Project no.:	**027657**
Project full title:	**Perception, Action & Cognition through learning of Object-Action Complexes**
Project Acronym:	**PACO-PLUS**
Deliverable no.:	**D8.1.3**
Title of the deliverable:	**Second Demonstration Application I: Robot-vision system equipped with a sophisticated haptic machinery that does multisensory object categorisation and first examples of a generalization.**

Contractual Date of Delivery to the CEC: 15.3.2008
Actual Date of Delivery to the CEC: 21.3.2008
Organisation name of lead contractor for this deliverable: SDU

Author(s): Dirk Kraft, Norbert Krüger, Mila Popovic, Kai Welke, Tamim Asfour, Kai Hübner, Danica Kragic, Alex Bierbaum, Florentin Wörgötter, Nicolas Pugeault, Christopher Geib, Ron Petrick, Mark Steedman, Bernhard Hommel, Renaud Detry, Justus Piater, Rüdiger Dillmann
Participant(s): KTH, BCCN, AAU, JSI, UL, UEDIN, SDU, ULg, UniKarl

Work package contributing to the deliverable: WP1,WP2,WP3,WP4,WP5
Nature: R
Version: Final
Total number of pages: 5
Start date of project: 1st Feb. 2006 **Duration:** 48 month

Project co-funded by the European Commission within the Sixth Framework Programme (2002–2006)

Dissemination Level

PU	Public	X
PP	Restricted to other programme 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:

We present 7 videos showing our progress on developping a cognitive system that perform learning, categorisations, and planning on different levels. Video number 2 explicitely realises the three level hierarchy developped in PACOplus.

Keyword list: Cognitive Architecture, Grounding, Learning, Object-Action Complexes

Table of Contents

1. Grasping Reflex

(GraspingReflexUncompressed.m2v)

The video shows the performance of the 'grasping reflex' derscribed in [4, 1] on a complex scene. All objects that are graspable by the two finger gripper have been grasped at the end based on this very simple mechanism. Intermediate results on successful, unstable, or unsuccessful grasps become stored together in terms of triplets (features, grasp (type as well as pose), evaluation) that will become input on higher learning level processes. In this way, a large amount of ground truth for these learning processes can be generated.

2. Computation and Performance of two Plans for Putting Objects on a Shelf

(ExecutingPlanInThreeLevelArchitecture.m2v)

The video shows the computation and performance of two plans in the three level hierarchy (sensory level, mid-level amd planning level) developped within PACOplus (see [2]). In the both cases, objects are put on a shelf with the option to be stacked together. For this low-level information is transformed via a (still premature) mid-level stage to a discrete state space from which plans can be generated. For the performance of plans, high level commands need to be translated (also via the mid-level) to concrete actions on the sensory level that are inherently continuous. In the case of the first plan objects are assumed to be open. During the execution of the second plan this information is gathered during execution through sensing actions.

3. Finger equipped with tactile Sensors used for Object Categorisation

(TactileSensor.m2v)

The video shows the status of our work on tactile devices. A finger equipped densely with MicroNav sensors is shown as well as the sub-parts it consists of (sensors, actuators, PCPs, etc.) The tactile device is then used for the categorisation of an objects in terms of openness/closeness, fullness/emptyness as well as its elasticity (for details, see D8.1.4).

4. Acquiring Object-Action Knowledge by Exploration

(pushing.mpg)

The goal of this research is to study methods that enable the robot to learn how to act on an object so that the desired outcome is achieved. In particular, we investigated how to learn a pushing behavior, where the aim is to push an object in a certain direction. As little as possible prior object and action knowledge should be provided to the system. Here we use exploration to acquire the initial object-action knowledge. The robot starts by applying random pushing actions from different directions. By collecting the results of these pushing actions, the robot builds a knowledge base, which is later used to train a two-layer neural network that encodes the pushing behavior.

To test the proposed approach we first experimented with pushing of a rectangular object. For this purpose, the robot needs to find an appropriate point of contact and the angle of push so that the object moves in the desired direction. After learning the robot should be able push the object along the desired trajectory.

The included video shows the object, the pusher, and the desired moving direction of the object (blue line). The two movies in the upper half of the video show the situation as seen from two different cameras (the overhead view provides information to the robot), while the lower portion of the video shows the robot's view of the world. Each time when the pusher needs to change its position with regard to the object, the robot determines if the pusher will collide with the object or not. If yes, then the pusher has to move up in order to avoid collision with the object. The path from the starting point to the goal point, either for the actual push or for reaching a suitable starting point for pushing, is drawn with green color. In the video we can see that the robot automatically finds a suitable point of contact for pushing and that the object is pushed in the desired direction.

5. Improved Object Accumulation

(AccumulationKalman.m2v)

The video shows the results of the accumulation of object shape based on the self induced robot motion (see deliverable 4.1.2) applying corrections by Kalman filters of independent estimates for each 3D primitive. The video shows the accumulated representation with and without the correction. A clear improvement is visible which will facilitate the performance of other processes high-level processes that make use of the accumulated representations (e.g., pose estimation, object recognition as well as higher level reasoning processes on objects).

6. Computation of Optimal Viewpoints

(VideoExtraccionModelo3DPoliedrico.wmv)

The video depicts preliminary results on entropy-based view point planning for model acquisition. In this sequence, features are planar segments in the scene, which are extracted by a range sensing device. The Staubli arm chooses the best next viewpoint in building an object model, based on an entropy reduction strategy. This action selection mechanism is currently being migrated to the objects/features related to the PACO kitchen scenario, using SDU sequences to further improve the work on the 'Birth of the object' (see [3]).

7. Rule Learning

(7to0.mpg)

We have extended CSIC's rule learning system from a one goal to a multi-goal profile. Rules can now be used reactively, as in action policy behaviors, or deliberatively, to generate plans by using rules as planning operators. To permit rules to be applied either reactively or deliberatively they are coded as a set of preconditions, a sequence of actions, and the final expected outcome. The preconditions must be observed before the rule can be applied, and the expected outcome encodes the effects that will be obtained after the execution of the action sequence. The action sequence consists of a list of actions, each one expressed in turn as a cause-effect. The video shows an example of the rule learning system applied to the solution of a simple manipulation problem, consisting of a 3x3 grid world in which a marked box must be moved from one place to another in the grid. There is only one empty cell at a time, and the robot must move boxes around to clear a path for the marked box.

References

[1] D. Aarno, J. Sommerfeld, D. Kragic, N. Pugeault, S. Kalkan, F. Wörgötter, D. Kraft, and N. Krüger. Early reactive grasping with second order 3d feature relations. *The IEEE International Conference on Advanced Robotics, Jeju Island, Korea*, 2007.

[2] D. Kraft, E. Başeski, M. Popović, A. M. Batog, A. Kjær-Nielsen, N. Krüger, R. Petrick, C. Geib, N. Pugeault, M. Steedman, T. Asfour, R. Dillmann, S. Kalkan, F. Wörgötter, B. Hommel, Renaud Detry, and Justus Piater. Exploration and planning in a three level cognitive architecture. In *International Conference on Cognitive Systems (CogSys)*, 2008. (accepted).

[3] D. Kraft, N. Pugeault, E. Başeski, M. Popović, D. Kragic, S. Kalkan, F. Wörgötter, and N. Krüger. Birth of the object: Detection of objectness and extraction of object shape through object action complexes. *Special Issue on "Cognitive Humanoid Robots" of the International Journal of Humanoid Robotics*, 2008. (accepted).

[4] M. Popovic. An early grasping reflex in a cognitive robot vision system. In *Master Thesis at the Maesrk Institute at the University of Southern Denmark*. 2008.