
- Grasp plan takes into account different **visual and motor criteria** (contribution of the Basal Ganglia):
  - visual criteria: center of mass; grasping margin; curvature;
  - motor criteria: finger extension; force distribution;
- Criteria modulation is done with the **contribution of the ventral stream**, according to object identity and recognition confidence, considering weight, density, roughness, fragility, compliance, …
- Assume a general task, such as safe lifting of the target object

Grasp recipe

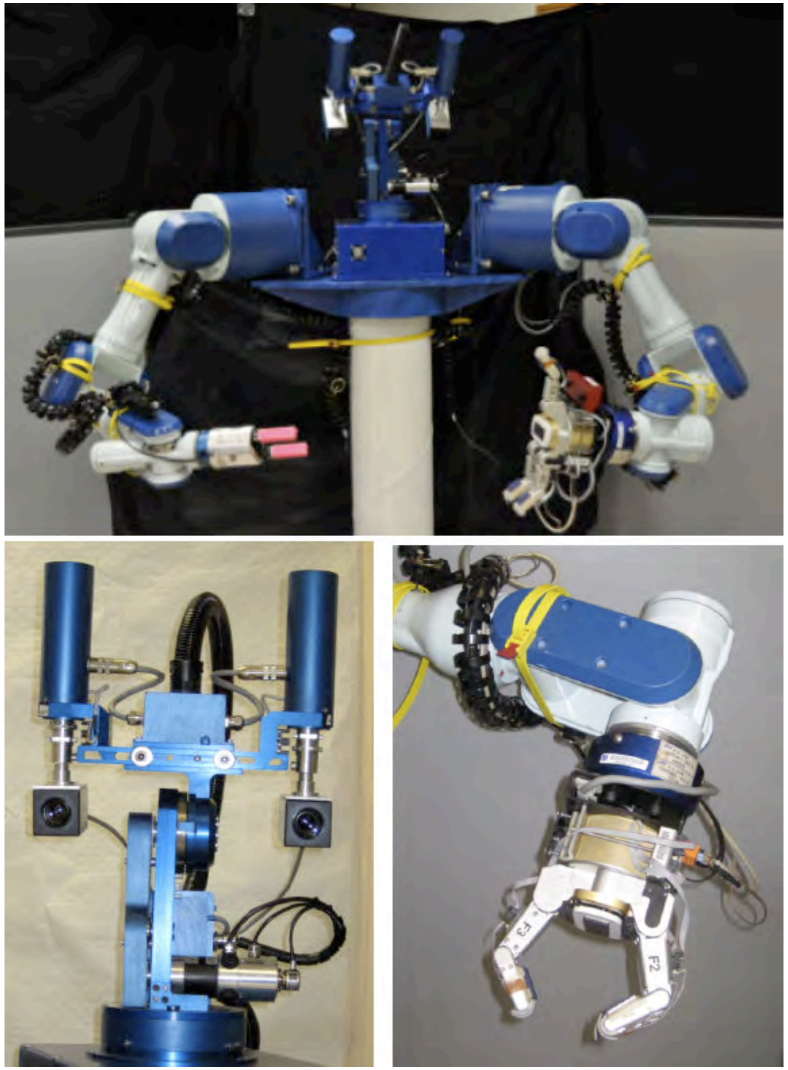
Define an opposition space between object and hand considering visual and proprioceptive information, balancing quality requirements and reaching comfort (movement cost)

Dominance of SOS or AOS activation affect approaching direction



Robot Setup

UNIVERSITAT JAUME•I

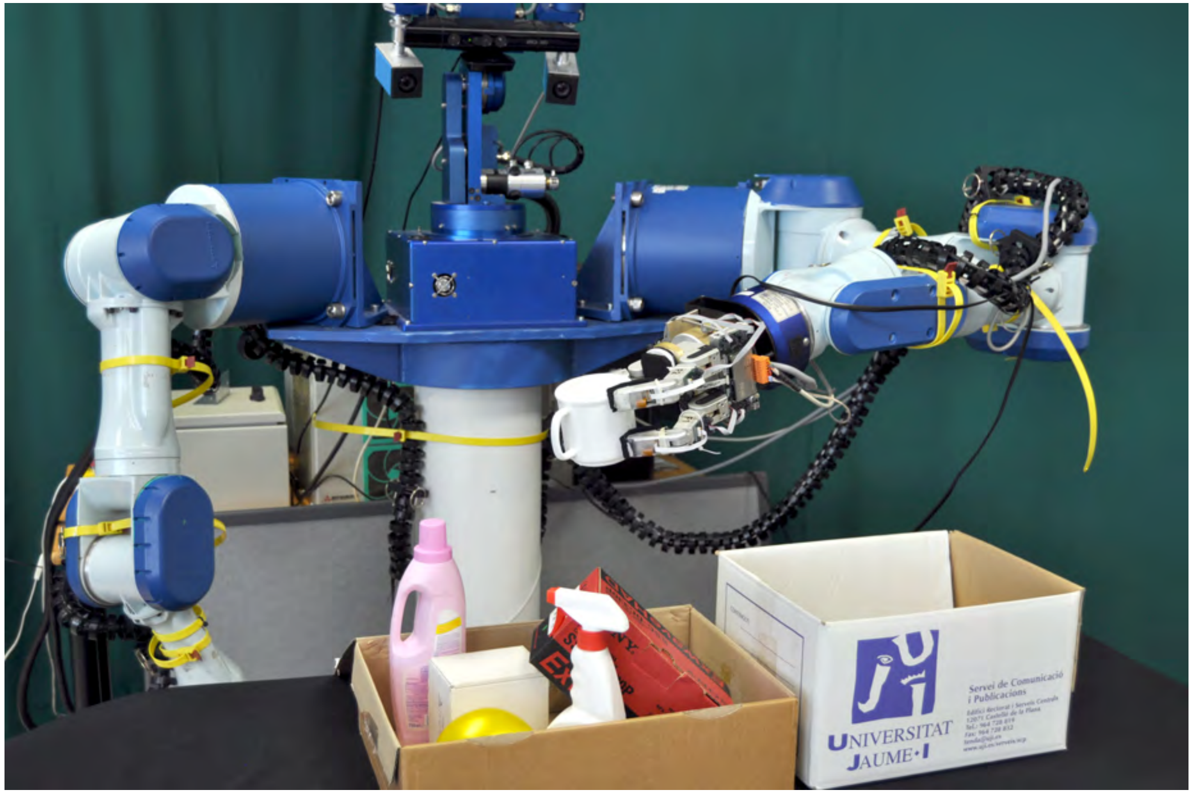


Grasp cooking

- The **robot extracts** all data required for the grasping action, and follows the above guidelines to **approach** the object and **shape** the hand
- Use of via-postures during reaching:
  - the first reaching phase brings the hand close to the object and in the right orientation;
  - during the second reaching phase the hand reaches the planned opposition axis and closes on the object;
- tactile feedback is used to stop action execution and assess its outcome


Video proof?

UNIVERSITAT
JAUME•I



Visit RobIn Lab channel at YouTube:

http://www.youtube.com/user/RobotIntelligenceLab

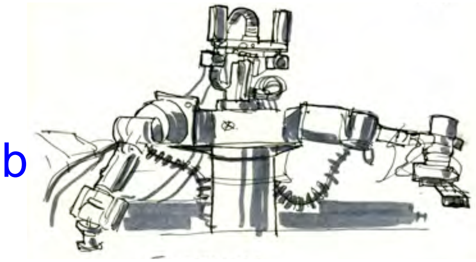
Conclusions

.

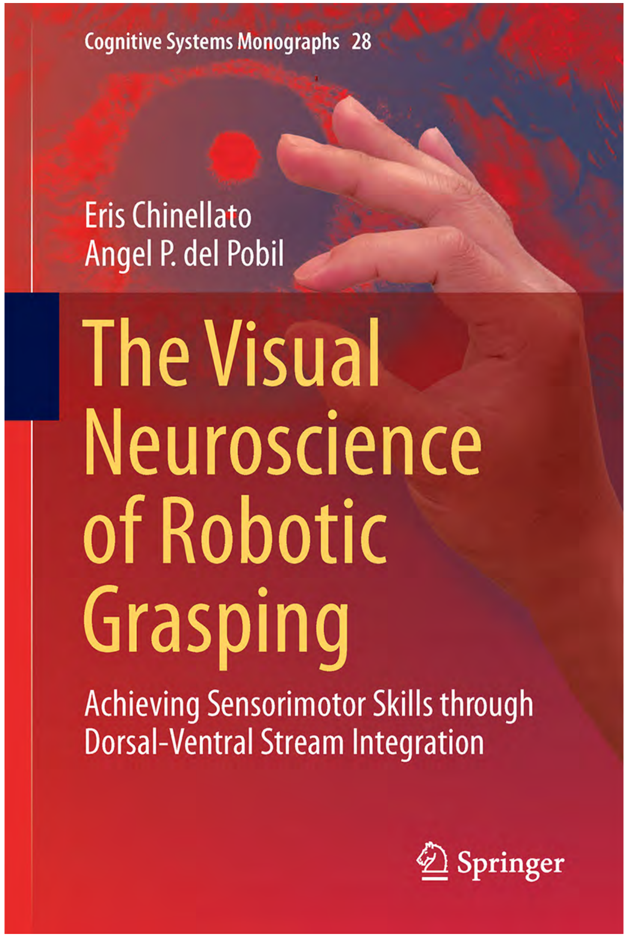
- An interdisciplinary approach to the task of object grasping, **bridging the gap between neuroscience and robotics at the computational level**.
- A **model of vision-based grasp planning and execution** strongly based on **neuroscience** data but with a **practical** stance:
  - findings are analyzed with the purpose of understanding actual **functions** and **data flow**;
  - practical implementation of grasping skills is considered;
  - special focus is put on the **integration** between the **dorsal** and **ventral** visual pathways;
- A robotic system in which previously unknown 3D objects can be grasped upon **visual estimation of their location, size and pose**, and in which grasp execution is reinforced by tactile feedback
  - the implementation builds on the integration between the streams as postulated by the proposed model.

Thank you!

UJI Robotic
Intelligence Lab



UNIVERSITAT
JAUME•I



E. **Chinellato**, A.P. **del Pobil**

The Visual Neuroscience of Robotic Grasping

Achieving Sensorimotor Skills through Dorsal-Ventral Stream Integration

Series: Cognitive Systems Monographs, Vol. 28

- **State of the art robotics and neuroscience research**
- **A novel point of view on the two visual streams theory, based on both theoretical and practical considerations, including robot implementations of computational models**
- **An insightful read both for life scientists and engineers**

ISBN: 978-3-319-20302-7
DOI: 10.1007/978-3-319-20303-4

Springer **springer.com**