Project no.: FP7-ICT-217077
Project full title: Heterogeneous 3-D Perception across Visual Fragments
Project Acronym: EYESHOTS
Deliverable no: D4.3b
Title of the deliverable: An embodied agent which learns to situate itself in the environment through active exploration

Date of Delivery:	21 March 2011	
Organization name of lead contractor for this deliverable:	UJI	
Author(s):	E. Chinellato, M. Antonelli, A.P. del Pobil	
Participant(s):	UJI, WWU	
Workpackage contributing to the deliverable:	WP4	
Nature:	Demonstrator	
Version:	1.1	
Total number of pages:	7	
Responsible person:	Angel P. del Pobil	
Revised by:	Katharina Havermann, Markus Lappe	
Start date of project:	1 March 2008	**Duration:** 36 months

Project Co-funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other program participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Abstract:

This deliverable describes the skills of the UJI humanoid robot in building a visuomotor awareness of its surrounding environment, and the experimental scenario in which the robot is expected to demonstrate such skills.
First, the robot learns the sensorimotor transformations required to perform coupled or uncoupled gazing and reaching movements in its peripersonal space. Then, through a purposeful exploration of the environment the robot is able to build a visuomotor memory of objects in space, in order to execute arm reaching movements to surrounding stimuli, either with or without gazing at them.

Contents

1 Executive summary

The behavioral skills of the robotic system according to the goals of WP4 will be presented in a live demonstration during the final review meeting. The experimental scenario is a *working-desk setup*, with simple objects, and full 3D movements for both eyes and arm. This setup will show the overall abilities of the robot in building a visuomotor awareness of its surrounding environment. How such abilities are improved and extended by integrating on the UJI humanoid robot the VVCA architecture of partners UG, K.U.Leuven and WWU/Chemnitz is the subject of D4.3c.

To better highlight the system capabilities and their relation to the neural model based on neuroscience data and insights, we briefly describe in this deliverable also some special experiments performed simulating the saccadic adaptation experimental paradigms used by partner WWU. Saccadic adaptation tests have been executed in simulation and with the real robot observing visual stimuli on a computer screen.

2 Introduction

The fundamental goal of EYESHOTS Work Package 4 was to design and develop a model for achieving visuomotor awareness of the environment by using eye and arm movements. The model, described in detail in deliverables D4.2 and D4.3a, and in the journal paper [2], has been now fully implemented on our "Tombatossals" humanoid robot setup [1]. Basic skills such as concurrent or decoupled gazing and reaching movements toward visual stimuli are available to the robot. As planned, the robot is able to show its visuomotor capabilities by performing oculomotor actions toward visual targets placed in its peripersonal space, or toward the location where its hand lies. Moreover, it is also able to perform arm reaching movements to visible objects, either with or without gazing at it. This is achieved through a purposeful exploration of the environment, which allows the robot to build a visuomotor memory of surrounding objects, described in Section 3. The robot skills are obtained by exploiting its sensorimotor coordination ability, provided by the transformations Visual⇒Oculomotor (V⇒O) and Oculomotor⇔Arm motor (O⇔A), implemented on a radial basis function framework. Such framework and the analysis of the robot behavior constitute a contribution to cognitive science research, as demonstrated by the saccadic adaptation experiments introduced in Section 4.

3 Showing visuomotor awareness

The skills that the robot is able to demonstrate in a real setup are described next. We follow a modular approach in which simple, atomic *tasks* like saccades or visual detections are composed into *behaviors* of different levels of complexity, such as learning, exploring or performing any specific action.

Training Visual⇒Oculomotor and Oculomotor⇔Arm motor networks

To be able to perform a complete visual exploration of the environment, the robot has to pass first through a learning stage designed to develop its own sensorimotor coordination. This is done by training the visual-oculomotor and the oculomotor-arm motor neural networks. It is important to remind that, to reduce the learning stage to a reasonable amount of steps, the networks are bootstrapped with the weights obtained by the model of the system trained in a similar fashion.

V⇒O NETWORK

- detection of visual target;
- do while target is not foveated:
  - compute oculomotor movement necessary to foveate on target;
  - execute corresponding saccadic movement;
  - detection of target;
  - update V⇒O network;

O⇔A NETWORKS

- random arm movement within range;
- train V⇒O network (with arm as visual target);
- update O⇒A and A⇒O networks;

When the networks have achieved a good performance in mapping the visuomotor movements required to perform all the desired actions, more complex behaviors can be executed. Nevertheless, learning can continue throughout the execution of other behaviors, and indeed the robot should keep learning in order to improve and adapt its visuomotor skills while doing any available action.

Visual exploration of environment

The first behavior after the initial training is the exploration of the environment to update the visuomotor memory of visible objects. Such behavior is also pre-defined, because it constitutes a step previous to dealing specifically with any surrounding objects.

- execute random saccadic movement;
- cycle through all objects visible on both eyes; detect ventral visual features and determine oculomotor location of each of them;
- check object memory for corresponding items and update accordingly; add new item if necessary;

This behavior allows thus to create a visuomotor memory of objects present in the environment, to be used in other specific tasks.

Custom actions

Building on the visuomotor memory acquired as described above, more custom actions of different complexity can be performed. Some of these actions require additional information provided by the human user, such as the identity of the object to gaze or reach at, or the number times exploratory movements should be repeated in overt or covert attention behaviors.

- look at the hand
- reach the gazing point
- show memory (color, optional position)
- gaze at a given object (either inside or outside the field of view)
  - if necessary, move eyes toward a different location, even out of usual range;
  - ask for target object (identify them with color names);
  - access memory by color and move to correspondent oculomotor coordinates;
- reach a given object (either foveated or not)

  - ask for target object (identify them with color names);
  - access memory by color and move to correspondent arm motor coordinates, irrespective of foveation;

- overt attention behavior (repeat n times)
  - select a random blob visible on both eyes; detect color and determine oculomotor location; execute foveation movement;
  - check object memory for corresponding items (in color/location/both - optional) and update accordingly;

- covert attention behavior (repeat n times)
  - optional random saccadic movement;
  - from current oculomotor position, cycle through all blobs; consider objects visible on both eyes;
  - check object memory for corresponding items (in color/location/both - optional) and update accordingly;

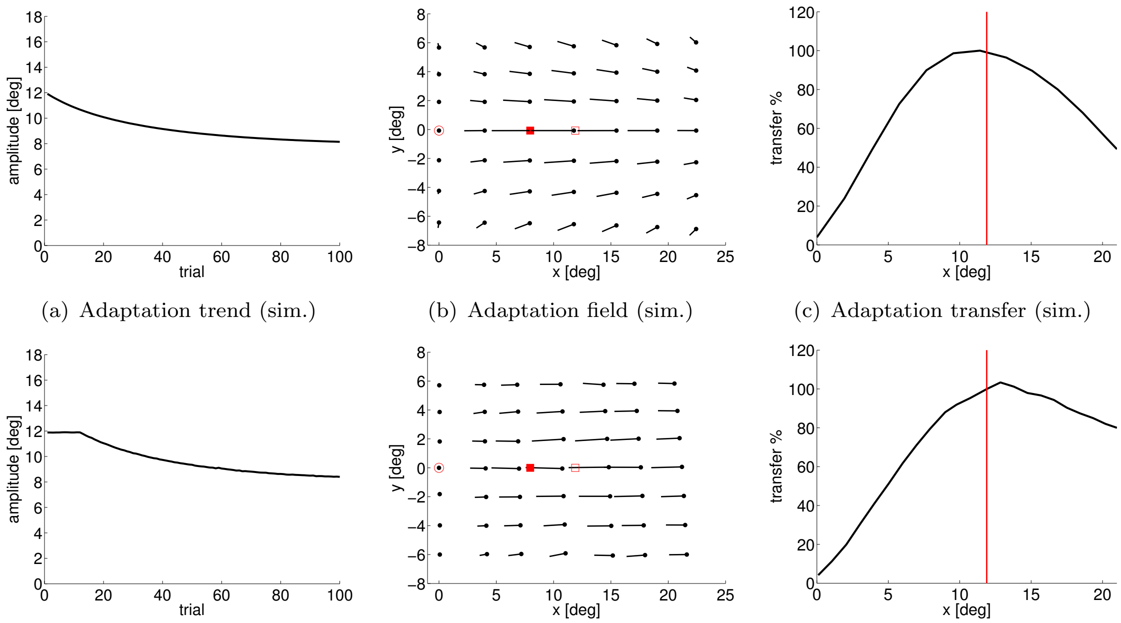
The behaviors presented in this last section constitute the global overall abilities of the system at this stage, but other behaviors can be easily generated by composing them in different ways according to the particular requirements of the final goal action.

4 Saccadic adaptation experiments

In order to check the underlying properties of the computational framework on which the robot behavioral abilities are built upon, a different experimental setup was established. We opted for a cognitive science setup similar to those used for the saccadic adaptation experiments performed by partner WWU [3, 4]. This consists of a computer screen placed within reaching distance, on which different visual stimuli associated to action signals are visualized. It is worthwhile to clarify that, although vergence varies just slightly in this setup, all transformations are tridimensional, and the robot acts as in a full 3D space.

We performed saccadic adaptation experiments emulating the protocols described in the human literature, applied to the robot model on a virtual setup first, and to the real robot on a computer screen later. The comparison between data obtained with human subjects, with the computational simulation and with the robot can provide insights on theoretical aspects related to visuomotor cognitive aspects and contextually allow us to validate our proposal. We tested each of the two experimental protocols, inward and outward adaptation, with two different distributions of the centers of the visual to oculomotor radial basis function network, logarithmic and uniform. An example of the sort of results we have obtained is shown in Fig. 1, for inward adaptation executed on the logarithmic network configuration.

Different properties observed in human saccadic adaptation studies were captured by our tests. Both the simulation and the robot experiments showed: plausible adaptation trends; slightly radial adaptation fields; features typical of the adaptation transfer on the horizontal component, such as asymmetry and late peak. It is interesting to observe that, in general, the robot results approximated the human data better than the simulated results. More details on the performed experiments and the results obtained will be provided in the EYESHOTS 3rd periodic report.



(a) Adaptation trend (sim.) (b) Adaptation field (sim.) (c) Adaptation transfer (sim.)

(d) Adaptation trend (robot) (e) Adaptation field (robot) (f) Adaptation transfer (robot)

Figure 1: Example of saccadic adaptation results for inward adaptation protocol and logarithmic network configuration, simulation (above) and robot data (below).

Bibliography

[1] M. Antonelli, E. Chinellato, and A.P. del Pobil. Implicit mapping of the peripersonal space of a humanoid robot. In *IEEE Symposium Series on Computational Intelligence - SSCI*, 2011.

[2] E. Chinellato, M. Antonelli, B.J. Grzyb, and A.P. del Pobil. Implicit sensorimotor mapping of the peripersonal space by gazing and reaching. *IEEE Transactions on Autonomous Mental Development*, In Press, 2011.

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

[4] F. Schnier, E. Zimmermann, and M. Lappe. Adaptation and mislocalization fields for saccadic outward adaptation in humans. *Journal of Eye Movement Research*, 3(3):1–18, 2010.