

## EYESHOTS

### Heterogeneous 3D Perception across Visual Fragments

S.P. Sabatini, G. Cannata, M. Lappe, F. Hamker, P. Fattori, C. Galletti, A. del Pobil, M. Van Hulle

Humans act in space. Sometimes, interactions in space 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 build representations of the 3D space, which are coherent and stable with respect to time. Such representations rely on a multisensory active exploration of the environment. In particular, purposive vision is an important source of information and provides a number of cues about the 3D layout of objects in a scene that could be used for planning and controlling goal-directed behaviours. "Seeing" is something we do rather than a sequence of hierarchical interpretative processes. From this perspective, the experience of "seeing" is not necessarily generated, but it expresses itself in the behaviour.

Following these premises, the project investigates the interplay existing between vision and motion control, and studies how to exploit this interaction to achieve knowledge of the surrounding environment that allows a robot to act properly. Robot perception will 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 behaviour, based on specific experiments, guide many of our envisaged solutions.

The research activity will result in: (1) A robotic system for interactive visual stereopsis [composed of: an anthropomorphic mechatronic binocular system; and software vision modules based on cortical-like population, to be used as an experimental platform]; (2) A model of a multisensory egocentric representation of the 3D space [constructed on binocular visual cues, signals from the oculomotor systems, signals about reaching movements performed by the arm]; (3) A model of human-robot cooperative actions in a shared workspace [relying on the concept of shared attention to understand the intention or goal of the communicating partner].

The approach follows the development of four Key Research Actions (KRAs):

*KRA 1: 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.

*KRA 2: 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.

*KRA 3: 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.

*KRA 4: 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.