PROJECT PERIODIC REPORT

Grant Agreement number: 217077

Project acronym: EYESHOTS

Project title: Heterogeneous 3-D Perception across Visual Fragments

Funding Scheme: Collaborative project

Date of latest version of Annex I against which the assessment will be made: 05/10/2007

Periodic report: 1st ☒ 2nd ☐ 3rd ☐

Period covered: from 01/03/2008 to 28/02/2009

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

Silvio P. Sabatini, Dr.

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

Declaration by the scientific representative of the project coordinator

I, as scientific representative of the coordinator of this project and in line with the obligations as stated in Article II.2.3 of the Grant Agreement declare that:

- The attached periodic report represents an accurate description of the work carried out in this project for this reporting period;
- The project (tick as appropriate):

  ☒ has fully achieved its objectives and technical goals for the period;

  ☐ has achieved most of its objectives and technical goals for the period with relatively minor deviations[1];

  ☐ has failed to achieve critical objectives and/or is not at all on schedule[2].

- The public website is up to date, if applicable.
- To my best knowledge, the financial statements which are being submitted as part of this report are in line with the actual work carried out and are consistent with the report on the resources used for the project (section 6) and if applicable with the certificate on financial statement.
- All beneficiaries, in particular non-profit public bodies, secondary and higher education establishments, research organisations and SMEs, have declared to have verified their legal status. Any changes have been reported under section 5 (Project Management) in accordance with Article II.3.f of the Grant Agreement.

Name of scientific representative of the Coordinator: Silvio Paolo Sabatini

Date: 06 April 2009

Signature of scientific representative of the Coordinator: ....

1 If either of these boxes is ticked, the report should reflect these and any remedial actions taken.

2 If either of these boxes is ticked, the report should reflect these and any remedial actions taken.

1. Publishable summary

EYESHOTS is a Collaborative Project funded by European Commission through its Cognitive Systems, Interaction, Robotics Unit (E5) under the Information and Communication Technologies component of the Seventh Framework Programme (FP7). The project was launched on the 1st of March 2008 and will run for a total of 36 months. 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)

which provide different expertise ranging from robotics, computer vision, neuroscience and experimental psychology.

1.1 Project's goal

The goal of EYESHOTS is to investigate the interplay existing between vision and motion control, and to study how to exploit this interaction to achieve a knowledge of the surrounding environment that allows a robot to act properly. Robot perception can be flexibly integrated with its own actions and the understanding of planned actions of humans in a shared workspace. The research relies upon the assumption that a complete and operative cognition of visual space can be achieved only through active exploration of it: the natural effectors of this cognition are the eyes and the arms.

Crucial but yet unsolved addressed issues are object recognition, dynamic shifts of attention, 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 avoids a central representation of the environment and rather uses specialized components that interact with each other and tune themselves on the task at hand.
In addition to a high standard in engineering solutions the development and application of novel learning rules enable the system to acquire the necessary information directly from the environment.

1.2 Specific objectives

The project aims to reach the following three specific objectives:

Objective 1: Development of a robotic system for interactive visual stereopsis. The function of the systems is to interactively explore the 3D space by active foveations. Benefits of the motor side of depth vision are expected to be bi-directional by learning optimal sensorimotor interactions.
Objective 2: Development of a model of a multisensory egocentric representation of the 3D space. The representation is constructed on (1) binocular visual cues, (2) signals from the oculomotor systems, (3) signals about reaching movements performed by the arm. Egocentric representations require regular updating as the robot changes its fixation point. Rather than continuously updating based on motor cues or a visual mechanism (i.e. optic flow), the model updates only the egocentric relationship and object-to-object relationships of those objects currently in the field of view. During motion, the model covertly and overtly shifts attention to objects in the environment to maintain the model's current awareness of the environment. The updating of the internal representation of spatial relations requires binding processes across the different visual fragments.
Objective 3: Development of a model of human-robot cooperative actions in a shared workspace. By the mechanism of shared attention the robot will be able to track a human partner's overt attention and predict and react to the partner's actions. This will be extremely helpful in cooperative interactions between the robot and a human.

1.3 Expected final results

By the end of the three years the following results will be achieved:

- Implementing strong "dynamic" and "pro-active" components in which the effect of eye movements and of arm reaching actions will express as joint visuo-motor features, patterns and relationships for a perceptual awareness of space;
- Building a contingent knowledge of the sensorimotor laws that govern the relation between possible actions and the resulting changes in incoming visual information.
- Binding of objects into a global workspace for cognitive task control.

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. From this perspective, we have been contacted by the international organization e-ISOTIS (Information Society Open To ImpairmentS, www.e-isotis.org), established and evolved with the scope to support the people with disabilities and elderly to overcome the existing barriers and have an independent living and quality of life, which is interested in the results of our project.

1.4 Work performed and main results achieved in the reporting period (01/03/08 – 28/02/09)

The diagram in Fig.1 illustrates the different project's components and their integration. The large greyed box represents the *EYESHOTS' Agent* engaged in the perception action cycle: The information flows circularly from the environment to sensory structures, to motor structures, back again to the environment, to sensory structures, and so on, during the processing/accomplishment of goal-directed behaviour. Robot perception is flexibly integrated with its own actions and with the understanding of planned actions of humans in a shared workspace. The study and models of human/primate behaviour, also based on specific experimental activities, provided the architectural guidelines for the organization of perceptual behaviour, steering many of our envisaged solutions. The stereo vision platform will be ultimately substituted by an anthropomorphic mechatronic system that emulates eye kinematics and actuation of the human eye.

During the first year, the project tackled the research issues related to (1) the development of the active vision system and the modelling of its behaviour (partners UG, K.U.Leuven, and WWU), and (2) the definition of models for action-oriented representation of reachable objects (partners UJI, UNIBO, and WWU). In doing so, we have, for the moment, considered the eyes and the arm as separate effectors. In perspective, the two subsystems will interact each other in order that the arm system will help the calibration of the vision system, and *vice versa*. Concerning the sensorimotor representation of the 3D space, specific experiments have been conceived by partners UNIBO, and preliminary analysis of the motor description of the visual fragments have been conducted by partner WWU, adopting the paradigm of saccade adaptation.

Essentially, the project is on schedule and, with respect to the objectives planned for the first period, the main results achieved are the following:

1. Analysis of the ideal problem of moving the vergence point over a smooth surface, from the kinematic control point of view. The requirements for its implementation (in particular in terms of feedback from vision modules) are derived.

2. A Virtual Reality simulator for binocular active vision system.

3. Assessment of the degree of adaptability of a population network of disparity detectors. Specifically, distinct specialization of the disparity detectors for vergence and depth perception have been analyzed and employed for vergence models.

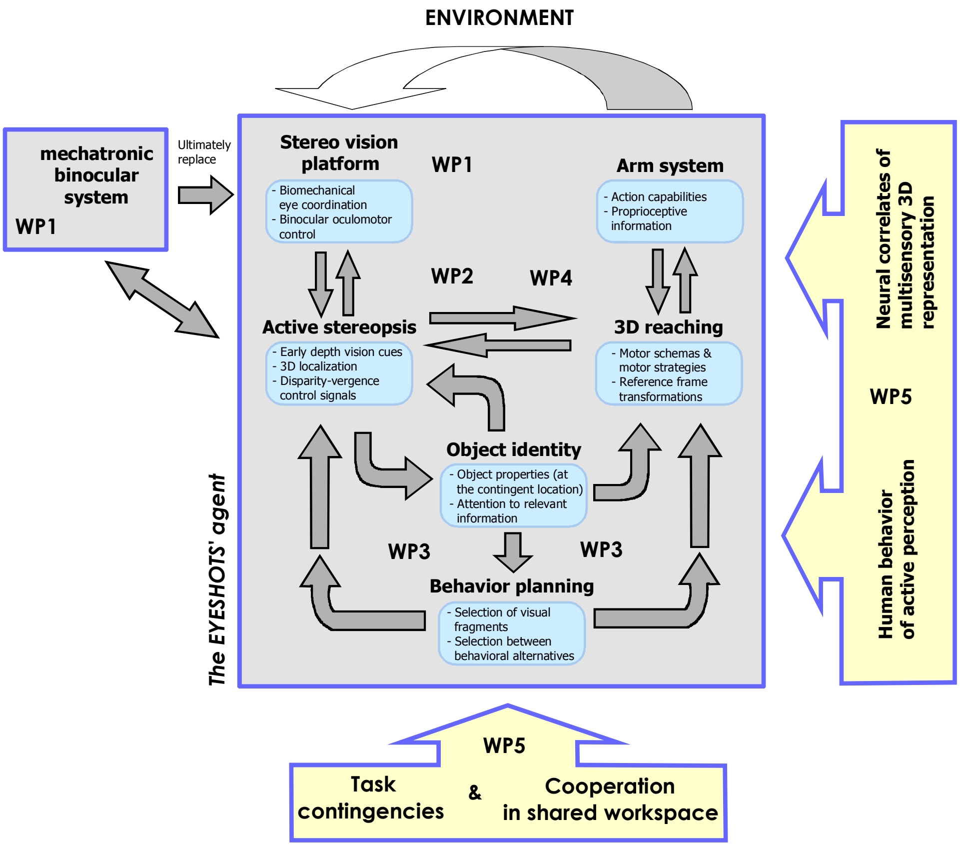


Figure 1: *The EYESHOTS project components and their integration.*

4. A method for iteratively improving the gaze estimate by simultaneously updating disparity and geometry estimates, thus compensating for motor vergence/version errors.

5. Learning of bidirectionally connected disparity tuned receptive fields.

6. The experimental set-up for neurophysiological experiments has been designed and implemented, including the device for reach-in-depth and it is in use for monkey training and neural recording.

7. Experimental evaluation of the motor contribution on the perceived location of objects via saccade adaptation

8. A conceptual framework for the modeling of ventral/dorsal interactions in reaching and grasping actions. Activities in collaboration with experimentalists (partners UNIBO and WWU) have been planned.

2 Project objectives for the period

2.1 Overview

The research activities planned in the EYESHOTS' project are grounded on three key concepts:

Embodiment – A complete and operative cognition of visual space needs a body, which poses structural constraints and allows active exploration (the natural effectors of this cognition are the eyes and the arms).

Fragmented vision – Every fixation provides a new piece of local information (fragment) separate across saccades. The subjective richness of visual representation of the world at any time is an illusion. Active fragmented vision represents a dynamic cognitive interpretation of the scene, which does not imply a real metrical 3D reconstruction, but instead a loose representation of objects that are actively bound on time for the task at hand.

Visuomotor descriptors of space – Motor signals (used for saccade execution and/or for reaching the fragment by an arm movement) are used for establishing awareness of the location of the sequence of visual fragments in space.

From this perspective, the project embraces four aspects:

(1) biomimetic control of eye movements,
(2) interactive visual stereopsis,
(3) selection and integration of visual information,
(4) visuomotor integration of perception.

The Work Program consists of 8 Work-packages (WPs). There will be five scientific and technological WPs (WP1-5), and three WPs, planned for: training (WP8), dissemination and exploitation of the project's results (WP7), and for general project coordination and management (WP6).

The workplan is organized to allow the *concurrent* development of these activities.
For each Workpackage we provide, from the Annex I of the GA, a synthesis of the objectives of the relates tasks for the reporting period.

WP1: Eye movements for exploration of the 3D space

(1) Study of the ocular mechanics and oculomotor control, for both single eye and conjugate movements.
(2) Study of the geometric and kinematic effects of eye movements on image flow for supporting the estimation of 3D information.

Task 1.1: *Binocular eye coordination* – Study of the geometric and kinematic effects of ocular movements on depth vision in the case of single eye, and conjugate binocular physiological motions. To this aim, also other phenomena, observed in humans and primates, as the dynamic orientation of Listing's plane will be considered.
Task 1.2: *Perceptual influences of non-visual cues* – Direct and indirect perceptual consequences of eye movements. First, we will derive the relationships between the version and vergence angles and the locations of the fixated point in the 3D space, as well maps of binocular correspondences (horizontal and vertical retinal disparities) around the fixation point. Second, we will introduce specific mechanisms (e.g., gain fields) to modulate the responses of the disparity detectors of the visual system on the basis of the vergence and version signals.
Task 1.3: *Control of voluntary eye movements* – Study of the interplay existing between the mechanics of the eye plant and the strategies implemented by the brain to drive typical biological ocular motions. The goal is to understand the control mechanisms adopted by the brain to coordinate the action of the extra-ocular muscles for the single eye and for conjugate ocular movements.

WP2: Active stereopsis

(1) Specialization of disparity detectors at different levels in a hierarchical network to see the effect of learning for the extraction of the binocular features.

(2) To learn a vergence motor strategy that optimizes the quality and efficiency of the feature-extraction for the specific task to be accomplished.
(3) To develop scanning strategies that accurately describe the head-centric disparity of a visual fragment so that it can be processed by precise, near-tuned disparity detectors.

Task 2.1: *Network paradigm for intelligent vergence control (reflex-like)* – To develop a convolutional network-based vergence control from a population of disparity-based detectors. As a first step, the network is trained off-line, using a set of binocular images obtained from an anthropomorphic robot eye system available by UG. In this phase, eye movements are simulated.
Task 2.2: *Interactive depth perception* – To develop a mechanism that renders the disparity-to-depth transformation robust to the limited accuracy of the motor system. Specifically, the mapping from disparity and eye position to 3D depth will be made insensitive to the effects of small vergence/version errors.

WP3: Selecting and binding visual fragments
(1) To derive information about object identity from a hierarchical representation of learned features.
(2) To learn distributed representations that actively bind and represent visual fragments for the task at hand. Reward-based learning approaches will be adopted.

Task 3.1: *Defining visual fragment: object identity* – Development of learning algorithms that allow us to encode object properties. The learning process starts from models of early visual areas, but it will then be applied at different levels of a hierarchical network to develop tuned receptive fields with increasing selectivity to complex features and disparity cues. As a first step, the network is trained off-line, using a set of binocular images obtained from an anthropomorphic robot eye system available by UG.
Task 3.2: *Selecting visual fragment* – Development of dynamic goal-directed attentional selection to bind object properties, on the basis of the momentarily existing task.
Task 3.3: *Selecting between behavioral alternatives* – We address the problem of learning the cognitive control of visual perception, in forms of visual-visual and visual-reward associations.

WP4: Sensorimotor integration
(1) Definition of an action-perception integrated representation of objects in the peripersonal space in a dynamic way.
(2) Generation of such representation through the interaction of the robotic system with the environment.

Task 4.1: *Merging perception-related and action-related visual information* – Define 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.
Task 4.2: *Generating visuo-motor descriptors of reachable objects* – Describe a perceptual framework for the purposeful building of the representation of Task 4.1 through active exploration of the peripersonal space. Visuomotor descriptors based on modelling of cortical functions, mainly from the parietal cortex.

WP5: Human behaviour and neural correlates of multisensory 3D representation
Definition and execution of specifically-designed psychophysical and neurophysiological experiments. The experiments are intended to provide architectural guidelines for the organization of perceptual interactions and will guide the production of artificial systems able to explore and interact with the 3D world. The psychophysical experiments will provide behavioral patterns (I/O specifications), while in-vivo experiments will provide architectural solutions (I/O + internal structural data).

Task 5.1: *Role of visual and oculomotor cues in the perception of 3D space* –To verify the role of non-visual and visual cues in the perception of the 3D peripersonal space in the medial parieto-occipital cortex.
Task 5.2: *Link across fragments* – To experimentally determine neural correlates of multisensory representation of 3D space obtained through active ocular and arm movements.
Task 5.3: *Motor description of fragment location* – To experimentally determine motor contributions of eye movements to fragment location via saccade adaptation.
Task 5.4: *Predicting behavior and cooperation in shared workspace* – To study specific aspects of human behavior in the combination of allocation of attention and direction of gaze that can be used for prediction in human-robot interaction.

WP6: Project coordination and management

To implement and maintain an effective administrative and management infrastructure of the project, including:

(1) Development and maintenance of the EYESHOTS project web-site (with a public section and a private section with restricted access to the consortium's members).
(2) Set-up and maintenance of e-Services for sharing and broadcasting documents and data.

WP7: Knowledge management, dissemination and use, synergies with other projects

To make the project results known to the Community of interested researchers and automation industry as one of the potential developers of next-generation robotic systems.

Task 7.1: Regular publication of research news, events, research results, and demos on the project website.
Task 7.2: Organization of brainstorms events – as open workshop sessions – where, in particular, young researchers are invited to present their current work related to the project.
Task 7.3: Journal publications, participation to workshops, conferences, and other forum and events. Set-up of a mailing-list that will be used instead of a Newsletter to disseminate results to interested parties.
Task 7.4: Synergies with other FP6 or FP7 projects.

WP8: Training, education and mobility

(1) To make local education, training activities and knowledge of the partners accessible for the entire consortium.
(2) To foster the exchange of personnel and to promote collaboration at every level of the consortium.

Task 8.1: To compile an extensive bibliography list and source/access information of the basic and relevant literature from computer science, biocybernetics, sensing and motor control as well as learning that will provide a common basis for teaching and education of students.
Task 8.2: Student's half-yearly seminars.
Task 8.3: Medium- and long-term visiting periods by young researchers and short-term visits of principal investigators.
Task 8.4: Organization of a Summer School.

As a whole, for what concerns the main S&T issues, the project's objectives for the reporting period were:

1. The development of models and architectures for the oculomotor control and the active stereopsis (vergence and depth vision), cf. Tasks 1.1, 2.1, 2.2, and 3.1 (UG, K.U.Leuven, WWU).
2. The definition of a conceptual framework (UJI) for the sensorimotor representation of the 3D space including (i) a planning of specific neurophysiological experiments to be conducted by partner UNIBO, and (ii) the interpretation of saccade adaptation psychophysical experiments, performed by partner WWU, in terms of motor descriptors of visual fragments.

For what concerns the other, management, issues the project's objectives were:

1. The launch of the project web-site, its maintenance and continuous update.
2. The creation of the Literature Database.

All of these objectives have been achieved.

Concerning the detail of the individual objectives, they are well documented in the individual workpackages sections and summarized in section 3.2.

2.2 Follow-up of previous review

Not applicable

3 Work progress and achievements during the period

Note: a "code" and a "number" label the publications of the Consortium in the reporting period: the codes 'P' and 'C' refer to journal paper and conference contribution, respectively. For the full list of publications please see section 5.2.

3.1 Progress overview and contribution to the research field

With reference to the overall architecture of the EYESHOTS' agent, Fig.2 depicts the status of the project. The research activities that have characterized the first reporting period are highlighted in yellow and they can be grouped in three parts (basically related to the three main specific objectives of the project):

- Part 1: Disparity for vergence and depth vision (WP1, WP2, WP3)
- Part 2: Sensorimotor representation of the 3D space (WP4, WP5)
- Part 3: Cooperative actions in a shared workspace (WP4, WP5).

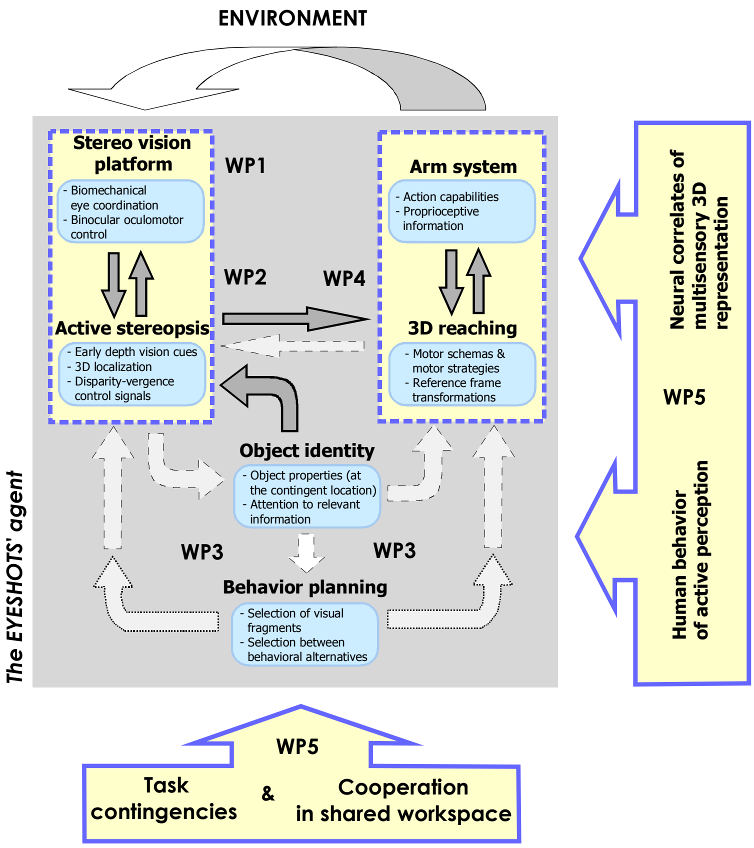


Figure 2: *Status of the project at the end of the 1st period. The research activities that have characterized the period are highlighted.*

The progress of work for each part can be summarized as follows:

Part 1: Disparity for vergence and depth vision [WP1, WP2, WP3]

- ***Binocular eye coordination***

  The research activity concerned the complex task of keeping the two eyes in alignment. This is an important function of the oculomotor system since a poor alignment produces large retinal disparities and a degradation of stereopsis. Then, it is also desirable to keep the lines of sight of the two eyes converged on an object or a surface of interest in order to increase the accuracy of the perception of all its properties.

  The work performed by partner UG focused mostly on the investigation of coordinated ocular motion strategies that succeed in replicating these behaviors. The goal has been that of designing a control mechanism to keep and to move the fixation point over a smooth surface (provided suitable visual feedback) within the peripersonal space.

  We restricted the analysis to the horizontal and vertical dimension, only, without considering the possibility of rotating around the direction of sight. This was object of investigation of the second part of the work. In fact, it is well established that oculomotor system follows precise laws on near and far fixation and that eye movements are characterized by torsional components proportional to the vergence and that depend on the elevation of gaze (the so called Extended Listing's Law, or L2) (Tweed & Vilis, 1990; Mok et al 1992; Porril et al 1999). The analysis of the relative posture of the eyes under static fixation conditions has led to a comparative characterization of three different ocular kinematics, one that corresponds to a classical pan/tilt system, and two that replicate the behavior of the human eyes on near and far fixation.

  <u>Comparative assessment with the current state of the art:</u> The study of the perceptual consequences of Listing’s Law and its family of motor constraints has a long and rich history, dating back to Donders and von Helmholtz. Since the establishment of the validity of Listing’s Law in far vision, several studies on the binocular control of eye movements have focused on the deviations from Listing's law in near vision (Mok et al, 1992; Van Rijn & Van der Berg, 1993; Minken & Van Gisbergen, 1994). All these studies agree that eyes movements have a torsional component that varies with vergence and gaze elevation to reduce the cyclovergence and restricts the motion of the epipolar line, thus permitting stereo matching to work with smaller search zones (Schreiber et al, 2001). Then, it has been demonstrated recently that the control of ocular torsion can be changed by a cyclodisparity stimulus. This suggests a view where ocular torsions are dynamically controlled to optimize binocular image alignment and simplify the perception of slanted surface (Schor et al., 2001; Maxwell et al., 2001; Steffen et al., 2002; Misslisch et al., 2001). Notwithstanding the several lines of evidence of complex binocular motor control strategies of human eyes, their perceptual consequences still remain an open question considering that, in principle, the brain could solve the binocular visual correspondence problem by using 3D feedback signals for the orientation of both eyes (Crawford et al., 2003). The advantages of such strategies could be fully understood only if one jointly analyses and models the problem of neural computation of stereo information, and if one takes into account the limited accuracy of the motor system. Unfortunately, models in this joint field are very seldom (Theimer & Mallot, 1994; Read & Cumming, 2006; Hansard & Horaud, 2008) and rarely address all the computational issues. In absence of such models, so far in robot vision, rectification techniques simply remove the problem by searching for correspondences along the epipolar lines or disregarding vertical disparities[3], but removing, in this way, any cognitive value related to active 3D eye movements in purposive vision. Hence, the ‘computational principles’ pointed out by the analysis of the abstract problem conducted in the first period by UG should be and will be further developed concurrently with the models defined in WP2 (Task T2.1 and Task 2.2) in the next period, in collaboration with K.U.Leuven.

- ***Distributed architectures for stereo processing***

  By using phase-based computational paradigms, local binocular disparity values can be related to local phase differences in the stereo image pairs. On this basis, consolidated *direct* phase-based measure techniques imposed themselves (e.g., Sanger, 1988; Fleet et al., 1991), to which correspond

[3] Zero vertical disparity condition is far away from the reality under fixation of close points in the peripersonal space.

distributed coding approaches (Fleet et al., 1996; Chen & Qian, 2004) based on populations of binocular energy units (Ohzawa et al., 1990, 1997; Qian, 1994) in which the output from receptive fields in both eyes is linearly combined by V1 simple cells and this sum is then passed through an output nonlinearity. The equivalence between phase-based techniques and energy-based models has been formally demonstrated (Qian & Mikaelian, 2000). Yet, the latter prove themselves more robust to noise and more flexible, since they can intrinsically embed adaptive mechanisms both at coding and decoding levels of the population code. In particular, already at a static level (no temporal integration of the receptive fields), partner UG has (1) obtained forms of specialization of the disparity detectors (Gibaldi et al., 2009a [C7]; Gibaldi et al., 2009b [C1]), and (2) derived a parametric generalization of the network's units that takes into account the statistical mean disparity patterns, which are predictable by the position of the eyes. In this way, the distributed code of binocular disparity can embed fixation constraints at an architectural level (Chessa et al., 2009b [C2]). Collaborative work between UG and K.U.Leuven is considering how to adaptively tune the disparity detectors (phase-shifts + position-shifts) to achieve optimal vergence, disparity decoding and their interactions.
Extensions of the approach to include dynamic receptive field properties (motion and motion-in-depth sensitivity) will be tackled in the second period.

<u>Comparative assessment with the current state of the art</u>: Binocular energy units are now consolidated models of complex cells in area V1 as demonstrated by the numerous recent works that propose architectural variants to enrich their functionality or that adopt them to describe complex perceptual behaviors (Read, 2002; Read et al., 2002; Tanabe & Cumming, 2008; Bridgel & Cumming, 2008; Haefner & Cumming, 2008; Read & Cumming, 2006; Serrano-Pedraza & Read, 2009; Nishimoto et al., 2006; Sanada & Ohzawa, 2006; Miura et al., 2008).
Similarly, in the ICT community there are several examples of neuromorphic (i.e., distributed) approaches profitably used to challenge conventional solutions to computer vision problems, by introducing sophisticated interpretation of biologically plausible operations (e.g., Tsang & Shi, 2007; Tsang & Shi, in press; Bayerl & Neumann, 2007).
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 specific design approach followed to implement the distributed architecture developed by UG demonstrated that it is possible to implement a 'neuromorphic' solution that is characterized by an *affordable computational cost*, to be efficiently employed in closed-loop robotic applications. A pilot GPU-based implementation of the distributed architecture for the computation of 2D disparity, using the Nvidia CUDA Library yielded encouraging results.

- ***Learning adaptive behaviors through brain-style architectures***

  Most of conventional vergence control models (Horng et al., 1998; Theimer & Mallot 1994; Patel et al. 1996; Hung et al. 1986; Hung 1997; Krishnan 1977) are based on the minimization of the horizontal disparity. We propose to avoid implicit computation of the disparity map and extract the vergence control signal directly from the population responses. A neural network paradigm has been chosen for this type of conversion/extraction procedure. K.U.Leuven and UG plan to adopt an increasing complexity strategy in the learning process: starting from the simplest one-unit architecture we will increase the number of units/layers until an acceptable level of generalization error is reached. Similar to the disparity statistics computation, we are going to use a simulator to create training datasets. Each sample in the training dataset will contain stereo image, actual vergence angle, actual gaze orientation, and the desired (for this particular case) vergence angle ("ground truth"). Using this dataset it will be possible to train and evaluate the proposed neural network based vergence controller.

  <u>Comparative assessment with the current state of the art</u>: Recent progress in the area mainly focuses on finding correspondences between existing models and experimental data**.**
  The Patel vergence model (Patel et al., 2001) predicts that the static vergence error (fixation disparity) is the result of an asymmetric dynamic responsiveness of the disparity vergence mechanism in the convergent and divergent direction. This model is recently re-investigated by

Jaschinski and colleagues (Jaschinski et al., 2008). The study supports the above-mentioned hypothesis and confirms a relation between static vergence and asymmetric dynamic vergence, which both are idiosyncratic vergence parameters.

Recently, the influence of depth perception on vergence has regained interest, because of the possibility to probe perception through vergence. In several studies, various depth cues were used to induce depth and the influence of perception on vergence was studied both without and with binocular disparity present. The question whether depth perception *per se* contributes to vergence has been investigated in the recent work of Wismeijer and colleagues (Wismeijer et al., 2008). They have shown that depth cues rather than perceived depth govern vergence that accompanies saccades. Moreover, in the study it has been suggested that monocular and binocular cues are weighted differently for perception and vergence. In (Tsang et al., 2008) an active stereo system with gaze and vergence control based on population responses of disparity neurons has been presented. The system uses but not simulates vergence by changing subwindow positions of left and right camera views. Although the model only requires a population of neurons in a single scale and sensitive to only three position disparities (near, fixation and far), the system demonstrated the tracking of targets in complex environment.

- ***Learning 3D visual descriptors for object identity and goal-directed attentive selection.***

  The goal is to learn receptive fields of the visual cortex by the statistics of the visual inputs using Hebbian learning principles and allow the visual information to be accessible by attentive selection for goal-directed action. This research direction extends the focus of WP2 to combine depth information with object identity.

  In the reporting period, partner WWU further revised their previous model of learning receptive fields in V1 (Hamker & Wiltschut, 2007) to improve the match of the spatial frequency tuning to V1 data. In this study they showed that more efficient codes improve the similarity of model data and experimental data and provided a novel learning algorithm (Wiltschut & Hamker, 2009 [J1]). Partner WWU further demonstrated that this algorithm can be adapted to lead to the development of cells with selective disparity tuning when stereo images are used as input material. Furthermore, the work started, slightly in advance, on a model of the Basal Ganglia for learning to select between behavioral alternatives within the context of a task by unspecific reward. This effort led to a review article (Vitay et al., in press [J6]). Preliminary results of the model have been published in a conference article (Vitay & Hamker, in press [C10]).

  <u>Comparative assessment with the current state of the art:</u> Goal-directed behavior requires a sensorimotor system to partially recognize relevant objects within its operating space allowing to direct attention to the particular object of interest. Attending a particular object then allows to dynamically bind object properties and to define the appropriate actions towards the object. Changes in the direction of gaze are achieved by saccades, and shifts of the depth of fixation by vergence. With respect to joint vergence control and saccades, depth information of the object of interest is already gained before the eye fixates the target object to generate the appropriate vergence command (Chaturvedi & Van Gisbergen, 1998). There is considerable overlap between the object recognition pathway and vergence control, e.g. by disparity information: V1 cells often show a tuning to one preferred absolute disparity (Fischer & Poggio, 1977). V2 neurons begin to show sensitivity to relative disparity and seem to be selective for disparity-defined edges (Qiu & von der Heydt, 2005). Neurons in V4 are sensitive to the orientation of disparity defined planes (Hinkle & Connor, 2005). So far, models of visual cortex encoding disparity have primarily be constructed by hand based on data (e.g. Read, 2004; Sabatini and Solari, 2004), but little work has been done on developing learning algorithms that lead to general purpose receptive fields similar as observed in the brain (but see Lippert et al., 2000; Lippert & Wagner, 2002).

Part 2: Sensorimotor representation of the 3D space

- For what concerns the neurophysiological experiments (***joint visuo-motor features in the parietal cortex***):

Since the starting date of EYESHOTS project, the activities of partner UNIBO have been focused on experimental set-up preparation and on monkey training. This required a huge effort and to this purpose several people have been hired.
At month 10 the experimental device for reach-in-depth task was ready, and the new binocular eye-tracking system was installed, tested and ready to be used for monkey training. This set up is essential to study the role of the medial parieto-occipital cortex in mastering the 3D peripersonal space.
At the same time, the surgery necessary for electrophysiological experiments was performed, and the training of the monkey for reach-in-depth and fix-in-depth tasks began. The monkey is now able to perform both tasks keeping a steady fixation of the target. The eye tracking system allows us to monitor the performance of the monkey in fixating the correct target, or in other words to differentiate the different vergence angles required in the tasks.
A pilot study has been already performed using a prototype of the fix-in-depth device. The neural activity of about 50 cells was recorded while the monkey was fixating targets placed at 5 different depths in front of him. Data will be analyzed in the next months.
Other analyses are being performed at the moment, in order to allow other partners to perform their tasks: in collaboration with partner UG, 2D gaze-dependent modulations in medial parieto-occipital cortex have been preliminarily analyzed. In collaboration with partner UJI, UNIBO is analyzing the role of visual and proprioceptive guidance of reaching movements in the medial parieto-occipital cortex. To study the first aspect, UNIBO selected the data from about one hundred neurons, which were tested with a fixation paradigm. With regard to the second aspect, data from 75 neurons tested with a delayed arm-reaching task both in darkness and in full light are available. These data have been object of joint publications with the other partners and the results will be presented in International Conferences (Breveglieri et al., 2009 [C6]; Chinellato et al., 2009a [C3]; Chinellato et al., 2009b [C4]).

<u>Comparative assessment with the current state of the art:</u> No major changes with respect to the state-of-the-art already presented in Annex I.

- For what concerns the psychological experiments (***construction of peripersonal space across eye movements***):

  Partner WWU have collected and analyzed data on the influence of motor and visual parameters on object localization obtained from saccade adaptation data. This data is needed for milestone M4 at month 12. The milestone was reached as expected. The hypothesis that saccade adaptation modifies perceived location of saccade goals was confirmed by the experiment. The experiments revealed differences between different conditions (reactive, stationary) and probe types (flashed vs. stationary targets) that show a clear influence of motor planning on visual localization. Partner WWU have also collected interesting data in a fixation condition. The results of the experiments are currently prepared for publication and will be reported in deliverable D5.3a at month 15. The research activity is on track. These and future experiments will help in clarifying modeling details for the representation of the reachable space of Task 4.3 (UJI). Fundamental issues are for example those about the factors that affect gazing strategy, such as the time available before and during the action, and the possible employment of peripheral visual information. The robot should finally be able to autonomously build the space representation according to such strategies, and suitably update it during its interaction with a human subject. Concerning the problem of how to adapt the experimental paradigm for the neurophysiological experiments on fragment location in monkeys, several discussions took place between partners UNIBO and WWU, which eventually led to a viable solution. The experiment will use a touch-screen set-up that will be constructed and programmed by the Münster group (WWU), and when ready implemented in Bologna to run the experiment. Data analysis will be done jointly between the UNIBO and WWU groups.

  <u>Comparative assessment with the current state of the art:</u> In the last year a number of new publications were released in the area of saccadic adaptation. An interesting new focus is hereby the difference between forward and backward adaptation and their functional connection to the emerging differences in saccadic trajectories (Catz et al., 2008; Golla et al. 2008). Computational

models including a cerebellar mechanism via a forward model and a “cortical mechanism” predict differences in saccade trajectories and target remapping (Ethier et al., 2008) and seem to connect well to psychophysical studies focusing on arm pointing to perceived targets (Hernandez et al.2008), to anti-saccades (Panouilleres et. al., 2009), or to visual localization (our study, see WP 5.3). Therefore, extensions of WP 5.3 to the area of forward and backward adaptation seem reasonable.

- From a modeling point of view (***integrated visuo-motor representation in reaching tasks***):
  Partner UJI have conceived a general conceptual framework to obtain a description of objects in the peripersonal space of a subject that includes two kinds of concepts, related to on-line, action-related features, and memorized, conceptual ones, respectively.
  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.
  To pursue the above goal, the UJI partner has developed a bibliographic review of the neuroscience findings related to the task of vision-based reaching and grasping. Neuroscience concepts were studied and interpreted in order to build a coherent and comprehensive model of the integration between the two kinds of visual data. 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. Most of the above has been summarized in deliverable D4.1. During the first year, UJI has been actively interacting with other partners of EYESHOTS. Single-cell data from the UNIBO partner are being analyzed and modeled by UJI with the goal of reproducing neuroscience theories on the robotic setup. This led to a joint publication [C3] [C4]. At the same time UJI members have been discussing with the colleagues from cognitive sciences about future experimental protocols useful for both groups (WP3). Then, a common experimental framework has been discussed with the researcher from WWU in order to coordinate research on bio-inspired robots and with human subjects (WP5). Finally, a UJI graduate student has been staying at the UG partner and collaborated to the integration goals of WP1 and WP2 with WP4. The next step for partner UJI, more directly related to Task 4.2, is to further develop and implement, first computationally and then on a robotic setup, some parts of the model outlined so far. Special focus will be initially put on the integration between stereoscopic retinal data with somatosensory information about object and arm state, in order to estimate object position and devise a reaching action plan as performed by area V6A in the dorsal stream.

  <u>Comparative assessment with the current state of the art:</u>
  The model we propose (detailed in Deliverable 4.1) extends the available framework of the functional task distribution between the two streams giving special attention to the dorsal stream subdivision, and proposing a novel approach to obtain an integrated representation.
  In perspective, for what concerns Task 4.2 and 4.3, (Thompson & Henriques, 2008) offer new insights on the process of reference frame transformation and spatial memory update during gazing and arm reaching movements. The study is aimed at checking whether remembered locations of pointing targets are updated following smooth-pursuit eye movements, as they are following saccades, and also investigated the role of visual information in estimating eye-movement amplitude for updating spatial memory. Results suggest that the oculomotor and arm-motor systems may rely on different sources of information for spatial updating. This is an important aspect to take into account when modeling the mutual modulation between gaze and arm movements in Tasks 4.2 and 4.3

Part 3: Cooperative actions in a shared workspace
Partner WWU have developed the setup and began data collection in a first experiment in which observers had to identify the target of a reaching movement by an actor. Observers had to look at the to be pointed object as soon as possible. Two factors were manipulated: the presence of gazing behavior and the visibility of the to be pointed objects. The gaze triggered a rapid and accurate response on the target object. Observers were able to identify the target objects when the arm was still at the beginning of its trajectory, with fewer saccades and more accurately when they had access to the gazing behavior of the actor. When gaze information was not available, the observers' gaze still leaded the hand movements of the actor, but was

comparatively slower in identifying the target object. The visibility of the target objects had an ameliorating effect on the spatial accuracy. These findings support the hypothesis that other's gaze direction is an essential predictive cue about the final location of a pointing movement. This is work towards milestone M9. Subgroup meeting between WWU and UJI took place for discussing issues of integration between human and robot setups.
Comparative assessment with the current state of the art: No major changes with respect to the state-of-the-art already presented in Annex I.

----- o -----

In general, there is a growing attention for moving the active vision perspective from systems in which just the effects of action influence the perception, to 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. The special issue of the Int. Journal of Humanoid Research on the "Active Vision of Humanoids" (Yiannis Aloimonos and Giulio Sandini, eds.), to appear in 2009, which followed the homonymous Workshop that happened in Pittsburgh, PS, USA in November 2007 during the Conference on Humanoid Robotics specifically addressed these challenging problems. "The peculiarity of humanoid vision stems from its purpose of supporting the action of an anthropomorphic body as opposed to 'just' being pattern recognition or image understanding. Thus, in some sense, the humanoid active vision constitutes an evolution from Active Vision of the 90's to the 'Action Vision' of the new millennium. Action Vision considers the motor system of the humanoid as an integral part of its perceptual machinery".

From this perspective, recent progress in vision research, computer vision, robotics and experimental psychology closely relates to sensorimotor paradigms at different levels of intergration, specifically suggesting advantages of a mutual calibration of the vision and the arm systems (Van Pelt & Medenport, 2008; Klier & Angelaki, 2008; Dankers et al., in press; Mossio & Taraborelli, 2008; Hansard & Horaud, 2008; Hamker et al., 2008; Georg & Lappe, 2009). See also the ESF-EMBO Symposium on "Three Dimensional Sensory and Motor Space: Perceptual Consequences of Motor Action", Sant Feliu de Guixols (Costa Brava), Spain 6-11 October 2007, http://www.fbw.vu.nl/~JSmeets/ESF_abstractbook.pdf .

Specifically, in stereoscopic vision research, attempts of assessing the disparity cues in natural settings during active vision tasks yield to several studies on the disparity statistics (Liu et al., 2008; Jansen et al., 2009). Though, systematic analyses of this issue in the peripersonal space (i.e., large vergence angles) and in a real-world situations (i.e., 3D exploration of a *real* enviroment) are still lacking.

Summarizing, the collaborative work of UG and K.U.Leuven settled down the conceptual bases for a flexible cortical-like architecture for both vergence control and depth perception that incorporate adaptive tuning mechanisms of the disparity detectors. Such adaptation will eventually depend on *priors*, which will be related both to the posture/positions of the eyes in the orbits and to the attentional signals (based on object properties) that might guide intentional exploration of the selected object. Both K.U.Leuven and WWU adopt learning paradigms for specialization of the disparity detectors in WP2 and WP3, respectively. Though, the learning processes have distinct and complementary objective functions: from one side, the generation of robust and efficient vergence control signals to support interactive stereopsis and, from the other side, the generation of specific (higher-order) visual descriptors to support object identity.

With respect to the interactive binocular exploration strategies, the work in WP1 has contributed, from a kinematic control point of view, to the definition a "computational theory" for the visually-based control of vergence on a smooth surface.

The active collaborations among the partners UJI, WWU and UNIBO in WP4 and WP5 have settled down the basis for the definition and the formulation of a contingent representation of the 3D workspace during active exploration.

The evolution of the model developed in WP3, by combining hierarchically the output of V1 receptive fields in V2 and then in V4, to determine object identity and any other "ventral" information useful for WP4 will be, in the next period, a link between the binocular vision strategies (fixation scans) and the action planning of reaching tasks, which will eventually jointly provide the sensorimotor (i.e., heterogeneous) 3D awareness of the peripersonal workspace.
An increment of the collaborative effort between partners UJI and WWU on the activities to be conducted in WP4 and WP5 is expected in the next period.

3.2 Workpackage progress

Here we recall the objectives for the tasks of the first period. Quick statements concerning the status are attached; the actual work performed will be detailed for each work package in the WP descriptions below.

WP1: Eye movements for exploration of the 3D space
Task 1.1: Binocular eye coordination
Analysis of binocular coordinated motion strategies for short distance performed, including relative ocular postures under static fixation conditions.
Scheduling: (month 1-12)
Planned and performed steps:

Step1: Co-ordinated motions of a couple of stereo cameras (computational point of view)

Performed actions: Work by UG. A study about how to specify a control model that allows us to coordinate the rotation of the cameras so that the vergence point is maintained most of the time over the surface of an object assumed within the peri-personal space of the robot.
Results: The ideal problem of moving the vergence point over a smooth surface in space is analysed from the kinematic control point of view. The requirements for its implementation (in particular in term of feedback requirements from vision modules) are derived.

Step2: Biological constraints on torsional ocular postures

Performed actions: UG analysed the properties of a particular class of static geometric postures that are known in the literature as Listing's Law configurations
Results: Under the assumption of static configurations, constraints on torsional ocular postures provide additional degrees of freedom on the eyes/cameras movements. Comparative analysis between different stereo-head systems: (i) a classical Tilt/Pan system, (ii) a system following Listing's Law, and (iii) a system that follows the binocular extension of Listing's Law (L2). A system implementing L2 has behaviour very close to that of a classical tilt/pan stereo-head.
The advantages of dynamically controlling ocular torsions to optimize binocular image alignment and thus simplify the perception of slanted surfaces are envisaged.

Status: successfully completed by month 12. Milestone M1 (low-level automatic servos based on primary disparity information) was reached as planned on month 6.
Documentation: Deliverable D1.1 and student report by Ester Martinez, on her stay in UG.
Publications: --
Revised planning: none

Task 1.2: Perceptual influences of non-visual cues.
The objective is to model eye position gain fields supporting head-centric depth perception.
Scheduling: (month 6-30)
Performed actions: UG started addressing the problem of the assessment of the functionalities of the existing gain field models and their suitability in the EYESHOTS framework. Preliminary analysis on gaze sensitivity in area V6A of the medial parieto-occipital cortex of the macaque performed, by cooperation between UNIBO and UG.
Results: A simple gain-field model for encoding head-centric depth information that uses the disparity tuning curves of the population network (developed in WP2) has been tested. We have verified that the peak-shaped

gaze fields reported by UNIBO in area V6A are not in contrast with the gain field models developed in the theoretical neuroscience literature.
<u>Status:</u> The work has started and proceeds as planned.
<u>Documentation:</u> --
<u>Publications:</u> Breveglieri et al., 2009 [C4]
<u>Revised planning:</u> none

Task 1.3: Control of voluntary eye movements in 3D.
The goal of the worktask is to model the action of the extraocular muscles to achieve correct ocular motions.
<u>Scheduling:</u> (month 6-30)
<u>Performed actions:</u> The work so far has focused on the modelling of the ocular mechanics in 2D and under extension to the 3D case.
<u>Status:</u> The work has started as planned at month 6.
<u>Documentation:</u> --
<u>Publications:</u> --
<u>Revised planning:</u> none

WP2: Active stereopsis
Task 2.1: Network paradigm for intelligent vergence control
The objective is to develop a convolutional network-based vergence control from a population of disparity-based feature detectors (cooperation between K.U.Leuven and UG).
<u>Scheduling:</u> (month 1-30)
<u>Planned and performed steps:</u>

<u>Step1: Network architecture design</u>

<u>Performed actions:</u> UG worked on the characterization of the population network and on its generalization. K.U.Leuven focussed on the extension of the LeNet convolutional network.

<u>Results:</u> (1) The degree of adaptability of the population network has been assessed; (2) A software module in MATLAB, which can be used for creation, training and testing of convolutional networks has been developed.

<u>Step2: Design adjustable binocular energy models to compensate the expected disparity field induced by eye movements</u>

<u>Performed actions:</u> K.U.Leuven generated a large set of virtual scenes with ground truth disparities at different fixation points and computed the statistics. UG computed the binocular energy and disparity with the population code and compared the results embedding fixation constraints.

<u>Results:</u> We have tested and verified the hypothesis that there is an expected disparity map as a function of image, given a stereo setup and its fixation point. This hypothesis is important for deriving a pre-emptive vergence control towards the new fixation point, prior to the eye movement.

<u>Step3: Vergence control</u>

<u>Performed actions:</u> The vergence is controlled using a convolutional network arranged in a closed loop. The vergence control signal is proportional to the disparity averaged/pulled over the “foveal” region (K.U.Leuven).

<u>Results:</u> The first version of the vergence simulator is ready. The simulator has been developed as a MATLAB-based platform for vergence control strategies development, modeling and testing.

<u>Step4: Specialization of disparity detectors for vergence control</u>

<u>Performed actions:</u> UG worked on a neural network paradigm to derive specialized disparity detectors for directly extracting disparity-vergence responses without explicit calculation of the disparity map.

<u>Results:</u> A model for reading out binocular energy population codes for short-latency disparity-vergence eye movements. Specific model’s characteristics: (1) wide working range with a reduced number of resource (single scale), (2) linear servos with fast reactions and precision.

<u>Status:</u> The work has started and proceeds as planned.

Documentation: Technical meeting notes by Manuela Chessa and Nicolay Chumerin (Leuven, 24-27 June, 2008).
Publications: Chessa et al., 2009b [C2], Gibaldi et al., 2009a [C7], Gibaldi et al., 2009b [C1].
Revised planning: Having verified the inaccuracy of measures obtained with an up-to-date robotic head (i-Cub, www.robotcub.org) available at UG, the training data sets (containing stereo image pairs, vergence angles and gaze directions) for learning disparity-vergence responses will be generated by the vergence simulator developed in Step3 and the VR tool developed in Task 1.1.

Task 2.2: Interactive depth perception.
The objective is to develop scanning strategies that accurately describe the headcentric disparity of a visual fragment so that it can be processed by precise, near-tuned disparity detectors.
Scheduling: (month 6-36)
Performed actions: K.U.Leuven concentrated in the first year on strategies how to perform disparity estimation and gaze refinement using inaccurate gaze measurements
Results: Method for iteratively improving the gaze estimate by simultaneously updating disparity and geometry estimates, thus compensating for motor vergence/version errors.
Using graphical processing units, K.U.Leuven has also obtained real-time performance on simultaneous multiscale phase-based disparity and motion estimation (35Hz @ 640x512).
Status: The work has started and proceeds as planned.
Documentation: --
Publications: --
Revised planning: none

WP3: Selecting and binding visual fragments
Task 3.1: Defining visual fragment: object identity
The objective is to model eye position gain fields supporting head-centric depth perception.
Scheduling: (month 1-24)
Performed actions: work by partner WWU. Development of an Hebbian network for learning multiple feature selective units.
Results: Learning of binocular receptive fields based on Hebbian learning. The network's cells learn localized, oriented, disparity tuned and bandpass filtering receptive fields, comparable to those in area V1 of the primate brain.
Status: The work has started and proceeds as planned. Milestone M3 (learning algorithm for bidirectionally …) was reached as planned on month 12.
Documentation: Deliverable D3.1a
Publications: Wiltschut and Hamker, 2009 [J1]
Revised planning: none

Task 3.2: Selecting visual fragment
Development of dynamic goal-directed attentional selection to bind object properties.
Scheduling: (month 7-30)
Performed actions: Models of receptive field (RF) dynamics before saccade onset. Differences from RF remapping.
Results: From the computational point of view, the model of receptive field dynamics that WWU is currently developing, is consistent with present observations and offers an interesting alternative compared to modelling attention as a simple spotlight.
Status: The work started as planned.
Documentation: --
Publications: --
Revised planning: none

Task 3.3: Selecting between behavioral alternatives
Learning of the cognitive control of visual perception.
Scheduling: (month 13-36)

Performed actions: work by WWU.
Results: Started the development of a model of Basal Ganglia for action selection and cognitive control. Preliminary model for memory retrieval property.
Status: Started earlier than planned, because WWU had already the capacities available.
Documentation: --
Publications: Vitay and Hamker, 2008 [C10]; Vitay et al., in press [J6].
Revised planning: none

WP4: Sensorimotor integration

Task 4.1: Merging perception-related and action-related visual information

The objective is to define 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.
Scheduling: (month 1-9)
Performed actions:

- Bibliographic review of types of visual representations, their features and interactions, with special attention to the dorsal and ventral visual cortical pathways.
- Functional study of brain areas and structures underlying those representations, mainly focusing on the dorsal substreams.

Results:

- Definition of a model framework including action-oriented and perception-oriented aspects.
- Advancement of hypothesis on the interactions between representations and the possible ways of building them.

Status: The work has started and completed as planned. Milestone M2 (merging of action and perception) was reached as planned on month 9.
Documentation: Deliverable 4.1
Publications: Chinellato et al. 2009a [C3]
Revised planning: none

Task 4.2: Generating visuo-motor descriptors of reachable objects

The objective is to describe a framework for the purposeful building of the representation of Task 4.1 through active exploration of the peripersonal space. Visuomotor descriptors are based on the modelling of cortical functions, mainly from the parietal cortex.
Scheduling: (month 7-30)
Performed actions:

- Study the integration between retinal and sensorimotor information in the posterior parietal cortex, especially regarding area V6A and reaching movements.
- Analysis of single-cell data aimed at modelling the processes that allows to modulate visuomotor action/perception processes: (i) statistical analysis of neural preferred gaze/reaching directions in different conditions and action epochs; (ii) principal component analysis of neural response in various conditions aimed at extracting potential basic neural representations for modelling.
- Study of the possible alternatives for active construction of a dynamic representation of close environment features (e.g. SOM, GNG, …).

Status: The work has started and proceeds as planned.
Documentation: --
Publications: Chinellato et al. 2009b [C4]
Revised planning: The interaction between partners UJI, WWU and UNIBO, and with WP5 suggests the convenience of increasing the research efforts preferentially toward dorsal mechanisms and structures and simplify the ventral aspects.

WP5: Human behavior and neural correlates of multisensory 3D representation

Task 5.1: Role of visual and oculomotor cues in the perception of 3D space.

The objective of this task is to verify the role of non-visual as well as visual cues in the perception of the 3D peripersonal space in the medial parieto-occipital cortex.

To this end, two preliminary activities are needed: laboratory set-up and monkey training.

Laboratory set-up:
Scheduling: (months 1-8)
Performed actions: A new video-based eye tracking system was purchased and installed in the lab. The system was set in a binocular configuration and it allows to monitor the performance of the monkey in fixating the correct target during task execution.
Results: Using a custom-made algorithm, eye position signals are used to calculate on-line version and vergence angles and monitor the performance of the monkey.
Status: Work started as planned at month 1. Milestone M6.ante (experimental set-up) was reached as planned on month 8.

Monkey training:
Scheduling: (months 9-12)
Performed actions: One macaque monkey has been successfully trained for the fix-in-depth task.
Results: At the end of the reporting period, the monkey is able to execute the task performing a saccadic eye movement toward the correct position and keeping a steady fixation of the target for 2-2.5 s.
Status: training ended.
Documentation: --
Publications: Breveglieri et al., 2009 [C6]
Revised planning: none

Task 5.2: Link across fragments.
The objective of this task is to experimentally determine neural correlates of multisensory representation of 3D space obtained through active ocular and arm movements.

To this end, two preliminary activities are needed: laboratory set-up and monkey training.

Laboratory set-up:
Scheduling: (months 1-8)
Performed actions: a reach-in-depth device was designed and manufactured.
Results: The device allows us to present the monkey with a set of fixation points and reaching targets placed at different distances, thus varying version/vergence eye signals and reach trajectory and distance, in order to assess the role of arm and oculomotor cues in the interaction with the 3D space.
Status: Work started as planned at month 1. Milestone M6.ante (experimental set-up) was reached as planned on month 8.

Monkey training:
Scheduling: (month 9-12)
Performed actions: The monkey training for reach-in-depth task was performed immediately after the training for the fixation-in-depth task. It ended successfully.
Results: Electrophysiological recording sessions will start at the beginning of the second reporting period, a couple of months earlier than planned.
Status: training ended.
Documentation: --
Publications: Fattori et al., 2009 [C5]
Revised planning: none

Task 5.3: Motor description of fragment location.
The objective of this task is to experimentally determine motor contributions of eye movements to fragment location via saccade adaptation.
Scheduling: (month 1-30)

Performed actions: Two behavioural studies using saccadic adaptation were conducted. The first study focused on the differences between reactive and scanning saccades to investigate the interactions between visual localization and saccade targeting. In the second study visual localization was decoupled from the motor execution of saccades.
Results: Two studies were completed. The first study is under review for a scientific publication, whereas the second study is currently in preparation. In addition, a third study was finalized during this period and recently the paper was published (Georg & Lappe, 2009 [J4]).
Status: Work started as planned at month 1. Milestone M4 (experimental data on fragment location in humans) was reached as planned on month 12.
Documentation:
Publications: Georg & Lappe, 2009 [J4]
Revised planning: none

Task 5.4: Predicting behaviour and cooperation in shared workspace.
The objective of this task is to study specific aspects of human behaviour in the combination of allocation of attention and direction of gaze that can be used for prediction in human-robot interaction.
Scheduling: (month 13-36)
Performed actions: WWU developed two setups for conducting behavioural experiments in single-subject and two-subjects settings. A first study employing the single-subject setup was started, data were collected and analysed. Tests are currently conducted on the two-subjects setup.
Results: The development of two setups is completed. First results with the single-subject setup are analysed.
Status: Setup development started at month 8 and data collection started at month 12 as planned.
Documentation: --
Publications: --
Revised planning: none

WP6: Project coordination and management
Scheduling: ongoing
Performed actions: See section 5.1
Results: The website of the project (www.eyeshots.it) has been launched and it is continuously updated (deliverable D6.1). The access key that will allow the reviewers to access the restricted area of the website is [username: eyeshots, password: 2toh23y3]
Revised planning: possible structural changes in the website will be brought about if/when necessary to include new sections.

WP7: Knowledge management, dissemination and use, synergies with other projects
Task 7.1: Regular publications of webpages
Scheduling: ongoing
Performed actions: See section 5.2. Deliverable D7.1
Task 7.2: Internal workshop sessions
Scheduling: ongoing
Performed actions: Tutorial sessions in the programme of the kick-off meeting (Bologna, 7-8 March 2008) have served as a first key event for training and ensuring a smooth set-up of the interdisciplinary consortium.
Task 7.3: External dissemination
Scheduling: ongoing
Performed actions: See section 5.2
Results: We have by now published 10 conference contributions and 6 journal papers.
Task 7.4: Synergies with other projects
Scheduling: ongoing
Performed actions: Presentation of the EYESHOTS project to prof. Mark Greenlee (coordinator of the FP6 project "Decisions-in-motion"), prof. Radu Horaud (coordinator of the FP6 project "Perception on Purpose"), and prof. Giulio Sandini (coordinator of the FP6 project "Robotcub").

Potential common interests were preliminary discussed with prof. Greenlee and prof Heiko Neumann (partners of "Decisions-in-motion"). Direct comparative evaluation of the performances of i-Cub stereo head (www.robotcub.org) with our envisaged solutions will be possible considering that the i-Cub head is available by partner UG.
Revised planning: none

WP8: Training, education and mobility
Task 8.1: Literature database
Scheduling: ongoing
Performed actions: Database created and launched. Deliverable D8.1.
Task 8.2: Student's seminars
Scheduling: planned to start in the second year in the month 18 meeting
Task 8.3: Personnel exchange
Scheduling: ongoing
Performed actions: Several subgroup meetings took place at the fringes of the periodic plenary councils. In addition: a 6-months stay of a PhD student from UJI at the UG partner, collaborating to the integration goals of WP1 and WP2 with WP4 ("Experience on phase-based techniques for stereo processing" and "Virtual reaching tasks based on stereo visual cues"), and a visit of a PhD student from UG at the K.U.Leuven for discussing the integration of learning paradigms in the cortical-like architecture of disparity detectors.
Task 8.4: Summer school
Scheduling: ongoing
Performed actions: The summer school was planned for the second year for Sep 2009 in Spain, within the scheme of the International UJI Robotics Summer School (IURS) of the UJI partner. Basically, it will be dedicated (but not exclusively reserved) to students from the EYESHOTS' consortium for deepening mutual knowledge among partners. We are planning to invite one or two external speakers.

In the following, we provide a detailed description of the progress of work for each work package -- except project management, which will be reported in section 5.

WP1: Eye movements for exploration of the 3D space

Leader: Giorgio Cannata (UG)
Contributors and planned/actual effort (PMs) per participant:
UG (14/15), UJI (2/2) and K.U. Leuven (0/0).
Planned/actual Starting date: Month 1/1

Workpackage objectives

The major goals of the workpackage are the study of ocular mechanics and oculomotor control, for both single eye and conjugate movements, as well as the specification of ocular motion strategies which could improve the capabilities of vision to perceive depth information. In particular, the target is to investigate the role of the ocular mechanics with respect to the strategies implemented by the brain to drive typical biological ocular movements (including saccades and vergence). A second objective is the study of the geometric and kinematic effects of ocular motions on image flow, for supporting the estimation of 3D information from ocular motions. The final goal of WP1 is the development of a bio-inspired stereoscopic robot system capable to emulate the ocular motions to be used during the planned experimental tests.

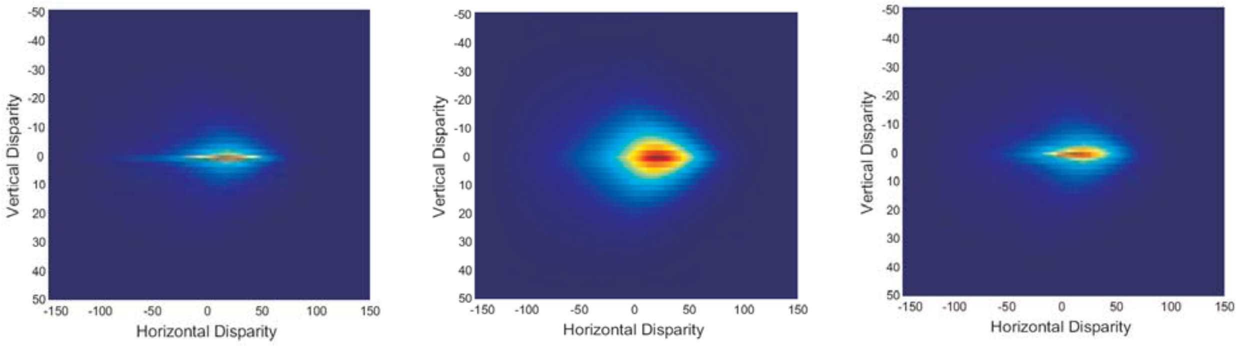
Progress towards objectives

Task 1.1: ***Binocular eye coordination***
This part of the work has focused on the investigation of ocular motion strategies and on the analysis of how they influence in terms of perception and estimation of of 3D information (depth).
This activitiy is mostly related to geometric and kinematics analysis of the vision system, while vision and image processing methods represent, depending on the case, *inputs* or *outputs* to/from the control modules.
The rationale adopted in Task 1.1 has been the following. We have given for granted that large range ocular movements between *discrete* salient features (in a *quasi static environment*) obey to *high speed* transient ocular motion strategies (*saccades*) where motion dynamics strongly dominates, thus limiting the effects of visual feedback to the depth perception.
On the other hand, we have conjectured that between short saccades, as they occur in the pareafoveal region during attentive visual scrutiny, *small range* and *low speed smooth* ocular movements might occur (or be implemented at least in the case of an artificial robotic vision system) with the goal of improving *locally* the perception of depth.
To this aim we have first analysed, from a computational point of view, the ideal problem of *moving* the vergence point over a smooth surface in space. This task has been tackled mostly from the kinematic control point of view to understand the requirements for its implementation (in particular in term of feedback requirements from vision modules). Furthermore, although the analysis refers to movements of *large* amplitude, it should be interpreted as an idealization of small controlled motions enabling the enhancement of depth perception.
As a limit case we have eventually considered the static conditions of the cameras when fixating an object within the *peri-personal* space.
This investigation has led to a critic study of the relative posture of the two cameras with respect to the perceived disparity information. In particular, the study aimed, on one hand to provide a perceptual interpretation of the *extended Listing's Law*, and on the other to analyze how different types of kinematic models (e.g. *pan-tilt* etc.) affect or limit its implementation.

The major results of Task 1.1 can be summarized as follows.

- A control strategy for moving the fixation point (from arbitrary initial conditions) over a surface S.
- A control strategy for moving the fixation point along a smooth surface S.
- A comparative analysis of the different kinematic ocular models and their quantitative effect, under static fixation conditions, on the torsional components of the eyes.

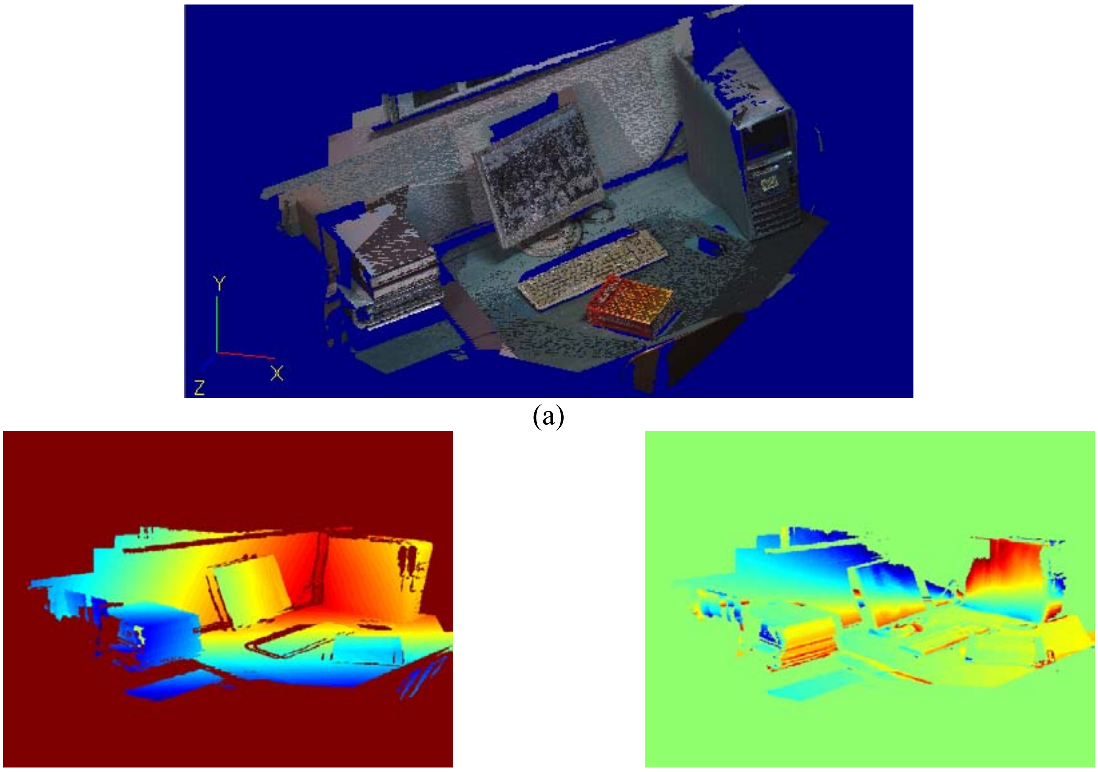
As far as it concerns the first two points, the idea was to investigate the required geometric and kinematic conditions to keep the binocular fixation over a given fixed surface S (within the peripersonal space) under static and dynamic conditions (sliding). Although the problem is in fact unique, we have for sake of clarity splitted it into two sub-problems: being the first the problem of controlling the vergence over the surface S, and the second the motion control of the fixation point over S. We have investigated these problems and proposed candidate control strategies. In particular we have shown, through simulative tests, that the proposed algorithms ensure the asymptotic tracking of smooth reference motions over the sliding surface S. To accomplish this task suitable geometric data are assumed to be computed by the vision modules. In particular, an algorithm to estimate the normal vector to the surface S has been proposed (and evaluated in a simplified simulation environment).
With respect to the third point we have shown that for near fixation points the *pan-tilt* kinematic model agrees with the *extended Listing's Law, L2,* (although it fails for far fixations). The analysis based on simulative investigation shows that actual ocular postures involves a *torsional* component required to improve the disparity perception over the whole visual field. This fact implies that a specific torsional active control component is a requirement to emulate the actual posture of the eyes. To this aim with reference to the first two points discussed above, it is interesting to remark that the *surface sliding* discusses above, and the torsional component control are not coupled. In particular, it is always possible to ensure the fulfilment of the *surface sliding* conditions as well as the torsional requirements (as specified by L2).



(a) (b) (c)

Figure 3: *Comparative analysis of the probability distribution of horizontal and vertical disparities for the three different stereo-head systems considered: (a) a classical Tilt/Pan system; (b) a system following Listing's Law; (c) a system following L2.*

In addition to the planned work, we have also realized a VR simulator for binocular active vision systems. The novelty of the approach is the use of VR as a tool to simulate, in closed perception/action loop, the behaviour of a binocular vision system that observes the scene, rather than just rendering the 3D perceptual illusion of the scene to a human observer. The simulator is implemented in C++, using OpenGL libraries and the Coin3D toolkit (www.coin3d.org). To obtain a stereoscopic visualization of the scene useful to mimic an active stereo vision system we have modified the `SoCamera` node of the Coin3D toolkit. In this way, we have obtained a fast tool, capable of handling the commonly used 3D modelling formats (e.g., VRML and OpenInventor) and the data acquired by a 3D laser scanner (Konica Minolta Vivid 910), specifically purchased by UG for benchmarking the active vision strategies developed in EYESHOTS with real-world conditions with known ground truth. The tool allows us to access the buffers used for the 3D rendering of the scenes. Figure 4a shows an indoor scene, acquired by the 3D laser scanner. The 3D data and the textures have been loaded in the virtual simulator, then the left and right projections, the horizontal and the vertical ground truth disparity maps, can be obtained, *for each possible fixation point* (e.g., see Fig.4b-c. The developed tool is currently being used to create a database of stereo image pairs with data about the vergence points and the ground truth disparities (see, www.pspc.dibe.unige.it/Research/vr.html), to be used in WP2 and WP3 for learning adaptive vergence strategies, as well as 3D visual descriptors for object identity. In general, the tool can be used both for algorithmic and behavioural benchmarks for the whole duration of the project.
The simulator and the general approach followed have been presented at the VISAPP'09 Conference (Chessa et al., 2009 [C8]).



(a)

(b) (c)

Figure 4: *Results obtained with the VR simulator. (a) The 3D model of an indoor scene, acquired by the laser scanner. (b-c) Horizontal and vertical ground truth disparity maps, for a given fixation point.*

Task 1.2: ***Perceptual influences of non-visual cues***

The work started on month 6. We separated the problem of version and vergence by considering (i) the validation of existing models for reference frame transformation in vision-based reaching tasks and, then, (ii) the study of pure vergence modulation of the disparity tuning curves available from the population network to encode the absolute depth of the fixated object. Specifically, we have referred to a model where the multiplicative gain modulation is obtained by a recurrent network fed by feed-forward (i.e., additive) combination of visual and non-visual (e.g., gaze position) signals (Salinas and Abbott, 1995). For the sake of simplicity, instead of learning the recurrent synaptic weights, we used pre-wired fixed synaptic connection kernels modeled by Differences of Gaussians. Moreover, for what concerns the remapping of target location in head-centered 2D coordinates (cf. pure version), we analyzed the gaze sensitivity data on area V6A of the medial parieto-occipital cortex of the macaque, available by UNIBO from experiments conducted before the starting date of the EYESHOTS project. On the basis of the experimental observations of a large number of cells that exhibit peak-shaped (rather that planar) gaze fields, we have verified that the peak-shaped gaze fields reported by UNIBO studies are not in contrast with the gain field models developed in the theoretical neuroscience literature (Salinas and Abbott, 1995; Pouget and Sejnowski, 1997). Rather, the use of peak-shaped (i.e., non monotonic) gaze fields even improves the efficiency of the coding scheme by reducing the number of units that are necessary to encode the target position.

Preliminary results on this analysis will be presented at the IWINAC'09 conference (Breveglieri et al., 2009 [C6], joint publication between UNIBO and UG).

Task 1.3: ***Control of voluntary eye movements in 3D***

This part of the work has so far focused on the investigation of the bio-mechanical models of the ocular plant including extra-ocular muscles dynamics. The starting point is the model of Bahill et al. (1980) and applied to minimum-time saccadic control by Enderle et al. (Enderle et al., 1984; Enderle & Wolfe, 1987). In particular, the work has focused on the extension of these models originally developed in 2D to the three dimensional case by extending of the eye model proposed in (Cannata and Maggiali, 2008) where it is shown that physiologically correct saccadic motions could be generated by any action of the four recti extra-ocular muscles. In view of the achievements of Task1.1, the model will require an extension also in view of the L2 requirements, in particular this will also reflect on Task1.4. The complexity arising, and not yet solved at the present time, refers to the coupling between the horizontal and vertical rectii. This issue is relevant for *continuous* visual feedback tasks as those discussed in Task1.1, but are also highly relevant for the computation of minimum-time saccades, where the control system is operating in open loop with respect to vision modules.

Deviations from the project workprogramme

The work has proceeded has planned no corrective actions have been necessary. Additional effort has been devoted to the development of the virtual active stereo vision simulator, which was originally unplanned. The simulator was necessary to generate data sets of stereo image pairs under controlled vergent axes conditions, and, in perspective, to provide a way for validating the behaviour of the active vision system.

WP2: Active stereopsis

Leader: Marc Van Hulle (K.U.Leuven)
Contributors and planned/actual effort (PMs) per participant:
UG (9/11) and K.U.Leuven (12/4)
Planned/actual Starting date: Month 1/1

Workpackage objectives

This Workpackage is devoted to the specialization of disparity detectors at different levels in a hierarchical network architecture to see the effect of learning (higher-order disparity detectors) for the extraction of the binocular features stereo and stereomotion. A vergence motor strategy is learned that, combined with the sparse detectors, optimizes the quality and efficiency of the feature-extraction for the specific tasks (guided by the attention signal) considered here. In a second task, scanning strategies are developed that enable this learned architecture to accurately describe the 3D structure (or headcentric disparity) of a visual fragment. This strategy will be designed to be robust to the limited accuracy of the motor system.
The extraction of binocular features occurs through a cortical-like population network, which has developed by UG, in its preliminary form, before the starting date of the project (Sabatini and Solari, 2004, cf. also Sabatini et al., 2007). The network provides a harmonic (i.e., amplitude and phase) representation of the visual signal, operated by a set of “simple cell” units (S-cells).
At the level of S-cells, the “totipotency” of the representation contains all the necessary basic components to differentiate into several classes of visual descriptors. Stereo and stereomotion percepts emerge in layers of disparity energy “complex cell” units (C-cells) that gather S-cells outputs according to specific architectural schemes.
Though these computations can be supported by neuromorphic architectural resources organized as hierarchical arrays of interacting nodes, the network does not give completely the way to a pure connectionist approach, rather it allows us to design an architecture in which, at every layer, the “connectionist paradigm” is steeped in “signal processing”. On this basis, learning processes could develop the high degree of variability of the cell’s responses towards the specialization of disparity detectors for the control of vergence. The selected learning paradigm is inspired by LeNet5 (LeCun et al., 1998), since it is expected to have a good performance being such a network optimized at every level of the hierarchy.

Progress towards objectives

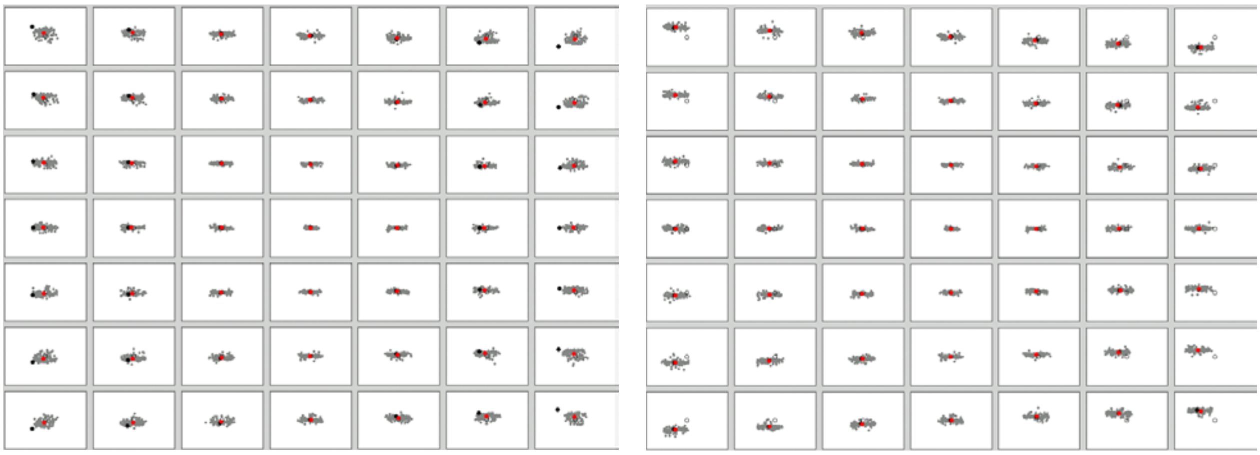
In general, starting from this ground, (1) the degree of adaptability of the cortical-like architecture have been assessed, limited to its static properties, for what concern both the *coding* and *decoding* issues (disparity range, number and distribution of cells for each orientation, ocular dominance, distribution of orientation channels, and reading-out mechanisms for static disparity), (2) the LeNet architecture has been extended to increase its flexibility and including new functionalities.

Specific progress on the tasks worked is reported as follows.

Task 2.1: ***Network paradigm for intelligent vergence control.***

Step 1: Embedding fixation constraints into binocular energy-based models of depth perception.
In natural viewing conditions, the disparity distributions (horizontal and vertical) critically depend on the orientation of the eyes. Over relatively large visual angles, the retinal disparity patterns experienced by a binocular vergent system engaged in natural viewing, present predictable components related to the positions of the eyes in the orbits. The 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. Although, from a conceptual point of view, the oculomotor parameterization of active stereopsis is a well-established issue (e.g., cf. Jenkin, 1996; Hansard & Horaud, 2008), mapping the oculomotor constraints into the neural population coding and decoding strategies is still

an open problem. We have analyzed the influence of changes in the fixation location and of the 3D structure of the environment on the distribution of disparity. In the simulation we have computed the statistics of the disparity distribution by varying the fixation point.



(a) (b)

Figure 5: *The subplots represent a grid of 7x7 retinal images for different fixation points.For different points of the left retina (open circles) we estimated the corresponding point for the right retina (grey dots) by a population of disparity detectors. The disparity relates to point image projections of randomly positioned objects in the peripersonal space, when the fixation point is in the primary position. The red dots represent the mean of the disparities, whereas the black dots represent the true mean disparity. Distribution of the estimated disparities between the left and the right retina are depicted (a) without and (b) with the compensation of the predictable components of the disparity pattern.*

As expected, the disparities for each retinal position are characterized by a *bias,* or offset, due to the epipolar geometry of the binocular system, that tends to orient the disparity along a restricted range of directions. 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 where phase-shifts and position-shifts play a different role: position-shifts are used to compensate the global components of the disparity pattern with respect to a fixation point in a primary position; phase-shifts are used to estimate the residual 2D disparity. We have verified that (see Fig.5), in this way, we improve the reliability of the disparity representation (in terms of accuracy and coverage) while keeping the amount of necessary resources to a minimum (Chessa et al., 2009b [C2]).

Step 2: Disparity map priors from epipolar geometry in vergence motor strategy

We further analyzed the effect of the *expected* disparity information perceived (at the target point), given a stereo setup and a starting fixation point, in a vergence motor strategy. Hereto, we developed an application where we can simulate different realistic stereo set-ups (including fixation points) and scene lay-outs (Fig.6). We set realistic parameters of the modelled stereo setup: baseline, focus length, work space position, number of objects, and so on, and generate a large number of synthetic scenes. For each scene, we make a number of saccades with different fixation points (Fig.7). For each fixation point, using a ray-tracing approach, we render vertical and horizontal disparity maps. We accumulate the obtained data in such a way that statistics (across all synthetic scenes) can be easily calculated for each fixation point (Fig.8). Finally, we can now use in a vergence control strategy, the obtained mean disparity maps as a prior to the saccade.

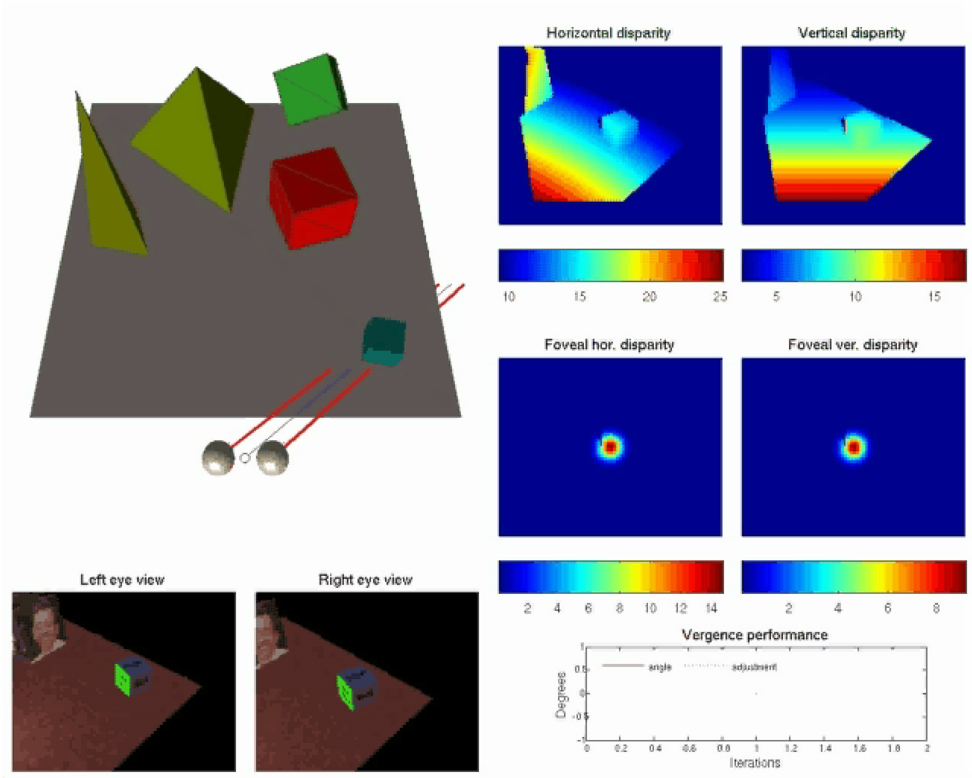


Figure 6: *Top left panel: realistic scene lay-outs with stereo set-up (eyes) and fixation point in 3D. Bottom left two panels: views obtained after ray-tracing for the two eyes. Top right two panels: horizontal and vertical disparities corresponding to the fixation point. Middle right two panels: distribution of vertical and horizontal disparities when performing vergence. Bottom right panel: vergence control as a function of time.*

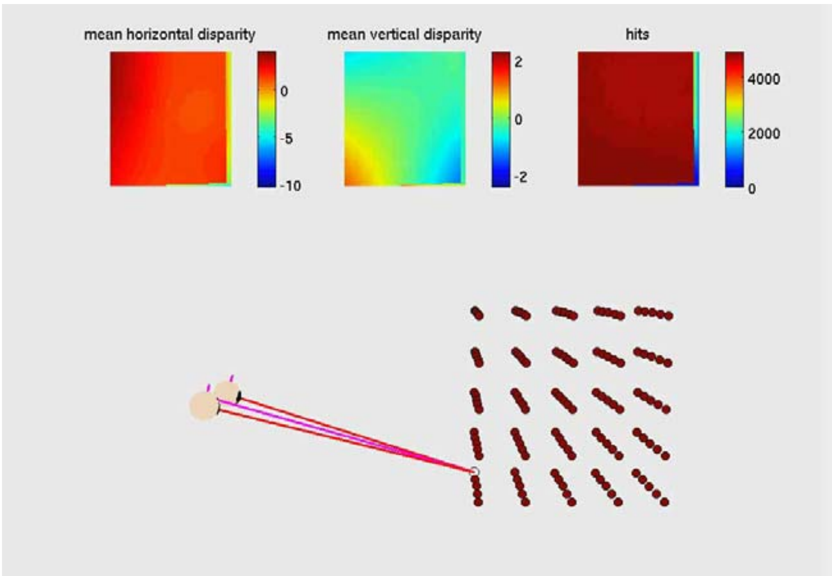


Figure 7: *Simulation environment showing the fixation points generated in 3D.*

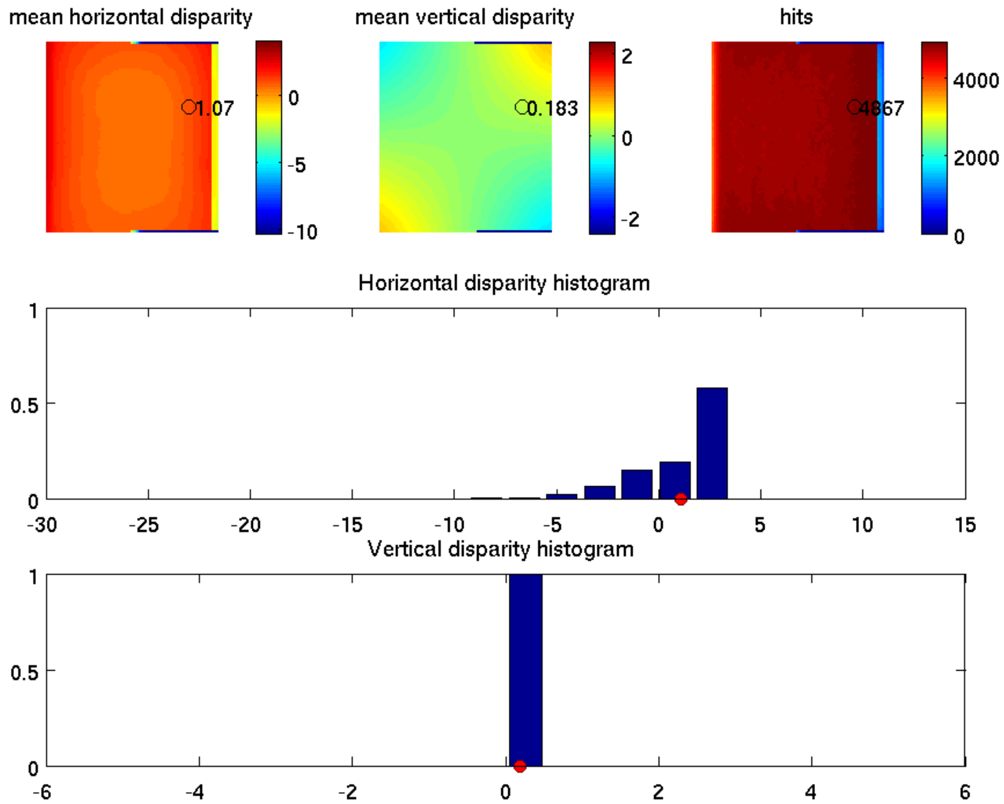


Figure 8: *Distribution of horizontal and vertical disparities (bottom 2 panels) for a given fixation point position in retinal coordinates. The maps in the top left and middle panels show the distribution of the mean horizontal and vertical disparities in retinal coordinates; the top right panel shows the number of times the 3D fixation point expressed in retinal coordinates corresponds to a non-occluded case ("hits").*

The vergence is controlled using a convolutional network arranged in a closed loop (Fig.9-top). The vergence control signal is proportional to the disparity averaged/pulled over the "foveal" region (Theimer and Mallot, 1994; Patel et al., 1996). The performance of this vergence control has been demonstated using our simulated environment. As a network paradigm, we use LeNet (LeCun et al., 1998) (Fig.9-bottom). LeNet has been used for digit/character recognition, generic object recognition, terrain classification for robot navigation, *etc*. We have introduced some new features in LeNet: during training: any feature map of any layer can be switched to a fixed (non-trainable) mode and back. Secondly, we have introduced the M-layer, to replace the S-layer. With one parameter we can control the way the M-layer operates: when the parameter a is large positive, then a max operation is performed, when it is large negative, a min operation, and when a=0 the usual averaging takes place. This feature will add up to the computational power of the convolutional network.

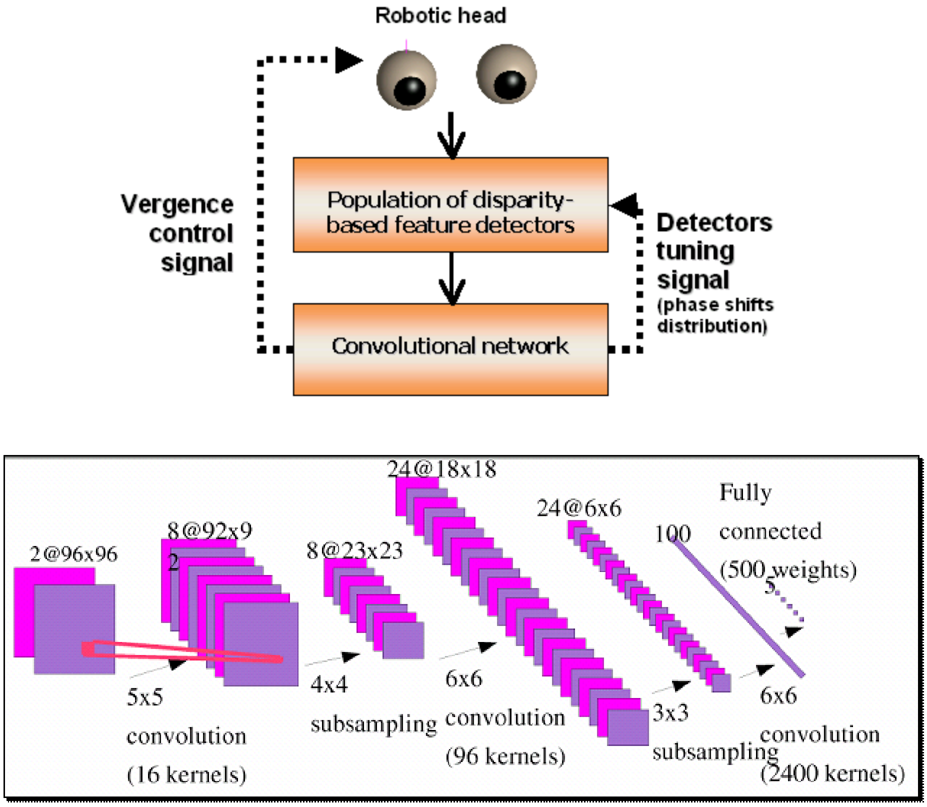


Figure 9: *(Top) Vergence control strategy. (Bottom) LeNet convolutonal network.*

Step 3: Binocular vergence control without explicit calculation of disparity
Experimental evidences show that, although depth perception and vergence eye movements are based on the activity of complex cells of the primary visual cortex, the brain adopts specific and separate mechanisms to combine binocular information and carry out the two distinct tasks.
Vergence control models that are based on a distributed population of disparity detectors, usually require first the computation of the disparity map, thus limiting the functionality of the vergence system inside the sensitivity range of the population of cells specialized for depth perception.
As for the control of vergence larger disparities have to be discriminated while keeping a good accuracy around the fixation point for allowing finer refinement and achieving stable fixations, alternative strategies might be employed.
We (Gibaldi et al., 2009a [C7], Gibaldi et al., 2009b [C1]) 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 taking a decision, but extracting a disparity-vergence response that allows us to nullify the disparity in the fovea, even if the stimulus presented is far beyond the disparity sensitivity range. The disparity-vergence response is 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. 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 additive 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 in a virtual environment achieving stable fixation and small response time to a wide range of disparities. The vergence movements produced are able bring and to keep the fixation point both on a steady and on a moving stimulus (e.g., see Fig.10).

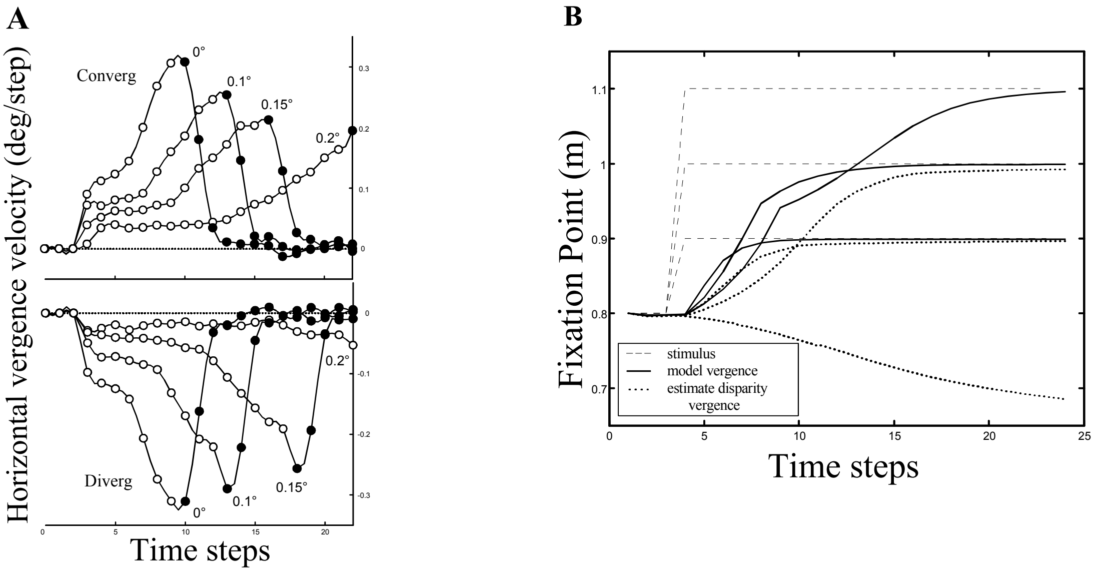


Figure 10: *(A) The velocity convergence (top) and divergence (bottom) response of the modelled cells to a disparity step 0.8 deg, in the presence of different values of vertical disparities (from 0 deg to 0.2 deg). The complementary use of the LONG and SHORT signals is evidenced: white circles denote LONG whereas black circles denote SHORT. (B) The comparison between the responses obtained by the vergence cells (solid line) and those obtained by explicit disparity decoding of the population activity (dotted line).*

Task 2.2: ***Interactive depth perception.***

Accurate gaze estimates are required for disparity estimation (image rectification), for depth estimation (disparity integration) and for vergence control (Task 2.1). The gaze refinements are estimated from vector disparity (horizontal and vertical disparity). We have developed an iterative algorithm that performs a simultaneous gaze and disparity estimation. Phase interpolation across orientations is done to compensate for 2D orientation differences resulting from the current gaze. To illustrate the performance of the algorithm, we take twice the same image, but rotate one, and consider the two as the binocular inputs (Fig.11). The rotation introduces a disparity (Fig.12, left panel), as the ground truth shows for the horizontal component (Fig.12, right panel). Our algorithm handles the rotation and yields a result similar to the ground truth (Fig.12, middle panel). The convergence of the algorithm is shown in Fig. 3. In perspective, this technique can be used for compensating rotation occurring in vergence situations.

Figure 11: *Example of an image and a rotated version to simulate the stereo camera output in case of a rotation error.*

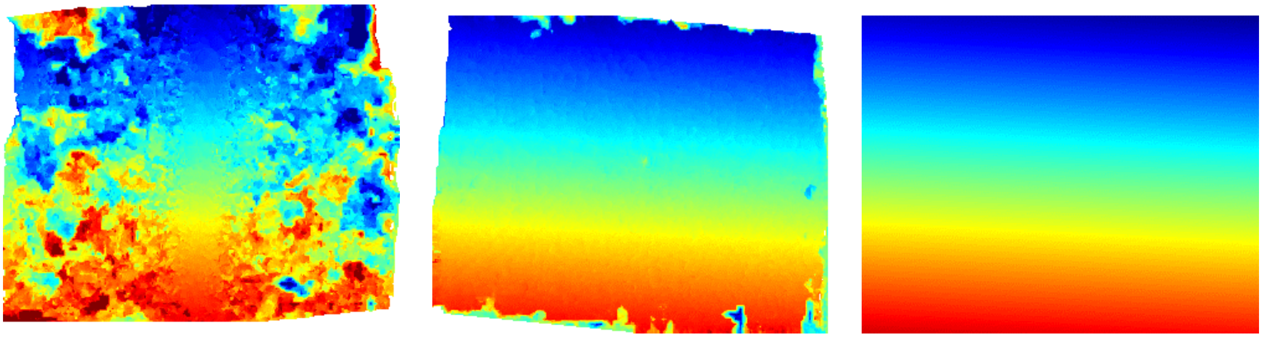


Figure 12: *Horizontal components of disparity due to rotation. Horizontal disparity estimates in case no correction is performed (left panel), the correction in our case (middle panel), and the ground truth (right panel).*

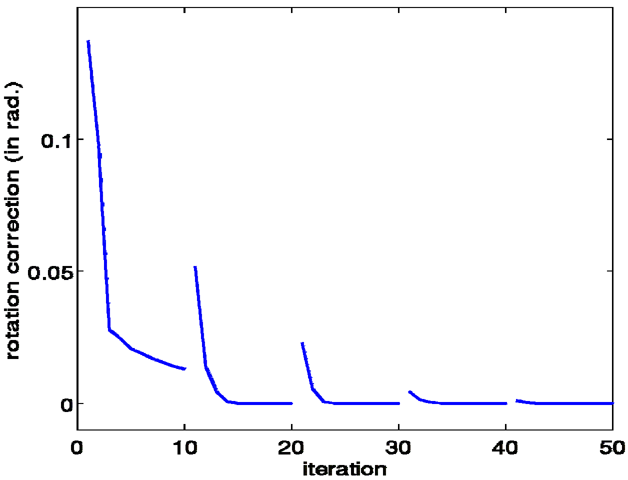


Figure 13: *Convergence of the algorithm estimating iteratively gaze and disparity.*

Deviations from the project workprogramme

None

WP3: Selecting and binding visual fragments

Leader: Fred Hamker (WWU)
Contributors and planned/actual effort (PMs) per participant:
WWU (18/18) and UG (2/2)
Planned/actual Starting date: Month 1/1

Workpackage objectives

This workpackage is devoted to develop novel concepts of selecting and binding within a fragmented 3D scene representation. One of those fragments is object identity. Object identity will be obtained from a bi-directional, hierarchical representation of learned feature detectors. The development of the appropriate learning rules will be an essential part of this project, since the learned connections will be used for the selection of a fragment. In the first period we aimed at learning V1-like feature detectors form stereo images. Beyond object identity, a distributed representation requires to actively bind and represent the relevant visual fragments for the task at hand. Thus, we study how attentional dynamics allow to actively bind features and build task relevant representations. Moreover, we will develop a novel framework for the task relevant binding of fragments in a global workspace using reward-based learning.
Starting point has been a model for learning V1 receptive fields (Hamker and Wiltschut, 2007) and models of attentional dynamics (Hamker, 2005, Hamker and Zirnsak, 2006, Hamker et al., 2008).

Progress towards objectives

Task 3.1: *Defining visual fragment: object identity*
Task 3.1 is devoted to develop novel concepts of learning to encode a fragmented 3D scene representation. Object identity will be encoded within a network of bi-directional, hierarchical learned feature detectors.
Since the early studies of receptive field properties in primary visual cortex, a major issue in neural coding has emerged, dealing with the question of why neurons have a particular receptive field structure. Since V1 neurons respond well to edges, edge detection has been considered as a useful operation of early vision emphasizing the important structural properties of a visual scene. However, this does not answer the questions about optimal edge detectors and particularly why edge detectors should emerge and not any other potentially useful detector. Important progress has arisen from the efficient coding hypothesis (Attneave, 1954, Barlow, 1961, Field, 1994). Particularly, recent contributions in this respect have shown that algorithms seeking for a statistical independence of the neural responses converge to localized, oriented, bandpass filters (Olshausen & Field, 1996; Bell & Sejnowski, 1997; van Hateren & van der Schaaf, 1998). However, despite this great success, a more close comparison with neural data revealed that the learned receptive fields do not capture the full frequency distribution as observed in experimental data (van Hateren & van der Schaaf, 1998, Ringach et al., 2002).
While many studies have demonstrated some relationship between neural receptive field properties and aspects of efficient coding, we investigated the quantitative influence of particular aspects of efficient coding on the similarity between the modelled receptive fields and the experimental data (Wiltschut & Hamker, 2009). We systematically varied critical model parameters and measured information theoretic properties of efficient coding in these different instances after learning. We then analyzed if these measurements of efficient coding correlate with the similarity between model and biological data -- the distribution of spatial frequency tuning (Ringach et al., 2002). We have further shown that this model with non-linear lateral competition, learned by an anti-Hebbian principle, and Hebbian learning of feedforward and feedback connections, develops receptive fields which are not only similar to V1 macaque data, but which match the distribution of the spatial frequency tuning, particularly if the neural code is efficient. We also compared our results to Independent Component Analysis (ICA), a standard linear method for learning receptive fields from natural scenes. We used the fast fix-point algorithm (Hyvärinen et al., 2001). Hoyer & Hyvarinen (2000) have shown that this algorithm can produce oriented bandpass filters and disparity tuned neurons similar to those in V1. This particular ICA algorithm shows similar deficits than the results from Lewicki et

al. (1999) and more mild deficits compared to those Ringach et al. (2002) has reported for Bell & Sejnowski (1997) ICA and the Olshausen & Field (1996) sparse coding algorithm with respect to the variety of receptive field properties. In all of these studies ICA only captures in part the whole distribution of receptive field properties.

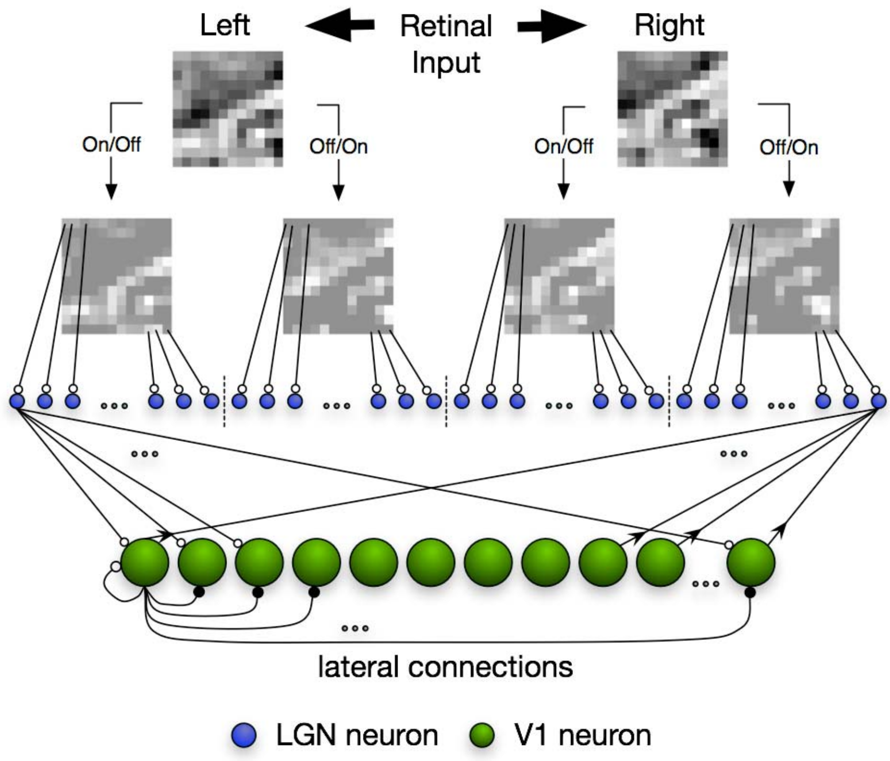


Figure 14: *The network consists of two layers. The neural dynamics implement a feedforward/ feedback system, where feedback strengthens the representation of the predicted features in the input. If other top-down signals were available, this network could be used to implement feature-based attention. The on/off responses are determined from the retinal input from the left and right eye represented by whitened/lowpass filtered image patches. The firing rate of these cells provides the input to layer LGN cells, which are subject to feedback control. The layer LGN cells represent by their activity the content of the visual scene with an additional small increase in firing rate if the input matches the expectation from V1. V1 is supposed to learn a more efficient representation of the visual content. The feedforward and feedback weights are learnt simultaneously.*

We further extended the network to learn binocular receptive fields from a set of stereo images (Fig.14). The set of stereo images was taken from a two-camera system mounted in our lab. For the initial testing of the algorithm we used a small set of images from indoor scenes. Figure 15 shows typical examples of tuning curves of binocular cells in the new model using indoor scenes for learning. We observe a large number of cells with disparity tuning. The first cell shown in Fig.15 is tuned to near objects, the other ones are tuned to far objects. However, the third cell shows quite broad tuning characteristics. An in depth comparison to V1 data with respect to disparity tuning will be reported in deliverable D3.1b using natural outdoor scenes. The feedback connections learned will be particularly important for Task 3.2 where they will be used to attentively bind visual fragments on demand.

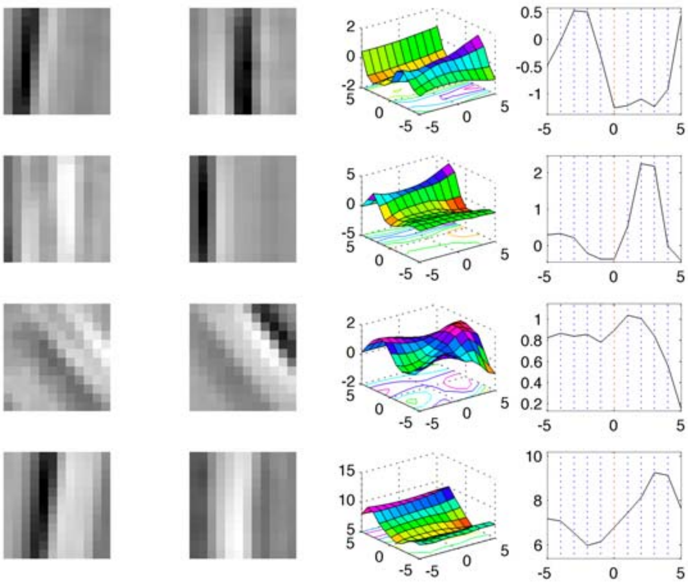


Figure 15: *Each row shows the tuning properties of a binocular cell after learning. The two images on the left show the receptive field of the left and right view. The third image shows the horizontal and vertical disparity tuning (in pixels) of the cell and the last image depicts the horizontal disparity tuning properties at zero vertical disparity.*

Task 3.2: ***Selecting visual frament***

Task 3.2 aims at understanding and developing mechanisms of attentional selection. We have earlier proposed that the feedback of an eye movement plan back into extrastriate visual areas leads to an increase of the number of receptive fields (RFs) processing the object at the saccade target location. This prediction appears consistent with observations made in V4 (Tolias et al., 2001). However, in other brain areas researchers have reported a remapping of receptive fields, i.e. cells respond to stimuli presented in their so called future receptive field before the eye starts to move (Duhamel et al., 1992; Sommer & Wurtz, 2006). A simple conclusion drawn from these observations could be that V4 cells have different peri-saccadic receptive field dynamics than other areas that show a remapping of receptive fields (Fig.16). However, on basis of our previously developed model (Hamker et al., 2008), we investigated to which extend both receptive field dynamics could be caused by the same mechanism.

To assess the change in RF size in our model (Fig.17C) we applied the same respone rule which was used by Tolias et al. (2001). That is, for a given neuron we determined its RF in both the pre-saccadic and peri-saccadic condition by stimuli, in the following referred to as probes, eliciting a response which has to be higher than half of the maximum response of the neuron.

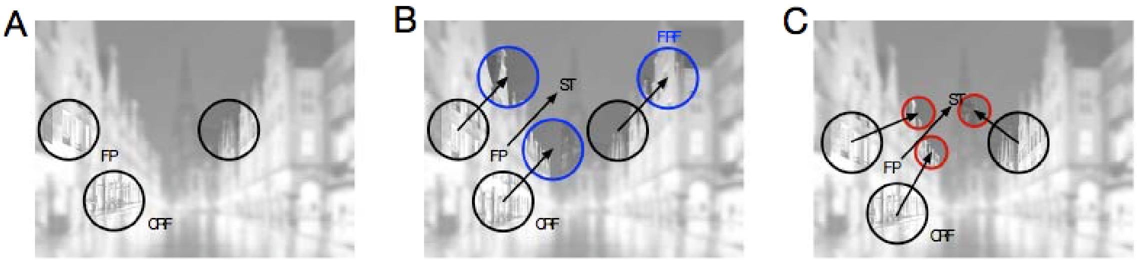


Figure 16: *Illustration of the observed RF dynamics. A) Three hypothetical examples of Current RFs (CRF) during fixation (FP) long before a saccade. B) Predictive remapping. Before saccade onset RFs are shifted to the location of the Future RF (FRF). C) Before saccade onset RFs shrink and are shifted towards the Saccade Target (ST).*

To compare the model RFs with remapping RFs we applied two tests. The first test is the most common method in the literature. It consists of placing one probe in the center of the FRF of a given cell (Fig.17A). Whereas the FRF is the region in visual space where the so called current RF (CRF) would be located after the eye movement. If the cell gets activated by the probe even before the eye movement it is classified as a predictive remapping cell. Note that the exact criteria used is less restrictive but requires that the latency of a neuron relative to saccade onset to a probe in the FRF must be shorter than the latency relative to the onset of a probe in the CRF in a fixation task. In the following we will refer to this test as the one probe test. The second test was introduced by Sommer and Wurtz (2006) and was designed to differentiate between remapping and the V4 RF dynamics. It consists of two probes. As in the one probe test the first probe is presented in the center of the FRF. The second probe is placed close to the saccade target (Fig.17B). If the RF translates parallel to the saccade the first probe should evoke a stronger activation of a neuron as compared to the second probe and vice versa. We will refer to this test as the two probe test.

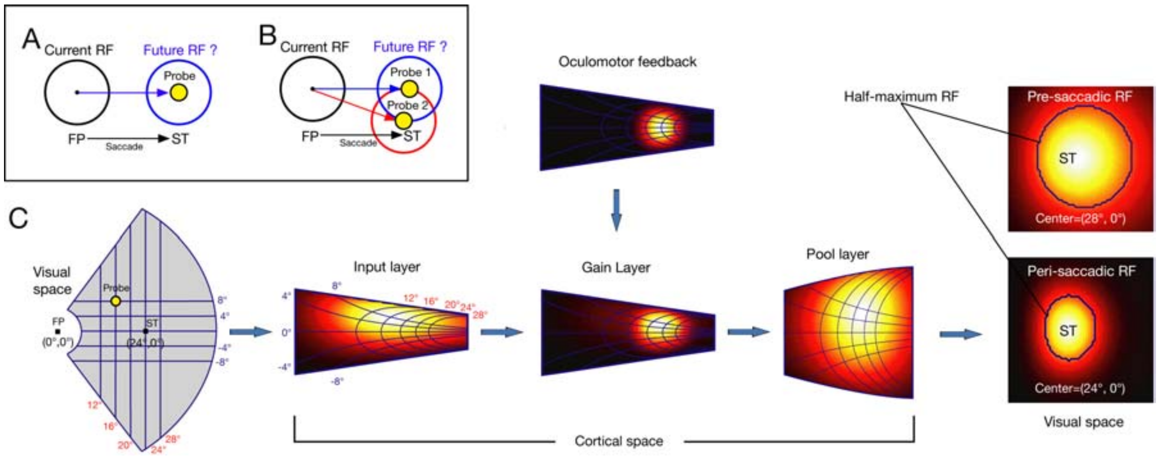


Figure 17: *Illustration of the model and the applied tests and measures. A) One probe test. B) Two probe test. C) The visual space, described in spherical coordinates, is mapped into cortical space according to cortical magnification factors. A visual area in the model consists of an input layer, a gain layer, and a pool layer. The input layer describes the unmodulated population activity caused by a probe presented in visual space. If there is no feedback signal the gain layer will be equal to the input layer, which is true for the pre-saccadic condition. In the peri-saccadic condition the population of the gain layer gets distorted towards the saccade target by the impact of the occulomotor feedback signal, which also leads to a change in the RFs of cells in the pool layer as it is illustrated on the right by the half-maximum profiles.*

We find that the model is consistent with both observations, the reported V4 RF shifts and predictive remapping as it is indicated by the one probe test (Fig.18 upper row), it is also evident from the two probe test that the latter is true only for a certain part of visual space (Fig. 3 lower row). Interestingly, these regions seem to correspond with the RF positions reported in the remapping studies. Since most remapping studies used mainly the one probe test, there is room for several implications under the assumption that the model is true. According to our results the most obvious implication is that, instead of representing a global phenomenon, predictive remapping might indeed be limited to a certain region of visual space, which encompasses roughly the region between the FP and the ST. However, it might still be that the reported V4 shifts and predictive remapping exist in parallel but take place at different levels of the processing hierarchy. In order to differentiate between those possibilities and their functional implications further more detailed measurements of peri-saccadic RF dynamics in both the ventral and dorsal pathway are required.
From the computational point of view, our model of receptive field dynamics is consistent with present observations and offers an interesting alternative compared to modelling attention as a simple spotlight.

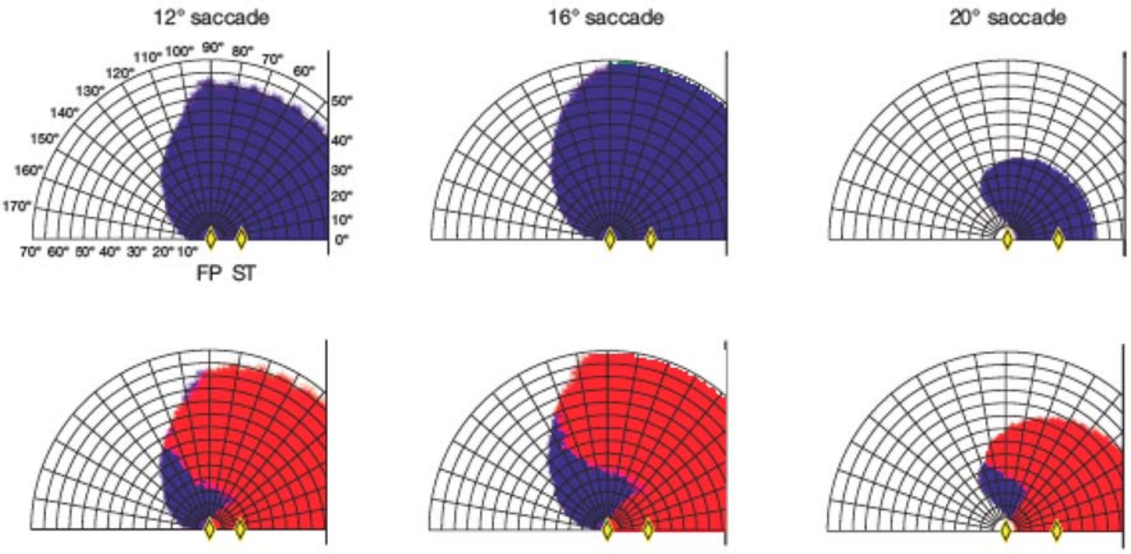


Figure 18: *Results of the one probe (upper row) and two probe test (lower row). The blue area consists of the positions of RF centers in visual space where cells are classified as remapping cells.*

Task 3.3: ***Selecting between behavioral alternatives.***

From an autonomous agent's point of view, active exploration requires that visual perception is as well an active process: the agent has to select which part of the incoming information has to be processed in priority to assure its survival. Visual attention is an example of the various cognitive processes involved in perception. Although it has a bottom-up component induced by the physical properties of the visual objects, visual attention has also a cognitive top-down component which guides the way visual information is processed, depending on the context or tasks requirements (Hamker, 2005). This goal-directed behaviour needs internal representations of the objects that are expected in order to favour their perception. In Task 3.3 we address the fundamental questions of how these internal representations can be learned, maintained and recalled when needed, relying on concepts of reinforcement learning (RL).

In neuroscience, specific attention has been called to RL since the seminal studies of Schultz (1998) who observed that dopaminergic (DA) neurons in the animal's midbrain show similar patterns in the context of Pavlovian conditioning as the error signal in the temporal-difference (TD) algorithm (Sutton & Barto, 1998). The integration of these DA neurons into a functional pathway lead to various computational models of basal ganglia (BG) that mimic the classic actor/critic architecture in order to explain BG functioning in various reward-dependent tasks such as action-selection, motor control or working memory. We reviewed the links of the TD algorithm to brain function and alternative models of the Basal Ganglia intended for action selection and cognitive control (Vitay et al., submitted). Particularly with respect to the cognitive guidance of visual perception we have started developing a novel model (Vitay & Hamker, in press), as illustrated in Fig.19.

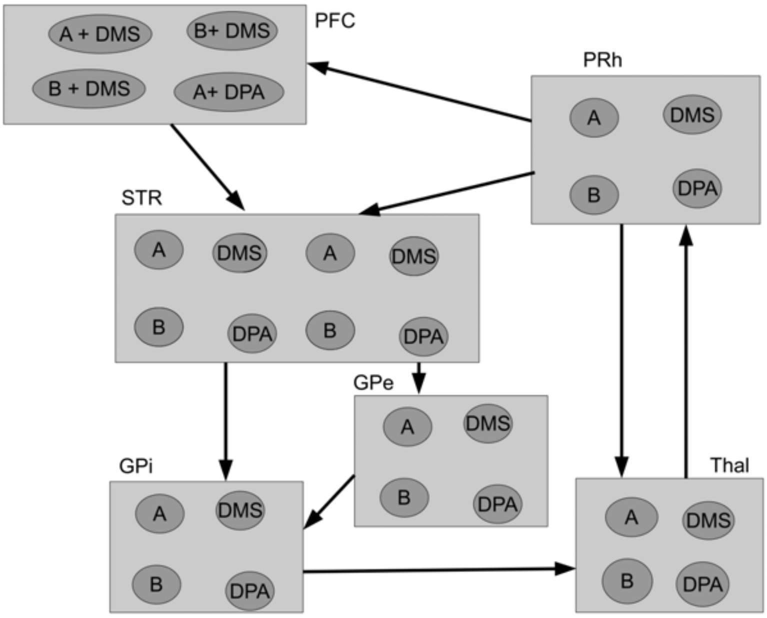


Figure 19: *Schematic architecture of the BG model for cognitive control of visual perception. PRh provides the striatum (STR) and prefrontal cortex (PFC) with a distributed representation of the identity of the visual object (A, B, DMS or DPA). Self-organization in PFC and STR allows to create a representation of both the identity of the object and the current task. Disinhibition of thalamus (Thal) can then retrieve the content of the memorized or associated object.*

Visual information is encoded in PRh, which is reciprocally connected to thalamic nuclei, so that coordinated self-organization in these two structures leads to segregated recurrent loops for each cluster. Thalamus is under tonic inhibition of the output structure of BG (GPi) to control the thalamic stimulation of PRh. GPi itself is inhibited by the striatum (STR) and other BG nuclei (GPe or STN (subthalamic nuclei, not included here)), meaning that activation of these structures can selectively disinhibit one or more thalamocortical loops in PRh and retrieve the content of a cluster. PRh projects to the striatum that will learn to associate PRh representations to the disinhibition of the corresponding thalamocortical loops thanks to the reward-prediction system represented by phasic DA releases to the striatum. These DA bursts can favour either LTP or LTD in corticostriatal connections depending on the polarization state of the striatal cells. We also consider a prefrontal area (probably located in orbitofrontal cortex) that will receive information about the identity of the objects cues, and exhibit compound representations of objects in the context of the current task. This PFC area projects on the same striatal cells as PRh. Self-organization of striatal cells through lateral competition creates separate representations of objects according to the task requirements, allowing different thalamocortical loops to be disinhibited depending on the task. The interplay between direct and indirect pathways in the BG circuitry favours this selection.

This functional model is not fully implemented yet, but a preliminary version (Fig.20). Each area is represented by a set of mean-firing rate units whose activity evolves through time following a differential equation. Our main hypothesis is that the content of working memory is not located in PFC or in BG, but can be retrieved through external activation of multimodal areas like PRh. We tested this idea by using an intermix of delayed match-to-sample (DMS) and delayed pair-association (DPA) tasks. In the DMS task, the model is presented with a visual object (A or B), creating a distributed representation in PRh. After a certain delay, a cue representing the DMS task is shown, and the representation of A or B in PRh is reset due to competition. The goal of the model in the DMS task is to retrieve the content of the previously shown item in order to obtain reward. In the DPA task, the system must on the contrary retrieve the content of the associated item to obtain reward.

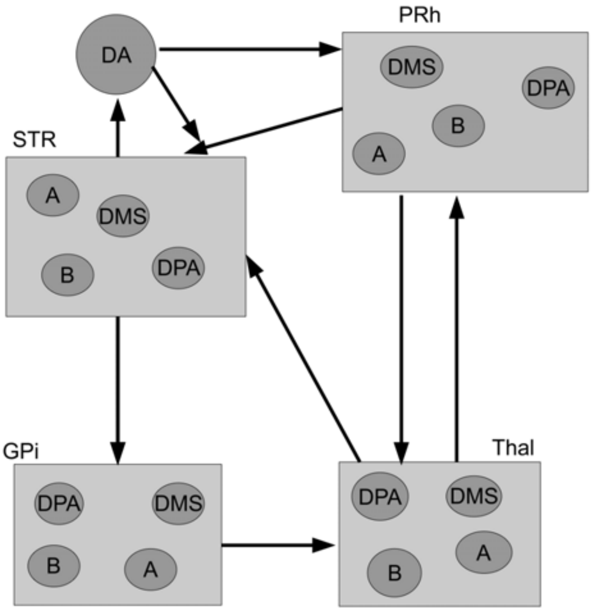


Figure 20: *Model for memory retrieval property. PRh provides STR with a distributed representation of the identity of an object (A, B, DMS or DPA). Self-organization in STR allows to create a different and reduced representation of this object. The disinhibition of thalamus by GPi can then retrieve the content of the memorized or associated object. DA is phasically activated by new patterns in STR and modulates the learning of the corticostriatal connections.*

PRh is stimulated by four different objects representing the A, B, DMS and DPA cues. It projects to the striatum in an all-to-all manner. The striatum inhibits GPi with a Gaussian connectivity kernel, as well as the inhibitory connections from GPi to thalamus. PRh and the thalamus are reciprocally connected in an all-to-all manner. The thalamus also projects to the striatum with a Gaussian connectivity kernel for reasons explained later. The reciprocal connections between PRh and thalamus are learned according to the same learning rule we use for the lateral connections in PRh.

Each time an object is presented, it creates some new activity in striatum, which phasically activates the dopamine cell for approximately 200 ms. DA then modulates the homeostatic learning of the corticostriatal connections, allowing LTP only at the time of a DA burst. Competition in the striatum and in thalamus ensures self-organization in these structures, so that only a few cells represent the same object. The topological projection from thalamus to striatum ensures that the two representations of the same object are similar but not identical, since thalamus has four times less cells than striatum in this model. Learning then converges to three different representations of the same object in PRh, striatum and thalamus, with a decreasing number of active cells.

During learning, GPi is always inhibited by striatal activity to ensure that thalamic cells can become active and learn their reciprocal connections with PRh. When the model has to learn DMS and DPA tasks, this inhibition will not be systematic anymore. Also, striatal cells are kept in the "down-state" during learning, i.e. they only respond to cortical stimulation but do not show sustained activities. Figure 21 highlights the properties of this model showing the simulation of a sequence of events. The object A is first shown for 200 ms. The corresponding cluster exhibits sustained activation after the disappearance of the object. We then artificially set the striatal cells into their "up-state", by self-stimulating them. Another object then appears (DMS, for example) which erases the sustained activation in PRh. When it disappears again, the striatal cells still fire and the thalamic cells corresponding to object A are disinhibited. This activates the thalamocortical loop and the cluster in PRh corresponding to object A shows high activity again, without any cortical input. Thus, a sustained activation of striatal cells allows memory retrieval in PRh.

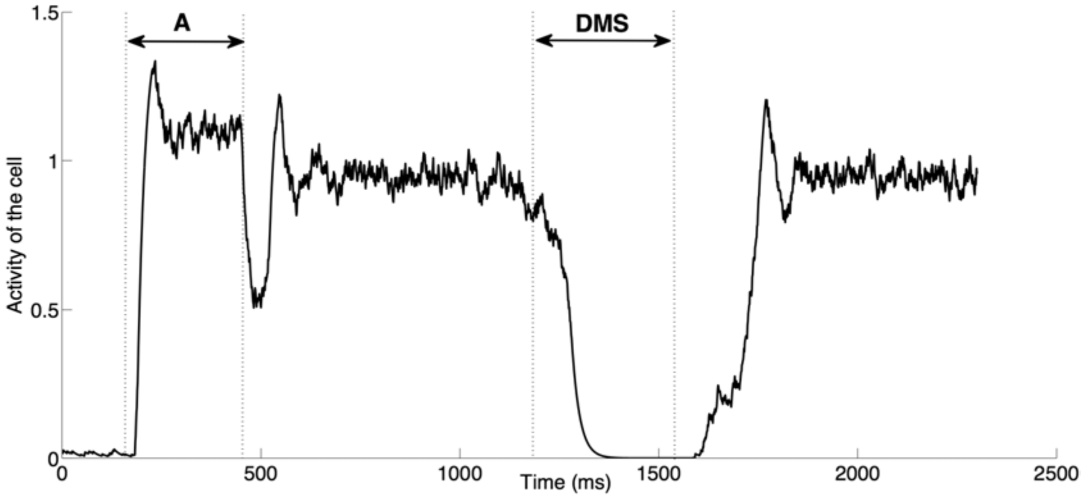


Figure 21: *Time course of the activity of a PRh cell. At approx. 200 ms after the start of the trial, the preferred object of this cell is presented for 200 ms. When the stimulation ends, the cluster exhibits sustained activation until a new object is presented at approx. 1200 ms. This new object disappears at 1600 ms and the cell becomes active again due to memory retrieval.*

Deviations from the project workprogramme

Task 3.3 is ahead of time since partner WWU had already the capacities available. Though, since this task does not need to wait for any outcomes from other tasks, the change in its schedule does not conflict/impact on the progress of work of the whole project.

WP4: Sensorimotor integration

Leader: Angel del Pobil (UJI)
Contributors and planned/actual effort (PMs) per participant: [planned]
UJI (12/12), UNIBO (3/3), UG (1/1)
Planned/actual Starting date: Month 1/1

Workpackage objectives

This workpackage is devoted to define (task 1) and generate (task 2) an action-perception integrated representation of objects in the peripersonal space in a dynamical way. Such integrated representation is achieved through the practical interaction of the robotic system with the environment, using both visual input and proprioceptual data concerning eye and arm movements. Task 1 has begun at month 1, while Task 2 has begun at month 7, after bibliographic review and definition of the model theoretical framework. Task 3 will start during the second year of the project and its final goal is to provide the robot with the skills of constructing a global awareness of its peripersonal space integrating visual data and head and arm movements.

Progress towards objectives

Task 4.1 ***Merging perception-related and action-related visual information***
Task 4.1 was aimed at defining an integrated visuomotor representation of the peripersonal space, which includes both action-related and perception-related aspects. This constitutes the first step toward a robotic system highly-skilled in its capacity of exploring the nearby space. The current knowledge on the neuroscience of vision-based reaching and grasping 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 is 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 use of binocular data and proprioceptive information regarding eye position, critical in the transformation of sensory data into appropriate motor signals. Moreover, compared with related approaches and previous works of EYESHOTS partners in the field, the model framework devised for Task 4.1 is especially focused on the duality of the dorsal visual stream. In fact the dorsal stream is composed of two parallel but communicating sub-streams, the dorso-medial pathway, mainly dedicated to proximal joints movement planning and reaching actions, and the dorso-lateral pathway, in charge of programming and monitoring distal joint movements required for grasping actions.
Fig.1 depicts the model framework defined in Task 4.1 and described in detail in the public Deliverable 4.1, which was obtained through the contribution of the UJI partner (mainly experienced in computational modelling for robotic purposes) and that of the UNIBO partner (highly knowledgeable of the neuroscience of the parietal cortex). The framework of Fig.1 depicts the main areas of the primate cortex dedicated to the interaction of a subject with his/her peripersonal space, and the fundamental connections between such areas. Next step is to achieve in a dynamic, active way, the sensorimotor knowledge of the environment postulated by the model.

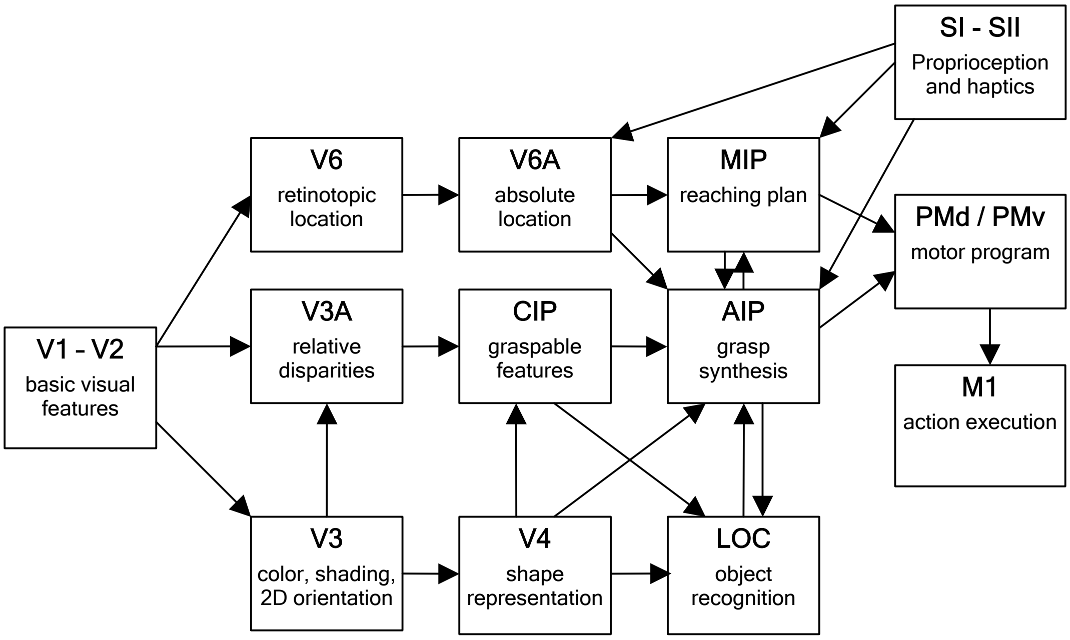


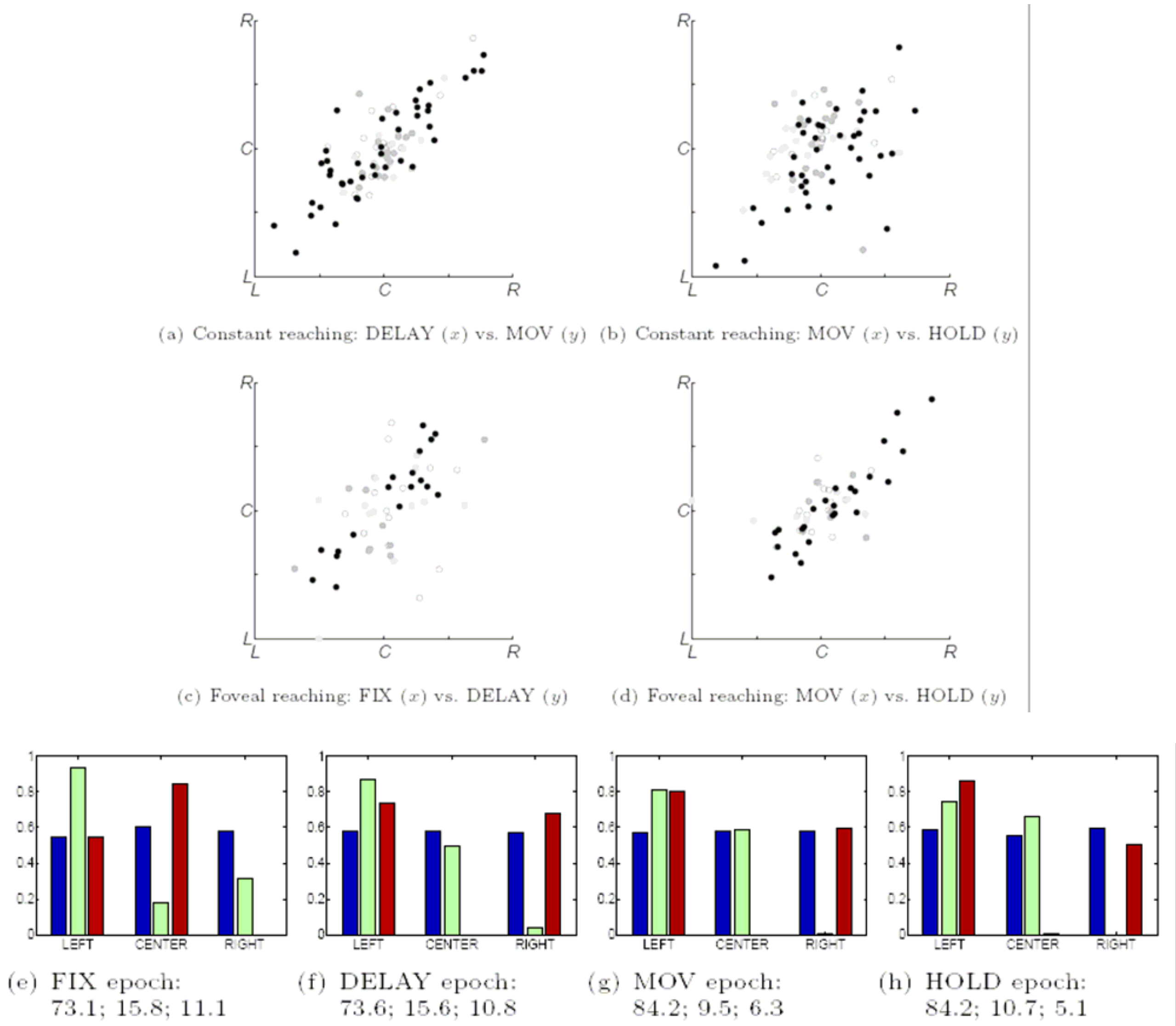
Figure 22: *Global model framework. The different information streams can be observed: ventral stream V3-V4-LOC, dorso-medial stream V6-V6A-MIP and dorso-lateral stream V3A-CIP-AIP.*

Task 4.2: Generating visuo-motor descriptors of reachable objects

The first goal in Task 4.2 is to achieve the dynamic, integrated sensorimotor representation described in Task 4.1. To this purpose, special focus has to be put on the integration between stereoptic retinal data with somatosensory information about object and arm state, in order to estimate object position and devise a reaching action plan as performed by area V6A in the dorsal stream. Ventral stream contribution at this stage regards mainly the experience of previous actions, as object identity recognition will be executed at a later stage in the project.

UJI has been actively interacting with other partners of EYESHOTS to achieve the above goal. The UNIBO partner provided single-cell data regarding experiments on monkeys reaching and gazing at different visual targets. These data are being analyzed and modelled by UJI (following UNIBO advice) 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. An example of the results obtained is shown in Fig.23. The correlation between different movement epochs in two experimental conditions is depicted in Fig.23 (a-d), whilst Fig.23 (e-h) shows the results of the Principal Components Analysis performed on four epochs of interest. The principal components obtained in this analysis constitute a first approximation for modelling the job of V6A neurons. Starting from such components, a population of artificial neurons can be generated which is able to emulate the sort of transformation and modulation between visual data and gaze and arm movements performed by the dorso-medial stream. The different properties captured in this work will be used to tune the behaviour of the neural population with various input sets corresponding to the different experimental conditions.

UJI members have been also devising, in strict collaboration with the colleagues of WWU, a common experimental framework useful for both WP4 and WP5 in order to coordinate research on bio-inspired robots and with human subjects. Such common experimental setup helps in guiding the development of the skills related to Task 4.2 and, later, 4.3, 5.3 and 5.4.



(a) Constant reaching: DELAY (x) vs. MOV (y) (b) Constant reaching: MOV (x) vs. HOLD (y)

(c) Foveal reaching: FIX (x) vs. DELAY (y) (d) Foveal reaching: MOV (x) vs. HOLD (y)

(e) FIX epoch: 73.1; 15.8; 11.1 (f) DELAY epoch: 73.6; 15.6; 10.8 (g) MOV epoch: 84.2; 9.5; 6.3 (h) HOLD epoch: 84.2; 10.7; 5.1

Figure 23: *Examples of the first results obtained by analysing data on reaching and gazing movements; (a-d) correlation analysis on neuron preferred direction (Left/Center/Right): comparison of preferred directions during one task (Constant or Foveal reaching) in different epochs (FIX/DELAY/MOV/HOLD), showing how neural responsiveness is related to action timing and conditions, and suggesting possible shared processes between epochs; (e-h) principal components analysis: normalized representation of the three eigenvectors obtained for each epoch during Foveal reaching protocol across conditions, with relative weights of the eigenvectors, showing a similar representation pattern for epochs DELAY, MOV and HOLD as compared to FIX.*

Deviations from the project workprogramme

The only significant change with respect to the original workplan is the tighter interaction that UJI is having with WWU, for what concerns the development of Tasks pertaining to both WP4 and WP5. Such interaction is being very helpful to the two groups in order to guide the first stages of the project taking into account from the beginning the future requirements of the robot skills on one side, and on human behavioural experiments on the other.

For what concerns the work planned for year 2, concerning Task 4.2 and the beginning of Task 4.3, it looks likely that more attention in the sensorimotor representation will be put on dorsal stream aspects. The ventral stream contribution will regard mainly memory issues, whilst object recognition will come in later stages of the project, also in accordance with the advancement of the interaction between WP4 and WP3.

WP5: Human behaviour and neural correlates of neural multisensory 3D representation

Leader: Patrizia Fattori (UNIBO)
Contributors and planned/actual effort (PMs) per participant:
UG (1/1), WWU (14/14), UNIBO (24/24), UJI (1/1)
Planned/actual Starting date: Month 1/1

Workpackage objectives

This Workpackage is devoted to the definition and the execution of specifically-designed neurophysiological and psychophysical experiments to study the human behavior of active perception and to find neural correlates of multisensory 3-D representation. Specific results of the different WP5 tasks will be used to implement computational models developed in other WPs, providing architectural guidelines for the organization of perceptual interactions and the production of artificial intelligent systems able to explore and interact with the 3D world.
In the first year the foreseen activity was the preparation of the experimental set-up and the realization of monkey training to perform fix-in-depth and reach-in-depth in controlled conditions.

Progress towards objectives

Task 5.1: ***Role of visual and oculomotor cues in the perception of 3D space.***
UNIBO group activities have been focused on experimental set-up preparation and on monkey training. This required a huge effort and to this purpose several people have been hired.
A new video-based eye tracking system was purchased and installed in the lab. The system was set in a binocular configuration and it allows to monitor the performance of the monkey in fixating the correct target during task execution. Using a custom-made algorithm, eye position signals are used to calculate on-line version and vergence angles and monitor the performance of the monkey.

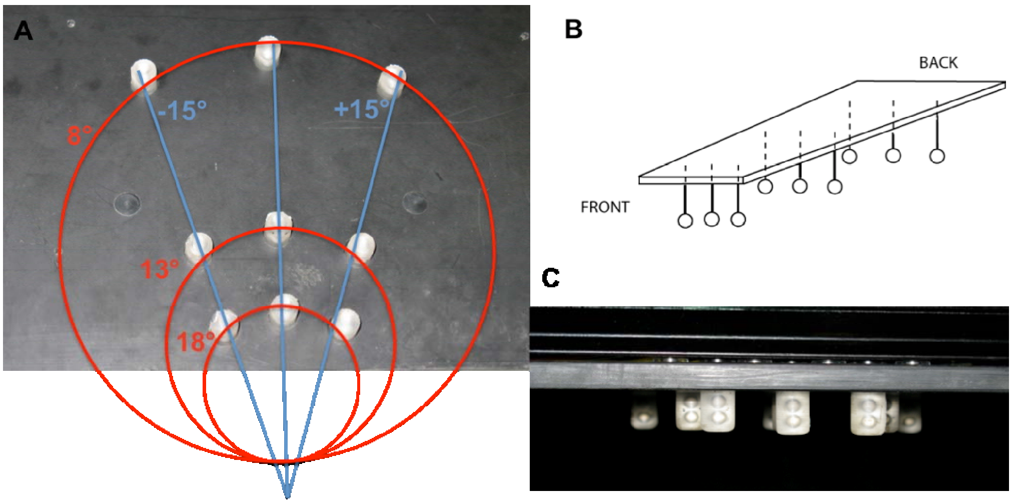


Figure 24: *The fixation-in-depth/reach-in-depth device. Each of the 9 targets contains a green/red bicolor LED use to instruct the monkey, and an embedded microswitch used to check target pressing. A) Bottom view of the device superimposed with isoversion (blue) and isovergence (red) lines. B) General prospect of the device. C) Frontal view of the device (as seen from the monkey side).*

A technician was hired to design and manufacture a fixation-in-depth and reach-in-depth device. The device allows us to present the monkey with a set of fixation points placed at different distances, thus varying version/vergence eye signals, in order to assess the role of these oculomotor cues in the perception of the 3D space. The fixation/reaching targets are placed in a 3x3 grid on a horizontal plane, at eye level of the

monkey. The position of the targets, shown in Fig.24, was chosen in collaboration with the UG group taking into account some physical constraints (maximal distance reachable by the monkey, vergence angle not too demanding). The device was ready at month 10.
Some technical trials were done in order to identify a protocol for visual stimulations.
At the same time (month 10), the surgery necessary for electrophysiological experiments was performed, and the training of the monkey for fixation-in-depth task began. At the end of the reporting period, the monkey is able to execute the task performing a saccadic eye movement toward the correct position and keeping a steady fixation of the target for 2-2.5 s.
A pilot study on the influence of eye movement signals in 3D space on the discharge of neurons in the medial parieto-occipital cortex has been performed using a prototype of the fixation-in-depth device. The neural activity of about 50 cells was recorded while the monkey was fixating targets placed at 5 different depths in front of him. Figure 25 shows the responses of some of the recorded neurons. Cumulative peri-stimulus time histograms (PSTHs) of neural activity of a neuron in medial parietal occipital cortex for nearest to farthest fixation position (bottom to top) are shown in Fig.25A. PSTHs are aligned on the beginning of fixation. It is evident that the neuron is differently activated during fixation, depending on the gazed position. In particular, the firing rate of the neuron was higher when the monkey was fixating at medium-far distances. In Fig.25B the responses of other three neurons are shown. The responses corresponding to the five fixation distances are superimposed and represented as spike density functions. Dark to light color tones represent near to far positions. The neuron in the top was not modulated by the task, the second one showed a phasic activation mainly for far distances, the third one showed both a phasic and a tonic activation mainly for near fixation distances. Quantitative data analysis at single cell and population levels will take place in the second year of the project.

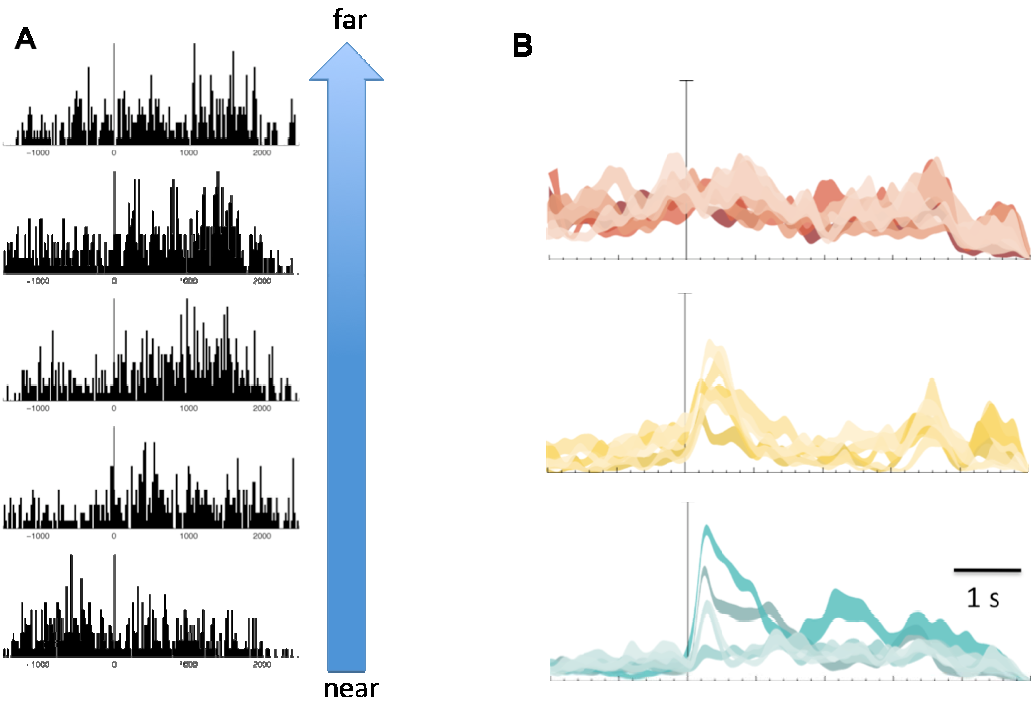


Figure 25: *Responses of neurons in the medial parietal occipital cortex during a fixation task with five different fixation distances. See text.*

At the end of this reporting period the training of the monkey in the fixation-in depth task has been almost completed, this will permit to begin electrophyisological recording slightly in advance with respect to workplan.

Other analyses have been performed in the first year of EYESHOTS, in order to allow other partners to perform their tasks: in collaboration with UG, 2D gaze-dependent modulations in medial parieto-occipital cortex have been analysed and modelled. To study this aspect, we selected the data from about one-hundred neurons, which were tested with a fixation paradigm. These data have been accepted as a joint publication with partner UG at the IWINAC 2009 conference (Breveglieri et al., 2009 [C6]).

Task 5.2: ***Link across fragments.***

This task is aimed at studying neural correlates of multisensory representation of 3D space obtained through active ocular and arm movements.
Most of the preparatory activities for this task (experimental set-up, monkey training) were performed in common with task 5.1, thus the description of these activities can be found in the previous section.
The monkey training for reach-in-depth task was performed immediately after the training for the fixation-in-depth task. The time sequence of the reach-in depth task is sketched in Fig.26. The monkey sits in a primate chair in front of the reach-in-depth device. The monkey presses the start button placed near its belly, outside its field of view. After a delay, one of the target lights up green, and the monkey has to perform a saccadic eye movement towards the target and to adjust its vergence. After a variable fixation period, the fixation target turns red. This is the go signal for the monkey to release the start button and perform a reaching movement toward the fixated target. The monkey has to push the target, and to keep its hand on it until the fixation LED switches off. The monkey releases the target and performs a backward movement toward the start button to be rewarded.
Electrophysiological recording sessions will start at the beginning of the second reporting period, a couple of months earlier than planned.

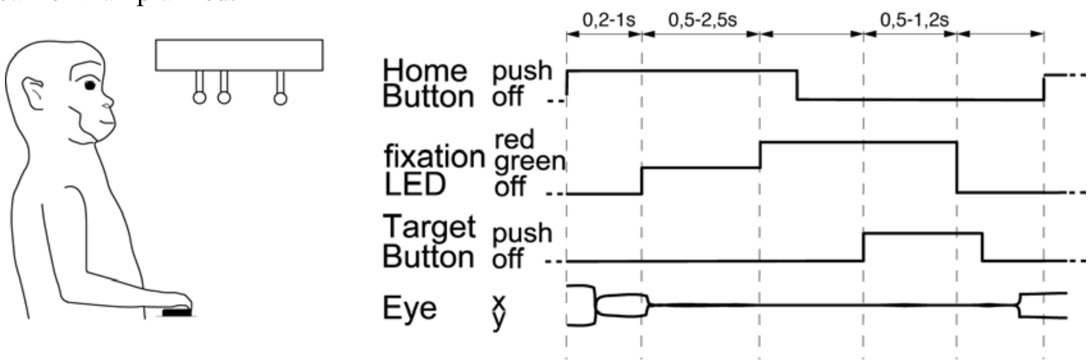


Figure 26: *Timing sequence of the reach-in-depth task.*

In addition, in collaboration with UJI we are analysing the role of visual and proprioceptive guidance of reaching movements in the medial parieto-occipital cortex. Data from 75 neurons tested with a delayed arm-reaching task both in darkness and in full light are available, analysis are going on. These data will be presented at international conferences during the second year of EYESHOTS. Moreover, we collaborated with UJI group in the analysis of single-cell data regarding reaching experiments with different spatial/retinotopic position of the targets. The results of these analysis have been accepted as joint publications with partner UJI at the IWINAC 2009 conference (Chinellato et al., 2009a [C3]; Chinellato et al., 2009b [C4]).

Task 5.3: ***Motor description of fragment location***
This work package uses saccadic adaptation to measure the contribution of motor parameters to fragment location. In saccadic adaptation, the relationship between saccade motor parameters and saccade target location are changed. The saccade target is systematically displaced during execution of the saccade. This displacement induces a visual error after the saccade, which, over the course of successive trials, leads to a change of the amplitude of the primary saccade. We study how the perceived location of objects surrounding the saccade target are affected by the adaptation.

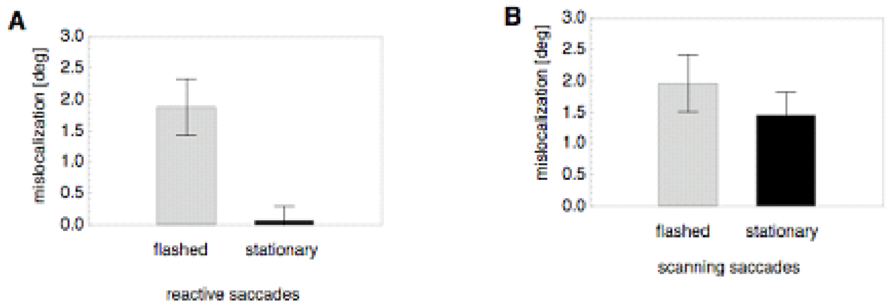


Figure 27: *Results from study I on mislocalization of targets for reactive and scanning saccades. A mislocalization of flashed and stationary targets after adaptation of reactive saccades B Analog result for scanning saccades.*

A first study determined the transfer of spatial information provided several hundred milliseconds before a saccade to the perceived space after the saccade. In this temporal window, a mislocalization of targets in the spatial neighborhood of the saccade target occurs. We used the differences between reactive and scanning saccades to investigate the interplay between visual localization and saccade targeting. Reactive saccade are elicited by suddenly appearing targets. Scanning saccades are executed within a group of targets which are constantly visible. These two types of saccades are performed by partially different brain pathways. To determine the influence of the pathways of saccadic adaptation on mislocalization, corresponding localization stimuli were used, i.e. flashed stimuli for reactive saccades and stationary stimuli for scanning saccades. Thus, adaptation experiments were conducted separately with reactive and scanning saccades, and, in both cases, mislocalization was tested with flashed and stationary targets. As depicted in Fig.27, reactive saccade adaptation induces mislocalization for flashed but not for stationary probes, whereas scanning saccade adaptation induces mislocalization for both flashed and stationary probes. A strong influence of the pathways of adaptation on mislocalization could therefore be demonstrated. Detailed results from this study will be reported in deliverable D5.3a and in an associated scientific publication that has been submitted in Feb. 2009.

In a second study, the visual localization was decoupled from the motor execution. Human subjects first underwent saccadic adaptation. Afterwards, the adaptation induced mislocalization of flashed targets was tested while subjects kept fixation. In this case, the localization task itself did not contain a motor component, and influences of motor learning (adaptation) on perceived space could be measured in isolation. The results show that motor adaptation has indeed an influence on the representation of spatial location during fixation. The strength of this influence depends on adaptation direction (forward vs. backward) and on the size of the visual error after the adapting saccade. Since the dependence on adaptation direction is very interesting also in conjunction with recent publication on adaptation (see update on state-of-the-art) we plan to continue with dedicated studies on forward adaptation. The results of the present study are currently prepared for a scientific publication and will be reported in detail in deliverable D5.3a.

With respect to the adaptation experiment in the monkey, which is planned to begin in the middle of the next reporting period, the Bologna and Münster groups jointly decided on the setup and procedure. The experiment will use a touchscreen setup that will be constructed and programmed by the Münster group, and when ready implemented in Bologna to run the experiment. Data analysis will be done jointly between the groups.

Also, in the reporting period a paper on saccade adaptation (Georg & Lappe, 2009 [J4]) appeared. While most of this data was collected before the project officially started some parts of the data analysis and one additional experiment was performed during the reporting period.

In summary, in Task 5.3 we have collected and analyzed data on the influence of motor and visual parameters on object localization obtained from saccade adaptation data. This data is needed for milestone M4 at month 12. The milestone was reached as expected. Our hypothesis that saccade adaptation modifies perceived location of saccade goals was confirmed by the experiments. Detailed results will be reported in deliverable D5.3a at month 15.

Task 5.4: ***Predicting behaviour and cooperation in shared workspace***

This task concerns the understanding of the sequence of allocation of attention, direction of gaze, and movement of the arm of a human cooperation partner. The understanding of these issues will allow the anticipation of particular actions based on the partner's behavior. In order to achieve this objective, we started developing two setups on which a series of experiments can be conducted.

The first setup is based on the Eyelink II eye tracker system (SR Research Ltd., Mississauga, Ontario, Canada) and with this setup we already started collecting data. The experiment itself requires the participant to watch a series of movies in which an actor performs gaze movements and reaching arm movement towards a set of targets. The participants have to identify the target of the actor's gazing and reaching movements. The participants had to look at the to be pointed target as soon as possible and these eye movements were recorded with the Eyelink eye tracker. Two factors were manipulated in this experiment: the presence of the gazing behaviour (actor's eyes visible/hidden) and the visibility of the target objects (targets visible/hidden). The data show that the actor's gaze can trigger a rapid and accurate response towards the target object. Participants were able to identify the target objects when the arm was still at rest at the beginning of its trajectory. On the other hand, when the actor's gaze information was not available, the participants' gaze still leaded the hand movements of the actor, but was comparatively slower in identifying the target object. In this case, an increment in the number of saccades was also observed. The second factor, the visibility of the targets, had mainly an effect on the spatial accuracy: when the targets were visible, the end-points of the saccades were centered on the targets, whereas in the opposite case, the end-points were more scattered. Figure 28 represents the temporal advantage (Δt) in identifying the pointed target when the actor's gaze is available to the participants. These preliminary findings support our hypothesis that other's gaze direction is an essential predictive cue about the final location of a pointing movement.

The second setup is based on two ViewPoint eye trackers systems (Arrington Research Inc., Scottsdale, AZ). At the moment the development of this setup is in a testing phase. This setup will allow the simultaneous recording of eye movements of two interacting participants.

These two parallel work streams are directed towards the achievement of the milestones M9.ante and M9 at months 18 and 27.

In addition, a meeting was organised with UJI (Feb 2009) to discuss the issues related to the objective of human-robot interaction. Preliminary decisions were made about which eye tracker to interface with the robot, and several different tasks that could be used were discussed.

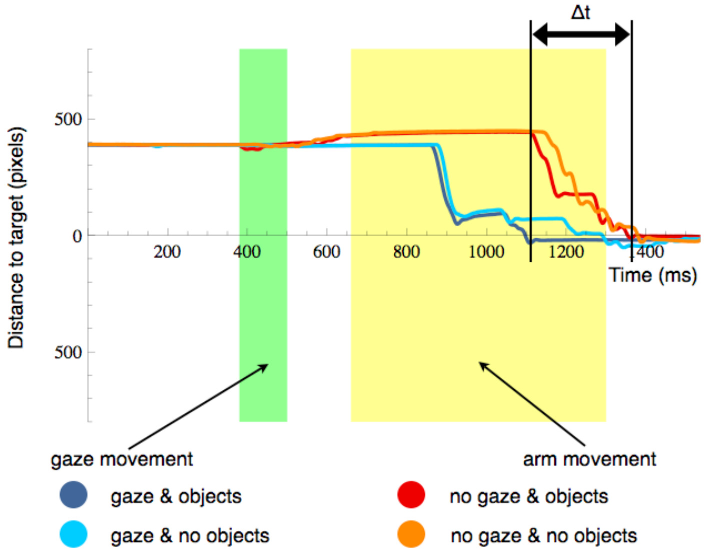


Figure 28: *Time to identify the correct target object. The continuous lines indicate the distance of the participant's eye position from a specific target object. When the line is close to zero it means that the gaze is located on the target object. The different colors specify the different conditions. The green rectangle shows the time slot in which the actor's gaze movement was performed, whereas the yellow rectangle shows the time slot in which the actor's arm movement was performed.*

Deviations from the project workprogramme

None

4 Deliverables and milestones tables

Deliverables (excluding the periodic and final reports)

TABLE 1. DELIVERABLES									
Del. no.	**Deliverable name**	**WP no.**	**Lead beneficiary**	**Nature**	**Dissemination level**	**Delivery date from Annex I (proj month)**	**Delivered Yes/No**	**Actual / Forecast delivery date**	**Comments**
D6.1	Launch of the project web-site	WP6	UG	D	PU	3, updates	Yes	31-May-08	
D7.1	Creation, composition and publication of web pages	WP7	UG	D	PU	3, updates	Yes	31-May-08	
D8.1	Literature database	WP8	UG	D	PU	3	Yes	31-May-08	
D4.1	Description of integrated representation	WP4	UJI	R	PU	9	Yes	02-Dec-08	
D1.1	Binocular eye coordination and its role in depth vision	WP1	UG	R	PU	12	Yes	7-Mar-09	
D3.1a	Demonstration of learning disparity-tuned feature selective cells	WP3	WWU	O	PU	12	Yes	3-Mar-09	

Milestones

TABLE 2. MILESTONES							
Milestone no.	**Milestone name**	**Work package no.**	**Lead beneficiary**	**Delivery date from Annex I**	**Achieved Yes/No**	**Actual / Forecast achievement date**	**Comments**
M1	Low-level automatic servos based on primary disparity information	WP1	UG	6	Yes	September 2008	
M2	Merging of action and perception	WP4	UJI	9	Yes	November 2008	
M3	Learning algorithm for bi-directionally connected disparity tuned feature-selective cells	WP3	WWU	12	Yes	February 2009	
M4	Experimental data on fragment location in humans obtained	WP5	WWU	12	Yes	February 2009	
M6.ante	Experimental set-up	WP5	UNIBO	8	Yes	October 2008	

5 Project management

5.1 Management activities

No unforeseen management issues have arisen. The management of the project is in good order and cooperations among partners are excellent.
All the planned activities are in time, and in some case, in advance with respect to the planned schedule.
No deviations from the planned milestones and deliverables occurred.

Coordination activities during the period

- Development and maintenance of the EYESHOTS web-site (with a public area and a private area with restricted access to the consortium's members).
- Set-up and maintenance of a repository for sharing documents, data, and software tools in the private area of the web-site. Material from meetings - agenda, presentations, minutes etc is made available to the consortium using this facility, as well.
- Creation of a Literature Database with bibliography list and source/access information of the basic and relevant literature from computer science, vision, sensing and motor control as well as learning. The EYESHOTS literature database features an ever-growing collection of references related to the research activities conducted in the project and provides a common basis for teaching and education of students.
- Communication with and between partners to ensure on-time schedule of the project and request of approvals for specific formal decisions. For general communication an email list is hosted by UG (eyeshots@unige.it).
- Collection of brief notes on subgroup meetings (available on the private area of the EYESHOTS website, http://www.eyeshots.it/private/minutes.php) to ensure a continuous monitoring of cooperative work towards the objectives.
- Communication, through the mailing list and the web-site, of forthcoming events, Call for Papers, News and Highlights from the scientific community on topics related to the project's research activities.
- Periodically, publications from the consortium have been collected and published in the website, together with the public deliverables due in the period.
- Evaluation of synergies with other projects.
- Regular reporting to the Commission as planned.
- Participation to the 1st FP7/ICT Coordinator Day 17 December, 2008 Bruxelles.

List of project meetings and cross-visits

The work has included a rich set of meetings (which started with a warm-up event before the official starting date of the project) to bring the consortium together to address the common research plan, and a rich set of partner-partner collaborations.

The following plenary project meetings were organized to discuss the progress across partners and work packages:

- ***Warm-up meeting*, 30-31 January 2008, Genoa.**
  Participants: Giorgio Cannata, Angel del Pobil, Patrizia Fattori, Claudio Galletti, Markus Lappe, Silvio Sabatini, Fabio Solari, Marc van Hulle, Andrea Canessa, Manuela Chessa, Nick Chumerin, Giulia Gastaldi, Agostino Gibaldi, Nicoletta Marzocchi, Karl Pauwels, Jayathu Samarawikrama.
- ***Kick-off meeting*, 7-8 March 2008, Bologna.**
  Participants: Silvio Sabatini, Giorgio Cannata, Manuela Chessa, Andrea Canessa, Agostino Gibaldi, Andrea Trabucco, Fred Hamker, Markus Lappe, Mark Voss, Eckart Zimmermann, Katharina Havermann; Claudio Galletti, Patrizia Fattori, Nicoletta Marzocchi, Rossella Breveglieri, Michela Gamberoni; Angel del Pobil, Ester Martinez, Beata Joanna Grzyb, Eris Chinellato ; Marc van Hulle, Nick Chumerin, Karl Pauwels.
  The objectives of the project were reviewed and the expectations for the first year were detailed. In addition, each principal investigator gave a tutorial presentation of the different research background

of the project and to provide a view of the competencies available. The meeting also involved early discussions of collaborations and joint meetings.
The tutorial sessions in the programme of the kick-off meeting have served as a first key event for training.

- ***month 6 periodic meeting*, 29-30 Sepetember 2008, Castellon de la Plana.**
Participants: Silvio Sabatini, Giorgio Cannata, Manuela Chessa, Andrea Canessa, Agostino Gibaldi, Yang Zhang, Fred Hamker, Markus Lappe, Mark-André Voss, Katharina Havermann, Robert Volcic, Claudio Galletti, Patrizia Fattori, Nicoletta Marzocchi, Rossella Breveglieri, Michela Gamberoni, Angel del Pobil, Ester Martinez, Beata Joanna Grzyb, Eris Chinellato, Nick Chumerin, Karl Pauwels.
- ***month 12 periodic meeting*, 9-10 February 2009, Muenster.**
Participants: Giorgio Cannata, Eris Chinellato, Patrizia Fattori, Fred Hamker, Markus Lappe, Silvio Sabatini, Marc van Hulle. Katharina Havermann, Robert Volcic, Frederik Beuth, Andrea Canessa, Agostino Gibaldi, Manuela Chessa, Nicoletta Marzocchi, Ester Martinez, Karl Pauwels, Nick Chumerin, Beata Grzyb.

In addition, subgroup meetings took place at the fringes of the periodic plenary councils or through specific cross-visits. We believe that the easiest way to transfer results from one institution to another is through an exchange of persons for a shorter or longer period including transfer of software as part of this.

- ***WP4/WP5 Specific subgroup meetings***

Bologna, 06/03/2008
UNIBO: Patrizia Fattori, Nicoletta Marzocchi, Rossella Breveglieri, Claudio Galletti
UJI: Angel del Pobil, Eris Chinellato, Beata Grzyb, Ester Martinez.
Discussed topics: common research framework, possible future experimental setup, coordinated data analysis.
Decisions: UNIBO approach to reaching-in-depth experiments; UJI analysis of previous and future single-cell data.

Münster, 11/02/2009
WWU: Robert Volcis, Markus Lappe.
UJI: Eris Chinellato.
Discussed topics: final setup of human-robot interaction experiments; interfacing issues; next and future WWU experiments; UJI subgoals to reach final robot skills.
Decisions: possible final collaborative setup; future tasks and coordinated work plan for both WWU and UJI in order to achieve such setup.

- ***WP1/WP5 Specific subgroup meeting***

Castellon de la Plana, 30/09/2008
UG: Giorgio Cannata, Silvio Sabatini, Andrea Canessa
WWU: Markus Lappe, Katharina Havermann
Discussed topics: definition of specific experiments for measuring the vergence/version components during short saccades in depth.
Decisions: WWU will start devising a proper experimental set-up.

The following visits took place during the first reporting period:

- Ester Martinez (UJI) stayed in UG from Jun 2008 to Dec 2008, working on integration between WP1 and WP2 with WP4.
- Manuela Chessa (UG) visited K.U.Leuven on 23-28 Jun 2008 to work on Tasks 2.1 and 2.2
- Silvio Sabatini (UG) visited K.U.Leuven on 31 March 2008 for WP2 activity coordination.
Silvio Sabatini (UG) visited UJI on 19 June 2008 for WP1-WP4 activity coordination.

5.2 Dissemination and use of the knowledge

Development of the project web-site

The project web-site (www.eyeshots.it) was launched on month 3 and it continuously updated. The early availability of a web site was essential for community wide dissemination. The public area of the web-site features, among others:

- An overview of Project's objectives (goals and basic ideas of the project).
- A description of the consortium and its expertise
- A list of the papers published by the consortium on results from the project's research activities
- An access to the bibliography database (read-only access, but not to the PDFs). The database is powered by a GPL web-based bibliography management system (Aigaion, www.aigaion.nl).

Specific pages in the public area the project web-site are dedicated to publish "research news" to keep the site up-to-date on breaking results from the different labs of the consortium. These "research dispatches" would be considered as a shop-window of the project.

List of publications

A total of 6 journal papers and 10 conference contributions (3 of them co-authored) have been published.

Journal papers [J]

[J1] Wiltschut J., Hamker F.H. Efficent coding correlates with spatial frequency tuning in model of V1 receptive field organization. Visual Neuroscience, Vol 26:21-34, 2009

[J2] Lappe M. What is adapted in saccadic adaptation? Journal of Physiology. Vol 587(1), 2009.

[J3] Hamker F.H., Zirnsak M. About the influence of post-saccadic mechanisms for visual stability on peri-saccadic compression of object location. Journal of Vision. Vol 8(14), 2008.

[J4] Georg K., Lappe M. Effects of saccadic adaptation on visual localization before and during saccades. Experimental Brain Research. Vol 192:9-23, 2009.

[J5] Hamker F.H., Zirnsak M., Calow D., Lappe M. The Peri-Saccadic Perception of Objects. PLoS Computational Biology. Vol 4(2), 2008.

[J6] Vitay, J., Fix, J., Beuth, F., Schroll, H., Hamker, F. H. Biological Models of Reinforcement Learning. Accepted for publication in the journal KI.

Conference papers [C]

[C1] Gibaldi A., Chessa M., Canessa A., Sabatini S.P., Solari F. Reading binocular energy population codes for short-latency disparity-vergence eye movements. 13th Int. Conference on Cognitive and Neural Systems, Boston, MA, USA, May 27-30, 2009.

[C2] Chessa M., Canessa A., Gibaldi A., Solari F., Sabatini S.P. Embedding fixation constraints into binocular energy-based models of depth perception. 13th Int. Conference on Cognitive and Neural Systems, Boston, MA, USA, May 27-30, 2009.

[C3] Chinellato E., Grzyb B.J., Fattori P., del Pobil A.P. (2009) Toward an Integrated Visuomotor Representation of the Peripersonal Space. International Work-conference on the Interplay between Natural and Artificial Computation, Santiago de Compostela, Spain June, 22-26, 2009

[C4] Chinellato E., Grzyb B.J., Marzocchi N., Bosco A., Fattori P., del Pobil A.P. (2009) Eye-hand coordination for reaching in dorsal stream area V6A: Computational lessons. International Work-conference on the Interplay between Natural and Artificial Computation, Santiago de Compostela, Spain June, 22-26, 2009

[C5] Fattori P., Bosco A., Breveglieri R., Marzocchi N., Galletti C. Visual and somatosensory guidance of reaching movements in the medial parieto-occipital cortex of the macaque. VSS2009 meeting Vision Sciences Society, Naples, FL, USA, May 2009.

[C6] Breveglieri R., Bosco A., Canessa A., Fattori P., Sabatini S.P. Evidence for Peak-shaped Gaze Fields in Area V6A: Implications for Sensorimotor Transformations in Reaching Tasks. International Work-conference on the Interplay between Natural and Artificial Computation, Santiago de Compostela, Spain June, 22-26, 2009

[C7] 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), Bruges, Belgium, April 2009.
[C8] Chessa M., Solari F., Sabatini S.P. A Virtual Reality Simulator for Active Stereo Vision System. International Conference on Computer Vision Theory and Applications 2009,VISAPP'09, Lisbon, Portugal 5-8 February 2009.
[C9] Pauwels K., Van Hulle M.M. Realtime Phase-based Optical Flow on the GPU, in: Proceedings of the CVPR'08 Workshop on Computer Vision on GPU, Anchorage, Alaska, June 23, 2008.
[C10] Vitay J., Hamker F.H. Binding objects to cognition: A brain-like systems approach to the cognitive control of visual perception. International Conference on Cognitive Systems (CogSys 2008), Karlsruhe, Germany, (in press).

Other dissemination activities

- Presentation of the project at the CogSys 2008 Conference (Karlsruhe, Germany, 2-4 April 2008).
- Preparation of flyer of the project (July 2008):
  ftp://ftp.cordis.europa.eu/pub/ist/docs/cognition/eyeshots-flyer_en.pdf

Events organized by the project or planned for the next period

- Partner UJI has co-organized a special session on "Robotics and Neuroscience" within the 3rd International Work-conference on the Interplay between Natural and Artificial Computation (IWINAC'09), Santiago de Compostela, Galicia, (Spain) June 22-26, 2009.
- A summer school was planned for the second year for Sep 2009 in Castellon (Spain), within the scheme of the International UJI Robotics Summer School (IURS) of the UJI partner.

List of exploitable results

None in the first period.

6 Explanation of the use of the resources

6.1 Justification of major cost items and resources

The following tables list the major cost items (one table by participant).

Table 6.1a Personnel, subcontracting and other major Direct cost items for Beneficiary 1 (UG) for the period

Work Package	Item description	Amount	Explanations
All	Personnel costs	€ 169,310.28	Costs of the worked person months for the permanent staff working on the project. Manuela Chessa also worked on the project, namely for 6 months, but she has her own funding.
	Subcontracting		
WP1, WP2, WP3	Major cost item 'a' (Equipment)	€ 1,000.00	Depreciation charge for: 3D laser scanner (Konica Minolta Vivid 910). Total cost of the equipment: € 30,000.00.
WP1	Major cost item 'b' (Consumable)	€ 1,476.00	ELMO motion control board
	Remaining direct costs	€ 15,375.21	
	TOTAL DIRECT COSTS	€ 187,161.49	

Table 6.1b Personnel, subcontracting and other major Direct cost items for Beneficiary 2 (WWU) for the period

Work Package	Item description	Amount	Explanations
WP3, WP5	Personnel costs	€ 143,408.48	Personnel costs include worked person months for the permanent personnel and postoctoral students
	Subcontracting		
WP5	Major cost item 'c' (Equipment)	€ 5,600.00	Depreciation charge for: Eyetracker. Total cost of the equipment: € 28.000,00
	Remaining direct costs	€ 17,124.05	
	TOTAL DIRECT COSTS	€ 166,132.53	

Table 6.1c Personnel, subcontracting and other major Direct cost items for Beneficiary 3 (UNIBO) for the period

Work Package	Item description	Amount	Explanations
WP4, WP5	Personnel costs	€ 63,259.91	Personnel costs include 28 person months, of which 7 of permanent personnel and 21 of temporary personnel devoted to the project
	Subcontracting		
WP5	Major cost item 'd' (Equipment)	€ 9,029.46	Depreciation charge for: Cyberkinetics Cerebus multi-channel data acquisition system. Total cost of the equipment: € 27088.38.
WP5	Major cost item 'e' (Equipment)	€ 5,481.75	Depreciation charge for: ETL - 200 Primate Eye Tracking lab with Binocular Upgrade to ETL-200. Total cost of the equipment: € 17160.26
WP5	Major cost item 'f' (Other costs)	€ 1,404.00	Clamps for micromanipulators (World Precision Instruments)
WP4	Major cost item 'g' (Other costs)	€ 1,938.77	Software (SER DATA)
WP5	Major cost item 'h' (Other costs)	€ 3,362.44	Electrodes (Thomas Recording)
WP5	Major cost item 'i' (Other costs)	€ 1,594.23	Lens (Translucent Technologies)
	Remaining direct costs	€ 13,059.13	
	TOTAL DIRECT COSTS	€ 99,129.69	

Table 6.1d Personnel, subcontracting and other major Direct cost items for Beneficiary 4 (UJI) for the period			
Work Package	Item description	Amount	Explanations
WP1, WP4, WP5	Personnel costs	€ 63,960.52	Costs of the person months for the permanent staff working on the project.
	Subcontracting		
WP1,WP4	Major cost item 'j' (travel)	€ 9,085.78	12 individual travels, only one over 1000euros (1028euros)
WP4	Major cost item 'k' (other costs)	€ 1,994.95	5 Conference Registrations
	Remaining direct costs	€ 0.00	
	TOTAL DIRECT COSTS	€ 75,041.25	

Table 6.1e Personnel, subcontracting and other major Direct cost items for Beneficiary 5 (K.U.Leuven) for the period			
Work Package	Item description	Amount	Explanations
WP2	Personnel costs	€ 23,935.56	N. Chumerin also worked on the project, namely for 8 months, but he has his own funding. K. Pauwels has been hired on the project from 1.11.08 onwards. His wage cost is charged to the project.
	Subcontracting		
	Remaining direct costs	€ 10,463.65	
	TOTAL DIRECT COSTS	€ 34,399.21	

6.2 Budgeted versus Actual Costs

Tabular overview of budgeted costs and actual costs, by beneficiary and by major cost item including personnel. The budgeted costs are taken from the Annex I.

Cost Budget Follow-up Table					*) total budget figures - not EC funding				
Contract N°:	217077	Acronym: EYESHOTS					Date:		
BENEFIC-ICIARY	TYPE of EXPENDITURE (as defined by participants)	BUDGET	ACTUAL COSTS (EUR)					Pct. spent	Remaining Budget (EUR)
			Period 1	Period 2	Period 3	Period 4	Total	Total	
		e	a1	b1	c1	d1	e1	(a1+b1+c1+d1)/e	e-e1
Part. 1	**Total Person-month**	**117**	38.44	0.00	0	0	**38.44**	33%	**78.56**
UG	Personnel costs	**414,029.88**	169,310	0.00	0.00	0.00	**169,310**	41%	**244,719.88**
	Travel costs	**36,000.00**	9,390	0.00	0.00	0.00	**9,390**	26%	**26,610.00**
	Durable Equipment (Depreciation)	**33,930.28**	4,005	0.00	0.00	0.00	**4,005**	12%	**29,925.28**
	Subcontracting/Audit	**1,863.38**	0	0.00	0.00	0.00	**0**	0%	**1,863.38**
	Other costs ('the rest')	**39,788.89**	4,456	0.00	0.00	0.00	**4,456**	11%	**35,332.89**
	Overheads	**314,249.43**	112,296	0.00	0.00	0.00	**112,296**	36%	**201,953.43**
	Total Costs	**839,861.86**	**299,457**	**0.00**	**0.00**	**0.00**	**299,457**	**36%**	**540,404.86**
Part. 2	**Total Person-month**	**103.00**	32	0.00	0.00	0.00	**32.00**	31%	**71.00**
WWU	Personnel costs	**453,200.00**	143,409	0.00	0.00	0.00	**143,409**	32%	**309,791.00**
	Travel costs	**29000**	8,130	0.00	0.00	0.00	**8,130**	28%	**20,870.00**
	Durable Equipment (Depreciation)	**38000**	10,144	0.00	0.00	0.00	**10,144**	27%	**27,856.00**
	Subcontracting/Audit	**2,000.00**	0	0.00	0.00	0.00	**0**	0%	**2,000.00**
	Other costs ('the rest')	**2200**	4,450	0.00	0.00	0.00	**4,450**	202%	**-2,250.00**
	Overheads	**313,440.00**	99,679	0.00	0.00	0.00	**99,679**	32%	**213,761.00**
	Total Costs	**837,840.00**	**265,812**	**0.00**	**0.00**	**0.00**	**265,812**	**32%**	**572,028.00**
Part. 3	**Total Person-month**	**96.00**	28.00	0	0	0	**28**	29%	**68.00**
UNIBO	Personnel costs	**196,800.00**	63,260	0.00	0.00	0.00	**63,260**	32%	**133,540.00**
	Travel costs	**26,360.00**	2,392	0.00	0.00	0.00	**2,392**	9%	**23,968.00**
	Durable Equipment (Depreciation)	**26,000.00**	18,015	0.00	0.00	0.00	**18,015**	69%	**7,985.00**
	Subcontracting/Audit	**0.00**	0	0.00	0.00	0.00	**0**	0%	**0.00**
	Other costs ('the rest')	**36,000.00**	15,462	0.00	0.00	0.00	**15,462**	43%	**20,538.00**
	Overheads	**171,096.00**	59,477	0.00	0.00	0.00	**59,477**	35%	**111,619.00**
	Total Costs	**456,256.00**	**158,606**	**0.00**	**0.00**	**0.00**	**158,606**	**35%**	**297,650.00**
Part. 4	**Total Person-month**	**52.00**	15.66	0.00	0.00	0.00	**15.66**	30%	**36.34**
UJI	Personnel costs	**257,712.00**	63,961	0.00	0.00	0.00	**63,961**	25%	**193,751.00**
	Travel costs	**27,000.00**	9,086	0.00	0.00	0.00	**9,086**	34%	**17,914.00**
	Durable Equipment (Depreciation)	**20,000.00**	0	0.00	0.00	0.00	**0**	0%	**20,000.00**
	Subcontracting/Audit	**2,979.00**	0	0.00	0.00	0.00	**0**	0%	**2,979.00**
	Other costs ('the rest')	**4,888.80**	1,995	0.00	0.00	0.00	**1,995**	41%	**2,893.80**
	Overheads	**185,760.48**	45,023	0.00	0.00	0.00	**45,023**	24%	**140,737.48**
	Total Costs	**498,340.28**	**120,065**	**0.00**	**0.00**	**0.00**	**120,065**	**24%**	**378,275.28**
Part. 5	**Total Person-month**	**47.50**	4.00	0.00	0.00	0.00	**4.00**	8%	**43.50**
K.U. Luven	Personnel costs	**309,552.00**	23,936	0.00	0.00	0.00	**23,936**	8%	**285,616.00**
	Travel costs	**10,000.00**	6,289	0.00	0.00	0.00	**6,289**	63%	**3,711.00**
	Durable Equipment (Depreciation)	**5,000.00**	4,175	0.00	0.00	0.00	**4,175**	84%	**825.00**
	Subcontracting/Audit	**1,000.00**	0	0.00	0.00	0.00	**0**	0%	**1,000.00**
	Other costs ('the rest')	**9,782.00**	0	0.00	0.00	0.00	**0**	0%	**9,782.00**
	Overheads	**200,600.40**	20,639	0.00	0.00	0.00	**20,639**	10%	**179,961.40**
	Total Costs	**535,934.40**	**55,039**	**0.00**	**0.00**	**0.00**	**55,039**	**10%**	**480,895.40**
TOTAL	**Total Person-month**	**415.50**	118.10	0.00	0.00	0.00	**118.10**	28%	**297.40**
	Personnel costs	**1,631,293.88**	463,876	0.00	0.00	0.00	**463,876**	28%	**1,167,417.88**
	Travel costs	**128,360.00**	35,287	0.00	0.00	0.00	**35,287**	27%	**93,073.00**
	Durable Equipment (Depreciation)	**122,930.28**	36,339	0.00	0.00	0.00	**36,339**	30%	**86,591.28**
	Subcontracting/Audit	**7,842.38**	0	0.00	0.00	0.00	**0**	0%	**7,842.38**
	Other costs ('the rest')	**92,659.69**	26,363	0.00	0.00	0.00	**26,363**	28%	**66,296.69**
	Overheads	**1,185,146.31**	337,114	0.00	0.00	0.00	**337,114**	28%	**848,032.31**
	Total Costs	**3,168,232.54**	**898,979**	**0.00**	**0.00**	**0.00**	**898,979**	**28%**	**2,269,253.54**

6.3 Planned versus Actual effort

Tabular overview of planned person-months and actual person-months, by beneficiary and by work package. The planned person-months for the period are gathered from the Annex I.

TABLE 6.3 : PERSON-MONTHS STATUS TABLE									
CONTRACT N°:	217077		**Partner - Person-month per Workpackage**						
ACRONYM:	EYESHOTS								
PERIOD:	01.03.2008 - 28.02.2009								
			TOTALS	**Coord. UG**	**WWU**	**UNIBO**	**UJI**	**K.U.Leuven**	
Workpackage 1:	Eye movements for exploration of the 3D space	**Actual WP total:**	**17.0**	15	0	0	2	0	
		Planned WP total:	**16.0**	14	0	0	2	0	
Workpackage 2:	Active stereopsis	**Actual WP total:**	**15.0**	11	0	0	0	4	
		Planned WP total:	**21.0**	9	0	0	0	12	
Workpackage 3:	Selecting and binding visual fragments	**Actual WP total:**	**20.0**	2	18	0	0	0	
		Planned WP total:	**20.0**	2	18	0	0	0	
Workpackage 4:	Sensorimotor integration	**Actual WP total:**	**16.0**	1	0	3	12	0	
		Planned WP total:	**16.0**	1	0	3	12	0	
Workpackage 5:	Human behavior and neural correlates of multisensory 3D representation	**Actual WP total:**	**40.0**	1	14	24	1	0	
		Planned WP total:	**40.0**	1	14	24	1	0	
Workpackage 6:	Project coordination and management	**Actual WP total:**	**7.0**	7	0	0	0	0	
		Planned WP total:	**5.0**	5	0	0	0	0	
Workpackage 7:	Knowledge Management, Dissemination, Synergies with other projects	**Actual WP total:**	**1.8**	0.94	0	0.5	0.33	0	
		Planned WP total:	**1.3**	0.5	0	0.5	0.33	0	
Workpackage 8:	Training, Education and Mobility	**Actual WP total:**	**1.3**	0.5	0	0.5	0.33	0	
		Planned WP total:	**1.3**	0.5	0	0.5	0.33	0	
Total Project Person-month		**Actual total:**	**118.1**	**38.4**	**32**	**28**	**15.7**	**4**	
		Planned total:	**120.7**	**33**	**32**	**28**	**15.7**	**12**	

7 Financial statements – Form C and Summary financial report

The summary financial report that consolidates the claimed Community contribution of all the beneficiaries in an aggregate form, and the separate financial statements from each beneficiary are provided in the following pages.

FP7 - Grant Agreement - Annex VI - Collaborative project

Summary Financial report - Collaborative project

Project acronym	EYESHOTS	Project nr	217077	Reporting period from	01/03/2008	to	28/02/2009	Page	1/1

Funding scheme	CP

				Type of activity								Total				
				RTD (A)		Demonstration (B)		Management (C)		Other (D)		Total (A+B+C+D)				
Benef. nr	If 3rd Party, linked to benef.	Adjustment (Yes/No)	Organisation Short Name	Total	Max EC Contrib.	Total	Max EC Contrib.	Total	Max EC Contrib.	Total	Max EC Contrib.	Total	Max EC Contrib.	Req. EC Contrib.	Receipts	Interest
1		No	UG	276,560	207,420	0	0	22,897	22,897	0	0	299,457	230,317	230,317	0	0
2		No	WWU	265,812	199,359	0	0	0	0	0	0	265,812	199,359	199,359	0	
3		No	UNIBO	158,606	118,954	0	0	0	0	0	0	158,606	118,954	118,954	0	
4		No	UJI	120,065	90,048	0	0	0	0	0	0	120,065	90,048	90,048	0	
5		No	K.U.Leuven	55,039	41,279	0	0	0	0	0	0	55,039	41,279	41,279	0	
Total				876,082	657,060	0	0	22,897	22,897	0	0	898,979	679,957	679,957	0	

2009-04-07 08:43

FP7 - Grant Agreement - Annex VI - Collaborative project

Form C - Financial Statement (to be filled in by each beneficiary)

Project Number	217077	Funding scheme	Collaborative project
Project Acronym	EYESHOTS		

Period from	01/03/2008	Is this an adjustment to a previous statement ?	No
To	28/02/2009		

Legal Name	UNIVERSITA DEGLI STUDI DI GENOVA	Participant Identity Code	999976687
Organisation short Name	UG	Beneficiary nr	1

Funding % for RTD activities (A)	75	If flate rate for indirect costs, specify %	60

1. Declaration of eligible costs/lump sum/flate-rate/scale of unit (in €)

	Type of Activity				
	RTD (A)	Demonstration (B)	Management (C)	Other (D)	Total (A+B+C+D)
Personnel costs	157,378	0	11,932	0	169,310
Subcontracting	0	0	0	0	0
Other direct costs	15,472	0	2,379	0	17,851
Indirect costs	103,710	0	8,586	0	112,296
Total costs	276,560	0	22,897	0	299,457
Maximum EC contribution	207,420	0	22,897	0	230,317
Requested EC contribution					230,317

2. Declaration of receipts

Did you receive any financial transfers or contributions in kind, free of charge from third parties or did the project generate any income which could be considered a receipt according to Art.II.17 of the grant agreement ?	No
If yes, please mention the amount (in €)	

3. Declaration of interest yielded by the pre-financing (to be completed only by the coordinator)

Did the pre-financing you received generate any interest according to Art.II.19 ?	No
If yes, please mention the amount (in €)	

4. Certificate on the methodology

Do you declare average personnel costs according to Art.II.14.1 ?	No
Is there a certificate on the methodology provided by an independent auditor and accepted by the Commission according to Art.II.4.4 ?	No

Name of the auditor		Cost of the certificate (in €), if charged under this project	

5. Certificate on the financial statements

Is there a certificate on the financial statements provided by an independent auditor attached to this financial statement according to Art.II.4.4 ?	No

Name of the auditor		Cost of the certificate (in €)	

6. Beneficiary's declaration on its honour

We declare on our honour that:

- the costs declared above are directly related to the resources used to attain the objectives of the project and fall within the definition of eligible costs specified in Articles II.14 and II.15 of the grant agreement, and, if relevant, Annex III and Article 7 (special clauses) of the grant agreement;
- the receipts declared above are the only financial transfers or contributions in kind, free of charge, from third parties and the only income generated by the project which could be considered as receipts according to Art.II.17 of the grant agreement;
- the interest declared above is the only interest yielded by the pre-financing which falls whithin the definition of Art.II.19 of the grant agreement;
- there is full supporting documentation to justify the information hereby declared. It will be made available at the request of the Commission and in the event of an audit by the Commission and/or by the Court of Auditors and/or their authorised representatives.

FP7 - Grant Agreement - Annex VI - Collaborative project

Form C - Financial Statement (to be filled in by each beneficiary)

Project Number	217077	Funding scheme	Collaborative project
Project Acronym	EYESHOTS		

Period from	01/03/2008	Is this an adjustment to a previous statement ?	No
To	28/02/2009		

Legal Name	WESTFAELISCHE WILHELMS-UNIVERSITAET MUENSTER	Participant Identity Code	999853691
Organisation short Name	WWU	Beneficiary nr	2

Funding % for RTD activities (A)	75	If flate rate for indirect costs, specify %	60

1. Declaration of eligible costs/lump sum/flate-rate/scale of unit (in €)

	Type of Activity				
	RTD (A)	Demonstration (B)	Management (C)	Other (D)	Total (A+B+C+D)
Personnel costs	143,409	0	0	0	143,409
Subcontracting	0	0	0	0	0
Other direct costs	22,724	0	0	0	22,724
Indirect costs	99,679	0	0	0	99,679
Total costs	265,812	0	0	0	265,812
Maximum EC contribution	199,359	0	0	0	199,359
Requested EC contribution					199,359

2. Declaration of receipts

Did you receive any financial transfers or contributions in kind, free of charge from third parties or did the project generate any income which could be considered a receipt according to Art.II.17 of the grant agreement ?	No
If yes, please mention the amount (in €)	

4. Certificate on the methodology

Do you declare average personnel costs according to Art.II.14.1 ?			No
Is there a certificate on the methodology provided by an independent auditor and accepted by the Commission according to Art.II.4.4 ?			No
Name of the auditor		Cost of the certificate (in €), if charged under this project	

5. Certificate on the financial statements

Is there a certificate on the financial statements provided by an independent auditor attached to this financial statement according to Art.II.4.4 ?			No
Name of the auditor		Cost of the certificate (in €)	

6. Beneficiary's declaration on its honour

We declare on our honour that:

- the costs declared above are directly related to the resources used to attain the objectives of the project and fall within the definition of eligible costs specified in Articles II.14 and II.15 of the grant agreement, and, if relevant, Annex III and Article 7 (special clauses) of the grant agreement;
- the receipts declared above are the only financial transfers or contributions in kind, free of charge, from third parties and the only income generated by the project which could be considered as receipts according to Art.II.17 of the grant agreement;
- the interest declared above is the only interest yielded by the pre-financing which falls whithin the definition of Art.II.19 of the grant agreement;
- there is full supporting documentation to justify the information hereby declared. It will be made available at the request of the Commission and in the event of an audit by the Commission and/or by the Court of Auditors and/or their authorised representatives.

FP7 - Grant Agreement - Annex VI - Collaborative project

Form C - Financial Statement (to be filled in by each beneficiary)

Project Number	217077	Funding scheme	Collaborative project
Project Acronym	EYESHOTS		

Period from	01/03/2008	Is this an adjustment to a previous statement ?	No
To	28/02/2009		

Legal Name	ALMA MATER STUDIORUM - UNIVERSITA DI BOLOGNA	Participant Identity Code	999993953
Organisation short Name	UNIBO	Beneficiary nr	3

Funding % for RTD activities (A)	75	If flate rate for indirect costs, specify %	60

1. Declaration of eligible costs/lump sum/flate-rate/scale of unit (in €)

	Type of Activity				
	RTD (A)	Demonstration (B)	Management (C)	Other (D)	Total (A+B+C+D)
Personnel costs	63,260	0	0	0	63,260
Subcontracting	0	0	0	0	0
Other direct costs	35,869	0	0	0	35,869
Indirect costs	59,477	0	0	0	59,477
Total costs	158,606	0	0	0	158,606
Maximum EC contribution	118,954	0	0	0	118,954
Requested EC contribution					118,954

2. Declaration of receipts

Did you receive any financial transfers or contributions in kind, free of charge from third parties or did the project generate any income which could be considered a receipt according to Art.II.17 of the grant agreement ?	No
If yes, please mention the amount (in €)	

4. Certificate on the methodology

Do you declare average personnel costs according to Art.II.14.1 ?	No
Is there a certificate on the methodology provided by an independent auditor and accepted by the Commission according to Art.II.4.4 ?	No

Name of the auditor		Cost of the certificate (in €), if charged under this project	

5. Certificate on the financial statements

Is there a certificate on the financial statements provided by an independent auditor attached to this financial statement according to Art.II.4.4 ?	No

Name of the auditor		Cost of the certificate (in €)	

6. Beneficiary's declaration on its honour

We declare on our honour that:

- the costs declared above are directly related to the resources used to attain the objectives of the project and fall within the definition of eligible costs specified in Articles II.14 and II.15 of the grant agreement, and, if relevant, Annex III and Article 7 (special clauses) of the grant agreement;
- the receipts declared above are the only financial transfers or contributions in kind, free of charge, from third parties and the only income generated by the project which could be considered as receipts according to Art.II.17 of the grant agreement;
- the interest declared above is the only interest yielded by the pre-financing which falls whithin the definition of Art.II.19 of the grant agreement;
- there is full supporting documentation to justify the information hereby declared. It will be made available at the request of the Commission and in the event of an audit by the Commission and/or by the Court of Auditors and/or their authorised representatives.

FP7 - Grant Agreement - Annex VI - Collaborative project

Form C - Financial Statement (to be filled in by each beneficiary)

Project Number	217077	Funding scheme	Collaborative project
Project Acronym	EYESHOTS		

Period from	01/03/2008	Is this an adjustment to a previous statement ?	No
To	28/02/2009		

Legal Name	UNIVERSITAT JAUME I DE CASTELLON	Participant Identity Code	999882985
Organisation short Name	UJI	Beneficiary nr	4

Funding % for RTD activities (A)	75	If flate rate for indirect costs, specify %	60

1. Declaration of eligible costs/lump sum/flate-rate/scale of unit (in €)

	Type of Activity				
	RTD (A)	Demonstration (B)	Management (C)	Other (D)	Total (A+B+C+D)
Personnel costs	63,961	0	0	0	63,961
Subcontracting	0	0	0	0	0
Other direct costs	11,081	0	0	0	11,081
Indirect costs	45,023	0	0	0	45,023
Total costs	120,065	0	0	0	120,065
Maximum EC contribution	90,048	0	0	0	90,048
Requested EC contribution					90,048

2. Declaration of receipts

Did you receive any financial transfers or contributions in kind, free of charge from third parties or did the project generate any income which could be considered a receipt according to Art.II.17 of the grant agreement ?	No
If yes, please mention the amount (in €)	

4. Certificate on the methodology

Do you declare average personnel costs according to Art.II.14.1 ?	No
Is there a certificate on the methodology provided by an independent auditor and accepted by the Commission according to Art.II.4.4 ?	No

Name of the auditor		Cost of the certificate (in €), if charged under this project	

5. Certificate on the financial statements

Is there a certificate on the financial statements provided by an independent auditor attached to this financial statement according to Art.II.4.4 ?	No

Name of the auditor		Cost of the certificate (in €)	

6. Beneficiary's declaration on its honour

We declare on our honour that:

- the costs declared above are directly related to the resources used to attain the objectives of the project and fall within the definition of eligible costs specified in Articles II.14 and II.15 of the grant agreement, and, if relevant, Annex III and Article 7 (special clauses) of the grant agreement;
- the receipts declared above are the only financial transfers or contributions in kind, free of charge, from third parties and the only income generated by the project which could be considered as receipts according to Art.II.17 of the grant agreement;
- the interest declared above is the only interest yielded by the pre-financing which falls whithin the definition of Art.II.19 of the grant agreement;
- there is full supporting documentation to justify the information hereby declared. It will be made available at the request of the Commission and in the event of an audit by the Commission and/or by the Court of Auditors and/or their authorised representatives.

FP7 - Grant Agreement - Annex VI - Collaborative project

Form C - Financial Statement (to be filled in by each beneficiary)

Project Number	217077	Funding scheme	Collaborative project
Project Acronym	EYESHOTS		

Period from	01/03/2008	Is this an adjustment to a previous statement ?	No
To	28/02/2009		

Legal Name	KATHOLIEKE UNIVERSITEIT LEUVEN	Participant Identity Code	999991334
Organisation short Name	K.U.Leuven	Beneficiary nr	5

Funding % for RTD activities (A)	75	If flate rate for indirect costs, specify %	60

1. Declaration of eligible costs/lump sum/flate-rate/scale of unit (in €)

	Type of Activity				
	RTD (A)	Demonstration (B)	Management (C)	Other (D)	Total (A+B+C+D)
Personnel costs	23,936	0	0	0	23,936
Subcontracting	0	0	0	0	0
Other direct costs	10,464	0	0	0	10,464
Indirect costs	20,639	0	0	0	20,639
Total costs	55,039	0	0	0	55,039
Maximum EC contribution	41,279	0	0	0	41,279
Requested EC contribution					41,279

2. Declaration of receipts

Did you receive any financial transfers or contributions in kind, free of charge from third parties or did the project generate any income which could be considered a receipt according to Art.II.17 of the grant agreement ?	No
If yes, please mention the amount (in €)	

4. Certificate on the methodology

Do you declare average personnel costs according to Art.II.14.1 ?			No
Is there a certificate on the methodology provided by an independent auditor and accepted by the Commission according to Art.II.4.4 ?			No
Name of the auditor		Cost of the certificate (in €), if charged under this project	

5. Certificate on the financial statements

Is there a certificate on the financial statements provided by an independent auditor attached to this financial statement according to Art.II.4.4 ?			No
Name of the auditor		Cost of the certificate (in €)	

6. Beneficiary's declaration on its honour

We declare on our honour that:

- the costs declared above are directly related to the resources used to attain the objectives of the project and fall within the definition of eligible costs specified in Articles II.14 and II.15 of the grant agreement, and, if relevant, Annex III and Article 7 (special clauses) of the grant agreement;
- the receipts declared above are the only financial transfers or contributions in kind, free of charge, from third parties and the only income generated by the project which could be considered as receipts according to Art.II.17 of the grant agreement;
- the interest declared above is the only interest yielded by the pre-financing which falls whithin the definition of Art.II.19 of the grant agreement;
- there is full supporting documentation to justify the information hereby declared. It will be made available at the request of the Commission and in the event of an audit by the Commission and/or by the Court of Auditors and/or their authorised representatives.

8 Certificates on the financial statements

List of Certificates that are due for this period, in accordance with Article II.4.4 of the Grant Agreement.

Beneficiary	Organisation short name	Certificate provided? yes / no	Any useful comment, in particular if a certificate is not provided
1	UG	no	Expenditure threshold not reached
2	WWU	no	Expenditure threshold not reached
3	UNIBO	no	Expenditure threshold not reached
4	UJI	no	Expenditure threshold not reached
5	K.U.Leuven	no	Expenditure threshold not reached

9 References

Attneave, F. Some informational aspects of visual perception. *Psychol. Rev*, **61**:183–193, 1954.

Bahill, A.T., Latimer, J.R. and Troost, B.T. Linear homeomorphic saccadic eye movement. *IEEE Trans. Biomed. Eng.*, **27**:631-639, 1980.

Barlow, H.B. Possible principles underlying the transformation of sensory messages. Rosenblith, W.A., Ed., *Sensory Communication*, pages 217–234, Cambridge, MA. MIT Press, 1961.

Bayerl, P. and Neumann, H. A fast biologically inspired algorithm for recurrent motion estimation. *IEEE Trans. on PAMI,* **29**(2): 246-260, 2007.

Bell, A. J. and Sejnowski, T.J. The "independent components" of natural scenes are edge filters. *Vis. Res.*, **37**:3327–3338, 1997.

Bridge1, H. and Cumming, B.G. Representation of binocular surfaces by cortical neurons *Current Opinion in Neurobiology*, **18(**4):425-430, 2008.

Cannata, G., Maggiali, M. Models for the Design of Bioinspired Robot Eyes, *IEEE Trans. on Robotics*, **24**:27-44, 2008.

Catz, N., Dicke, P.W., and Thier, P. Cerebellar-dependent motor learning is based on pruning a purkinje cell population response. *Proc. Nati. Acad. Sci.*, **105**(20):7309–7314, 2008.

Chaturvedi, V. and Van Gisbergen, J.A.M. Shared Target Selection for Combined Version-Vergence Eye Movements. *J. Neurophysiol.*, **80**:849–862, 1998.

Chen, Y. and Qian, N. A Coarse-to-fine Disparity Energy Model with both Phase-shift and Position-shift Receptive Field Mechanisms, *Neural Computation*, **16**:1545-1577, 2004.

Crawford, J.D., Martinez-Trujillo J.C. and Klier, E.M. Neural Control of Three-dimensional Eye and Head Movements, *Curr. Opinion in Neurobiology*, **13**:655-662, 2003.

Dankers, A. A., Barnes, N., Bischof, W. F. and Zelinsky, A. Humanoid vision resembles primate archetype. *Springer Tracts in Advanced Robotics*, in press.

Duhammel, J.R., Colby C.L., and Goldberg, M.E. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, **255**:90-92, 1992.

Enderle, J.D. and Wolfe, J.W. Time-Optimal Control of Saccadic Eye Movements, *IEEE Trans. Biomed. Eng.*, **34**:43-55, 1987.

Enderle, J.D., Wolfe, J.W. and Yates, J.T. The linear homeomorphic saccadic eye movement model A modification, *IEEE Trans. Biomed. Eng.*, **31**:717-720, 1984.

Ethier, V., Zee, D. S., and Shadmehr, R. Changes in control of saccades during gain adaptation. *J. Neurosci.*, **28**(51):13929–13937, 2008.

Field, D.J. What is the goal of sensor y coding? *Neural Comput*, **6**:559–601, 1994.

Fischer, B. and Poggio, G.F. Binocular interaction and depth sensitivity in striate and pestriate cortex of behaving rhesus monkey. *J. Neurophysiol.*, **40**:1392–1405, 1977.

Fleet, D.J., Jepson, A.D. and Jenkin, M.R.M. Phase-based disparity measurement. *CVGIP: Image Understanding*, **53**(2):198–210, 1991.

Fleet, D.J., Wagner, H. and Heeger, D. J. Neural encoding of binocular disparity: Energy models, position shifts and phase shifts. *Vis. Res.*, **36(**12):1839-1857, 1996.

Georg, K., and Lappe, M. Effects of saccadic adaption on visual localization before and during saccades. *Exp. Brain Res.*, **129**:9–23, 2009.

Golla, H., Tziridis, K., Haarmeier, T., Catz, N., Barash, S., and Thier, P. Reduced saccadic resilience and impaired saccadic adaptation due to cerebellar disease. *Eur. J. Neurosci.*, **27**(1):132–144, 2008.

Haefner, R.M. and Cumming, B.G. Adaptation to natural binocular disparities in primate V1 explained by a generalized energy model. *Neuron*, **57**(1):147-58, 2008.

Hamker, F. H. and Wiltschut, J. Hebbian learning in a model with dynamic rate-coded neurons: an alternative to the generative model approach for learning receptive fields from natural scenes. *Network, Computation in Neural Systems*, **18**:249-266, 2007.

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, 2005.

Hamker, F. H., Zirnsak, M. V4 receptive field dynamics as predicted by a systems-level model of visual attention using feedback from the frontal eye field. *Neural Networks*, **19**:1371-1382, 2006.

Hamker, F.H., Zirnsak, M., Calow, D., Lappe, M. The peri-saccadic perception of objects and space. *PLOS Computational Biology*, **4(2)**:e31, 2008.

Hansard, M. and Horaud, R.P. Cyclopean Geometry of Binocular Vision. *JOSA A*, **25**(9):2357–2369, 2008.

Hernandez, T. D., Levitan, C. A., Banks, M. S., and Schor, C. M. How does saccade adaptation affect visual perception? *JOV*, **8**(8):1–16, 2008.

Hinkle, D.A. and Connor, C.E. Quantitative characterization of disparity tuning in ventral pathway area V4. *J. Neurophysiol.,* **94**:2726-37, 2005.

Hoyer, P.O. and Hyvarinen, A. Independent component analysis applied to feature extraction from colour and stereo images. Network, **11**:191–210, 2000.

Horng, J, Semmlow, J., Hung, G.K., and Ciuffreda, K. Initial Component in Disparity Vergence: A Model-Based Study. *IEEE Trans. Biomed. Eng.*, **45**(2): 249-257, 1998.

Hyvärinen, A., Karhunen, J., and Oja, E. Independent Component Analysis. *Wiley*, New York, 2001.

Hung, G.K., Semmlow, J.L., Ciufferda, K.J. A Dual-Mode Dynamic Model of the Vergence Eye Movement System. *IEEE Trans. Biomed. Eng.*, **33**(11):1021-1028.

Hung, G.K. Dynamic model of the vergence eye movement system: simulations using MATLAB/SIMULINK. *Computer Methods and Programs in Biomedicine*, **55**:59–68, 1998.

Jansen, L., Onat, S., Koenig, P. Influence of disparity on fixation and saccades in free viewing of natural scenes. *J. of Vision* **9:**1-19, 2009

Jaschinski, W., Svede, A., and Jainta, S. Relation between fixation disparity and the asymmetry between convergent and divergent disparity step responses. *Vis. Res.*, **48**(2):253-263, 2008.

Jenkin, M.R.M. Stereopsis near the Horoptor. *International Conference on Control, Automation, Robotics and Vision*, Singapore, 1996.

Klier, E.M. and Angelaki, D.E. Spatial updating and the maintenance of visual constancy. *Neuroscience* **156**(4):801-818, 2008

Krishnan, V.V., and Stark, L. A heuristic model for the human vergence eye movement system, *IEEE Trans Biomed Eng*, **24**:44-9, 1977.

LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, **86**(11):2278-2324, 1998.

Lewicki, M.S., Hughes, H., and Olshausen, B.A. Probabilistic framework for the adaptation and comparison of image codes. *JOSA A*, **16**:1587–1601, 1999.

Lippert, J., Fleet, D., and Wagner, H. Disparity tuning as simulated by a neural net. *Biol. Cybern.,* **83**:61-72, 2000.

Lippert, J. and Wagner, H. Visual depth encoding in populations of neurons with localized receptive fields. *Biol. Cybern*., **87**:249-61, 2002.

Liu Y., Bovik, A.C. and Cormack L.K. Disparity statistics in natural scenes. *J. of Vision*, **8**:1-14, 2008.

Maxwell, J.S., Graf, E.W. and Schor, C.M. Adaptation of torsional eye alignment in relation to smooth pursuit and saccades. *Vis. Res.* **41**:3735–3749, 2001.

Minken, A.W.H., and Van Gisbergen, J.A.M. A three dimensional analysis of vergence movements at various level of elevation. *Exp. Brain Res.*, **101**(2):331–345, 1994.

Misslisch, H., Tweed, D. and Hess, B.J. Stereopsis outweighs gravity in the control of the eyes. *J. Neurosci.* **21**(3):126-130, 2001.

Miura, K., Sugita, Y., Matsuura, K., Inaba, N., Kawano, K. and Miles, F.A. The Initial Disparity Vergence Elicited With Single and Dual Grating Stimuli in Monkeys: Evidence for Disparity Energy Sensing and Nonlinear Interactions. *J. Neurophysiol*., **100**:2907-918, 2008.

Mok, D., Ro, A., Cadera, W., Crawford, J.D., and Vilis., T. Rotation of listing's plane during vergence. *Vision Res.*, **32**:2055–2064, 1990.

Mossio, M. and Taraborelli, D. Action-dependent perceptual invariants: From ecological to sensorimotor approaches. *Consciousness and Cognition* **17**:1324-1340, 2008.

Nishimoto, S., Ishida, T. and Ohzawa, I. Receptive Field Properties of Neurons in the Early Visual Cortex Revealed by Local Spectral Reverse Correlation. *J. Neurosci.*, **26**(12):3269-3280, 2006.

Ohzawa, I., DeAngelis, G.C., Freeman, R.D. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science,* **249**:1037-1041, 1990.

Ohzawa, I., DeAngelis, G.C., and Freeman, R.D. Encoding of binocular disparity by complex cells in the cat's visual cortex. *J. Neurophysiol*. **77**:2879-2909, 1997

Olshausen, B.A. and Field, D.J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, **381**:607–609, 1996.

Panouilleres, M., Weiss, T., Urquizar, C., Salemme, R., Munoz, D. P., and Pelisson, D. Behavioral evidence of separate adaptation mechanisms controlling saccade amplitude lengthening and shortening. *J. Neurophysiol.*, **101**(3):1550–1559, 2009.

Patel, S.S., Jiang, B.C. and Ogmen, H. Vergence Dynamics Predict Fixation Disparity. *Neural Computation*, **13**(7):1495-1525, 2001.

Patel, S.S., Ögmen, H., White, J.M. and Jiang, B.C. Neural Network Model of Short-Term Horizontal Disparity Vergence Dynamics. *Vis. Res.*, **37(**10):1383-1399, 1996.

Porrill, J., Ivins, J.P., and Frisby, J.P. The variation of torsion with vergence and elevation. *Vision Res.*, **39**:3934–3950, 1999.

Pouget, A., and Sejnowski, T.J. Spatial Transformations in the Parietal Cortex Using Basis Functions. *J. Cog. Neurosci*. **9**(2):222-237, 1997.

Qian., N. Computing stereo disparity and motion with known binocular cell properties. *Neural Computation*, **6**(3):390–404, 1994.

Qian, N. and Mikaelian, S. Relationship between Phase and Energy Methods for Disparity Computation , *Neural Computation*, **12**:279-292, 2000.

Qiu, F.T. and von der Heydt, R. Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules. *Neuron,* **47**:155-166, 2005.

Read, J.C. and Cumming, B.G. Does depth perception require vertical-disparity detectors? *JOV*, **6**(12):1323-55, 2006.

Read, J.C.A. A simple model accounts for the response of disparity-tuned V1 neurons to anti-correlated images. *Vis. Neurosci*., **19**:735-753, 2002.

Read, J.C.A. Early computational processing in binocular vision and depth perception. *Progress in Molecular Biology and Biophysics*, **87**:77-108, 2004.

Read, J.C.A., Parker, A.J. and Cumming, B.G. Testing quantitative models of binocular disparity selectivity in primary visual cortex. *J. Neurophysiol*., **90**:2795-2817, 2002.

Ringach, D.L., Bredfeldt, C.E., Shapley, R.M., and Hawken, M.J. Suppression of neural responses to nonoptimal stimuli correlates with tuning selectivity in macaque V1. *J. Neurophysiol*., **87**:1018–1027, 2002.

Sabatini, S.P. and Solari, F. Emergence of motion-in-depth selectivity in the visual cortex through linear combination of binocular energy complex cells with different ocular dominance. *Neurocomputing,* **58**–**60**:865-872, 2004.

Sabatini, S.P., Gastaldi, G., Solari, F., Diaz, J., Ros, E. Pauwels, K., Van Hulle, M., Pugeault, N. and Krueger, N. Compact and accurate early vision processing in the harmonic space. 2nd International Conference on Computer Vision Theory and Applications, 8-11 March, 2007, Barcelona, Spain.

Salinas, E., and Abbott, L.F. Transfer of coded information from sensory to motor networks. *J. Neurosci.*, **15**:6461-6474, 1995.

Sanada, T.M. and Ohzawa, I. Encoding of Three-Dimensional Surface Slant in Cat Visual Areas 17 and 18. *J. Neurophysiol*., **95**:2768-2786, 2006.

Sanger, T.D. Stereo disparity computation using Gabor filters. *Biol. Cybern.,* **59**:405–418, 1988.

Schor, C.M., Maxwell, J.S. and Graf, E.W. Plasticity of convergence-dependent variations of cyclovergence with vertical gaze. *Vis. Res.* **41**:3353–3369, 2001.

Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.*, **80**:1–27, 1998.

Serrano-Pedraza, I. and Read, J.C.A. Read Stereo vision requires an explicit encoding of vertical disparity. *JOV,* **9**(4):1-13, 2009.

Schreiber, K.M., Tweed, D.B., and Schor, C.M. The extended horopter: Quantifying retinal correspondence across changes of 3d eye position. *J. Vis.*, **6**:64–74, 2006.

Sommer, M.A., and Wurtz, R.H. Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, **444**:374-377, 2006.

Steffen, H., Walker, M. and Zee, D.S. Changes in Listing's plane after sustained vertical fusion. *Invest. Ophthalmol. Vis. Sci.* **43:**668–672, 2002.

Sutton, R.S. and Barto, A.G. Reinforcement Learning: An Introduction. *MIT Press, Cambridge*, MA. A Bradford Book, 1998.

Tanabe. S, and Cumming, B.G. Mechanisms underlying the transformation of disparity signals from V1 to V2 in the macaque. *J Neurosci*., **28**(44):11304-11314, 2008.

Theimer W.M. and Mallot H.P. Phase-Based Vergence Control and Depth Reconstruction Using Active Vision, *CVIGP*, **60(**3):343-358, 1994.

Thompson, A.A. and Henriques, D.Y.P. Updating Visual Memory Across Eye Movements for Ocular and Arm Motor Control. *J. Neurophysiol.*, **100**:2507–2514, 2008.

Tolias, A.S., Moore, T., Smirnakis, S.M., Tehovnik, E.J., Siapas, A.G., and Schiller, P.H. Eye movements modulate visual receptive fields of V4 neurons. *Neuron*, **29**:757-767, 2001.

Tsang, E.K.C. and Shi, B.E. Disparity Estimation by Pooling Evidence from Energy Neurons, *IEEE Tran. on Neural Networks*, in press.

Tsang, E.K.C. and Shi, B.E. Normalization Enables Robust Validation of Disparity Estimates from Neural Populations. *Neural Computation*, **20**(10):2464-2490, 2007.

Tsang, E.K.C., Lam, S.Y.M., Meng, Y. and Shi B.E. Neuromorphic implementation of active gaze and vergence control. *ISCAS*, pages 1076-1079, IEEE, 2008.

Tweed, D., and Vilis, T. Geometric relations of eye position and velocity vectors during saccades. *Vis. Res.*, **30**(1):111–127, 1990.

Van Pelt, S., and Medenport, W.P. Updating Target Distance Across Eye Movements in Depth. *J. Neurophysiol*. **99**: 2281-2290, 2008

van Hateren, J. H. and van der Schaaf, A. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. Biol. Sci.*, **265**:359–366, 1998.

Van Rijn, L.J., and Van den Berg, A.V. Binocular eye orientation during fixations: Listing's law extended to include eye vergence. *Vision Res.*, **33**(5/6):691–708, 1993.

Vitay, J. and Hamker, F. H. (in press). Binding objects to cognition: A brain-like systems approach to the cognitive control of visual perception. *International Conference on Cognitive Systems,* Karlsruhe, Germany, 2008.

Vitay, J., Fix, J., Beuth, F., Schroll, H., Hamker, F. H. Biological Models of Reinforcement Learning. Submitted to KI.

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.

Wismeijer, D.A., van Ee, R. and Erkelens, C.J. Depth cues, rather than perceived depth, govern vergence. *Exp. Brain Res*., **184**(1):61-70, 2008.