SEVENTH FRAMEWORK PROGRAMME

Project Acronym:	Xperience
Project Type:	IP
Project Title:	Robots Bootstrapped through Learning from Experience
Contract Number:	270273
Starting Date:	01-01-2011
Ending Date:	31-12-2015

XPERIENCE.ORG

Deliverable Number:	D5.2.1
Deliverable Title:	Scenario 1 demonstration (I): Robot demonstration on learning of object specific affordances
Type (Internal, Restricted, Public):	PU
Authors:	D. Kraft, N. Krüger, A. Ude, T. Asfour, D. Schiebener
Contributing Partners:	SDU, UIBK, JSI, KIT

Contractual Date of Delivery to the EC: 31-01-2012
Actual Date of Delivery to the EC: 06-02-2012

Contents

1 Executive Summary **3**

2 Content of the Deliverable **4**

2.1 Affordance based grasp exploration based on two grasping OACs 4

2.2 Stirring and cutting 4

2.3 Object exploration by pushing 4

2.4 Blind grasping 5

3 Conclusion **7**

3.1 Links to other Workpackages 7

Chapter 1

Executive Summary

This Deliverable deals with the demos in the context of WP5.2. The execution of a variety of OACs and the associated learning is shown on different platforms at SDU, KIT, JSI, and UGOE.

The videos associated with Section 2.1 and Section 2.2 will be delivered 2 weeks before the review.

Chapter 2

Content of the Deliverable

A learning agent needs to be able to explore its environment and seamlessly acquire data that can later be used for fast learning by structural bootstrapping. Reliable explorative processes are needed for this purpose. In the attached videos we demonstrate our first implementations of such explorative processes.

2.1 Affordance based grasp exploration based on two grasping OACs

This subpart shows the execution of the two explorative grasping behaviors described in [7] AgnoGrasp (grasping without object knowledge, see [9]) and ObjGrasp (grasping utilizing object knowledge, see [6]). Grasping is done with the SDH dexterous three finger hand. We show how the two processes are used on the MARVIN platform at the SDU environment to build up grasping experience in terms of 'experiments'. Grasping attempts are shown and internal processes are visualized, e.g., storing the episodic memory and the development reliability measure M. This demo will show the application of the OAC formalism [7]) in a cognitive architecture in which 'experiments' are permanently produced in the exploration process delivering data as an 'outside–in' processes. This data is then the basis for the structural bootstrapping as an 'inside–out' processes.

2.2 Stirring and cutting

We demonstrate two actions: stirring and cutting, executed using Semantic Event Chains (SECs) [1]. We supplement the relational information of SECs by information required for execution, that is: pose, trajectory and force information where required. This way the abstract representation of a SEC turns into the concrete realization of an OAC (Object Action Complex,[13, 7]). We show how execution of actions generalizes to several different situations, dependent on sensory information. This shows that the resulting OAC adheres to the affordance principle, where actions generalize across (similar) objects. The approach is strictly generative and there are no exploratory components involved. This is done on purpose to avoid lengthy learning procedures. The setup is based on a KUKA light-weight robot arm available at UGOE.

2.3 Object exploration by pushing

In the next two videos we show how data about unknown objects can be acquired by applying explorative pushing actions. The pushing behavior (or the pushing OAC) shown in the first video has been acquired by kinesthetic guiding. It is not updated in the course of the experiment. The presented work has been primarily done at JSI and was tested in part in collaboration with ATR, Kyoto, Japan. An implementation for a humanoid robot at KIT is currently under way.

In video **ObjectLearningByPushing** a robot acquires information about unknown objects that afford pushing. The video shows a sequence of pushing actions generated by the robot with the aim of detecting new objects and accumulating data about them. It also shows the results of visual processing. The explorative process starts by extracting visual features like SIFTs and MSERs, which are used as a basis for the generation of initial hypotheses about the objects on the table. After selecting the most promising hypothesis, the robot initializes its pushing OAC (encoded by dynamic systems) and attempts to push the ensemble of hypothetical object features. The hypothesis is assumed confirmed if the hypothetical object features move as a rigid body. The robot then generates additional pushes and collects more information about the object. The accumulated object features are used to learn an object representation based on a state of the art bag-of-features model. The subsequent active object classification process is similar to the model acquisition process; the robot first generates initial hypotheses about possible objects in the scene and then attempts to push the most promising hypothetical object. Based on the resulting image motion, the hypothesis is either confirmed or rejected. If the hypothesis is confirmed, additional object features are segmented from the scene. The segmented features are then used for classification based on the bag-of-features approach. We have shown that the proposed paradigm is successful at autonomous, explorative object learning and classification.

Video **InteractiveObjectLearning** demonstrates that the robot can be aided by a human coach, who guides the learning process by performing the pushing actions (or other types of manipulation actions) instead of the robot. This significantly reduces the complexity of the task because the pushing OAC is not needed. Based on the interface showing the results of hypothesis generation, the coach can select the most promising one and move the object. Since the system does not have any expectation about the induced object motion, this generates the same information visible to the robot as if it pushed the object by itself. In this way data accumulation is simplified and also becomes faster because a human coach can normally perform the movements faster and in a more robust way than the robot.

This work has been published in [10] and [12] (attached to D2.3.1 and D.4.1.1).

2.4 Blind grasping

In our previous work, we investigated methods for the dexterous haptic exploration of unknown objects based on a dynamic potential field approach (see [4, 3, 5]). Based on 3D point clouds acquired from contact information during guiding the robot hand along the surface of the region on interest, grasp hypotheses are calculated based 1) on geometric features extracted from the 3D point cloud and 2) a post-processing step to verify the executability of grasp hypotheses by the robot hand. While the main efforts in our previous work on haptic exploration were devoted to generate a complete 3D object model and as many grasp hypotheses as possible, our strategy in Xperience is slightly different. The goal is to build rich multi-sensory object representations by a tightly coupling of visual and haptic information. Therefore, we are working on an integrated approach, by which unknown objects are explored and grasped in haptic-guided way in a first step. In the second step, objects in the robot's hand are manipulated in front of the robot's eyes to enrich the object representation and generate multiple object views.

First results toward this goal are reported in Chapter 3 of Deliverable 2.1.1.

In the first year of the project, we investigated grasping of unknown objects based on haptic exploration and corrective hand/arm movements. We evaluated the approach on the humanoid robot ARMAR-IIIb in the context of grasping unknown objects placed on the table in front of the robot. Based on visually generated height map of the scene, promising grasping positions are determined. Contact detection and estimation of contact position is determined using multi-sensory information acquired from the 6D force-torque sensors in the wrists, the tactile sensors in the finger tips and the joint encoder values of the finger joints. Upon contact, corrective movements are executed until contact with the palm is detected. In this case, we assume that the object is inside the hand and the hand is closed trying to enclose the object. A grasp stability check is performed using the position and tactile sensor data. Only if the grasp is classified as stable, the object is lifted while continuously checking the grasp stability. The approach for grasp stability detection is based on temporal filtering of a support vector machine classifier output and estimates the stability continuously during the grasp attempt (see [2, 8]). The implementation and validation of the approach on ARMAR is presented in [11].

The video **BlindGrasping** demonstrates the approach of grasping unknown objects, mainly based on haptic information, on the humanoid robot ARMAR-IIIb.

This work has been reported in Chapter 3 of Deliverable 2.1.1. The grasp stability detection based on tactile sensor data and joint angle data of the finger joints is submitted to the IEEE BioRob-2012 conference (see [11], attached to D2.1.1).

Chapter 3

Conclusion

OACs and exploratory behaviors executed on different demonstrators are presented. They provide a basis for the integration efforts that will be performed in the next phase of the project.

3.1 Links to other Workpackages

Deliverable D5.2.1 is based on work performed in the WPs 1-4.

References

[1] E. E. Aksoy, A. Abramov, J. Dörr, K. Ning, B. Dellen, and F. Wörgötter. Learning the semantics of object-action relations by observation. *The International Journal of Robotics Research*, 30(10):1229–1249, 2011.

[2] Yasemin Bekiroglu, Janne Laaksonen, Jimmy Alison Jørgensen, Ville Kyrki, and Danica Kragic. Assessing grasp stability based on learning and haptic data. *Robotics, IEEE Transactions on*, 27(3):616 –629, 2011.

[3] A. Bierbaum, M. Rambow, T. Asfour, and R. Dillmann. A Potential Field Approach to Dexterous Tactile Exploration. pages 360 – 366, Daejeon, Korea, 2008.

[4] A. Bierbaum, K. Welke, D. Burger, T. Asfour, and R. Dillmann. Haptic Exploration for 3D Shape Reconstruction using Five-Finger Hands. pages 616 – 621, Pittsburgh PA, USA, Nov 29 - Dec 01 2007.

[5] Alexander Bierbaum, Matthias Rambow, Tamim Asfour, and Rüdiger Dillmann. Grasp Affordances from Multi-Fingered Tactile Exploration using Dynamic Potential Fields. pages 168 – 174, Paris, France, 2009.

[6] Renaud Detry, Dirk Kraft, Oliver Kroemer, Leon Bodenhagen, Jan Peters, Norbert Krüger, and Justus Piater. Learning grasp affordance densities. *Paladyn Journal of Behavioral Robotics*, (accepted), 2011.

[7] Norbert Krüger, Christopher Geib, Justus Piater, Ronald Petrick, Mark Steedman, Florentin Wörgötter, Aleš Ude, Tamim Asfour, Dirk Kraft, Damir Omrčen, Alejandro Agostini, and Rüdiger Dillmann. Object-action complexes: Grounded abstractions of sensory-motor processes. *Robotics and Autonomous Systems*, 59(10):740 – 757, 2011.

[8] J. Laaksonen, V. Kyrki, and D. Kragic. Evaluation of feature representation and machine learning methods in grasp stability learning. In *10th IEEE-RAS International Conference on Humanoid Robots*, pages 112–117, 2010.

[9] M. Popović, G. Kootstra, J.A. Jørgensen, D. Kragic, and N. Krüger. Grasping unknown objects using an early cognitive vision system for general scene understanding. In *Proceedings of the International Conference on Intelligent Robots and Systems (IROS), September 25-30, San Francisco, CA.*, 2011.

[10] D. Schiebener, A. Ude, J. Morimoto, T. Asfour, and R. Dillmann. Segmentation and learning of unknown objects through physical interaction. In *Proc. IEEE-RAS Int. Conf. Humanoid Robots, Bled, Slovenia*, pages 500–506, 2011.

[11] J. Schill, J. Laaksonen, M. Przybylski, V. Kyrki, T. Asfour, and R. Dillmann. Learning Continuous Grasp Stability for a Humanoid Robot Hand Based on Tactile Sensing. In *IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012. (submitted).

[12] A. Ude, D. Schiebener, H. Sugimoto, and J. Morimoto. Integrating visual processing and manipulation for autonomous learning of object representations. In *Proc. IEEE Int. Conf. Robotics and Automation, Saint Paul, MN*, 2012.

[13] Florentin Wörgötter, Alejandro Agostini, Norbert Krüger, N. Shylo, and Bernd Porr. Cognitive agents - a procedural perspective relying on the predictability of object-action-complexes (oacs). *Robotics and Autonomous Systems*, 57(4):420–432, 2009.