- EYESHOTS -
Heterogeneous 3D Perception across Visual Fragments

Heterogeneous 3D Perception across Visual Fragments

STREP Project funded by the European Commission, Unit E5 "Cognitive Systems, Interaction, Robotics", under the 7th Framework Program

EYESHOTS

Grant No.: 217077
Duration: 3 years
Starting date: 1.03.08

Budget: 2,400,000 Euro
Total effort: 415.5 PMs

Principal investigators & expertise:

Giorgio Cannata [1,2,3]
Angel del Pobil [1,3,4]
Patrizia Fattori [6,9,10]
Fred Hamker [4,7,8]
Markus Lappe [6,7,8,9]
Silvio P. Sabatini [1,5,7]
Marc Van Hulle [4,5,7]

[1] Robotics
[2] Biomechanical models
[3] Motor control
[4] Machine learning
[5] Computer vision
[6] Experimental neuroscience
[7] Theoretical neuroscience
[8] Cognitive psychology
[9] Psychophysics
[10] Neurophysiology

Keywords:
Bio-inspired robotics
Depth vision
Spatial cognition
Active stereopsis
Eye movements
Visual attention
Vision and action
Perceptual learning
Visuomotor coordination
Parietal area V6A

Contact:
Dr. Silvio P. Sabatini
DIBE - University of Genoa
☎ +39-010-3532092
✉ silvio.sabatini@unige.it

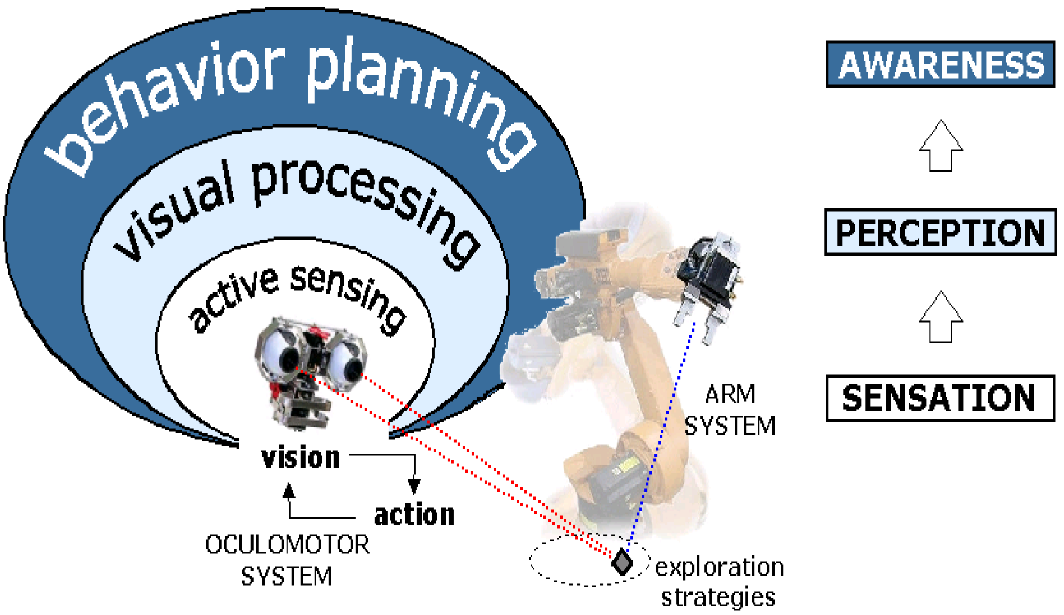
www.eyeshots.it

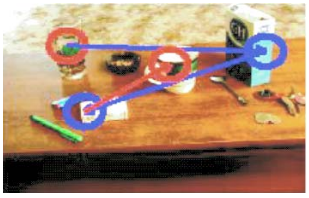
Rationale

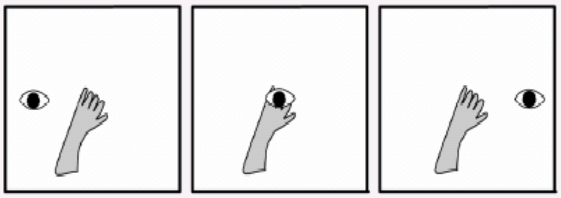
Humans act in space. Sometimes, interactions are explicit, as we point, reach or grasp the things around us. Other interactions are implicit, an awareness of where we are and what things surround us. In general, to interact effectively with the environment, it can be argued that humans might use complex motion strategies at ocular level (but possibly extended to other body parts, e.g. head and arms, so possibly using multimodal feedback) to extract information useful to gain 3D awareness of the peripersonal space as *contingent*, *head-centric, heterogeneous* representations, which are coherent and stable with respect to time. An important difference between human and artificial vision systems is that humans become aware of their visual environments by making both reflexive and exploratory eye-movements (seeing vs. looking). This awareness is a crucial factor in achieving truly anthropomorphic systems that are able to interact with humans and other anthropomorphic systems, and it is an important challenge in machine intelligence and artificial cognitive systems (EC, 7th Framework Program, ICT Challenge 2).
The steps that need to be taken for building the stable percept of the 3D environment are adaptive or subject to learning: the sensorimotor representations, the active stereopsis, and the binding of the visual fragments in the 3D scene that are visually scanned. ■

Goals

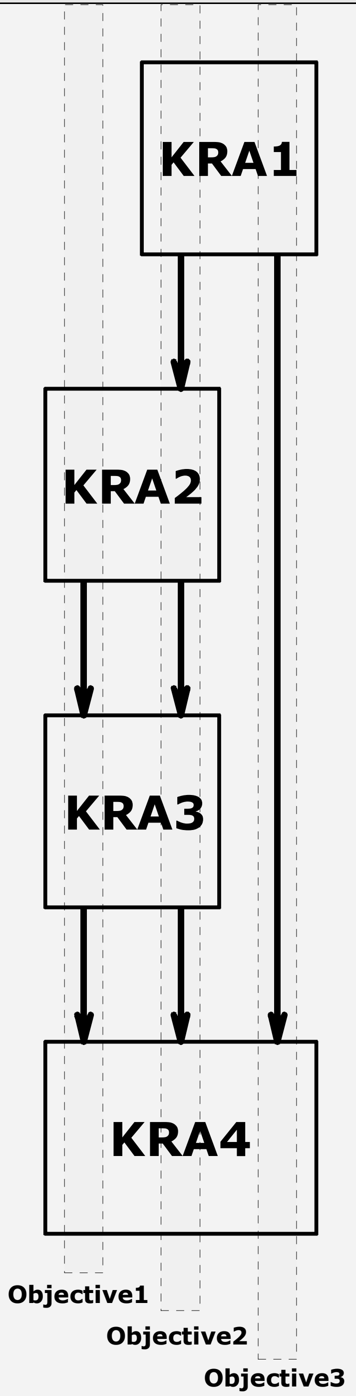
The research intends to investigate the interplay existing between vision and motion control, and to study how to exploit this interaction to achieve a knowledge of the surrounding environment that allows a robot to act properly. Robot perception can be flexibly integrated with its own actions and the understanding of planned actions of humans in a shared workspace. The research relies upon the assumption that a complete and operative cognition of visual space can be achieved only through active exploration of it: the natural effectors of this cognition are the eyes and the arms.
Crucial but yet unsolved issues we address are object recognition, dynamic shifts of attention, 3D space perception including eye and arm movements including action selection in unstructured environments. We propose a flexible solution based on the concept of visual fragments, which avoids a central representation of the environment and rather uses specialized components that interact with each other and tune themselves on the task at hand.
In addition to a high standard in engineering solutions the development and application of novel learning rules enables our system to acquire the necessary information directly from the environment.
The study and models of human/primate behavior, based on specific experiments, guide many of our envisaged solutions. ■







"Seeing" is something we do* rather than a sequence of hierarchical interpretative processes. From this perspective, *the experience of "seeing" in not necessarily generated, but it expresses itself in the behaviour. [O'Regan & Noe, 2001]



Key Research Actions

KRA1: Constructing visual perception of space by interactive stereopsis: The search for optimal visuomotor coordination to achieve robust and stable visual percepts is challenged. KRA1 provides an input to KRA2 contributing to the definition of a visual fragment of the observed scene.

KRA2: Recursive modulation of perception across visual fragments: Definition of a strategy to achieve a global perception of the 3D spatial relations and relative 3D motion for controlling spatially directed actions (e.g., reaching), and, in general, visually-guided goal-directed movements in the whole peripersonal workspace. KRA2 will focus on: (1) An attentional-based selection of visual fragments, (2) a construction of peripersonal space across eye movements.

KRA3: Visuospatial awareness and planning behaviour: The problem of constructing an action-minded representation of the 3D space is addressed. KRA3 will contribute to: (1) The definition of joint representation signals of eyes and hand movements in a 3D extrinsic coordinate frame, on which to base the 3D location of a visual target with respect to a point on the body surface; (2) the definition of shared attention behaviour in common workspaces.

KRA4: Human behaviour replicas by integration/interactive paradigms: This is a technical KRA concerned in the "translation" of the scientific achievements of KRAs1-3 into operative modules/subsystems (hw/sw robotic systems) characterized by perceptual/cognitive capabilities that emulate the human behaviour. ■

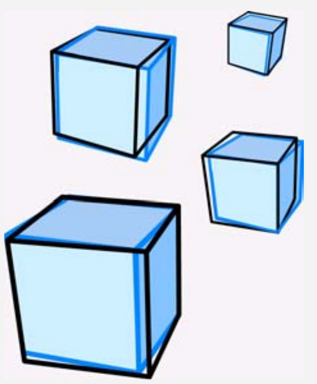
Specific objectives

Objective 1: Development of a robotic system for interactive visual stereopsis.
The function of the systems is to interactively explore the 3D space by active foveations. Benefits of the motor side of depth vision are expected to be *bidirectional* by learning optimal sensorimotor interactions

Objective 2: Development of a model of a multisensory egocentric representation of the 3D space.
The representation is constructed on (1) binocular visual cues, (2) signals from the oculomotor systems, (3) signals about reaching movements performed by the arm. Egocentric representations require regular updating as the robot changes its fixation point. Rather than continuously updating based on motor cues or a visual mechanism (i.e. optic flow), the model updates only the egocentric relationship and object-to-object relationships of those objects currently in the field of view. During motion, the model covertly and overtly shifts attention to objects in the environment to maintain the model's current awareness of the environment. The updating of the internal representation of spatial relations requires binding processes across the different visual fragments.

Objective 3: Development of a model of human-robot cooperative actions in a shared workspace.
By the mechanism of shared attention the robot will be able to track a human partner's overt attention and predict and react to the partner's actions. This will be extremely helpful in cooperative interactions between the robot and a human. ■

Consortium



UNIVERSITA' DEGLI STUDI DI GENOVA

ALMA MATER STUDIORUM - UNIVERSITA' DI BOLOGNA

WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER

UNIVERSITAT JAUME·I

KATHOLIEKE UNIVERSITEIT LEUVEN