PROJECT FINAL REPORT

Grant Agreement number: 217077

Project acronym: EYESHOTS

Project title: Heterogeneous 3-D Perception across Visual Fragments

Funding Scheme: Collaborative project

Period covered: from 01/03/2008 to 28/02/2011

Name, title and organisation of the scientific representative of the project's coordinator:

prof. Silvio P. Sabatini

Department of Biophysical and Electronic Engineering - DIBE

University of Genoa - UG

Tel: +39-010-353-2092 (2289)

E-mail: silvio.sabatini@unige.it

www.eyeshots.it

TABLE OF CONTENTS

1 FINAL PUBLISHABLE SUMMARY REPORT

1.1 Executive summary

The present report synthesises the major outcomes of the project EYESHOTS (*Heterogeneous 3-D Perception across Visual Fragments*, grant no. 217077), a Collaborative Project funded by the European Commission through its Cognitive Systems, Interaction, Robotics Unit (E5) under the Information and Communication Technologies component of the 7th Framework Programme (FP7), and conducted from March 2008 up to February 2011. The consortium is composed of 7 research units of 5 research centres:

University of Genoa, Italy	(UG)
Westfälische Wilhems-University Münster, Germany	(WWU)
University of Bologna, Italy	(UNIBO)
University Jaume I, Castellòn, Spain	(UJI)
Katholieke Universiteit Leuven, Belgium	(K.U.Leuven)

with expertise ranging from robotics, computer vision, neuroscience and experimental psychology.

At the end of the **first year**, numbered among the project's assets were a front-end vision module providing a cortical-like representation of the binocular visual signal for both vergence control and depth estimation, and a conceptual framework for modelling cortical ventral/dorsal interactions in reaching (and grasping) actions. At that stage, the eyes and the arm system were considered as separate effectors. In the **second year,** we more decisively addressed the problem of combining 3D space information obtained through active ocular and arm movements, with the final objective of controlling spatially directed reaching actions, and, in general, visually-guided goal-directed movements in the whole peripersonal workspace. To this end, a first level of integration has been achieved both for (1) the *visuomotor coordination of eye movements* (K.U.Leuven, UG, WWU) and for (2) the *visual/oculomotor/arm-motor coordination* (UJI). In the **third year**, the work focused on the integration of the different modules on the robot platforms to validate the approach in real-world conditions. Specific achievements were: (1) the development of working modules for stereopsis and oculomotor control; (2) the validation of interactive stereopsis behavior in real-world situations on an iCub platform; (3) the realization of a non-conventional tendon-driven mechatronic binocular system; (4) the integration of perceptual/visuomotor strategies on the eye/head-arm UJI robot.

Together with these integration activities, the analysis of the experimental data yielded significant results on the neural mechanisms used to link different fragments through visual, attentional, oculomotor, and arm-movement related cues. Neurophysiological findings of UNIBO showed that a large majority of cells in area V6A are modulated by ocular and/or reaching movements in 3D. These results have contributed to the definition of an integrated representation of visual, arm and eye sensorimotor information on which to base the 3D localization of visual targets with respect to the body. Concerning the motor influence of visual perception, the analysis of the experimental data of WWU on saccadic adaptation yielded remarkable evidences of the oculomotor components of visual target localization, which have been included in the final model.

All the processing modules are built on the basis of distributed representations in which sensorial and motor aspects coexist explicitly or implicitly. The models resort to a hierarchy of learning stages at different levels of abstraction, ranging from the coordination of binocular eye movements (e.g., learning disparity-vergence servos), to the definition of contingent saliency maps (e.g., learning of object detection properties), up to the development of the sensorimotor representation for bidirectional eye-arm coordination. On our opinion, this can be considered, an interesting methodological result of the EYESHOTS project. Through the distributed coding, indeed, it is possible to avoid a sequentialization of sensorial and motor processes, that is certainly desirable for the development of cognitive abilities at a pre-interpretative (i.e., sub-symbolic) level, e.g., when a system must learn binocular eye coordination, handling the inaccuracies of the motor system, and calibrate the active measurements of the space around it. Distributed computing is crucial to guarantee accessibility and interaction of the information at different levels of coding and decoding, by postponing decisions as much as possible.

The up-to-date results, videos and other dowloadable material on the project work are available on the project web-site www.eyeshots.it

1.2 Summary of the project context and objectives

The goal of EYESHOTS was to investigate the interplay that exists 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 relied 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 (see Fig.1).

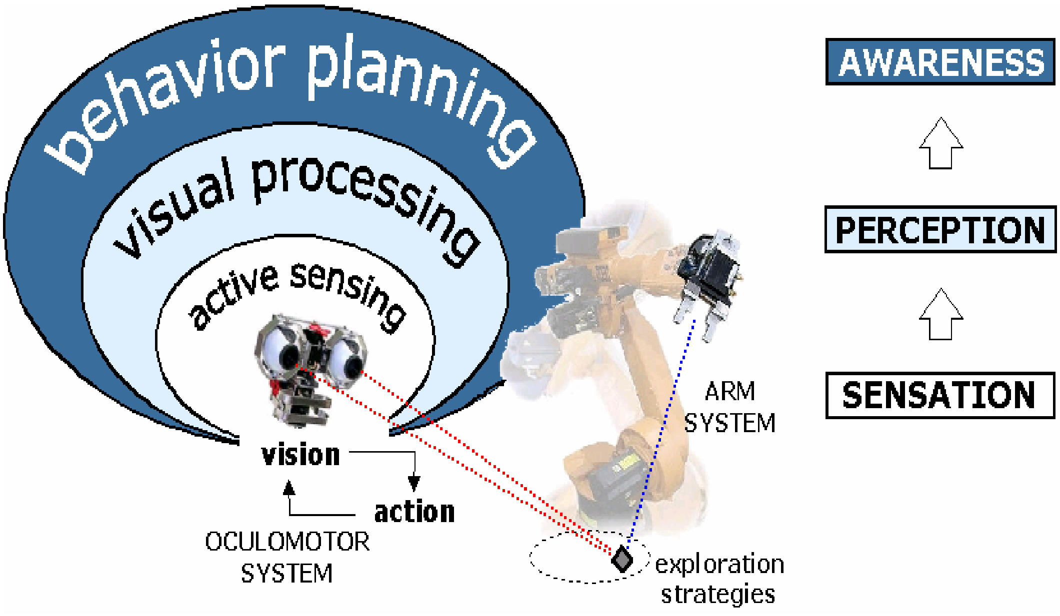

Figure 1: *The EYESHOTS perspective for 3D space perception.*

Crucial issues we addressed were object recognition, dynamic shifts of attention, three-dimensional (3D) space perception including eye and arm movements, and action selection in unstructured environments. The project proposes a flexible solution based on the concept of *visual fragments*, which wanted to avoid 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 enable the system to acquire the necessary information directly from the environment.

1.2.1 Specific objectives

The project aimed to reach the following **three specific goals**:

Goal 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 bi-directional by learning optimal sensorimotor interactions.

Goal 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.

Goal 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.

Having in mind these objectives, the project activities articulated in the concurrent development of the following major themes:

- **Design (and control) of an anthropomorphic mechatronic binocular system**

  The starting point of the project was to complete the understanding of the biomechanics of the ocular motions in humans and primates and to transfer these results into guidelines for the design of robotic eyes which could provide different solutions for the implementation of humanoid robots.

  Beside the specific robotic applications, it is assumed that the implementation of a bio-inspired robot eye (or robot head) is also the starting point for the analysis and the assessment of the motor control strategies implemented by the brain to drive the very high dynamics of ocular rotation.

  In this sense it was, and it is still considered, a key feature of EYESHOTS the target of developing a prototype of robot eye featuring bio-inspired concept and design, which are strongly different from the conventional *stiff* pan-tilt platforms. The basic idea is that *emulating* [ocular motions] *is different from simulating* [them]. In other terms, it is possible with a *conventional* robot system to obtain a desired target behavior by constraining it at control level; it is however, in general, not possible to achieve the same behavior as an emerging one due to the implicit characteristics of the plant.

  It is then reasonable to assess, from an engineering point of view that state-of-the-art conventional *stiff* robots can guarantee high accuracy and (reasonably) high speed, but they cannot allow us to perform experiments where the motion characteristics arise from the intimate nature of the mechanics of the plant itself.

- **Early perception-action cycles in binocular vision**

  According to the current trends of active, purposive vision systems, the motor system of a humanoid robot should be an integral part of its perceptual machinery. Traditionally, however, in robot vision systems, perception-action loops close at a "system level" (by decoupling *de facto* the vision modules from those dedicated for motor control and motor planning), and the exploitation of the computational effects of the eye movements on the visual processes are very seldom in artificial artifacts.

  The limit of the approach was that solving specific high-level tasks usually requires sensory-motor shortcuts at the system level, and specific knowledge-based rules or heuristic algorithms have to be included to establish behavioural consistency relationships among the extracted perceptual features and the desired actions. The risk is to abandon distributed representations of multiple solutions to prematurely construct integrated description of cognitive entities and commit the system to a particular behaviour.

  Conversely, our claim was that early/complex interactions between vision and motor control are crucial in determining the effective performance of an active binocular vision system with a minimal amount of resources and coping with uncertainties and inaccuracies of real systems. The complexity of integrating efficiently and with adequate flexibility the different aspects of binocular active vision prevented, indeed, till now a full validation of the visuomotor approaches to 3D perception in real world situations.

  The research moved from the belief that the advantages of binocular visuomotor strategies could be fully understood only if one jointly analyzes and models the problem of neural computation of stereo information, and if one takes into account the limited accuracy of the motor system. Defending an early visuomotor approach to 3D perception, we looked for instantiations of visuomotor optimization principles concurrently with the design of distributed neural models/architectures that can efficiently embody them.

- **Developing cognitive vision modules and oculomotor control strategies for fixating selected and memorized targets**

The ability to share a peripersonal workspace strongly relies on the processing of visual information, on a working memory, and on appropriately acting to the visual input. Moreover, vision has to be an active process and under high level cognitive control (humans normally expect a certain object to appear in their visual field given that they move their eyes). Our approach for object recognition was based on the concept of visual attention (Hamker, 2005a), which describes the ability of the visual cortex to focus processing resources on a certain object or visual fragment (feature-based attention) or on a certain location (spatial attention). The transition of feature-based attention into a spatially selective attention signal should be achieved by an oculomotor loop via the frontal eye field (FEF) which selects the location of a particular object for a saccadic eye movement.

Learning was expected to be a very crucial property for flexible and adaptive object recognition. The learning of appropriate receptive fields occurs indeed in the whole hierarchy of the visual stream: in early visual areas (e.g. V1), it is possible to learn simple visual fragments (e.g. edges) so that higher visual area neurons can learn to represent a single view of an object.

Finally, vision should be accompanied by executive functions that allow for a selection between different behavioral alternatives as well as to use and hold previously visible information in working memory (WM) if beneficial for the task. These functions would relate to the development of a model of cortical-basal-ganglia-loops. This model should learn the target object for a particular action – the object which will lead to reward. Knowledge about the target object can then be used to guide attention or to maintain object information in WM and use it together with the present visual input for a final decision about motor action.

- **Neurophysiological evidences of joint visuomotor descriptors of the 3D space in the parietal cortex**

In order to look to or reach for objects of interest, information about the site of retinal activity must be combined with information about eye (and head) position. Thus, the twin issues of spatial representation and coordinate transformation within the central visual pathways have been the subject of vigorous investigation in the recent years and are eminently deserving of continued study. Considerable progress has been made in understanding the role of the parietal and frontal cortices in these computations, but most of the studies have been performed with targets located in frontal plane (parallel to the eye' plane). Our objective in EYESHOTS was to analyze the functioning of these systems in more natural situations, in which the animal look or reach for a stationary object located in positions in depth in the peripersonal space. These experiments are technically difficult since they need to be conducted in a more natural or realistic setting, while still maintaining appropriate experimental control.

Neurophysiological experiments had to be conducted in the medial parieto-occipital cortex, located in the caudal part of the superior parietal lobule. This is a very crucial node of brain, at the boundary between areas that analyse information on passive sensory modalities and areas involved in coordinating active eye- and arm-movements (Galletti et al., 2003). The selected cortical area contains neurons responsive to visual stimuli (Galletti et al., 1996; Galletti et al., 1999), as well as cells modulated by somatosensory inputs, mainly from the upper limbs (Breveglieri et al., 2002), and arm movement-related neurons (Galletti et al., 1997; Fattori et al., 2001). Other previous works demonstrated that the monkey medial parieto-occipital cortex is involved in elaborating eye movement signals both for fixation (Galletti et al., 1995), for saccadic eye movements (Galletti et al., 1995; Kutz et al., 2003), and for reaching movements (Marzocchi et al., 2008). In EYESHOTS we aimed at studying whether this crucial node within the dorsal visual stream (Goodale and Milner, 1992) encodes target locations in 3D, through the concurrent use of information on vergence eye movements and arm movements in depth. This characterization wanted to provide indications on the role of non-visual cues (such as the eyes' version and vergence angles) on the perception of the 3D space as well as the role of visual cues in mastering the 3D peripersonal space.

- **Psychophysical evidences of motor-contingency of the peripersonal sensory space**

We hypothesized that motor signals used for saccade execution are also used for the awareness of the spatial locations of the sequence of visual fragments, and that, because these motor signals are plastic, perceptual awareness of peripersonal space is dynamic as well, in the sense that it can be adapted to the motor parameters. Using saccadic adaptation as an experimental paradigm we wanted to study the shaping of perceptual space by sensorimotor contingencies. A motor-contingency of peripersonal sensory space would allow a direct mapping between a fragment's location and the motor command to reach the fragment by an eye movement. Such a direct mapping would ease motor planning and would allow quicker goal-directed movements.

- **Integration of a bio-inspired sensorimotor representation of the reachable space on an anthropomorphic robot platform**

A fundamental contribution expected from EYESHOTS was to take inspiration from neurophysiological and psychophysical findings in order to design and implement on a humanoid robot a model for achieving visuomotor awareness of the environment by using eye and arm movements.

Starting from an integrated object representation which includes cognitive and visuomotor aspects of surrounding stimuli, the artificial agent (either a real robot or a simulation) was expected to be able to interact with stimuli in its peripersonal space by performing an active exploration of it. Such exploration would allow the agent to: 1) learn to coordinate and associate visual stimuli with oculomotor and arm motor movements; 2) build an implicit visuomotor representation of potential targets in its peripersonal space.

Taking as main inspiration the data and insights provided by the experiments performed by partners UNIBO, to which UJI was also expected to collaborate, the model would implement, in a biologically plausible fashion, the sensorimotor transformations performed by the primate posterior parietal cortex, with special emphasis on the role of the dorso-medial stream, and namely on area V6A.

The outcome of applying such a model on a real humanoid robotic platform should be a set of basic skills, such as concurrent or decoupled gazing and reaching movements toward visual stimuli. The robot was planned to be able to show its visuomotor capabilities by performing oculomotor actions toward visual targets placed in its peripersonal space, or toward the location of its hand. Moreover, it should also able to perform arm reaching movements to visible objects, either with or without gazing at them.

1.3 Main S&T results/foregrounds

1.3.1 Constructing visual perception of space by interactive stereopsis

Vision is the first source of information about the 3D space. The search for optimal visuomotor coordination to achieve robust and stable percepts does pose a major challenge.

1.3.1.1 Anthropomorphic robotics eye system

There exists the evidence of particular types of eye movements (typically saccades and smooth pursuit) which obey to a basic geometric principle known as Listing's Law. Listing's law specifies that the amount of torsion during saccades and smooth pursuit is zero. This property cannot be achieved by conventional pan-tilt mechanisms unless torsion is properly and actively controlled. However, as the kinematics supporting Listing's Law is not straightforward, the basic question arises how Listing compatible movements could be achieved on a *generic* kinematic structure unless very complex *control circuits* (in a neuro-control framework) or *control models* (in a robot control framework), possibly based on sophisticated sensing, are used.

We have proved that the origin of the characteristics of Listing compatible movements can be grounded on the geometric and mechanical characteristics of the oculo-motor plant. As a matter of fact a reasonably simple model can be defined to achieve Listing's compatible movements independently of the control actions generated by the actuators. This means that by implementing a non conventional robot following the guidelines specified by the models investigated throughout the project is possible to *naturally* achieve ocular behaviours that approximate to a large extent the motion of a real eye.

Therefore, we have pursued a *strong bio-inspired approach* trying to emulate in the mechanical implementation all the major features arising from the analytical models developed. Despite the simplicity of the model, its implementation on a robot testbed has not been straightforward, and anyway subject to various engineering and technological trade-offs.

The geometry of the robot prototype is based on a bio-mechanical model of the ocular plant including extra-ocular muscles dynamics, and a solution to emulate the mechanical properties (viscosity and elasticity) of the eye orbital soft tissues has been considered. The starting point was the model of Bahill et al. (1980) and applied to minimum-time saccadic control by Enderle et al. (1984, 1987). In particular the work has focused on the extension of these models originally developed in 2D to the three dimensional case by extending the eye model proposed in (Cannata & Maggiali, 2008) where it is shown that physiologically correct saccadic motions could be generated by any action of the four recti extra-ocular muscles. The relevant subsystems are: the eyeball and the vision system, the support structure, and the actuation system.

Robot Eye Design

The robot eye prototype must emulate the mechanical structure and the movements of a human eye with a comparable working range. This robot eye has approximately the shape of a truncated cone where the larger diameter and smaller ones are 34 and 45 mm, respectively with an overall lenght of 120 mm. The robot is designed on the assumptions that the eyeball is a sphere with three degrees of freedom about its center and the actuation system that drives the eyeball is a combination of linear motors, springs and tendons. Each linear motor has two springs in parallel, the motor and the springs are connected to the eyeball through the tendons. Figure 2 shows the detailed schematics of the system.

The Eyeball

The eyeball is a precision machined (in house) PTFE sphere with a diameter of 28 mm. The sphere has been machined to host the vision system (a commercial CMOS microcamera with miniature optics) and to route the video signal cables to the external electronics.

Supporting structure

The structure designed to support the eyeball, the motors and the springs is composed of four distinct components (shown in Fig. 3 and described below):

- eyeball support: a low friction support designed to hold the eyeball and to implement the pointwise pulleys,
- frame: a structure cross, attached to the eyeball support, to hold the motors and the springs,
- front flange: an anterior flange to lock the eyeball on the eyeball support,
- rear flange: a posterior flange to lock the motors and the springs on the frame.

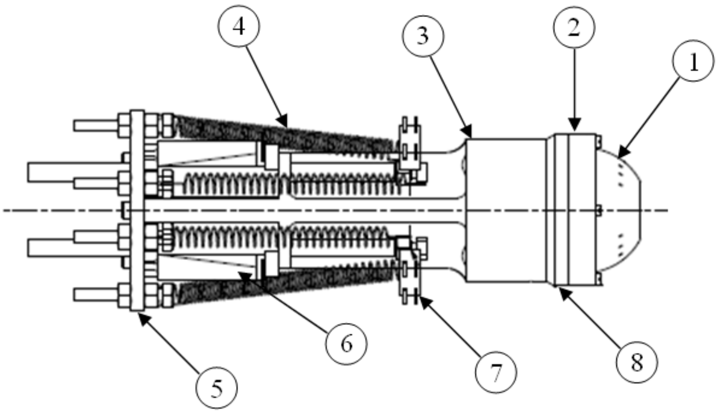

Figure 2: *Lateral view of the eye robot prototype and its modules:* ***1)*** *Eyeball,* ***2)*** *Front flange,* ***3)*** *Frame,* ***4)*** *Spring,* ***5)*** *Rear flange,* ***6)*** *Linear motor,* ***7)*** *Position sensor,* ***8)*** *Eyeball support.*

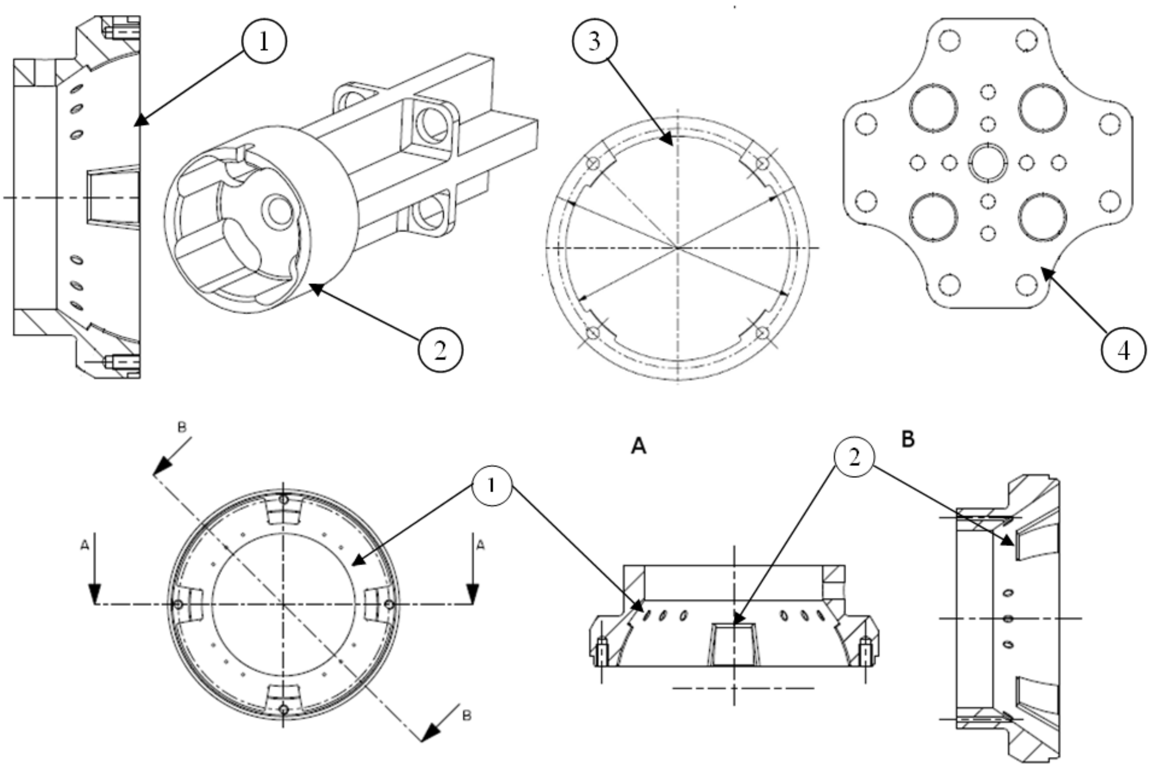

Figure 3: ***(Top)*** *CAD model of the supporting structure:* ***1)*** *Eyeball support,* ***2)*** *Frame,* ***3)*** *Front flange,* ***4)*** *Rear flange.* ***(Bottom)*** *CAD model of the eyeball support: posterior and two lateral view* ***(A,B)****:* ***1)*** *Pointwise pulley holes,* ***2)*** *Screw holes.*

The eyeball support module holds the eyeball and is the most critical part of the system. It is made of TEFLON and has the major function of implementing the pointwise pulleys (responsible to guarantee mechanical implementation of Listing's Law). The pulleys route the actuation tendons and they ensure the correct mechanical implementation of Listing's Law.

The position of the pointwise pulleys is symmetrical with respect to the position of the insertion points on the eyeball. On the eyeball support there are four groups of three pulleys: the central for the tendon attached to the linear motor and the two lateral for the tendons attached to the springs.

Figure 3 (bottom) shows a sketch of the posterior and two different lateral views of the eyeball support.

Actuation System

The actuation system is composed of four tendons, four force generator (DC brushless linear motors) and eight springs.

The tendons are thin stiff wires, connected to the rods of the motors and to the springs. The actuators pull the tendons and drive the movements of the eyeball. There are two springs in parallel to each motor that pull the tendons in the opposite direction with respect to the motor one. The main function of the springs is to emulate the elasticity of the orbit of the human eye and to restore the zero position (primary position) of the eyeball when the system is not actuated. A picture of the current prototype is shown in Fig. 4.

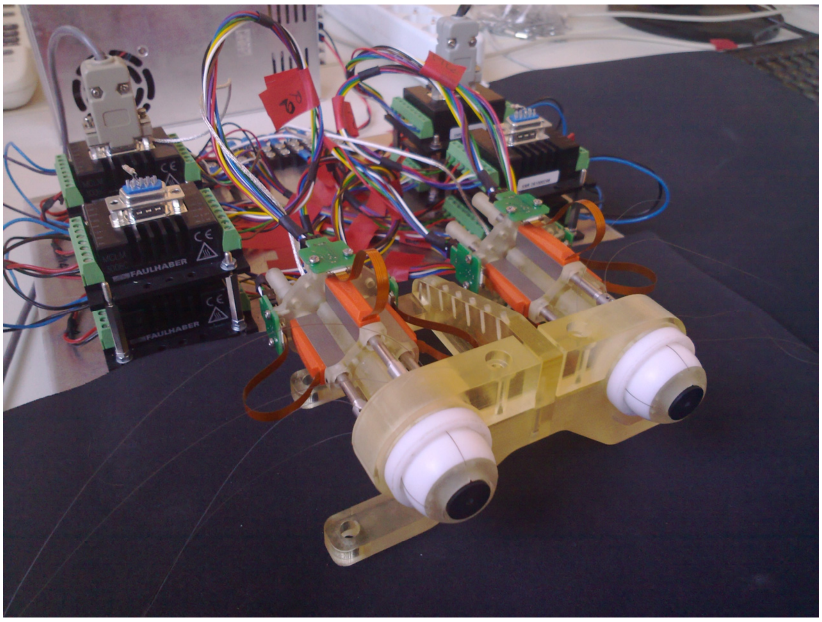

Figure 4: *Picture of the final robot prototype.*

1.3.1.2 Visuomotor paradigms and cortical-like architectures

Pushed by wide neurophysiological evidences of modulatory effects of motor and premotor signals on the visual receptive fields across several cortical areas, one of the EYESHOTS' goal concerned the design of systems where the acting itself, and even its planning, operate in parallel with perception, thus really closing the loops and taking full advantage of a concurrent/anticipatory perception-action processing. In this context, cortical-like architectures for both vergence control and depth perception (in the 3D peripersonal space), which incorporate adaptive tuning mechanisms of the disparity detectors have been investigated.

The extraction of binocular features relies upon a full (i.e., amplitude, orientation and phase) harmonic representation of the visual signal, operated by a set of "simple cell" units (S-cells). Such a representation allows us to reconsider and analyze the flexibility and robustness of the multi-channel perceptual coding (see Sabatini et al., 2010), adopted by the early stages of the mammalian visual cortex, for the "atomic" components of early vision. Oriented disparity tuning emerges in layers of binocular energy "complex cell" units (C-cells) that gathers S-cells outputs according to specific architectural schemes (Chessa et al., 2009a).

Two case-studies show how large-scale networks of cortical-like binocular cells can provide a flexible medium on which to base coding/decoding adaptation mechanisms related to sensorimotor schema.

At the coding level, the position of the eyes in the orbits can adapt the disparity tuning to minimize the necessary resources, while preserving reliable estimates (Chessa et al., 2009b). Over relatively large visual angles, the retinal disparity patterns experienced by a binocular vergent system engaged in natural viewing present, indeed, predictable components related to the positions of the eyes in the orbits. These predictable components may be used as *priors* to optimally allocate the computational resources to ease the recovery of the unpredictable components of disparity, which are dependent on the structure of the scene, only. The spatial relationships between target points on the left eye and the mean of the corresponding points for the right eye are embedded into a hybrid energy-based model (Fleet et al., 1996) where phase-shifts and position-shifts play a different role: (i) position-shifts are used to compensate the averaged epipolar disparity pattern over all the fixations, (ii) phase-shifts are used to estimate the residual 2D disparity. A statistical analysis of the disparity patterns for a fixating observer in a real-world environment highlighted differences between the different eye movement paradigms (i.e., a pan-tilt (i.e., Helmholtz) camera system, Listing's law, and its binocular extension L2), and suggests the possibility of a mutual "calibration" of the vision and the oculomotor system to compensate disparity components due to the epipolar geometry (see Fig. 5).

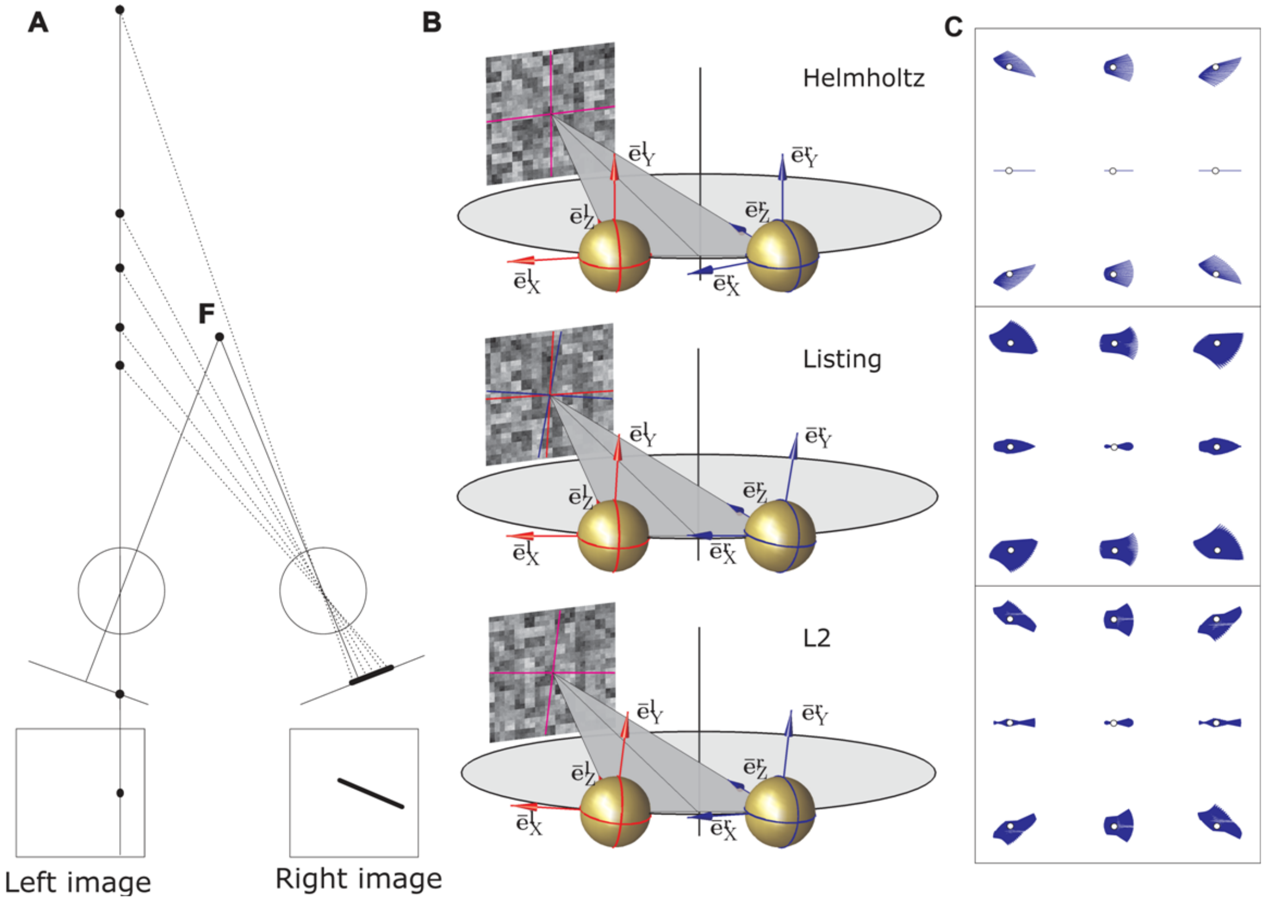

Figure 5: *(A) Given an image point on the left retina, all the possible corresponding points on the right retina lie on a line: the epipolar line (heavy black line). The epipolar lines change their orientation depending on the relative orientation of the two eyes.* ***(B)*** *In the three represented cases the eyes fixate the same point in the world but in different ways, following an Helmholtz (or Tilt-Pan), a Listing, and a L2 system. The surface is always orthogonal to the gaze line. The blue and red lines represent the projections of the horizontal and vertical meridians for the left and the right eye on the surface. In case of perfect alignment the two lines superimpose becoming purple.* ***(C)*** *For the three systems in considered, the epipolar lines for a grid of 3×3 retinal points of the left eye are depicted for different gaze direction. The reference point on the left eye are represented by an open circle, while the epipolar lines are represented by blue lines. The value of version and elevation angle varies in a range of ±40°. The epipolar lines tilt and move depending on the azimuth and the elevation of the gaze direction, spanning an area that defines the correspondence search zone. The spread of the search zone is related to the geometry of the system.*

At the decoding level, specialized read-out mechanisms can be obtained for directly extracting disparity-vergence responses, to gain linear servos with fast reaction and precision, and wide working range with a reduced amount of resources (Gibaldi et al., 2010). Short-latency disparity-vergence responses are then obtained by a weighted sum of the population response, where the weights are computed by minimizing a functional that embeds two very specific goals: (1) to obtain signals proportional to horizontal disparities, (2) to make these signals to be insensitive to the presence of vertical disparities. In particular, partner UG (Gibaldi et al., 2009, Gibaldi et al., 2010, Gibaldi et al., 2011) developed a model that, mimicking the behaviour of the cells in the Medial Superior Temporal area (Takemura et al., 2001), combines the population responses without an explicit calculation of the disparity map, but deriving a disparity-vergence response directly from the population response. That allows us to trigger the correct vergence movement and eventually nullify the disparity in the fovea even when the presented stimulus is far beyond the theoretical size-disparity correlation limit, which bounds the correct disparity estimation inside a range defined by the angular size of the receptive fields. The desired feature of the horizontal disparity tuning curve for vergence is an odd symmetry with a linear segment passing smoothly through zero disparity, which defines a critical servo range over which changes in the stimulus horizontal disparity elicit roughly proportional changes in the amount of the horizontal vergence angle. On the basis of the Dual Mode theory (Hung et al., 1986), the model provides two distinct vergence control mechanisms: a fast mode enabled in the presence of large disparities, and a slow mode enabled in the presence of small disparities. Moreover we extract an additional signal T0, sensitive to zero disparity, that automatically switches between LONG and SHORT, depending on the disparities present in the scene. We tested the proposed model both in virtual and real-world environments achieving stable fixation and small response time to a wide range of disparities. The vergence movements produced are able to drive and to keep the fixation point both on a steady and on a moving stimulus (e.g., see Fig. 6).

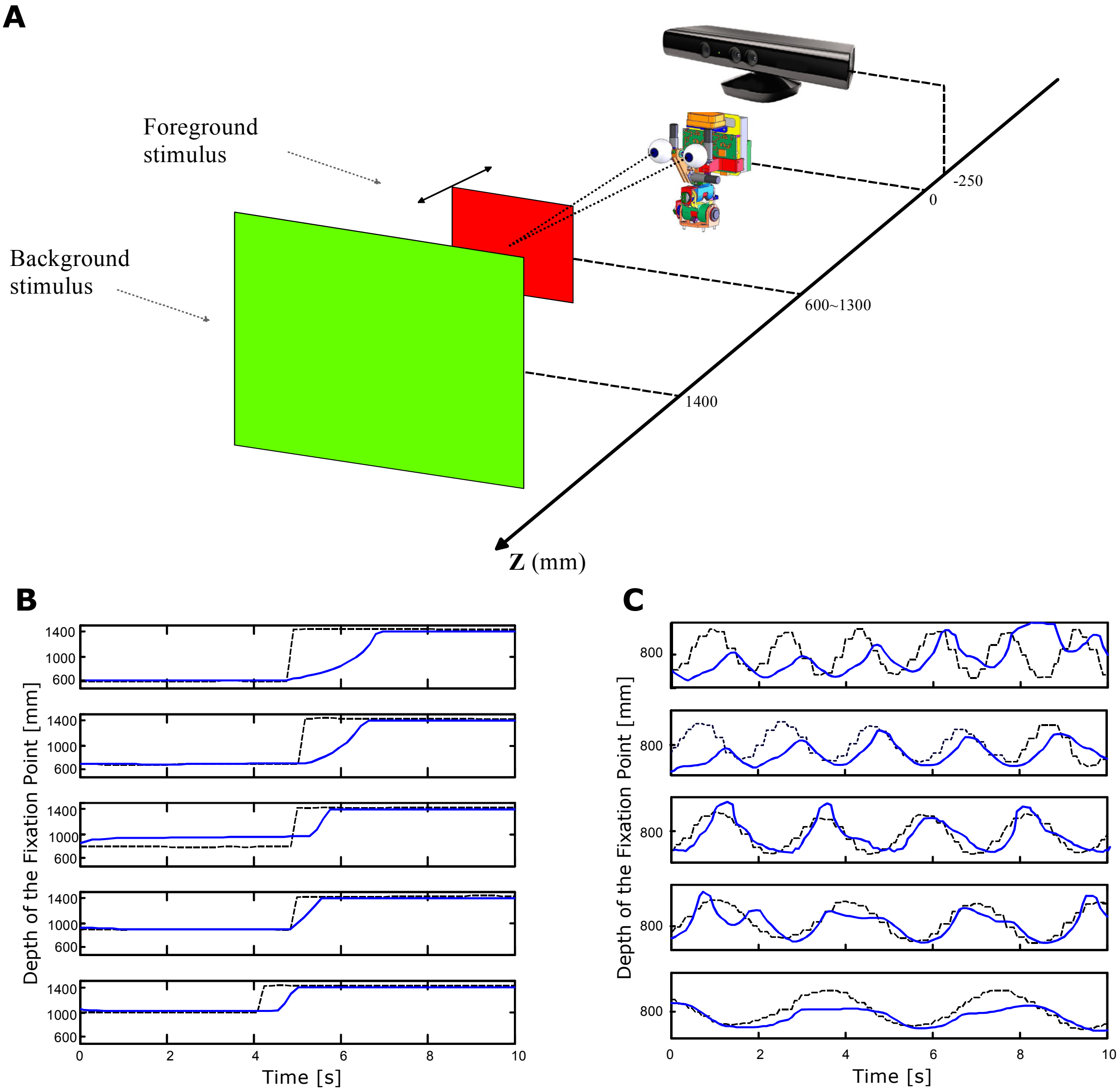

Figure 6: *(A) The experimental setup used to test the Dual-Mode vergence module on the iCub head. Trajectories of the fixation point (blue solid line) with respect to the depth of the stimulus (black dashed line), in case of a step in depth **(B)**, and of an oscillating stimulus **(C)**. The depth of the stimulus in computed from the middle point of the baseline of the iCub head.*

From a methodological point of view, the proposed approach points out the advantages and the flexibility of distributed and hierarchical cortical-like architectures against solutions based on a conventional systemic coupling of sensing and motor components, which in general poses integration problems since too heterogeneous and complex processes must be coupled. Through the distributed coding, indeed, it is possible to avoid a sequentialization of sensorial and motor processes, that is certainly desirable for the development of cognitive abilities at a pre-interpretative (i.e., sub-symbolic) level, e.g., when a system must learn (binocular) eye coordination, handling the inaccuracies of the motor system, and calibrate the active measurements of the space around it. The design strategy of these active visual cortical networks jointly involves three concurrent aspects: (i) signal processing, by defining the proper descriptive elements of the visual signal (in the Gibsonian sense) and the operators to measure them, (ii) geometry of the system and its kinematics, that directly relates to the embodiment concept, (iii) connectionism paradigms, that define neuromorphic architectural solutions for information processing and representation. The connectionism paradigm (i.e., hierarchical, distributed computing) is crucial to guarantee accessibility and interaction of the information at different levels of coding and decoding, by postponing decisions as much as possible.

For the simulated results, virtual reality fixations in a real-world peripersonal scene are considered. The scenes (stereo images and ground truth range data) are first captured by a 3D laser scanner (Konica Minolta Vivid 919) and then "explored" through a VR simulator. The developed tool has been used to create a database of real-world range data and stereo image pairs for a variety of fixations, in order to guide modeling and for algorithmic and behavioral benchmarks in real-world but fully measured environments. Benchmark images and sequences have been made available to the scientific Community (http://www.pspc.dibe.unige.it/Research/vr.html).

1.3.1.3 Learning vergence behaviors for interactive stereopsis

The objective was to develop a network-based vergence control from a population of disparity-tuned complex cells. To this end, we investigated the specialization of these disparity detectors at different levels in a hierarchical network architecture to see the effect of learning specific coding and decoding strategies for active vergence control and depth vision. The extraction of binocular features occurs through a cortical-like population network, developed by partner UG. This network provides a distributed disparity representation to the vergence control network (VC-net).

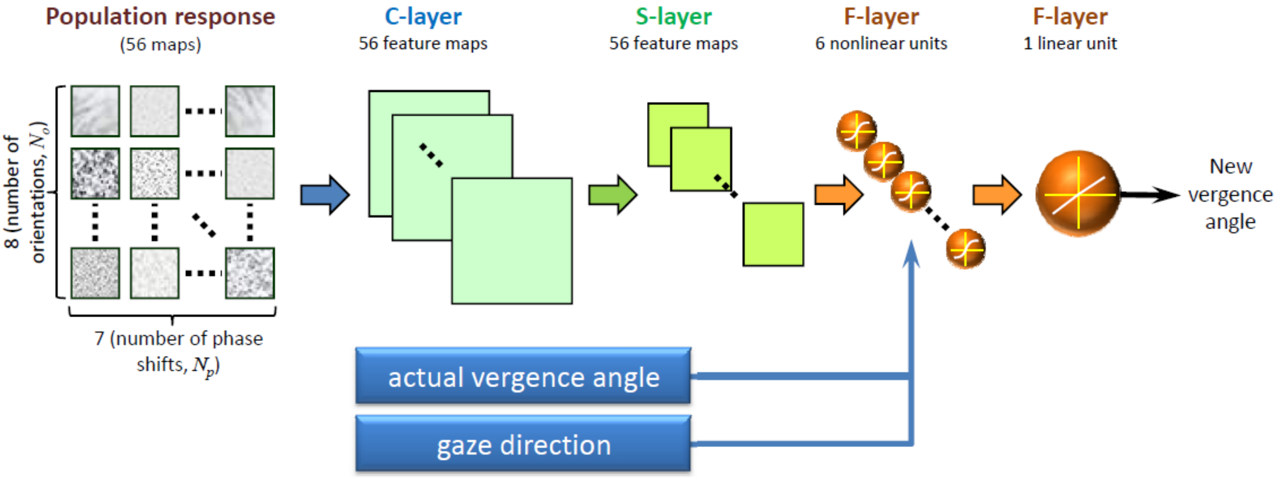

Figure 7: *Convolutional vergence control network and its inputs.*

Given the population responses, the proposed VC-net was trained to produce angular vergence control, which in turn is further executed by the oculomotor plant. Partner K.U.Leuven with UG has developed two types of VC-net paradigms: a linear- and a convolutional one (see Fig. 7). The former paradigm assumes that the gaze direction is orthogonal to its baseline and the object surface is a frontoparallel plane orthogonal to the gaze direction. The second paradigm goes beyond these assumptions, and operates reliably in the general case where all restrictions on the orientation of the gaze, as well as the stimulus position, type and orientation, are dropped (Chumerin et al., 2010).

The convolutional network has been further extended with the aim of including (i) kinematic (i.e., in terms of eye rotation velocity) vergence control, and (ii) attentional signals (based on object properties) that might guide intentional exploration of the selected object by performing version eye movements.

In the first case, a Robotic Head Model (RHM) was used (developed by partner UG) that takes as input rotational velocities for both eyes and provides the exact position and orientation of the both cameras (eyes) to the image renderer. The robotic head is composed of a bio-inspired ocular model and a pan-tilt platform commonly used in robot vision. The bioinspired oculomotor plant is composed of a Head block, an Eye block, and an Extra Ocular Muscles block. The pan-tilt system is composed of a Head block, a Pan-tilt block, and a joint velocities block. The Head block models the human head. It is assumed to be fixed with respect to the reference frame called world. The head is modeled like a rigid body regardless the mass, the dimensions and the inertia of the body. The Pan-tilt block models a pan-tilt system, which can be represented as a kinematic chain with two degrees of freedom. The oculomotor plant is composed of the head, the two eyeballs and the extra ocular muscles that drive each eye in a particular position (four recti muscles). The Joint velocities block computes the SVD (Singular Value Decomposition) of the pan-tilt kinematic chain.

In the second case, to achieve the version eye movements, Partners K.U.Leuven, WWU and UG have developed a Vergence-Version Control model with Attention effects (VVCA), see Fig. 8. The version is driven by an attention signal which is coming from the saliency map (frontal eyefield, FEF). Hereto an Object Recognition System (ORS) is used (provided by partner WWU). The main function of the ORS (in the context of VVCA) is to process stereo images produced by the renderer (or by the cameras) and compute the position of the object of interest (in the form of saliency maps), which then can be used for version (saccade) planning. The VVCA has some improvements with respect to the VC-net control paradigms. The VVCA provides kinematic eye movement control (i.e., control in terms of rotation velocities), which is common in robotics, it incorporates version control based on an object-related attention signal, and it is able to reproduce realistic eye movement trajectories.

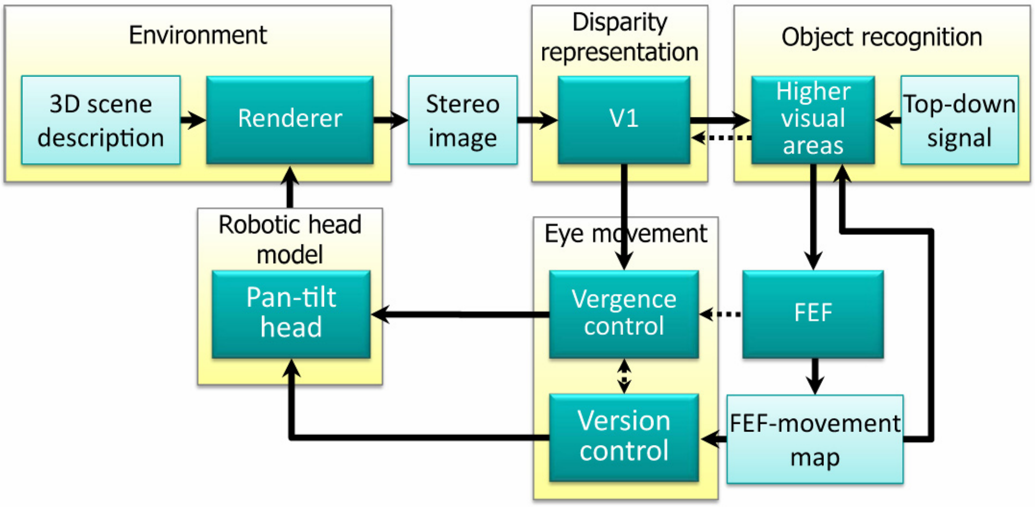

Figure 8: *The block-diagram of the proposed VVCA model. The stereo image generated by the simulator is processed by the disparity detector population, to produce the population response. Depending on which vergence control network is used, the population response is then directed to either the population response post-processing block, which is producing the post-processed population response (the linear VC-net case), or directly to the vergence control network module (the convolutional VC-net case). The (raw/post-processed) population response, together with the actual values of the gaze direction and the vergence angle, are fed into the vergence control network module, the main module of the model. The goal of the VC-net is to produce a new vergence angle, to get the fixation point onto the surface of the object of interest, without changing the gaze direction.*

The VVCA model is a scientific step forward in that it starts from a cortical-like population network of disparity detectors, and adopts a learning paradigm to achieve active vergence control from attention signals and performs interactive stereopsis. Technologically, it is a step forward in that it is also made compatible with pan-tilt camera systems, common in robot vision, taking into account their kinematics.

1.3.1.4 Attentive stereoscopic object recognition by learnable receptive fields

Visual perception is an active cognitive process. The visual cortex can focus processing resources on a certain object or visual fragment (feature-based attention; for its functional role see (Hamker, 2005b); for psychophysical correlates, see (Hamker, 2005a)) or on a certain location (spatial attention). While feature-based attention can involve simple features such as color or motion, we have expanded this concept to learned mid-level feature-detectors (Beuth et al., 2010). As a first demonstration the model has been integrated into a system of version/vergence control and tested in virtual reality. Second, a similar setup has been designed for real world object perception in cooperation with the partner UJI. In both systems, the responses from the stereoscopic edge detection model (V1 energy model, Chessa et al., 2009a, see also Sabatini et al., 2010) serve as input in a Higher Visual Area (HVA) representing views of objects. We used a weight sharing approach to analyze the whole visual scene in parallel, i.e. the detection of objects is independent of the location of the object in the visual scene on a 2D plane. A top-down "attention signal" can be used to guide perception to a target object by biasing the selection of a particular object (Fig. 9). An oculomotor loop via the frontal eye field (FEF) allows us to select the location of a particular object for a saccadic eye movement and in addition provides a spatially selective attention signal directed to the saccade target.

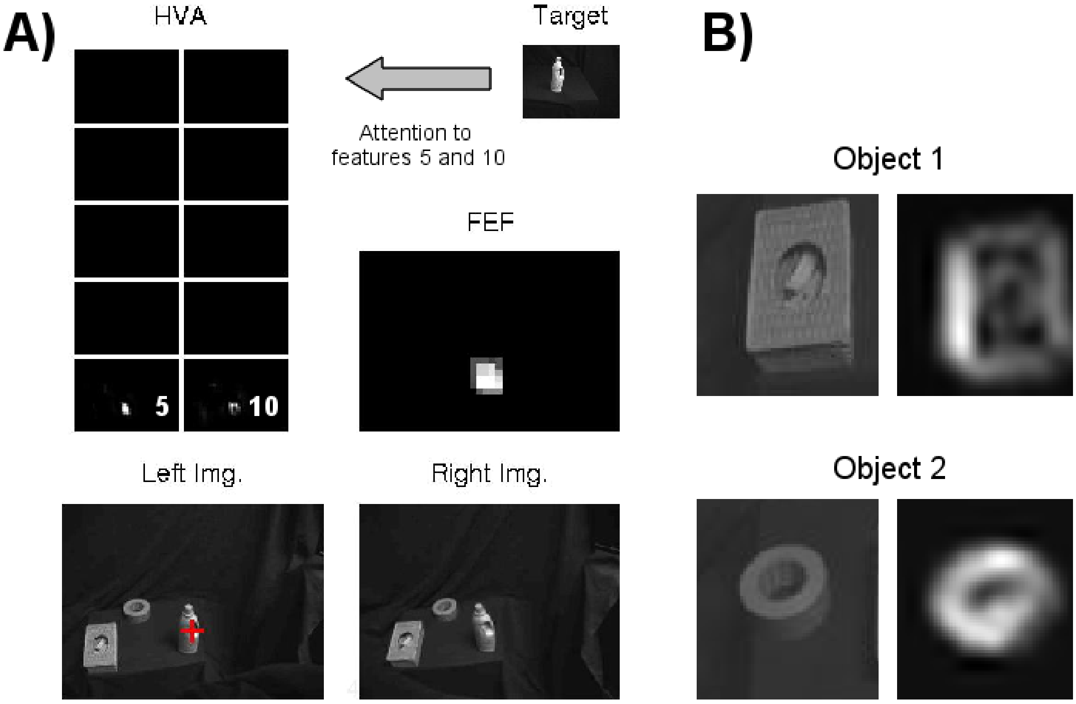

Figure 9: **A)** *Example of the successful localization of an object in a natural scene. Each box in HVA (higher visual area, like V4/IT) shows the activity of a single feature (visual fragment) in image coordinates. The x and y axis of all maps correspond to the spatial x and y axis of the images. The system has to search for the target object, what results in specific attention to the features 5 and 10. FEF (frontal eye field) encodes the spatial position of the object. A red cross (displayed only for left image) marked the position of the target object after successful recognition.* **B)** *Two examples of a visual fragment encoded by HVA cells which represents a single view of an object. More precisely, the figure shows the weights as a maximum over all features (phase, orientation, and disparity) in V1.*

<u>*Learning of receptive fields:*</u> Learning is a very crucial property for flexible and adaptive object recognition. We achieve the object selectivity by learning the feedforward weights (V1→HVA) with a biologically motivated unsupervised learning algorithm (Wiltschut and Hamker, 2009) and a trace rule using temporal continuity for the development of view-invariant representations of objects (like in Földiak, 1991; Rolls & Stringer, 2001; Wallis & Rolls 1997). The idea of trace learning uses the temporal correlations that are included in the visual input. Trace learning relies on the fact that on average, saccades in the vincinity of a target object are more likely than saccades to different objects. Thus, changes in objects are slower than changes in the exact retinal stimulation, so that the visual input is more likely to originate from different views of the same object, rather than from a different objects. To combine stimuli that are presented in succession to one another, activation of a pre-synaptic cell (V1) is combined with the post-synaptic activation (HVA) of the previous stimulus using the Hebbian learning principle. We simulated an appropriate temporal presentation protocol where the system mostly kept the fixation at one object and only rarely switched to another object.

<u>*Results:*</u> We benchmarked the object recognition system with ten objects in a simulated virtual reality (VR) (Beuth et al., 2010) and with three objects from the robot cameras. We obtained a recognition rate of 100% (VR) resp. of 98% (robot). Fig. 9a illustrates an example using the robot setup.
Among the recognition quality, we can inspect the nature of a visual fragment. A visual fragment can be defined as a neuron encoding statistically significant information, hence reacting to a specific stimulus at a certain location. Therefore, a fragment can be determined by the underlying learning algorithm of the receptive field. We have shown that a biological unsupervised learning algorithm creates early visual fragments (like edges) comparable to cells in V1 (Wiltschut et al., 2009) and that a similar algorithm develops view tuned cells in HVA (Beuth et al., 2010). Fig. 9b shows the learned target templates for two HVA cells using objects in real world scenes. This examination of a visual fragment has contributed to its definition which is necessary for recursive modulation of perception across visual fragments.

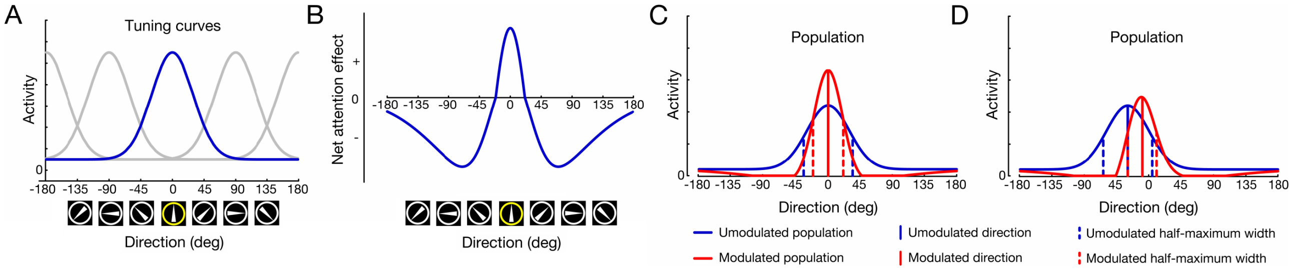

Figure 10: A) *In Zirnsak & Hamker (2010) we investigated the affect of feature based attention on the encoding of objects. Tuning curves are modeled by Gaussian functions, here for movement direction.* ***B)*** *The net effect of attention. Directions close to the attended direction are enhanced while directions farther away are suppressed.* ***C)****, Consistent with experimental data, the population response in the model is sharpened as indicated by the half maximum response when the stimulus direction is equal to the attended direction.* ***D)*** *For a stimulus with a direction shifted by -30° from the direction, the feature population response is distorted and the population vector of the neural activity is attracted towards the attended direction.*

1.3.2 Recursive modulation of perception across visual fragments

This contributed to the 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 peripersonal space.

1.3.2.1 Attentional-based selection of visual fragments and controlling of behaviour

Visual attention does not only guide visual perception it also can affect object representation. Our research has been focused on three questions: 1) we validated the developed mechanism of attention relation with biological data, 2) we investigated the selection of visual fragments by attention, and 3) we developed a model for the putative sources of the attention signal: the basal ganglia and prefrontal cortex which are the cognitive structures involved in action selection and working memory (WM).

<u>Relation to psychophysical data:</u> Our previous model (Hamker, 2007) predicts distortions through feature-based attention in a population of visual fragments. The general idea of this model is that the gain resulting from feature-based attention is not uniform across the whole population, but rather peaks around the attended feature, what in turn distorts the population response so that it is effectively shifted towards the attended feature. In Zirnsak & Hamker (2010), we were able to positively compare this effect with obtained psychophysical data. We observed that direction estimates of the static motion aftereffect (SMAE) drastically change when human observers attend to a stimulus whose motion direction differs from the one of the adaptor. We simulate this distortion effect in a feature-based attention model without learning, illustrated by Fig. 10. We expect that this observation can be generalized to other features or visual fragments.

<u>Concept of feature-based attention:</u> Feature-based attention is implemented by a feedback signal that biases specific visual fragments originating from the inferior temporal cortex. Attention operates as a multiplicative gain increase over the hierarchy of the visual stream (Fig 11, blue arrows). Thus, top-down attention can bias a particular object for selection and an oculomotor loop via the frontal eye field (FEF) can pick up this location for a saccadic eye movement. Moreover, the FEF provides a spatially selective attention signal just prior to the onset of a saccade. A subset of fragments is associated with each object and they must be bound together during recognition. The binding process operates continuously, but it can roughly be illustrated by two processes. One operates in parallel over all fragments and reinforces the ones that are relevant for the task at hand, independently of their location in the visual scene. The other is linked to action plans, here the eye movement, and binds together the fragments which are consistent with the action plan, typically by their location in the visual scene. The loop between V4/TEO and the frontal eye field (FEF) implements this idea (Fig 11, green arrows). Both processes use competition to decrease the activity of irrelevant features and locations in HVA.

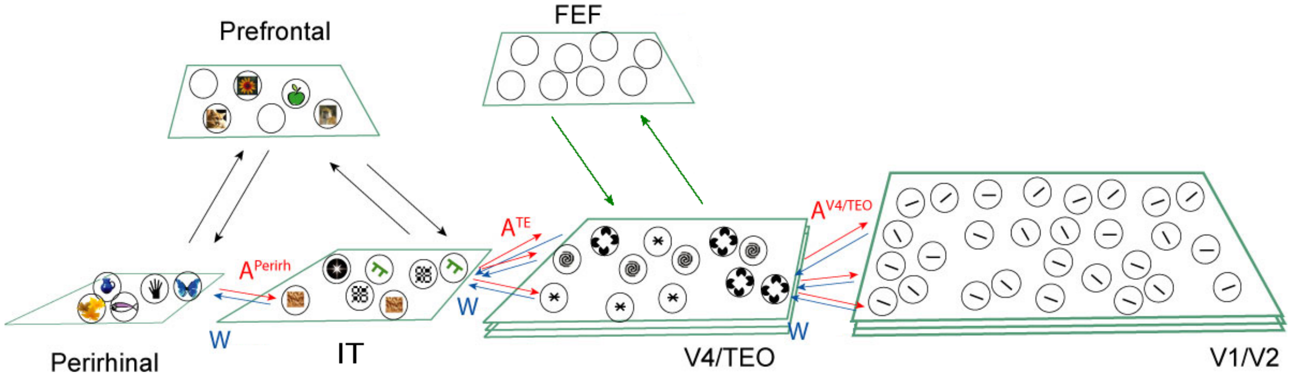

Figure 11: *A model of the ventral pathway and the frontal eye field (FEF) for attention and object recognition. Red arrows denoted feature-based attention, green ones spatial attention. IT: inferior temporal cortex.*

<u>Task based guidance of vision:</u> Vision is guided by visual attention. The appropriate control of attention for the current task takes place in higher level cortical areas and in the basal ganglia. These areas form thalamic-cortical-basal-ganglia loops (Alexander et al., 1986) which are essential for learning and maintaining working memory (WM) which includes selecting between behavioral alternatives. In Schroll et al., (submitted) we have shown a model based on this looped structure which is able to learn a) to store relevant stimuli in WM or to suppress them from it depending on the task requirements and b) to choose the correct behavior for the task (here a category). The learning in both cases is based on the received reward. The structures of the basal ganglia (BG) receive dopaminergic afferents from SNc (substantia nigra pars compacta) providing the BG with an error signal of reward prediction (Schultz et al., 1997; Hollerman & Schultz, 1998; Schultz, 2002): Dopamine bursts (above a tonic baseline level) occur from unexpected rewards while dopamine depletions (under this baseline) follow omissions of expected rewards. The working memory comprises loops via the prefrontal cortex and is learned based on this dopamine signal. The WM model learns to maximize the received reward for the whole task by estimating the expected reward for each symbol. If the model receives more reward than expected, the model reinforces specific dopamine-modulated connections in the loop, favorizing the memorization of the reward-predicting symbol in this loop. On the opposite, if the model receives less reward than expected, the model has remembered the wrong object, what

results in the depression of dopamine-modulated connections. Given the currently presented stimuli in inferior temporal cortex and the previously stored information in WM, the model learns to produce a category response. This category response was actually modeled in Schroll et al. (submitted) as a button press, but could also be a visual selection via a saccade (Vitay & Hamker, 2010). This decision process could be the source of an attention signal which in turn guides the vision. Generally spoken, the attention-based selection of visual fragments, the working memory and the stimulus-reward associations facilitates the awareness to the objects in the scene and provides information on how the system can interact with them.

1.3.2.2 Construction of peripersonal space across eye movements

Eye movements dissipate the stream of visual information received through the retina in different, seemingly unconnected fragments of visual information. Instead of degrading the power of the retinal input, the saccadic system is taking advantage of this discretization. The saccadic movements allow a direction of the high resolution fovea on fragments of interest, which greatly enhances visual acuity. To use this refinement in a stable and constructive way, the execution of the saccadic system is carefully controlled including all applicable inputs as well as a great specification of the saccadic space. Inputs to the saccadic system are all sources of information used to calculate motor modifications. The saccadic space describes the output space, in which the saccade is executed. In a generalized view, also the construction of the paradigm inducing saccades as well as its corrections can be included in this framework.

Our first and main objective was to investigate and describe the role of oculomotor space in visual space perception. To this end we studied perceptual localization changes induced by saccadic adaptation. The first localization study considered localization errors after modification of the saccadic amplitude while a saccade is executed (Zimmermann & Lappe, 2009). Scanning and reactive saccades were adapted. They differ in their transfer behavior, which means a modification of reactive saccades does not transfer as strong to scanning saccades as the transfer of scanning adaptation to reactive saccades. An analogy to this differences in reactive and scanning saccades was created regarding localization by the use of flashed and stationary stimuli. This proved the close interconnection of visual perception and the visuomotor system. But the power of the motor modifications induced by saccadic adaptation is demonstrated by the existence of localization errors even in the absence of saccades (Zimmermann & Lappe, 2010). With this publication the relevance of the sensitivity and specificity of the saccadic system for visual perception is demonstrated. Therefore, the presented work has shed light into the construction of peripersonal space via the recursive modulation of visual fragments. Saccadic eye movement bind visual fragments via the execution of saccades, which are modified in a spatially specialized way depending on a complex net of input information. The modifications are executed specific to saccadic circumstances. They modify visual perception in an active and recursive way.

We then considered specifics of the mechanisms of saccade and how these specifics relate to spatial representations. Until recently it was believed, that only the retinal input is considered for the control of the saccadic system. A similar view existed concerning the consideration of the output space of saccadic control. Only the relative target position was believed to play a role. A third aspect, which empowers saccadic control to act goal directed is a saccade type dependent control, which was ignored until recently as relevant factor for saccadic control. In a set of nine studies we showed, that non-consideration of these three factors, namely input parameters, the consideration of output space as well as the saccade type, is a severe simplification of saccadic control, which disregards the power and complexity of the saccadic control system with respect to the binding of fragments. Furthermore, two studies were conducted to demonstrate the impact of saccadic motor control on visual perception. The influence of motor modifications on visual localization closely associates perception and eye movement control.

We conducted two studies concerning the input parameters considered in saccadic motor control. The first study introduced the parameter of consistency into the learning procedure (Havermann & Lappe, 2010). Whereas the averaged retinal error was left equal, the consistency was varied in three degrees of consistency. It was shown that the consistency of a saccadic learning signal has a great effect on the learning behavior. Consistent errors are considered more important and thus more powerful by saccadic control. From this, we can conclude an inter-trial memory of saccadic errors in the manner of a bayesian estimator of the necessity of a motor correction.

In a second study, the retinal position error was generalized to a moving stimulus (Havermann et al., 2011a). In this study an adaptation to moving stimuli was shown. A trade off between pursuit and saccadic behavior led to long-term modifications of the saccadic amplitude. Therefore, motion can be considered as an additional input to the saccadic control system. In particular, in this study the legwork of another motor system could be shown to influence saccadic motor control, which redefines the implication of “input” to the saccadic system. Not only external signals have to be considered, also signals from other brain areas have an effect on saccadic motor control.

Regarding the specificity of the output parameters of the saccade a set of three studies showed a specificity of saccadic adaptation with respect to the initial eye position during adaptation (Havermann et al, 2011b; Wulff et al, 2011; Zimmermann & Lappe, 2011). This constitutes a drastic change in the view on the generalization behavior of motor modifications in saccadic space, which was believed to be complete. The eye position specificity was not only shown in human subjects, but also in monkeys. These results demonstrate the fact, that saccadic motor control can adjust eye movements specific to their execution in space. This greatly improves the sensitivity of motor control. Moreover, the eye position dependence of adaptation also extended to visual localization (Zimmermann & Lappe, 2011), thereby corroborating our view that motor contingencies contribute to spatial awareness.

A third general aspect of an eye movement is the modus, in which it is executed. Depending on the timing of target presentation, saccades are believed to be executed by partially different mechanisms. The saccade types distinguished are reactive, scanning saccades, memory guided as well as gap saccades are considered. Reactive saccades are executed to a suddenly appearing target, whereas scanning saccades are executed within a stable pattern of stimuli. In gap saccades the time of the presentation of the visual feedback after a saccade is delayed, whereas in memory guided saccades even the evoking target is extinguished before the execution of the saccades. The mechanistic differences of saccades answering the different stimuli described origin from the different needs of abilities for the execution. The execution of scanning saccades adds a spatial component concerning the relative position of the targets to the needs for reactive saccades; gap saccades and memory saccades need signal integration and target memory. Behaviorally the differences were shown by a varying amount of transfer of motor modifications.

A similar degree of mechanistic complexity is given regarding the direction of motor control, which can be shortening or lengthening. Although the existence of different mechanisms for the two directions as well as saccade types is well known, many characteristics of the saccadic system were considered to generalize over saccade type and adaptation direction. Two detailed studies were conducted, which clarified important aspects of these mechanistic differences. The first study considered the output parameter of saccadic amplitude and saccade direction with respect to its generalization concerning the direction of motor control. Collins et al., 2007 found an adaptation field for the modification of amplitude and direction for inward adaptation, which showed a sensitivity of motor control with respect to saccadic amplitude and direction. In an adaptation experiment only saccades of the same amplitude and direction were adapted to the same amount. Any deviation in direction or amplitude lowered the effect of amplitude modifications. A similar behavior is observed for outward adaptation (Schnier & Lappe, 2010). As for inward adaptation the specificity extends to spatial localization.

We also studied the transfer behavior between different types of adaptation and found a modification of the saccade metric especially in inward adaptation (Schnier & Lappe, 2011). The saccade shortening is achieved via a decrease in peak velocity and saccade duration, but only the decrease in peak velocity transfers to other saccade types. This points out the complexity of motor modifications and their specificity and generalization regarding saccade types.

1.3.3 Visuospatial awareness and planning behavior

The problem of constructing an action-minded representation of the 3D space has been addressed by developing a multisensory description of 3D space obtained through active ocular and arm movements.

1.3.3.1 A joint, implicit 3D representation of the peripersonal space built and accessed by eyes and hand movement signals

The first step performed in order to achieve the above goals was to define an integrated visuomotor representation of the peripersonal space, which includes both action-related and perception-related aspects.

This constitutes the basic concepts on which to develop a robotic system highly-skilled in its capacity of exploring the nearby space. The knowledge on the neuroscience of vision-based reaching in humans and other primates was described and analyzed, in order to establishing the computational bases for a robotic system able to achieve advanced skills in the interaction with close objects. A more detailed description of the concepts directly useful for the generation of the integrated representation was elaborated, starting from a real situation of an agent facing an environment within which it is expected to interact. The outcome was a model framework for representing the peripersonal environment that includes on-line, action-oriented visual information with perceptual knowledge about objects and memories of previous interactions.
The inspiration of such description comes from the distinction between sensorimotor and perceptual visual processing as performed by the two visual pathway of the primate cortex. The obtained description includes on-line, action-oriented visual information (dorsal stream) with knowledge about nearby object and memories of previous interaction experiences (ventral stream). Particular importance has been given to the dorso-medial stream, principally involved in proximal joint control and nearby space actions. We gave great importance to the use of binocular data and proprioceptive information regarding eye position, critical in the transformation of sensory data into appropriate motor signals. The next step toward the final goal was to devise a way to achieve, in a dynamic, active way, the sensorimotor knowledge of the environment postulated by the model.

To achieve the dynamic, integrated sensorimotor representation described above, special focus has been put on the integration between stereoscopic retinal data with somatosensory information about object and arm state, in order to estimate object position and devise gazing and reaching action plans as performed with the aid of area V6A of the dorso-medialal stream. Partner UNIBO provided single-cell data regarding experiments on monkeys reaching and gazing at different visual targets. These data have being analyzed and modelled by UJI in collaboration with UNIBO, with the goal of reproducing and assess neuroscience theories on the robotic setup. Interesting insights regarding the kind of representations maintained by V6A neurons and employed to perform the transformations required to achieve a mutual modulation between sensory stimuli to motor commands have been produced. The different properties captured in such analysis have been modelled in a Radial Basis Function (RBF) network framework. With our model we verified what computational advantages could be given by a responsiveness pattern such as that of V6A, including neurons having only visual response, neurons apparently involved mainly in motor actions and mixed neurons, activated in all phases of sensorimotor processes. The model itself constitutes the basic computational layer on which simulated and real robot behaviours are implemented.

The basic framework of the model is made of a central body-centered representation that is accessed and updated both by limb sensorimotor signals on the one hand, and by visual and oculomotor signals on the other hand. Indeed, this seems to be how the problem is solved in the brain, where different areas, or populations of neurons in the same areas are dedicated to different transformations. Most importantly, this approach is consistent with the findings related to area V6A. The global structure of the model follows these principles, and is thus modular, separated in the visual and oculomotor component and the arm sensorimotor component. We implemented three different neural networks. The first one transforms the visual input in an oculomotor reference frame. The second and third networks, whose structures and parameters are tightly related to each other for improving biological plausibility, perform the transformations from oculomotor to arm-motor reference frames and vice-versa. Learning the transformation from binocular visual data to eye position consists in identifying visual targets and foveating them with both eyes, in order to associate appropriate version and vergence movements to retinal locations. The second learning phase, in which arm movements are introduced, is further subdivided in two stages, respectively *free* and *goal-based*. The free exploration consists of random arm movements and subsequent saccades toward the final hand position, which allows the agent to learn the transformation from joint space to oculomotor space. In the goal-oriented exploration a target object in space has to be foveated and reached. During this process, the inverse transformation having arm joints in output is learnt.

As a complementary goal, we were also able to apply our model to emulate the saccadic adaptation experiments according to the guidelines of partner WWU, verifying that our model does exhibit saccadic adaptation, altering its ability to perform correct saccades according to the deceptive feedback. The analysis of error distributions around the target point and of error vectors has also provided interesting information that have been eventually compared with human data and with the results of matching experiments performed with the real robot.

Our simulation supports the hypothesis that a mixed population of neurons such as that observed in V6A is especially suitable to a cortical area that contextually codes for different reference frames. From a robotic point of view, through the use of basis function neurons whose configuration was set according to what suggested by neuroscience data, we were able to learn very accurately the transformations between oculomotor and joint spaces, in a way suitable to their application in the robotic setup.

1.3.3.2 Cortical visuomotor representation of the 3D reachable space

The medial parieto-occipital cortex of the primate brain is a region where visual information and motor signals are integrated, to provide the dorsal visual stream with information suitable to perform the control of ocular movements and arm movements directed to targets of interest. In EYESHOTS, we provided results that, altogether, indicate that within the medial parieto-occipital area V6A of the macaque, there is a representation of space that can provide a map useful to address eye movements and arm movements in the 3D space.

Interacting in the peripersonal space requires coordinated arm and eye movements to visual targets in depth. Area V6A is a key region engaged in the process of visual-to-motor signal transformations, because it jointly processes visual information, eye position and arm movement related signals (Galletti et al., 2003; Fattori et al., 2010). However, before EYESHOTS, there was no evidence in the medial parieto-occipital cortex of spatial encoding in three dimensions. Here, using single neuron recordings in behaving macaques, we studied the neural signals related to binocular eye position in a task that required the monkeys to perform saccades and fixate targets located near the body up to positions located far away from the body, well beyond the reachable space. A significant proportion of neurons were modulated by both gaze direction and depth, i.e., by the location of the foveated target in 3D space (Hadjidimitrakis et al., 2010). Interestingly, while the population activity of V6A neurons showed no preference for any gaze direction, the same population displayed a strong bias for near, peripersonal space. This preference for targets within reaching distance, evident during both target capturing and fixation, suggests that binocular eye position signals can be exploited to form a representation of the peripersonal/reachable space. This representation can be used to perform the sensori-to-motor transformations needed to perform successful reaching movements in depth (Hadjidimitrakis et al., 2011).

Using a set up where targets were located in depth along iso-vergent and iso-version curves (see Fig. 12A), we studied the neural signals related to depth encoding in a task that required the monkeys to perform eye movements and arm movements. The majority of neurons were modulated by both version and vergence angle, i.e., by the location of the foveated target in depth. The reach-related discharges were modulated by depth both during the execution of reaching movements towards the targets and during the return movement back to the body, and also by static holding of hand on different positions in depth. There was not a bias for a particular sector of the peripersonal space: the spatial fragments are sampled in the same way, with the same definition, across the entire reachable space. These data (see Fig. 12) suggest that depth signals are implemented functionally in V6A to direct eye movements and arm movements in the 3D space (Breveglieri et al., 2011). This modulation, in an area primarily involved in visuo-motor transformation for reaching, together with attentional-related signals already shown to be present in V6A (see following paragraph), may form a neural basis for linking eye movements and arm movements across fragments.

Link across single visual fragments can be obtained in many physiological situations. Commonly, in natural conditions, when we catch with vision a target of a potential reaching action, we move the eyes toward it and then the hand, or we can “capture” our target with the focus of attention. Attention is used to enhance neural processing of selected parts of a visual scene. It increases neural responses to stimuli near target locations and is usually coupled to eye movements. Covert attention shifts, however, decouple the attentional focus from gaze, allowing to direct the attention to a peripheral location without moving the eyes. We found that covert attention shifts modulate ongoing neuronal activity in the medial parieto-occipital area V6A. We performed single cell recordings from area V6A in monkeys trained to fixate straight-ahead, while shifting attention outward to a peripheral cue and inward again to the fixation point. We found that neurons in V6A are influenced by spatial attention. The attentional modulation occurs without gaze shifts and cannot be

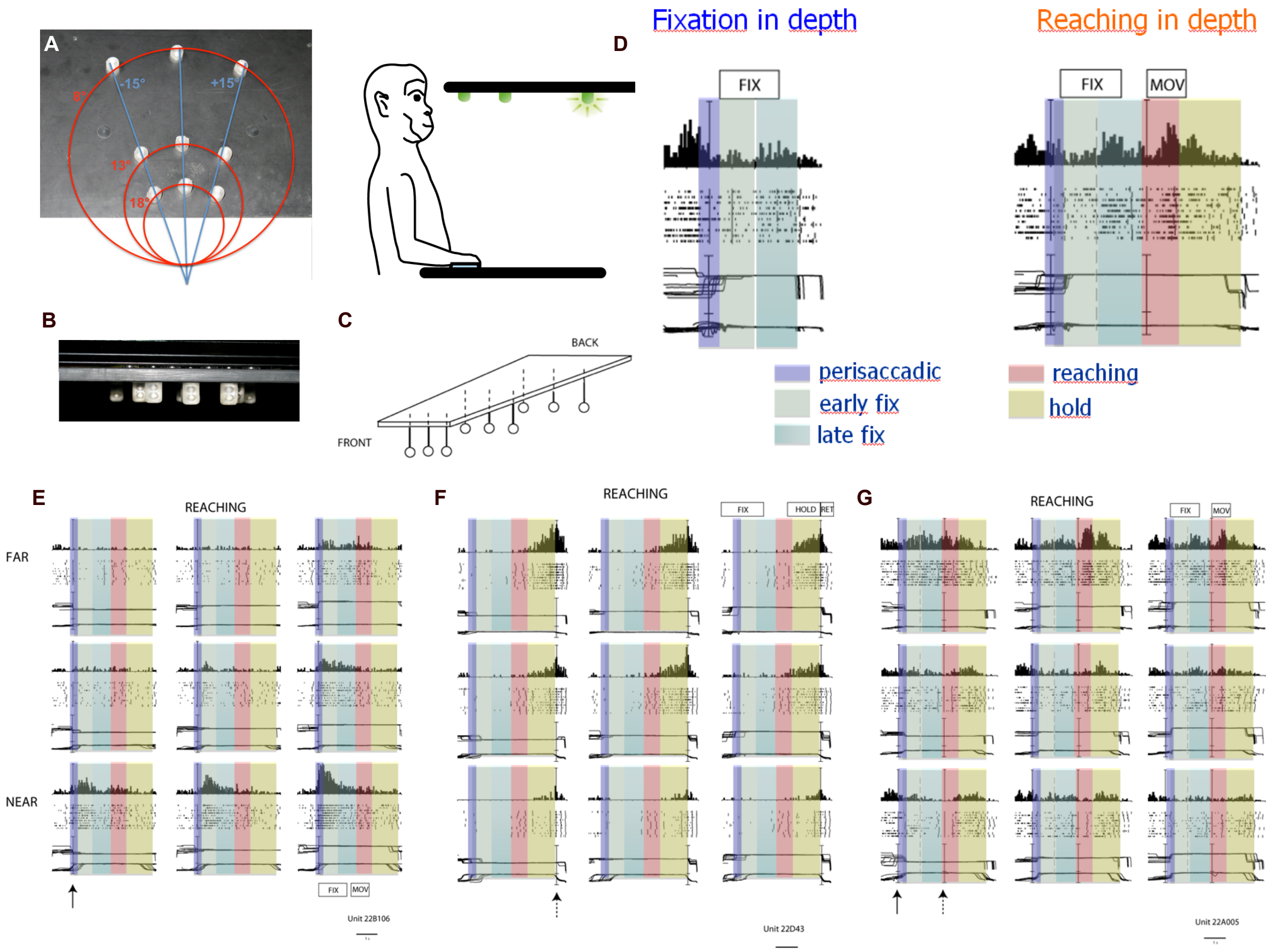

Figure 12: *Neurophysiological evidences of fixating and reaching in depth modulation in the posterior parietal cortex, area V6A.* ***(A)*** *The fixation-in-depth/reach-in-depth device. Each of the 9 targets contains a green/red bicolor LED used to instruct the monkey, and an embedded microswitch used to check target pressing. Bottom view of the device superimposed with isoversion (blue) and isovergence (red) lines.* ***(B)*** *Frontal view of the device (as seen from the monkey side).* ***(C)*** *Schematic view of the device.* ***(D)*** *The time epochs of the experiments for fixation in depth and reaching in depth indicated by different color shadings.* ***(E)*** *Example of a cell modulated by eye movements only.* ***(F)*** *Example of a cell modulated by arm movements only.* ***(G)*** *Example of a cell modulated by both eye- and arm-movements.*

explained by visual stimulations. Visual, motor, and attentional responses can occur in combination in single neurons. This neural behaviour can form a neural basis for coupling attention to the preparation of reaching movements (Galletti et al., 2010). Our results show that cortical processes of the generation of attention are related not only to eye-movement control, as many studies have shown, but also to the control of arm movements, a finding suggested by some previous behavioral findings. Therefore, the widely-held view that spatial attention is tightly intertwined with - and perhaps directly derived from - motor preparatory processes should be extended to a broader spectrum of motor processes than just eye movements.

The visuomotor representation of rechable space relies on different sensory inputs and on copy of motor commands of the ongoing movement. Previous work had shown that area V6A is rich in reach-related cells (Fattori et al., 2001) and that reaching activity starts before the earliest electromyographic activity (Galletti et al., 1997), such that it cannot be entirely explained by afferent somatic signals. Reaching activity is instead compatible with corollary discharge of motor commands that could reach V6A from the premotor cortex (Fattori et al., 2005; Gamberini et al., 2009). Recent work from EYESHOTS project has shown that V6A could participate in planning and executing reaching movements both when visual feedback of the moving arm is available, as well as when visual input is not available (Bosco et al., 2010). Monkeys were trained to execute reaching movements in two conditions: in darkness, where only the reaching target was visible, and in full light, where the monkey also saw its own moving arm and the environment. The vast majority of V6A neurons were significantly related to the task in at least one of the two conditions. The majority of task-related cells showed reach-related activity in both visual conditions, some were modulated only in light, while others only in dark. The sight of the moving arm often changed dramatically the cell's response to arm movements. In some cases the reaching activity was enhanced and in others was reduced, sometimes so strongly reduced that the cell response completely disappeared. These neuronal behaviors may represent differences in the degree to which cells are influenced by feedback control versus feedforward movement planning. Some recent models highlight the importance of a "state estimator" as a central controller of the motor act, that is a comparator of the expected state of the movement with the actual visual/somatosensory feedback evoked by the moving arm itself for the purpose of controlling arm movements (Grafton, 2010; Shadmehr et al., 2010). The presence in V6A of different types of neurons relying differently on the different visual/somato-sensory inputs available, as well as motor-like cells discharging during reaching movements, can suggest that V6A is able to encode reaching by flexibly weighting the motor goals according to the available sensory inputs, as modeled by Sabes and coworkers (Sober & Sabes, 2005; McGuire & Sabes, 2009) and as also suggested by very recent brain imaging studies for the human homolog of V6A (Bernier & Grafton, 2010). This possibility is in agreement with the current Bayesian models of flexible sensory integration for movement guidance (van Beers et al., 2002; Knill & Kersten, 2004), suggesting that this could be a mechanism serving to prioritize the most reliable sensory cues in a given situation. On these grounds, partner UNIBO suggested that area V6A might be able to compare anticipated and actual sensory feedback evoked by the moving arm (Bosco et al., 2010) and proposed that area V6A can act as a "state estimator" in the neural circuits controlling arm actions. Some V6A cells could rely on mechanisms switching off somatosensory activity when acting in the light, maybe because these cells rely more on vision than on proprioception. Conversely, other V6A neurons might rely more on proprioception rather than vision when arm actions are organized, and this could be the reason why these neurons show enhanced reaching activity in the dark when compared to the light. Other neurons in V6A can rely merely on corollary discharge signals, that is motor-like signals arriving to V6A mainly through direct connections with the premotor cortex (Gamberini et al., 2009) that could allow to monitor and online control the correct execution of the intended action (Gamberini et al., 2011).

1.3.3.3 Humanoid behavior replicas by integration/interactive paradigms

One of the principal goals of EYESHOTS was to provide the robot with advanced skills in its interaction with the environment, namely in the purposeful exploration of the peripersonal space and the contextual coding and control of eye and arm movements. The implementation on an actual, physical sensorimotor setup is a potential source of additional insights for the computational model, hardly achievable with simulated data.

During year two of the project, we acquired a pan-tilt-vergence stereo head, configured it and set it up on a new torso body, together with two multi-joint arms. Great efforts have been put by the Intelligence Robotic Lab in order to integrating it and interfacing it, both in hardware and software, to the previously available equipment, to the new computational modules developed according to the computational framework described above, and to advanced, bio-inspired visual modules provided by partners UG, K.U.Leuven and WWU.

We fully implemented in 3D the visual/oculomotor and oculomotor/arm-motor basis function networks which allow bidirectional transformations between retinotopic, head-centered and arm-centered reference frames. We ported the networks on the robot platform, enabling the humanoid robot *Tombatossals* to accurately learn the transformations between visual, oculomotor and arm joint spaces by actively interacting with its surrounding environment.

The use of robot hardware constitutes a possible complication in the realization of a model of cortical mechanisms, and some issues that would easily be solved in simulated environments have to be dealt with more accurately, considering the real world implementation. The robot is initially endowed with an innate knowledge of how to move in its environment, which is later developed and customized through exploration and interaction with visual and tactile stimuli. Following this idea, the network weights learned by the computational model are used as a bootstrap condition for the actual experimental learning process by the robot.

The building blocks of the final robotic system (see Fig.13) are thus the RBF transformation modules, visual processing, head and arm actuators, working memory, visuomotor memory and the three computational modules V1, VC-nets and ORS, deriving from the Vergence Version Control model with Attention effects (VVCA).

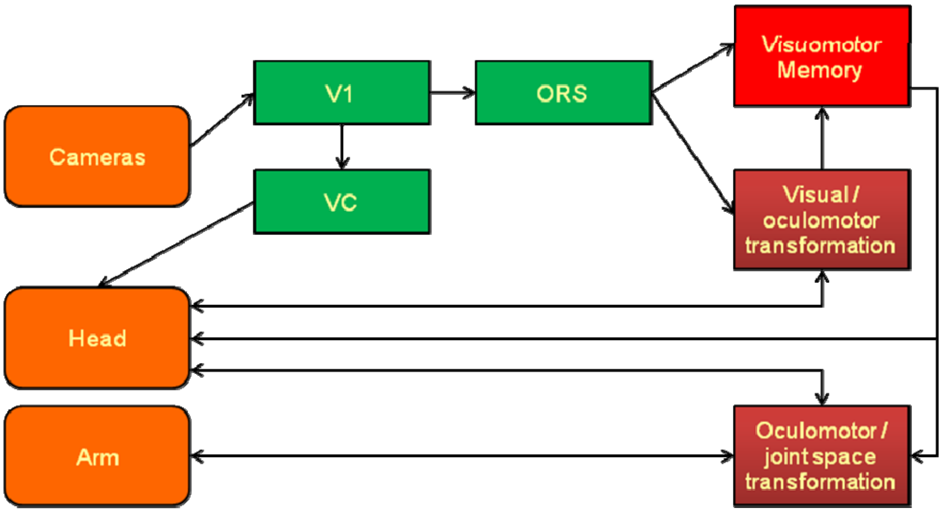

Figure 13: *Building blocks of the final robotic system.*

Basic skills such as concurrent or decoupled gazing and reaching movements toward visual stimuli are available to the robot. As planned, the robot is able to learn a visuomotor representation of nearby objects, and shows its capabilities by performing oculomotor actions toward visual targets placed in its peripersonal space, or toward the location where its hand lies. Moreover, it is also able to perform arm reaching movements to visible objects, either with or without gazing at them. This ability is achieved through a purposeful exploration of the environment, which allows the robot to build a visuomotor memory of surrounding objects.

A working-desk setup, with simple objects, and full 3D movements for both eyes and arm, allows the robot to show its fundamental visuomotor skills in its interaction with visible and reachable objects (see Fig. 14). In any case, to be able to perform a complete visual exploration of the environment, the robot has to pass first through a learning stage designed to develop its own sensorimotor coordination. This is done by training the visual-oculomotor and the oculomotor-arm motor neural networks. Basically, such a training is performed through gazing and reaching movements of three fundamental types: gazing toward an external stimulus, gazing toward the hand, concurrent movement of hand and eyes toward a same location. After the robot has gained a reasonably good precision in its visuomotor coordination, the appearance of external visual targets allow it to execute a more purposeful exploration of its environment.

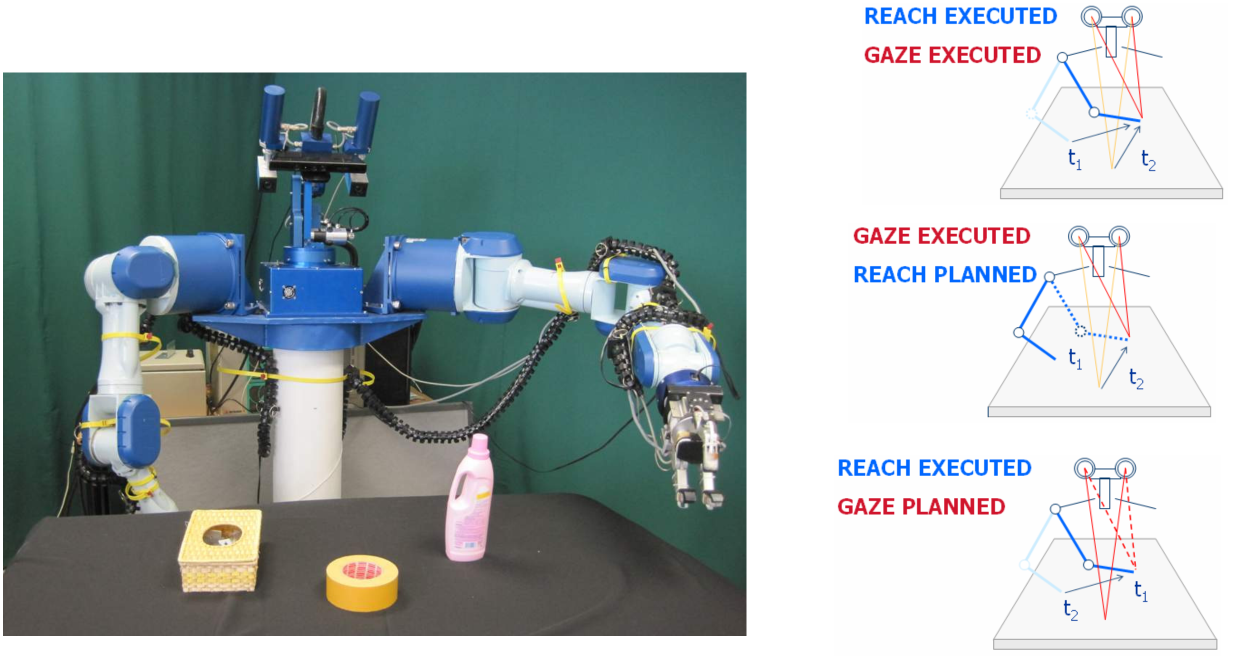

Figure 14: ***(Left)*** *The multi-object experimental setup used with the UJI humanoid robot* Tombatossals. ***(Right)*** *Concurrent gaze and reaching actions used to test the bilateral interaction between binocular oculomotor coordination and reaching tasks: saccades and reaching actions alternate sequentially both in the execution or planning phases. The robot employs the egocentric representation of peripersonal space it has gained, to interact with surrounding objects, recognize them and perform custom visuomotor and arm-motor actions, such as: foveate on the hand; reach the gazing point; show memory; foveate on a given object (either inside or outside the field of view); reach a given object (either foveated or not); execute a sequence of saccades by employing either covert or overt attention..*

In fact, in addition to building the sensorimotor skills necessary to gaze and reach at potential visual and motor targets in the nearby space, the construction of an integrated knowledge of the environment requires also the identification of actual objects or targets, and the use of a memory of previously observed/reached objects. In this associative learning schema the robot is required to use the learned visuomotor skills in order to estimate the eye motor movements necessary to fixate visible targets. Such potential oculomotor parameters are employed to build and maintain a visuomotor memory of the objects present in the robot peripersonal space.
After the learning process has reached a reliability threshold, the system can be required to fixate and/or reach a target given its identity, using the visuomotor associations it has previously learned. This can be done even on objects placed out of the field of view, if their location has been previously observed. As a final result, the robot achieves very good open-loop gazing and reaching skills toward new and memorized objects.

A second experimental setup is aimed at showing in more detail the properties of the radial basis function model applied to the real robot. This scenario sees the robot interacting with a computer screen similarly to what is done in typical psychophysics experiments. More exactly, we employed this setup to emulate the saccadic adaptation paradigms used by partner WWU. The comparison between data obtained with human subjects, with the computational simulation and with the robot has provided insights on theoretical aspects related to visuomotor cognitive aspects and contextually allowed us to validate our proposal.

1.3.3.4 Definition of shared attention behaviour in common workspaces

This task focusesed on the understanding of the sequence of allocation of attention, direction of gaze, and movement of the arm of a human cooperative partner. This research should allow the anticipation of particular actions based on the partner's behaviour.

We investigated the interplay between attention, direction of gaze, and movement of the arm of a human cooperative partner in two studies with single subjects and in one study in which two subjects interacted directly. *Study 1* investigated if gaze direction changes can be used to predict forthcoming pointing movements of another person seen in a controlled situation in a video. The experiment required the participant to watch a series of movies in which an actor performed gaze movements and reaching arm movements towards a set of targets. The participant had to predict the to-be-pointed-at target by looking at it. The movies could differ in two aspects. The gazing behaviour of the actor could be visible or occluded and the targets could be as well visible or occluded. Participants' responses were mainly influenced by the availability of the gazing behaviour of the actor. The visibility of the targets had an effect on the spatial accuracy only. Actor's gaze triggered rapid and accurate responses towards the target objects, which were accurately identified when the actor's arm was still at the beginning of its trajectory toward the target object. When the actor's gaze information was not available, the participants could still predict which target object was relevant in a particular trial by relying on the actor's hand movement only. In this case, participants' gaze still leaded the hand movements of the actor, but was comparatively slower and less accurate. The results provide evidence that other's gaze direction can be used advantageously as a predictive cue about the final location of a pointing movement and can be complemented by the kinematic cues provided by the hand movement.

The second single actor experiment (*Study 2*) was designed to further explore the relation between gaze behaviour and arm movements and its influence on the allocation of attention. Ample evidence supports the idea that social signals, such as eye gaze, influence our voluntary eye movements. We used a gaze cueing paradigm to study the overt orienting (involving eye movements) of attention. Participants were asked to perform a speeded saccade to left or right of a centrally presented instruction cue. As distracter stimuli we used a gaze direction change and a small hand movement. Gaze and hand cueing were effective at triggering saccades in the opposite to the intended direction (directional errors). A stronger gaze cueing effect was, however, observed when the gaze and hand cue were presented in conjunction and the proportion of saccades following the gaze cue increased irrespective of the small hand movement direction. These findings suggest a process that prioritises potentially relevant actions (eye movements performed in conjunction with a hand movement) to which the visual system automatically responds. In addition, this finding can be implemented as one of the heuristics for an optimal human-robot interaction. This study is under review at Attention, Perception & Psychophysics.

Study 3 focused on human-human interaction in a cooperative task in shared workspace. Being able to predict outcomes of other's movements is crucial for action coordination (see Fig. 15). We addressed this issue by exploring the gaze behaviour of pairs of participants involved in a simple cooperative task. Each of them had to move an object in the vertical plane around an obstacle and make contact with the object of the other participant. A stereotypical gaze behaviour was observed: (1) at the start of each trial a fixation was directed towards the own object; (2) fixation was kept on a central location of the setup; (3) saccades were then regularly directed towards the partner's object in the terminal phase of the movement prior to the contact between objects. Our results thus suggest that the stereotypical gaze behaviour is necessary to establish a closed loop between the two participants that allows a coordinated fine-tuning of the joint interaction. The expectations that a human actor has about the cooperation partner influence the deployment of attentional resources. In a second condition one of the two participants had the freedom to determine the contact location between objects and the other participant had to comply with this behaviour in normal predictability and low predictability conditions. The timing of object-oriented saccades was modulated by the predictability of the contact location. The object directed saccades started significantly earlier in the low predictability condition than in the normal predictability condition. The necessity of gazing on the partner's object after this very short time lapse prevented the execution of any other gazing behaviour toward other objects in the environment. In sum, the first three experiments suggest, that gaze tracking can be used to predict allocation of attention and behaviour in shared workspace.

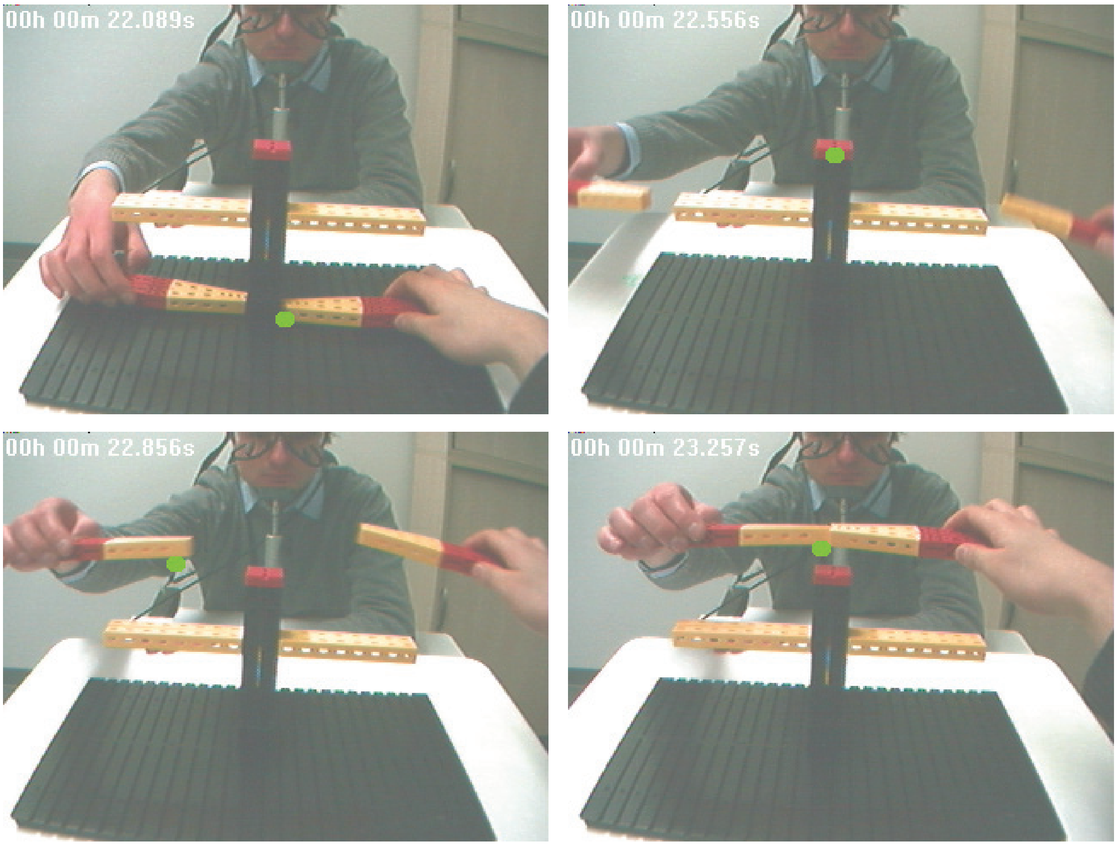

Figure 15: *Four video-frames showing the execution of the requested movement: the two participants had to move an object in the vertical plane around an obstacle and make contact with the object of the other participant. This video-frames were extracted from the recording of one of the participants. The green dot in each video-frame indicates the gazed location.*

After these experiments were finished we decided to explore possibilities to determine the degree of acceptance of a robot as an interaction partner under conditions, in which humans represent the actions of robots in a similar way to the actions of other conspecifics. In this study (*Study 4*) we tested if the *Social Simon Effect* (SSE) can be used as a benchmarking tool for the perceived humanness of a robot under conditions of shared workspace. The SSE is a psychological effect that provides a good index for action co-representation (e.g. Sebanz et al., 2003). It is also interesting because some studies seem to indicate that action co-representation is specific to biological agents, hence facilitating human social interactions only with conspecifics (e.g. Tsai & Brass, 2007; Tsai et al., 2008). However, other studies suggest that action co-representation may be explained by top-down attribution processes about the perceived animacy of the observed agent (Liepelt & Brass, 2010; Liepelt et al., 2010). We first investigated whether action co-representation, as measured by the SSE, can be found when interacting with a real robot that is perceived as human-like. Therefore we developed an experimental setup in which a Social Simon Task was shared between a human and the UJI humanoid robot *Tombatossals* (see Fig. 16). Prior to the experiment participants were told that the robot functions in a human-like manner, e.g. being able to actively decide when to respond on the basis of a neural network. In this experiment we found a significant SSE. This suggests that co-representation of non-biological agents can occur if an agent is perceived as human-like. In order to test if it is actually the perceived humanness of the robotic system that induced the SSE in the first experiment, we conducted the same Social Simon Task as in the first experiment, but now we introduced the robot to the participants as functioning in a non-human, and purely deterministic manner, e.g. being completely controlled by a computer program. Under these conditions we did not observe a SSE. This suggests that action co-representation as measured by the SSE does not seem to occur when interacting with a purely deterministic and non-human-like robot. Taken together, our results indicate that action co-representation is not exclusively tuned to biological agents. Instead, higher order cognitive processes, like the perceived humanness of a robot, seem to influence the amount of action co-representation in a top-down manner. Co-representation of robotic (non-biological) actions can occur if a robot is perceived as functioning in a human-like way. Further, the findings indicate that the SSE may be used as a benchmark-tool for the perceived humanness of a robotic system.

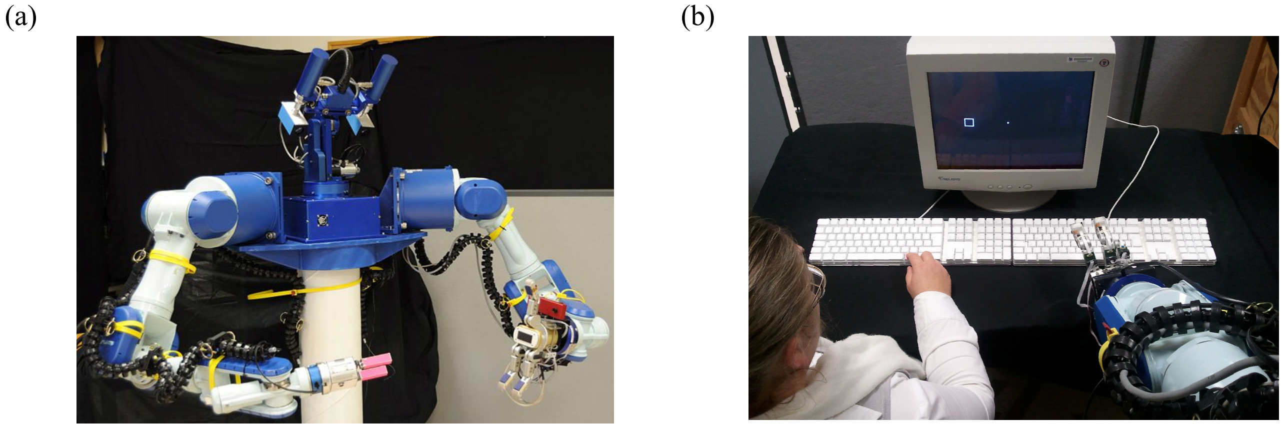

Figure 16: ***(a)*** *The figure shows the humanoid robot "Tombatossals" that was used for the human-robot interaction task.* ***(b)*** *The figure shows the experimental setup. The participant, seated on the left side of the monitor, responded to the square by pressing a button with the right hand. The robot, seated on the right side, responded to the diamond by pressing a button with its left hand.*

1.4 Potential impacts

Although the project EYESHOTS has an explorative, pre-industrial character, the innovative computational paradigms and the cognitive engineering solutions, devised to operate adaptively outside the manufactured environments as well as pragmatic application scenarios, are expected to have impact on service robotics.

Indeed, as an important new market is appearing, namely that of machines/robots that can learn and generalise their learning so as to be able to communicate and cooperate with people or each other, we need to create anthropomorphic machines that are aware of their environments, including the presence of humans and other anthropomorphic machines. This is thus prior to perceiving and understanding and adapting to the behaviours of the latter. In this way, this project is at the front-end of such a system by having units that analyse visual scenes, by processing their dynamic aspects, that evolve over time through integrating new aspects of the visual scenes (binding visual "fragments"), and that reorganize themselves collectively (such as active stereopsis and learning sensorimotor transformations). In summary, the originality of this project resides in that it grounded on models of cerebral mechanisms of attention and eye movement control, integrating the most recent results of visual neuroscience. It contributed to innovative technologies by developing system components that can be used in numerous high-tech robotic applications. And to close the circle, it contributed to visual neuroscience by helping improve current understanding of cerebral mechanisms of human attention, active stereopsis and eye movement control through its application to real-world examples provided by an operational anthropomorphic machine.

Specific potentially exploitable results obtained by the different partners are listed here below.

UG

Biomechanics of the oculomotor plant

The activity performed during the project has focused on the scientific aspects related to the analysis and modelling of the biomechanics of the oculomotor plant. The major impact is therefore expected within the Scientific Community. A significant engineering effort has been also put on the development of a blockset library (for SIMULINK) to support the development of models featuring the dynamic equations that describe a detailed model of the oculomotor plant (including muscle-tendon dynamics). This library, will be publicly accessible through the project's website and could provide support for researchers and scientists.

Furthermore a robot eye prototype has been developed to prove experimentally the feasibility of the implementation of a miniature tendon driven and bioinspired robot eye. The design documentation (required to build replicas of the prototype) will be made available to the Scientific Community shortly after the completion of the project.

Vision modules for interactive stereopsis

The efficacy of distributed (cortical-like) models has usually to cope with the efficiency of the hosting (non-neural) computational architectures. Indeed, although the performances of these models were promising, they have never been *largely* employed in real-world applications. This is mainly due to their high computational cost. The major impact of the work conducted by partner UG concerned the demonstration that, by following a specific design approach, it is possible to implement *neuromorphic* solutions for stereo (and motion) vision problems that are characterized by an affordable computational cost, to be efficiently employed in closed-loop robotic applications. The *higher flexibility* of the neural solution does not come at the price of lower performances, being in some cases even more effective than its "non-neural" counterparts (e.g., with respect to noise resistance). The current C++ implementation process up to 2 fps by using a population of 56 binocular complex cells, with receptive fields of 41×41 pixel and an image resolution of 640×480 pixels. A pilot GPU-based implementation of the distributed architecture for the computation of the full (i.e., 2D) disparity, using the Nvidia CUDA Library yielded encouraging results (up to 10 fps by using a population of 72 binocular cells, an image resolution of 512x512 pixel and 6 spatial scales). The use of the this module to control the binocular active movement of a stereo robot head made possible a system validation of the approach in real-world condition, to demonstrate that a bio-inspired solution can be used as a systematic alternative to computer vision, working on raw images and real video sequences.

More generally, cortical-like architectures as bio-inspired structural paradigms to solve computer vision tasks, can represent a viable solution for the next-generation robot vision systems, which should be capable of calibrating and adapting autonomously through the interaction with the environment.

The approach and the results have been presented at Conferences and Workshops on "Computer Vision", "Neural Modeling", "Visual Perception" and "System Neuroscience", and reported in peer-reviewed journals and edited book chapters (as detailed in the tables below).

WWU

Object recognition, generally implemented in a hierarchical bottom-up process in which the complexity of representation along with the receptive field size increases, leads to a strong overlapping of populations encoding features belonging to different objects. These ambiguities in cell populations encoding features within the same receptive field limit the use of these approaches for non-segmented scenes like natural images.

The closely linked paradigms of active vision, purposive vision and animate vision have proposed that bottom-up directed vision is an ill-posed-problem and suggested each task requires its own specific algorithm. In this regard, an universal, general vision is not possible. According to these paradigms, the fundamental problem of vision is the selection of the relevant information within the scene and the computation of an appropriate representation. An "active" vision system – in the sense of a visually selective device – is able to acquire the necessary information on demand by focusing on the relevant areas within the visual scene and taking different views from the same object.

We have developed a general multipurpose framework for such goal directed vision tasks. It includes learning rules for learning object selective cells, methods for implementing visual attention by means of a dynamic interactive process and the implementation of executive functions which are also subject to learning.

This work has been primarily made available to the public by publication in scientific journals with a high visibility and by presentation on international conferences.

UNIBO

The neurophysiological work conducted by UNIBO has been shared with all the partners of the EYESHOTS project, as documented by several joint publications and meeting presentations belonging to both the neuroscientific and robotic fields. In addition, through the extensive exchanges and collaborations within the project, UNIBO has biologically oriented scientists from other disciplines participating in the consortium and offered to control engineers new ways to control their systems. Exploitation of results has been achieved entirely within the context of EYESHOTS (see also "Neurophysiological evidences of joint visuo-motor descriptors of 3D space in the parietal cortex" at p.6 of this document).

Several dissemination activities have been performed by the constant presence of UNIBO members to National, European and International conferences on Robotics and Neuroscience. Two UNIBO members have actively participated in the summer school organized by EYESHOTS. UNIBO members have given lectures and organized symposia in European conferences, and been invited to present the results of the research in European research centers.

UJI

The visuomotor skills of the humanoid robot "Tombatossals" and the flexibility it is endowed with, thanks to its underlying biologically-inspired computational structure, allows it to act as a test subject for psychophysical experiments, as justified by the promising results obtained in the saccadic adaptation protocol. For example, we could investigate on the robot how saccadic adaptation transfers to arm reaching movements, and employ the outcome to propose new experiments with human subjects. Indeed, saccadic adaptation is surely not the only experimental paradigm that the robot is able to properly execute. The modular structure of the software permits to easily define fairly complex visuomotor behaviors including many different configurations of the workspace, and a number of possible tasks which involve either eye or arm movements or both. Possible research goals the robot could contribute to are the effect of different stereoptic configurations (in term of vergence and version values) on the properties of goal-directed reaching movements, or the study of the mechanisms behind short and medium-term visuomotor memories regarding objects in the peripersonal space.

K.U.Leuven

Future robots are expected to be able to actively and voluntarily explore their environments. To this end, a number of issues need to be tackled, such as the version/version control of the robot "eyes" in relation to the scene layout. The version control could be driven by attention given to interesting object properties or other aspects of the scene. The vergence angle of the robot "eyes" needs to be controlled without putting restrictions on the object surface angle and the gaze angle, and the disparities and gaze angle accurately estimated in the presence of inaccuracies in the eye motor control system.

We have developed an integerated approach to version/vergence control, based on a learning strategy, instead of being hardwired. This way, the robot learns to control the eyes, which can be part of a more encompassing learning behavior of the robot. The learning also takes into account the transformation of retinal disparities into an accurate 3D scene description, substantially going beyond what is possible given the accuracy of the motor system. The learning takes also into account the effect of the eye movements itself on the estimate, thereby solving the 2D correspondence problem.

All these advancements are step stones towards cognitive robots that can learn to perceive and explore their environments autonomously.

1.5 Project data and contact details

The official EYESHOTS project website: ***www.eyeshots.it***

EYESHOTS logo:

EYESHOTS
Heterogeneous 3D Perception across Visual Fragments

Project contact details:

Prof. Silvio P. Sabatini

DIBE – University of Genoa

Via Opera Pia 11a, I-16145 Genova, Italy
Phone: +39-010-3532092
Fax: +39-010-3532289
e-mail: silvio.sabatini@unige.it
URL : www.pspc.dibe.unige.it

1.6 References

Alexander, G.E., DeLong, M.R. and Strick, P.L., Parallel organization of functionally segregated circuits linking the basal ganglia and cortex. *Annual Review of Neuroscience*, **9**:357-381, 1989.

Bahill, A.T, Latimer JR, Troost BT. Linear homeomorphic model for human movement. *IEEE Trans Biomed Eng*. **27**(11):631-9, 1980.

Bernier, P.M., and Grafton, S.T. Human posterior parietal cortex flexibly determines reference frames for reaching based on sensory context. *Neuron* **68**:776-788, 2010.

Beuth, F., Wiltschut, J. and Hamker, F.H., Attentive Stereoscopic Object Recognition, in: Machine Learning reports of AG Computational Intelligence, University of Leipzig, 04/2010:41-48, 2010.

Bosco, A., Breveglieri, R., Chinellato, E., Galletti, C., and Fattori, P. Reaching activity in the medial posterior parietal cortex of monkeys is modulated by visual feedback. J Neurosci 30:14773-14785., 2010.

Breveglieri, R., Kutz, D.F., Fattori, P., Gamberini, M., and Galletti, C. Somatosensory cells in the parieto-occipital area V6A of the macaque. *Neuroreport* **13**:2113-2116, 2002.

Breveglieri, R., Hadjidimitrakis, K., Bosco, A., Sabatini, S., Galletti, C., and Fattori, P. Balanced sampling of visual fragments in the reachable space by parieto-occipital neurons. Manuscript in preparation, 2011.

Cannata G., Maggiali, M. Models for the Design of Bio-Inspired Robot Eyes. *IEEE Trans. on Robotics* **24**(1), 2008.

Chessa, M., Sabatini, S.P., and Solari, F. A fast joint bioinspired algorithm for optic flow and two-dimensional disparity estimation. In ICVS, pages 184-193, 2009a

Chessa, M., Canessa, A., Gibaldi, A., Solari, F. Embedding Fixation Constraints into Binocular Energy-based Models of Depth Perception. International Conference on Cognitive and Neural Systems, Boston, Massachusetts, 27-30 May, 2009, 2009b.

Chumerin, N., Gibaldi, A., Sabatini, S.P., Van Hulle, M.M. Learning eye vergence control from a distributed disparity representation. *International Journal of Neural Systems*, **20**(4):267-278, 2010.

Collins, T., Dore-Mazars, K., and Lappe, M. Motor space structures perceptual space: Evidence from human saccadic adaptation. *Brain Res.*, **1172**:32-39, 2007.

Enderle JD, Wolfe JW, Yates JT. Optimal control of saccadic eye movements. *Biomed Sci Instrum*. **20**:143-8, 1984.

Enderle JD, Wolfe JW. Time-optimal control of saccadic eye movements. *IEEE Trans Biomed Eng*. **34**(1):43-55, 1987.

Fattori, P., Gamberini, M., Kutz, D.F., and Galletti, C. 'Arm-reaching' neurons in the parietal area V6A of the macaque monkey. Eur J Neurosci 13:2309-2313, 2001.

Fattori, P., Kutz, D.F., Breveglieri, R., Marzocchi, N., and Galletti, C. Spatial tuning of reaching activity in the medial parieto-occipital cortex (area V6A) of macaque monkey. Eur J Neurosci 22:956-972., 2005.

Fattori, P., Raos, V., Breveglieri, R., Bosco, A., Marzocchi, N., and Galletti, C. The dorsomedial pathway is not just for reaching: grasping neurons in the medial parieto-occipital cortex of the macaque monkey. J Neurosci 30:342-349, 2010.

Fleet, D.J., Wagner, H. and Heeger. D.J. Modelling binocular neurons in the primary visual cortex. In *Computational and Biological Mechanisms of Visual Coding*, M. Jenkin and L. Harris (eds.), Cambridge University Press, pp. 103-130, 1996.

Földiak, P. Learning invariance from transformation sequences. Neural Computation, 3:194-200, 1991.

Galletti, C., Battaglini, P.P., and Fattori, P. Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey. Eur J Neurosci 7:2486-2501, 1995.

Galletti, C., Fattori, P., Battaglini, P.P., Shipp, S. and Zeki, S. Functional demarcation of a border between areas V6 and V6A in the superior parietal gyrus of the macaque monkey. *Eur J Neurosc.*, **8**:30-52, 1996.

Galletti, C., Fattori, P., Kutz, D.F., and Battaglini, P.P. Arm movement-related neurons in the visual area V6A of the macaque superior parietal lobule. Eur J Neurosci 9:410-413, 1997.

Galletti, C., Fattori, P., Kutz, D.F. and Gamberini, M. Brain location and visual topography of cortical area V6A in the macaque monkey. *Eur J Neurosc.*, **11**:575-582, 1999.

Galletti, C., Kutz, D.F., Gamberini, M., Breveglieri, R. and Fattori, P. Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Exp Brain Res.*, **153**:158-170, 2003.

Galletti, C., Breveglieri, R., Lappe, M., Bosco, A., Ciavarro, M., and Fattori, P. Covert shift of attention modulates the ongoing neural activity in a reaching area of the macaque dorsomedial visual stream. PLoS ONE 5:e15078, 2010.

Gamberini, M., Passarelli, L., Fattori, P., Zucchelli, M., Bakola, S., Luppino, G., and Galletti, C. Cortical connections of the visuomotor parietooccipital area V6Ad of the macaque monkey. J Comp Neurol 513:622-642, 2009.

Gamberini, M., Galletti, C., Bosco, A., Breveglieri, R., and Fattori, P. Is the medial posterior-parietal area V6A a single functional area? J Neurosci 31:5145-5157, 2011.

Gibaldi A., Chessa M., Canessa A., Sabatini S.P., Solari F. A neural model for binocular vergence control without explicit calculation of disparity. 11th European Symposium on Artificial Neural Networks (ESANN09), 2009.

Gibaldi, A., Chessa, M., Canessa, A., Sabatini, S.P., and Solari, F. A cortical model for binocular vergence control without explicit calculation of disparity. *Neurocomp*., **73**:1065-1073, 2010.

Gibaldi A., Canessa A., Chessa M., Solari F., Sabatini S.P. A cortical model for vergence control: advantages of space-variant geometry of the cortical domain. Cosyne11, Salt Lake City, Utah, February 24 - March 1, 2011.

Goodale, M.A., and Milner, A.D. Separate visual pathways for perception and action (Review). Trends Neurosci 15:20-25, 1992.

Grafton, S.T. The cognitive neuroscience of prehension: recent developments. Exp Brain Res 2010.

Hadjidimitrakis, K., Breveglieri, R., Placenti, G., Bosco, A., Sabatini, S.P., Galletti, C., and Fattori, P. Fix Your Eyes Where You Can Reach: Neurons in the Macaque Medial Parietal Cortex Prefer Gaze Positions in Peripersonal Space. submitted, 2011.

Hadjidimitrakis, K., Breveglieri, R., Bosco, A., Placenti, G., Bertozzi, F., Fattori, P., and Galletti, C. Fixation in depth reveals a preference for reachable space in the macaque medial parieto-occipital cortex. Program No 757 2010 Neuroscience Meeting Planner San Diego, CA: Society for Neuroscience, 2010 Online, 2010.

Hamker, F.H. The reentry hypothesis: the putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cerebral Cortex.*, **15**:431-447, 2005a.

Hamker, F.H., The emergence of attention by population-based inference and its role in distributed processing and cognitive control of vision. *Comp Vis Image Understand*, 100:64-106, 2005b.

Hamker, F. H. The mechanisms of feature inheritance as predicted by a systems-level model of visual attention and decision making. *Advances in Cognitive Psychology*, **3**:111-123, 2007.

Havermann, K. Lappe, M. The influence of the consistency of postsaccadic visual errors on saccadic adaptation. *J Neurophysiol*, **103**(6), 3302–3310, 2010.

Havermann, K., Volcic, R., Lappe, M. Saccadic adaptation to moving stimuli. In preparation, 2011a.

Havermann, K., Zimmermann, E., Lappe, M. Eye position effects in saccadic adaptation. *J Neurophysiol*. Under revision, 2011b.

Hollerman, J.R. and Schultz, W. Dopamine neurons report an error in the temporal prediction of reward during learning. Nature Neuroscience, 1(4):304-309, 1998.

Hung, G.K., Semmlow, J.L., and Ciuffreda, K.J. A dual-mode dynamic model of the vergence eye movement system. *IEEE Trans. Biomed. Eng.*, **36**(11):1021-1028, 1986.

Knill, D.C., and Kersten, D. Visuomotor sensitivity to visual information about surface orientation. J Neurophysiol 91:1350-1366. Epub 2003 Oct 1329., 2004.

Kutz, D.F., Fattori, P., Gamberini, M., Breveglieri, R., and Galletti, C. Early- and late-responding cells to saccadic eye movements in the cortical area V6A of macaque monkey. Exp Brain Res 149:83-95., 2003.

Liepelt, R., and Brass, M. Top-Down modulation of motor priming by belief about animacy. *Exp Psychol*, **57**:221-227, 2010.

Liepelt, R., Prinz, W., and Brass, M. When do we simulate non-human agents? Dissociating communicative and non-communicative actions. *Cognition,* **115**:426-434, 2010.

McGuire, L.M., and Sabes, P.N. Sensory transformations and the use of multiple reference frames for reach planning. Nat Neurosci 12:1056-1061, 2009.

Marzocchi, N., Breveglieri, R., Galletti, C., and Fattori, P. Reaching activity in parietal area V6A of macaque: eye influence on arm activity or retinocentric coding of reaching movements? *Eur J Neurosci* 27:775-789, 2008.

Rolls, E.T. and Stringer, S.M. Invariant object recognition in the visual system with error correction and temporal difference learning. Network, 12(2):111-129, 2001.

Sabatini S.P., Gastaldi G., Solari F., Pauwels K., Van Hulle M.M., Diaz J., Ros E., Pugeault N., and Kruger N.. A compact harmonic code for early vision based on anisotropic frequency channels. *Computer Vision and Image Understanding*, **114**(6):681–699, 2010.

Schnier, F., Zimmermann, E., Lappe, M. Adaptation and mislocalization fields for saccadic outward adaptation in humans. *Journal of Eye Movement Research*, 3(4), 2010.

Schnier, F. & Lappe, M. Differences in inter-saccadic adaptation transfer between inward and outward adaptation. *J. Neurophysiol.* Submitted.

Schroll, H., Vitay, J., and Hamker, F.H. Working memory and response selection: A computational account of interactions among cortico-basal ganglio-thalamic loops, Submitted.

Schultz, W., Dayan, P. and Montague, P.R., A neural substrate of prediction and reward. Science, 275:1593-1599, 1997.

Schultz, W., Getting formal with dopamine and reward. Neuron, 36:241-263, 2002.

Sebanz N., Knoblich, G., and Prinz, W. Representing others‘ actions: just like one‘s own? *Cognition*, **88**:B11-B21, 2003.

Sober, S.J., and Sabes, P.N. Flexible strategies for sensory integration during motor planning. Nat Neurosci 8:490-497, 2005.

Stenzel, A., Chinellato E., Tirado Bou, M.A., del Pobil, A.P., Lappe, M., and Liepelt. R. How humanoid robots become human-like partners in joint action. In preparation.

Takemura A, Inoue Y, Quaia C, Miles FA. Single-unit activity in cortical area MST associated with disparity-vergence eye movements: evidence for population coding, *J. Physiol*. **85** (5):2245–2266, 2001.

Tsai, C.-C., and Brass, M. Does the human motor system simulate Pinocchio’s actions? *Psychol Sci*, **18**:1058-1062, 2007.

Tsai, C.-C., Kuo, W.-J., Hung, D. L., and Tzeng, O. J. L. Action co-representation is tuned to other humans. *J Cognitive Neurosci*, **20**:2015-2024, 2008.

van Beers, R.J., Wolpert, D.M., and Haggard, P. When feeling is more important than seeing in sensorimotor adaptation. Curr Biol 12:834-837, 2002.

Vitay, J. and Hamker, F.H., A computational model of the inuence of basal ganglia on memory retrieval in rewarded visual memory tasks. *Frontiers in Computational Neuroscience,* **4**(13):1-18, 2010.

Volcic, R. and Lappe, M. Predictive eye movements in gaze and action observation. In preparation.

Volcic, R. and Lappe, M. Potentially purposeful actions divert overt attention. *Attention, Perception and Psychophysics*. Under revision

Volcic, R. and Lappe, M. Gaze behavior in cooperative action. In preparation.

Wallis, G. and Rolls, E. T. Invariant face and object recognition in the visual system. Prog Neurobiol, 51(2):167-194, 1997.

Wiltschut, J., and Hamker, F. Efficient coding correlates with spatial frequency tuning in a model of v1 receptive field organization. *Vis. Neurosci*., **26**:21-34, 2009.

Wulff, S., Havermann, K,, Lappe, M., Bosco, A., Galletti, C., Fattori, P. Eye position effects in saccadic adaptation in the macaque. In preparation.

Zimmermann, E., and Lappe, M. Motor signals in visual localization. *Journal of Vision*, **10**(6), 2. 2010.

Zimmermann, E. Lappe, M. Eye position effects in oculomotor plasticity and visual localization. *J. Neurosci*. In press.

Zirnsak, M. and Hamker, F.H. Attention alters feature space in motion perception. J. Neurosci., 30(20):6882-6890, 2010.

2 USE AND DISSEMINATION OF FOREGROUND

2.1 Section A: Dissemination of foreground

TEMPLATE A1: LIST OF SCIENTIFIC (PEER REVIEWED) PUBLICATIONS, STARTING WITH THE MOST IMPORTANT ONES

NO.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year of publication	Relevant pages	Permanent identifiers[1] (if available)	Is/Will open access[2] provided to this publication?
1	*A Cortical Architecture on Parallel Hardware for Motion Processing in Real Time.*	*Pauwels, K.*	*Journal of Vision*	*10(10):18*	*Association for Research in Vision and Ophthalmology*		*2010*	*1-21*	*http://www.journalofvision.org/content/10/10/18*	*Yes*
2	*Compact harmonic code for early vision based on anisotropic frequency channels*	*Sabatini, S.P.*	*Computer Vision and Image Understanding*	*114(6)*	*Elsevier*		*2010*	*681-699*		*Yes*
3	*Learning eye vergence control from a distributed disparity*	*Chumerin, N.*	*International Journal of Neural Systems*	*20(4)*	*World Scientific Publishing Co.*	*Singapore*	*2010*	*267-278*		*Yes*

---

[1] A permanent identifier should be a persistent link to the published version full text if open access or abstract if article is pay per view) or to the final manuscript accepted for publication (link to article in repository).

[2] Open Access is defined as free of charge access for anyone via Internet. Please answer "yes" if the open access to the publication is already established and also if the embargo period for open access is not yet over but you intend to establish open access afterwards.

	representation									
4	*Convolutional Network for Vergence Control.*	*Chumerin, N.*	*Proceedings 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2009)*	*Nov 24-27, 2009*	*IEEE*	*Bratislava, Slovak Republic*	*2009*	*1-6.*		*Yes*
5	*Implicit sensorimotor mapping of the peripersonal space by gazing and reaching*	*Chinellato, E.*	*IEEE Transactions on Autonomous Mental Development*	*3(1)*	*IEEE*		*2011*	*43-53*	*10.1016/j.neucom.2010.07.029*	*No*
6	*The Dorso-medial visual stream: From neural activation to sensorimotor interaction*	*Chinellato, E.*	*Neurocomputing*	*74(8)*	*Elsevier*		*2011*	*1203-1212*	*10.1109/TAMD.2011.2106781*	*No*
7	*Is the medial posterior parietal area V6A a single functional area?*	*Gamberini M.*	*Journal of Neuroscience*	*31(13), Mar*	*Society of Neuroscience*	*United States of America*	*2011*	*5145-5157*		*yes*
8	*Reaching activity in the medial posterior parietal cortex of monkeys is modulated by visual feedback*	*Bosco A.*	*Journal of Neuroscience*	*30(44), Nov*	*Society of Neuroscience*	*United States of America*	*2010*	*14773-85*		*yes*
9	*The Dorso-medial visual stream: from Neural Activation to*	*Chinellato E.*	*Neurocomputing*	*doi:10.1016/j. neucom.201*	*Elsevier*		*2010*			*yes*

	Sensorimotor Interaction			*0.07.029*						
10	*Covert Shift of Attention Modulates the Ongoing Neural Activity in a Reaching Area of the Macaque Dorsomedial Visual Stream.*	*Galletti C.*	*PloS ONE*	*5(11)*	*Bart Krekelberg, Rutgers University*	*United States of America*	*2010*	*e15078*		*yes*
11	*The cortical visual area V6 in macaque and human brains.*	*Fattori P.*	*Journal of Physiology-Paris*	*103 (2009)*	*Elsevier*	*Paris*	*2009*	*88 - 97*		*yes*
12	*Evidence for Peak-shaped Gaze Fields in Area V6A: Implications for Sensorimotor Transformations in Reaching Tasks.*	*Breveglieri R.*	*Series: Lecture Notes in Computer Science. Subseries: Theoretical Computer Science and General Issues*	*5602 ISBN: 978-3-642-02266-1*	*Mira, J.; Ferrandez, I. M.; Alvarez Sanchez, J.-R.; de la Paz, F.; Toledo, F. J*	*IWINAC 2009, Santiago de Compostela, Spain*	*2009*	*324-333*		*yes*
13	*Toward an Integrated Visuomotor Representation of the Peripersonal Space*	*Chinellato E.*	*Series: Lecture Notes in Computer Science. Subseries: Theoretical Computer Science and General Issues*	*5602 ISBN: 978-3-642-02266-1*	*Mira, J.; Ferrandez, I. M.; Alvarez Sanchez, J.-R.; de la Paz, F.; Toledo, F. J*	*IWINAC 2009, Santiago de Compostela, Spain*	*2009*	*314-323*		*yes*
14	*Eye-Hand Coordination for*	*Chinellato E.*	*Series: Lecture Notes in*	*5602 ISBN:*	*Mira, J.; Ferrandez, I.*	*IWINAC 2009,*	*2009*	*304-313*		*yes*

	Reaching in Dorsal Stream Area V6A: Computational Lessons		Computer Science. Subseries: Theoretical Computer Science and General Issues	978-3-642-02266-1	M.; Alvarez Sanchez, J.-R.; de la Paz, F.; Toledo, F. J	Santiago de Compostela, Spain				
15	A computational model of the influence of basal ganglia on memory retrieval in rewarded visual memory tasks	Vitay, Hamker	Frontiers in Computational Neuroscience	4-13	Frontiers		2010	Article 13		yes
16	Efficient Coding correlates with spatial frequency tuning in a model of V1 receptive field organization	Hamker, Wiltschut	Visual Neuroscience	26			2009	21-34		no
17	Attention Alters Feature Space in Motion Processing	Zirnsak, Hamker	J Neurosci	30(20)			2010	6882-6890		yes
18	Biological Models of Reinforcement Learning.	Vitay, Hamker	Künstliche Intelligenz	3			2010	12-18		no
19	Computational models of spatial updating in peri-saccadic perception	Hamker	Phil. Trans. R. Soc. B	366			2011	554-571		no
20	Attentive Stereoscopic Object Recognition	Beuth, Hamker	Machine Learning reports	04			2010	41-48		yes
21	Mislocalization of	Eckart	Journal of	29(35)			2009	11055-		yes

	flashed and stationary visual stimuli after adaptation of reactive and scanning saccades	*Zimmermann*	*Neuroscience*					*11064*		
22	*Motor signals in visual localization*	*Eckart Zimmermann*	*Journal of Vision*	*10(6):2*			*2010*	*1-11*	*http://www.journalisofvision.org/content/10/6/2*	*yes*
23	*Potentially purposeful actions divert overt attention*	*Robert Volcic*	*Attention, Perception& Psychophysics*	*Under revision*						
24	*Adaptation and mislocalization fields for saccadic outward adaptation in humans*	*Fabian Schnier*	*Journal of Eye Movement Research*	*3(3:4)*			*2010*	*1-18*		*yes*
25	*The influence of the consistency of post saccadic visual error on saccadic adaptation*	*Katharina Havermann*	*Journal of Neurophysiology*	*105*			*2010*	*1130-1140*		*yes*
26	*Eye position effects in saccadic adaptation*	*Katharina Havermann*	*Journal of Neurophysiology*	*under revision*						
27	*Eye position effects in saccadic adaptation in macaques*	*Svenja Wulff*	*In preparation*							
28	*Eye position effects in oculomotor*	*Eckart Zimmerma*	*Journal of Neuroscience*	*In press*						

	plasticity and visal localization	*nn*								
29	*Differences in inter-saccadic adaptation transfer between inward and outward adaptation*	*Fabian Schnier*	*Journal of Neurophysiology*	*submitted*						
30	*A cortical model for binocular vergence control without explicit calculation of disparity*	*Gibaldi, A.*	*Neurocomputing*	*73*			*2010*	*1065-1073*		
31	*Design strategies for direct multiscale and multiorientation visual processing in the log-polar domain*	*Solari, F.*	*Pattern Recoglition Letter*	*accepted*						
32	*Virtual Reality to Simulate Visual Tasks for Robotic Systems*	*Chessa, M.*	*Virtual Reality*		*InTech*		*2011*			
33	*Visuo-motor constraints in binocular eye coordination: optimization theories revisited*	*Canessa, A.*		*submitted*						
34	*Early perception-action cycles in binocular vision: Visuomotor paradigms and*	*Sabatini, S.P.*	*Developing and Applying Biologically-Inspired Vision Systems:*		*IGI Global*		*2011*			

	cortical-like architectures		*Interdisciplinary Concepts*							

TEMPLATE A2: LIST OF DISSEMINATION ACTIVITIES

NO.	Type of activities[3]	Main leader	Title	Date	Place	Type of audience[4]	Size of audience	Countries addressed
1	*Conference*	*K.U.Leuven*	*IDEAL conference*	*1-3 September 2010*	*University of West Scotland, Paisley, Scotland*	*Scientific Community*	*50*	*International*
2	*Conference*	*K.U.Leuven*	*2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2009)*	*November 24-27, 2009*	*Bratislava, Slovak Republic*	*Scientific Community*	*200*	*International*
3	*Conference*	*K.U.Leuven*	*Workshop on Visual Computer Vision on GPUs (in conjunction with CVPR 2008)*	*June 27, 2008*	*Anchorage, Alaska*	*Scientific Community*	*500*	*International*
4	*Conference*	*K.U.Leuven*	*7th International*	*8-10 July*	*St. Augustine,*	*Scientific*	*120*	*International*

---

[3] A drop down list allows choosing the dissemination activity: publications, conferences, workshops, web, press releases, flyers, articles published in the popular press, videos, media briefings, presentations, exhibitions, thesis, interviews, films, TV clips, posters, Other.

[4] A drop down list allows choosing the type of public: Scientific Community (higher education, Research), Industry, Civil Society, Policy makers, Medias ('multiple choices' is possible.

			Workshop on Self-Organizng Maps	2009	USA	Community		
5	Conference	K.U.Leuven	ICORR2009	23-26 June 2009	Kyoto, Japan	Scientific Community	200	International
6	Lecture	K.U.Leuven	Univ. Manchester	4 March 2008	Manchester, UK	Scientific Community (lab of Prof. H. Yin)	20	UK
7	Lecture	K.U.Leuven	Danish Technical University	19 March 2008	Lyngby, Denmark	Scientific Community (lab of Prof. J. Larsen)	20	DK
8	Conference	UJI	Implicit mapping of the peripersonal space of a humanoid robot.	April 2011	Paris (France)	Scientific Community	200	International
9	Conference	UJI	Hierarchical object recognition inspired by primate brain mechanisms	April 2011	Paris (France)	Scientific Community	200	International
10	Conference	UJI	An Integrated Virtual Environment for Visual-based Reaching	February 2011	Seoul (Korea)	Scientific Community	100	International
11	Conference	UJI	Visuomotor spatial awareness through concurrent reach/gaze actions	January 2010	Zurich (Switzerland)	Scientific Community	50	International
12	Conference	UJI	Toward an Integrated Visuomotor Representation of the Peripersonal Space	June 2009	Santiago de Compostela (Spain)	Scientific Community	100	International
13	Conference	UJI	Eye-Hand Coordination for Reaching in Dorsal Stream Area V6A: Computational	June 2009	Santiago de Compostela (Spain)	Scientific Community	100	International

			Lessons					
14	*Summer school*	*UJI*	*International UJI Robotics School IURS 2009 on Visuomotor Interaction*	*September 2009*	*Benicassim (Spain)*	*Scientific Community*	*40*	*International*
15	*Workshop*	*UJI*	*IWINAC 2009 Pre-Organized Session on Robotics and Neuroscience*	*June 2009*	*Santiago de Compostela (Spain)*	*Scientific Community*	*50*	*International*
16	*Conference*	*UNIBO*	*INSERM, unit 864*	*24 February 2011*	*Lyon, France*	*Scientific Community*		*International*
17	*Poster*	*UNIBO*	*WINTER SCHOOL, FENS*	*9-16 January, 2011*	*Obergurgl, Austria*	*Scientific Community*		*International*
18	*4 Posters*	*UNIBO*	*2010 Neuroscience Meeting Society for Neuroscience*	*13-17 November, 2010*	*San Diego, Ca, USA*	*Scientific Community*		*International*
19	*Conference*	*UNIBO*	*International Conference on Parietal Lobe Function, European Science Foundation*	*20-21 September, 2010*	*Artis Zoo, Amsterdam*	*Scientific Community*		*International*
20	*Poster*	*UNIBO*	*61 th Meeting of Italian Society of Physiology*	*15-17 September, 2010*	*Varese, Italy*	*Scientific Community*		*Italy*
21	*Conference*	*UNIBO*	*3rd Mediterranean Conference of Neuroscience "Neuronal processes of attention and action and their use in artificial intelligent systems"*	*13-16 December, 2009*	*Alexandria, Egypt*	*Scientific Community*		*International*
22	*2 Posters*	*UNIBO*	*2009 Neuroscience*	*17-21*	*Chicago, IL,*	*Scientific*	*10000*	*International*

			Meeting Society for Neuroscience	*October, 2009*	*USA*	*Community*		
23	*Conference*	*UNIBO*	*9th International UJI Robotics School on Visuomotor Interaction*	*14-18 September, 2009*	*Bonaire Hotel, Benicàssim, Spain*	*Scientific Community*	*30*	*International*
24	*Presentation*	*UNIBO*	*VSS2009 meeting Vision Sciences Society*	*8-13 May, 2009*	*Naples, Fl, USA.*	*Scientific Community*	*1500*	*International*
25	*Two presentations*	*UNIBO*	*2nd annual meeting Primate Neurobiology*	*23-25 March, 2009*	*Gottingen - Tubingen, Germany*	*Scientific Community*		*Germany*
26	*Conference*	*WWU*	*CogSys*	*April 2-4, 2008*	*Karlsruhe (Germany)*	*Scientific Community*	*600*	
27	*Conference*	*WWU*	*Bernstein Conference on Computational Neuroscience*	*September 30, 2009*	*Frankfurt (Germany)*	*Scientific Community*	*800*	*International*
28	*Conference*	*WWU*	*CNS, Computational Neuroscience*	*July, 18th 2009*	*Berlin (Germany)*	*Scientific Community*	*1500*	*International*
29	*Conference*	*WWU*	*Meeting of the Vision Science Society*	*May, 2010*	*Naples (USA)*	*Scientific Community*	*1500*	*International*
30	*Conference*	*WWU*	*New Challenges in Neural Computation*	*Septemer, 21st, 2010*	*Karlsruhe (Germany)*	*Scientific Community*	*60*	*Germany*
31	*Symposium*	*WWU*	*Fusing brain and mental domains*	*12 September 2008*	*Münster, Germany*	*Scientific Community*	*150*	*Germany*
32	*Retreat*	*WWU*	*Retreat of the Otto Creutzfeld Center*	*January 2009*	*Rotenberge, Germany*	*Scientific Community*	*50*	*Germany*
33	*Conference*	*WWU*	*Vision Science Society Meeting*	*8-13 May 2009*	*Naples, Florida, USA*	*Scientific Community*	*2000*	*international*
34	*Conference*	*WWU*	*European Conference on Visual Perception*	*24-28 May 2009*	*Regensburg Germany*	*Scientific Community*	*600*	*international*
35	*Conference*	*WWU*	*European Conference on Eye Movements*	*23-27 August 2009*	*Southampton, UK*	*Scientific Community*	*600*	*international*

36	*Conference*	*WWU*	*International UJI Robotics School on Visuomotor Interaction*	*14-18 September 2009*	*Benicassim, Spain*	*Scientific Community*	*30*	*international*
37	*Retreat*	*WWU*	*Retreat of the Otto Creutzfeld Center*	*Janurary 2010*	*Rotenberge, Germany*	*Scientific Community*	*50*	*Germany*
38	*Conference*	*WWU*	*Vision Science Society Meeting*	*7-12 May 2010*	*Naples, Florida, USA*	*Scientific Community*	*2000*	*international*
39	*Conference*	*WWU*	*Neuroscience*	*13-17 November 2010*	*San Diego, USA*	*Scientific Community*	*30000*	*international*
40	*Conference*	*WWU*	*Tagung experimentell arbeitender Psychologen*	*13-16 March 2011*	*Halle, Germany*	*Scientific Community*	*1000*	*Germany*
41	*Conference*	*UG*	*13th Int. Conference on Cognitive and Neural Systems*	*27-30 May 2009*	*Boston, MA, USA*	*Scientific Community*	*200*	*International*
42	*Conference*	*UG*	*13th Int. Conference on Cognitive and Neural Systems*	*27-30 May 2009*	*Boston, MA, USA*	*Scientific Community*	*200*	*International*
43	*Conference*	*UG*	*11th European Symposium on Artificial Neural Networks*	*April 2009*	*Bruges, Belgium*	*Scientific Community*	*400*	*European*
44	*Conference*	*UG*	*International Conference on Computer Vision Theory and Applications 2009*	*5-8 February 2009*	*Lisbon, Portugal*	*Scientific Community*	*300*	*International*
45	*Conference*	*UG*	*International Conference on Computer Vision Systems*	*October 2009*	*Liege, Belgium*	*Scientific Community*	*200*	*International*
46	*Conference*	*UG*	*3rd Mediterranean Conference of*	*13-16 December*	*Alexandria, Egypt*	*Scientific Community*	*50*	*International*

			Neuroscience	*2009*				
47	*Conference*	*UG*	*32nd European Conference on Visual Perception*	*August 2009*	*Regensburg, Germany*	*Scientific Community*	*400*	*International*
48	*Conference*	*UG*	*32nd European Conference on Visual Perception*	*August 2009*	*Regensburg, Germany*	*Scientific Community*	*400*	*International*
49	*Conference*	*UG*	*32nd European Conference on Visual Perception*	*August 2009*	*Regensburg, Germany*	*Scientific Community*	*400*	*International*
50	*Conference*	*UG*	*4th Int. Conf. On Cognitive Systems*	*27-28 January 2010*	*Zurich, Switzerland*	*Scientific Community*	*300*	*European*
51	*Conference*	*UG*	*4th Int. Conf. On Cognitive Systems*	*27-28 January 2010*	*Zurich, Switzerland*	*Scientific Community*	*300*	*European*
52	*Conference*	*UG*	*33th European Conference on Visual Perception*	*22-26 August, 2010*	*Lausanne, Switzerland*	*Scientific Community*	*400*	*European*
53	*Conference*	*UG*	*Cosyne 2011*	*February 24 - March 1, 2011*	*Salt Lake City, Utah, USA*	*Scientific Community*	*400*	*International*
54	*Conference*	*UG*	*34nd European Conference on Visual Perception*	*28th August – 1st September 2011*	*Tolouse, France*	*Scientific Community*	*400*	*International*
55	*Conference*	*UG*	*34nd European Conference on Visual Perception*	*28th August – 1st September 2011*	*Tolouse, France*	*Scientific Community*	*400*	*International*
56	*Conference*	*UG*	*34nd European Conference on Visual Perception*	*28th August – 1st September 2011*	*Tolouse, France*	*Scientific Community*	*400*	*International*
57	*Conference*	*UG*	*34nd European Conference on Visual*	*28th August – 1st*	*Tolouse, France*	*Scientific Community*	*400*	*International*

			Perception	*September 2011*				
58	*Conference*	*UG*	*34nd European Conference on Visual Perception*	*28th August – 1st September 2011*	*Tolouse, France*	*Scientific Community*	*400*	*International*
59	*Conference*	*UG*	*IEEE/RSJ International Conference on Intelligent Robots and Systems*	*25-30 September 2011*	*San Francisco, CA, USA*	*Scientific Community*	*200*	*International*
60	*Talk*	*UG*	*Centre de Recherche Cerveau & Cognition - CNRS-Université Paul Sabatier Toulouse 3*	*28 May 2010*	*Toulouse, France*	*Scientific Community*	*60*	*France*
61	*Talk*	*UG*	*Workshop on "Predictive models in humanoid gaze control and locomotion" at IEEE 2010 Robotics: Science and Systems Conference*	*June 2010*	*Saragozza, Spain*	*Scientific Community*	*60*	*Spain*

2.2 Section B: Exploitation plans

2.2.1 Part B1

No applications for patents, trademarks or registered designs have been submitted.

No IP rights have been claimed for any invention during this project. Upon completion of the project some of the results obtained are still sensible for publication. Papers related to material referring to the project's work are under preparation and will be submitted within 2011.

Concerning the robot eye development, no trade secret or other undisclosed information will be retained and all the software and design data required for building replicas of the robot eye developed during this project will be made publicly made available through the project's website (following the submission of the related papers quoted above).

2.2.2 Part B2

Type of Exploitable Foreground[5]	**Description of exploitable foreground**	**Confidential Click on YES/NO**	**Foreseen embargo date dd/mm/yyyy**	**Exploitable product(s) or measure(s)**	**Sector(s) of application**[6]	**Timetable, commercial or any other use**	**Patents or other IPR exploitation (licences)**	**Owner & Other Beneficiary(s) involved**
SOFTWARE LIBRARY (Simulink toolbox)	*EYESHOTS Blockset Library (*)*	*YES*	*31/12/2011*	*Software toolbox for robot eye movements and control analysis.*	*J62.0.9 - Other information technology and computer service activities*	*N/A*	*Open Source*	*Beneficiary:* *Giorgio Cannata and Andrea Trabucco (DIST – Universita' di Genova)*
Robot Design Drawings	*Design drawings and CAD models of the robot eye prototype (**)*	*YES*	*31/12/2011*	*Humanoid robot eye system*	*J62.0.9 - Other information technology and computer service activities*	*N/A*		*Beneficiary:* *Giorgio Cannata and Andrea Trabucco (DIST – Universita' di Genova*

(*) EYESHOTS Blockset Library

---

[19] A drop down list allows choosing the type of foreground: General advancement of knowledge, Commercial exploitation of R&D results, Exploitation of R&D results via standards, exploitation of results through EU policies, exploitation of results through (social) innovation.

[6] A drop down list allows choosing the type sector (NACE nomenclature) : http://ec.europa.eu/competition/mergers/cases/index/nace_all.html

Further research needed: the EYESHOTS Blockset Library is an open software collecting a number of dedicated blocks for modelling eye/head movements for the simulation of oculomotor systems. The library is focused on the the bioinspired models investigated throughout the EYESHOTS project, but includes also models of conventional pan-tilt systems for comparison or specific analysis.
The library contains also the modules required to set-up the control architecture of the actual robot prototype as well as camera image acquisition blocks to capture video sequences captured by the cameras on board the robot eye.
Extension of the library to include interface blocks with a VR environment for complete simulation of oculo-motor control and vision has been partially pursued during the EYESHOTS project. Additional work is required to complete this activity.

(**) Design drawings and CAD models of the robot eye prototype
Further research needed: The bioinspired robot prototype is completed. However, the electronics and the actuation system are (as planned) based on commercial components. The specific control models developed during the project suggest the implementation of *ad-hoc* electronics and sensing to complete the while system implementation and testing. New actuation technologies (e.g. piezo-electric motors) have been recently demonstrated. Despite their very high cost appear to be a promising solution to overcome some of the limitations posed by the linear brushless direct drive servo-motors currently used. The modular design of the robot mechanics should make possible the upgrade of the system .

Potential impact has to be evaluated.
The results of the project were originally (i.e. at the beginning of the project) not oriented towards direct industrial or commercial exploitation. The project had a strong scientific and methodological aim. Therefore, the direct exploitation of the results can be foreseen mostly within the Scientific Community.

3 REPORT ON SOCIETAL IMPLICATIONS

3.1

A General Information *(completed automatically when **Grant Agreement number** is entered.*

Grant Agreement Number:	217077
Title of Project:	Heterogeneous 3-D perception across visual fragments
Name and Title of Coordinator:	Silvio P. Sabatini, Prof.

B Ethics

1. Did your project undergo an Ethics Review (and/or Screening)? • If Yes: have you described the progress of compliance with the relevant Ethics Review/Screening Requirements in the frame of the periodic/final project reports? Special Reminder: the progress of compliance with the Ethics Review/Screening Requirements should be described in the Period/Final Project Reports under the Section 3.2.2 *'Work Progress and Achievements'*	***No***
2. Please indicate whether your project involved any of the following issues (tick box) :	***YES***
RESEARCH ON HUMANS	
• Did the project involve children?	
• Did the project involve patients?	
• Did the project involve persons not able to give consent?	
• Did the project involve adult healthy volunteers?	X
• Did the project involve Human genetic material?	
• Did the project involve Human biological samples?	
• Did the project involve Human data collection?	
RESEARCH ON HUMAN EMBRYO/FOETUS	
• Did the project involve Human Embryos?	
• Did the project involve Human Foetal Tissue / Cells?	
• Did the project involve Human Embryonic Stem Cells (hESCs)?	
• Did the project on human Embryonic Stem Cells involve cells in culture?	
• Did the project on human Embryonic Stem Cells involve the derivation of cells from Embryos?	
PRIVACY	
• Did the project involve processing of genetic information or personal data (eg. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)?	
• Did the project involve tracking the location or observation of people?	
RESEARCH ON ANIMALS	
• Did the project involve research on animals?	X
• Were those animals transgenic small laboratory animals?	
• Were those animals transgenic farm animals?	
• Were those animals cloned farm animals?	
• Were those animals non-human primates?	X
RESEARCH INVOLVING DEVELOPING COUNTRIES	
• Did the project involve the use of local resources (genetic, animal, plant etc)?	
• Was the project of benefit to local community (capacity building, access to healthcare, education etc)?	

DUAL USE	
• Research having direct military use	No
• Research having the potential for terrorist abuse	No

C Workforce Statistics

3. Workforce statistics for the project: Please indicate in the table below the number of people who worked on the project (on a headcount basis).

Type of Position	**Number of Women**	**Number of Men**
Scientific Coordinator	-	1
Work package leaders	1	5
Experienced researchers (i.e. PhD holders)	4	8
PhD Students	6	5
Other	-	1

4. How many additional researchers (in companies and universities) were recruited specifically for this project?	**1**
Of which, indicate the number of men:	1

D Gender Aspects

5. Did you carry out specific Gender Equality Actions under the project?	○ X	Yes No

6. Which of the following actions did you carry out and how effective were they?

		Not at all effective				Very effective
☐	Design and implement an equal opportunity policy	○	○	○	○	○
☐	Set targets to achieve a gender balance in the workforce	○	○	○	○	○
☐	Organise conferences and workshops on gender	○	○	○	○	○
☐	Actions to improve work-life balance	○	○	○	○	○
○	Other:					

7. Was there a gender dimension associated with the research content – i.e. wherever people were the focus of the research as, for example, consumers, users, patients or in trials, was the issue of gender considered and addressed?

○ Yes- please specify ☐

X No

E Synergies with Science Education

8. Did your project involve working with students and/or school pupils (e.g. open days, participation in science festivals and events, prizes/competitions or joint projects)?

○ Yes- please specify ☐

X No

9. Did the project generate any science education material (e.g. kits, websites, explanatory booklets, DVDs)?

X Yes- please specify: Slide shows and videos of the lectures of the summer school in Benicassim

○ No

F Interdisciplinarity

10. Which disciplines (see list below) are involved in your project?

○ Main discipline[7]: 1,1

○ Associated discipline[7]: 2.2 | ○ Associated discipline[7]: 3.1, 5.1

G Engaging with Civil society and policy makers

11a Did your project engage with societal actors beyond the research community? *(if 'No', go to Question 14)*	○ X	Yes No

11b If yes, did you engage with citizens (citizens' panels / juries) or organised civil society (NGOs, patients' groups etc.)?

○ No

○ Yes- in determining what research should be performed

○ Yes - in implementing the research

○ Yes, in communicating /disseminating / using the results of the project

---

[7] Insert number from list below (Frascati Manual).

11c In doing so, did your project involve actors whose role is mainly to organise the dialogue with citizens and organised civil society (e.g. professional mediator; communication company, science museums)?	○ ○	Yes No

12. Did you engage with government / public bodies or policy makers (including international organisations)

○ No
○ Yes- in framing the research agenda
○ Yes - in implementing the research agenda
○ Yes, in communicating /disseminating / using the results of the project

13a Will the project generate outputs (expertise or scientific advice) which could be used by policy makers?

○ Yes – as a **primary** objective (please indicate areas below- multiple answers possible)
○ Yes – as a **secondary** objective (please indicate areas below - multiple answer possible)
○ No

13b If Yes, in which fields?

Agriculture
Audiovisual and Media
Budget
Competition
Consumers
Culture
Customs
Development Economic and Monetary Affairs
Education, Training, Youth
Employment and Social Affairs

Energy
Enlargement
Enterprise
Environment
External Relations
External Trade
Fisheries and Maritime Affairs
Food Safety
Foreign and Security Policy
Fraud
Humanitarian aid

Human rights
Information Society
Institutional affairs
Internal Market
Justice, freedom and security
Public Health
Regional Policy
Research and Innovation
Space
Taxation
Transport

13c If Yes, at which level?

○ Local / regional levels
○ National level
○ European level
○ International level

H Use and dissemination

14. How many Articles were published/accepted for publication in peer-reviewed journals?	**27**
To how many of these is open access[8] provided?	**27**
How many of these are published in open access journals?	**18**
How many of these are published in open repositories?	**1**
To how many of these is open access not provided?	
Please check all applicable reasons for not providing open access:	
☐ publisher's licensing agreement would not permit publishing in a repository ☐ no suitable repository available ☐ no suitable open access journal available ☐ no funds available to publish in an open access journal ☐ lack of time and resources ☐ lack of information on open access ☐ other[9]: ……………	

15. How many new patent applications ('priority filings') have been made? *("Technologically unique": multiple applications for the same invention in different jurisdictions should be counted as just one application of grant).*	-

16. Indicate how many of the following Intellectual Property Rights were applied for (give number in each box).	Trademark	-
	Registered design	-
	Other	-

17. How many spin-off companies were created / are planned as a direct result of the project?	-
Indicate the approximate number of additional jobs in these companies:	

18. Please indicate whether your project has a potential impact on employment, in comparison with the situation before your project:

☐ Increase in employment, or	☐ In small & medium-sized enterprises
☐ Safeguard employment, or	☐ In large companies
☐ Decrease in employment,	☐ None of the above / not relevant to the project
X Difficult to estimate / not possible to quantify	

19. For your project partnership please estimate the employment effect resulting directly from your participation in Full Time Equivalent (*FTE = one person working fulltime for a year*) jobs:	*Indicate figure:* *11*

---

[8] Open Access is defined as free of charge access for anyone via Internet.
[9] For instance: classification for security project.

Difficult to estimate / not possible to quantify	☐

I Media and Communication to the general public

20. As part of the project, were any of the beneficiaries professionals in communication or media relations?

○ Yes X No

21. As part of the project, have any beneficiaries received professional media / communication training / advice to improve communication with the general public?

○ Yes X No

22 Which of the following have been used to communicate information about your project to the general public, or have resulted from your project?

☐ Press Release
☐ Media briefing
☐ TV coverage / report
☐ Radio coverage / report
X Brochures /posters / flyers
☐ DVD /Film /Multimedia

☐ Coverage in specialist press
☐ Coverage in general (non-specialist) press
☐ Coverage in national press
☐ Coverage in international press
X Website for the general public / internet
☐ Event targeting general public (festival, conference, exhibition, science café)

23 In which languages are the information products for the general public produced?

☐ Language of the coordinator
☐ Other language(s)

X English

Question F-10: Classification of Scientific Disciplines according to the Frascati Manual 2002 (Proposed Standard Practice for Surveys on Research and Experimental Development, OECD 2002):

FIELDS OF SCIENCE AND TECHNOLOGY

1. NATURAL SCIENCES
1.1 Mathematics and computer sciences [mathematics and other allied fields: computer sciences and other allied subjects (software development only; hardware development should be classified in the engineering fields)]
1.2 Physical sciences (astronomy and space sciences, physics and other allied subjects)
1.3 Chemical sciences (chemistry, other allied subjects)
1.4 Earth and related environmental sciences (geology, geophysics, mineralogy, physical geography and other geosciences, meteorology and other atmospheric sciences including climatic research, oceanography, vulcanology, palaeoecology, other allied sciences)
1.5 Biological sciences (biology, botany, bacteriology, microbiology, zoology, entomology, genetics, biochemistry, biophysics, other allied sciences, excluding clinical and veterinary sciences)

2 ENGINEERING AND TECHNOLOGY
2.1 Civil engineering (architecture engineering, building science and engineering, construction engineering, municipal and structural engineering and other allied subjects)
2.2 Electrical engineering, electronics [electrical engineering, electronics, communication engineering and systems, computer engineering (hardware only) and other allied subjects]
2.3. Other engineering sciences (such as chemical, aeronautical and space, mechanical, metallurgical and materials engineering, and their specialised subdivisions; forest products; applied sciences such as geodesy, industrial

chemistry, etc.; the science and technology of food production; specialised technologies of interdisciplinary fields, e.g. systems analysis, metallurgy, mining, textile technology and other applied subjects)

3. MEDICAL SCIENCES

3.1 Basic medicine (anatomy, cytology, physiology, genetics, pharmacy, pharmacology, toxicology, immunology and immunohaematology, clinical chemistry, clinical microbiology, pathology)
3.2 Clinical medicine (anaesthesiology, paediatrics, obstetrics and gynaecology, internal medicine, surgery, dentistry, neurology, psychiatry, radiology, therapeutics, otorhinolaryngology, ophthalmology)
3.3 Health sciences (public health services, social medicine, hygiene, nursing, epidemiology)

4. AGRICULTURAL SCIENCES

4.1 Agriculture, forestry, fisheries and allied sciences (agronomy, animal husbandry, fisheries, forestry, horticulture, other allied subjects)
4.2 Veterinary medicine

5. SOCIAL SCIENCES

5.1 Psychology
5.2 Economics
5.3 Educational sciences (education and training and other allied subjects)
5.4 Other social sciences [anthropology (social and cultural) and ethnology, demography, geography (human, economic and social), town and country planning, management, law, linguistics, political sciences, sociology, organisation and methods, miscellaneous social sciences and interdisciplinary , methodological and historical S1T activities relating to subjects in this group. Physical anthropology, physical geography and psychophysiology should normally be classified with the natural sciences].

6. HUMANITIES

6.1 History (history, prehistory and history, together with auxiliary historical disciplines such as archaeology, numismatics, palaeography, genealogy, etc.)
6.2 Languages and literature (ancient and modern)
6.3 Other humanities [philosophy (including the history of science and technology) arts, history of art, art criticism, painting, sculpture, musicology, dramatic art excluding artistic "research" of any kind, religion, theology, other fields and subjects pertaining to the humanities, methodological, historical and other S1T activities relating to the subjects in this group]