

KSOM

STARTPAGE

PEOPLE  
MARIE CURIE ACTIONS

International Outgoing Fellowships (IOF)  
Call: FP7-PEOPLE-2011-IOF

PART B

"KSOM"

## Contents

1 RESEARCH AND TECHNOLOGICAL QUALITY	1
2 TRAINING	10
3 RESEARCHER	11
4 IMPLEMENTATION	17
5 IMPACT	20
6 ETHICS ISSUES	23



Figure 1: Robot simulating a shopping task and making use of KSOM generated semantic object map.

## 1 RESEARCH AND TECHNOLOGICAL QUALITY

### 1.1 Research and technological quality, including any interdisciplinary and multidisciplinary aspects of the proposal

Environment models are resources for enabling robots to perform their tasks more reliably, efficiently, and competently by using information about the environments. The proposed project KSOM (Knowledge-enabled Semantic Object Maps) investigates a new generation of environment models that are to enable robots to perform everyday manipulation tasks in human environments effectively and efficiently. To this end, KSOM proposal extends state-of-the-art models such as Semantic Object Models with probabilistic spatio-temporal information: they probabilistically represent where objects typically are, where objects belong, the state of objects depending on their location (perishable objects are typically to be found in refrigerators), the dependence of activity contexts on object locations (if pots are on the oven then the activity context is probably meal preparation). Using these representational mechanisms KSOM project provides semantic object maps that can characterize objects and places in the environment with respect to the role they play in activities (affordances). These are necessary for autonomous robots to do the right actions to the right objects in the right way: when setting the table, the robot should retrieve the cups from the cupboard because they are clean. When cleaning the table the robot should put the dirty cups into the dishwasher to clean them. Semantic object maps in KSOM project enable the robot to infer this kind of commonsense knowledge. Technologically, KSOM will be realized as semantic object models combined with relational probabilistic models that relate objects to activities, places, time, roles, etc. KSOM semantic object maps are (semi-)automatically acquired through the combination of semantic object model acquisition, the observation of everyday activities, the use of web instructions on everyday activities and statistical relational learning. Usefulness of the KSOM project will be demonstrated in the context of setting the table, loading the dishwasher and simulating the shopping task (see Figure 1, [video](#)). The joint probability distributions of semantic object maps in KSOMare to represent the dynamic aspects of the environment in the context of everyday activities, which include:

- *the typical locations of objects*: cups and plates can typically be found on the table, in the dishwasher, and in the cupboards;
- *the home/storage location of objects*: cups and plates are stored in a particular cupboard before

they are used and after they are cleaned;

- *typical arrangements of objects*: some objects are typically used together in certain configurations, for example at lunch;
- *action-related places*, the places where actions and activities are performed;
- *role/affordance-based object models*: the roles that objects play in activities;
- *time-space-state representations*: the changes of objects' state and position over time.

**Dejan: SHORTEN**

According to the research topics of the KSOM project, we will now briefly review relevant research directions in the areas of semantic object maps, object perception, knowledge representation for robots, statistical relational models, spatio-temporal learning, and entity resolution. The project will explicitly not deal with the research on perception of humans and their activities. Where needed we will either apply manual labeling techniques or make use of the existing frameworks such as e.g. OpenNI<sup>1</sup>.

### 1.1.1 Semantic Object Maps: Acquisition and Use

As semantic object maps include models of the environment structure and the static components of the environment including the furniture, appliances, etc. we will first consider the state of the art in the acquisition and use of 3D semantic object maps for everyday manipulation.

In Wuenstel and Moratz (2004), Wünstel and Moratz use a graph representation to detect chairs, but the relation descriptions are manually estimated, and thus it is unclear whether the proposed method scales. Posner et al. Posner et al. (2008) use probabilistic graphical models such as Markov random fields to label planar patches in outdoor urban datasets. Their work is based on that of Anguelov et al. Anguelov et al. (2005) and Triebel et al. Triebel et al. (2007), which define point-based 3D descriptors and classify them with respect to object classes such as chairs, tables, screens, fans, and trash cans in the former, respectively wires, poles, ground, and scatter in the latter.

Iocchi and Pellegrini Iocchi and Pellegrini (2007), Mozos et al. Mozos et al. (2007), and Vasudevan et al. Vasudevan et al. (2007) use 2D laser sensors to create a map used for navigation and additional semantics are acquired through the use of vision. An object-based approach for cognitive maps is used to recognize objects and classify rooms in different categories in Vasudevan et al. (2007), while in Mozos et al. (2006, 2007), places are semantically labeled into doorways, kitchens, corridors, rooms. The advantages of these representations are straightforward: they keep the computational cost low enough and base their localization and pose estimation on the well-known 2D SLAM (Simultaneous Localization and Mapping) problem, while the problem of place labeling is solved through the usage of feature descriptors and machine learning. However, by reducing the dimensionality of the mapping to 2D, most of the world geometry needed for manipulation is lost. Also, the label categories need to be learned a priori through supervised learning and this makes it unclear whether these representations scale well.

For modeling the static part of the environment Nüchter and Hetzberg Nüchter and Hertzberg (2008) classify 3D sensed data from a laser sensor into walls, floor, ceiling, and doors based on angular thresholds. Static objects including standing humans are detected in two steps, first a hypothesis is extracted from a depth image using an appearance based method, then 3D matching to a model is performed to evaluate the classification. This however requires 2D models for all poses of each object to be detected, as well as the complete 3D models. The same holds true for postures in the case of humans.

**Include my mapping paper.**

### 1.1.2 Object Detection, Recognition, and Reconstruction for Pick-and-Place Tasks

Semantic object maps in KSOM must also include models of the objects of daily use, which are to be manipulated by the robot. Thus, we now consider recent work in object detection, categorization, localization, and reconstruction for robotic object perception.

---

<sup>1</sup><http://www.openni.org/>

A computer vision- and machine learning-based method is used by Saxena et al. Saxena et al. (2008) to train classifiers that can predict the grasping points in an image. This is then applied to images of unseen objects. To obtain 3D positions of grasping points, the authors use stereo cameras, but their approach works reliably only to the extent provided by the training data. Another issue is the segmentation of objects, since grasp points are provided with no information about what objects are in the scene and to which of them do the identified points correspond. Bone et al. Bone et al. (2008) use an accurate line laser and a camera to build models and identify grasping points for novel objects with very encouraging results. However the system was tested only on two objects, thus its scalability is not clear.

In purely computer vision based approaches, features like the ones described by Lowe Lowe (2004) or Lepetit and Fua Lepetit and Fua (2006) are used to find matches between parts of a scene and a database of object images. The problem with these kinds of approaches is that they only work for objects that are in the database, and since no knowledge about the 3D information is known, the system can easily make mistakes and return false positives (e.g., a cereal box containing a picture of a beer bottle printed on it will get recognized as a bottle of beer). Some of the solutions adopted involve the offline creation of complete 3D models for the targeted objects and finding feature spaces to match the partial views with models in the database, as presented by Collet et al. Collet et al. (2009). Another approach to obtain 3D information directly from camera images is to project CAD models from a database to the image and search for good matches in the edges domain, as in Ulrich et al. Ulrich et al. (2009) for example. While this is a more direct method, it is still dependent on a database of CAD models.

Thrun and Wegbreit Thrun and Wegbreit (2005) describe a method for detecting and verifying symmetries in point clouds obtained from a single viewpoint which works very well for nicely segmented objects, however the problem of under- or over-segmented objects remains.

Object categorization goes hand in hand with segmentation and is usually performed using a single sensing device. Given a large set of training values containing all possible views, most approaches try to abstract the problem by using features like the ones proposed by Quack et al. Quack et al. (2007), Yan et al. Yan et al. (2007), and Savarese and Fei-Fei Savarese and Fei-Fei (2007), which work best on low scale and texture variance. The scaling variance can be reduced significantly by a previous segmentation.

To improve the segmentation, Lai and Fox Lai and Fox (2009) take a randomized approach by classifying a “soup of segments” generated by different combinations of clusters, and use the results to get a final segmentation. They also explore different domain adaptation techniques in order to incorporate synthetic data in training their classifier. Though they are working with outdoor data, the same techniques can be applied to segment and classify indoor objects using low resolution 3D data, e.g. from time-of flight cameras.

include my both humanoids papers

### 1.1.3 Symbolic Knowledge for KSOM

For performing high-level tasks, robots need large amounts of semantic information from their environment model, like the types, locations and properties of objects. Very few systems exist that offer this kind of deep semantic environment information.

At the highest level of abstraction, one considers the recognized objects, whose semantic meaning is only implicitly represented: Humans immediately associate various properties with something called a “cupboard”, while robots usually do not have this kind of knowledge. Without an explicit knowledge representation, different robots or even different parts of the same robot may have a very different notion of an object.

Deeper semantic representations, which also describe object properties, such as the point at which to grasp an object or location of the opening of a container like a bottle, are used by Okada et al. Okada et al. (2007) but are mainly hand-coded and do not leverage the power of hierarchical, abstract knowledge representations. Galindo et al. Galindo et al. (2008) present a system for automatically building maps that combine a *spatial hierarchy* of local metric and global topological maps with a *conceptual hierarchy* that describes some semantic properties of rooms and objects. In this respect, their approach is similar to ours, but the conceptual hierarchy is much simpler and the spatial description much coarser.

include my IROS 2011 paper

#### 1.1.4 Statistical Relational Models

In order to declaratively represent highly complex domains, in which there are relations between a variable number of relevant entities and which are furthermore governed by uncertainty, one requires a representation formalism that combines statistical with relational components, abstracting away from concrete entities to compactly represent general principles about the relevant aspects of the real world. In statistical relational learning, a number of such formalisms have been proposed, as presented by Getoor and Taskar Getoor and Taskar (2007) and De Raedt De Raedt (2008). Statistical relational learning methods have countless applications, including collective classification Neville and Jensen (2003), link prediction Taskar et al. (2003) and object identification Singla and Domingos (2006). In particular, the methods of statistical relational learning allow to represent full-joint distributions over logical propositions about a changing set of entities in a concise, declarative manner, and they provide a fully integrated framework for learning and inference. Therefore, they are ideally suited to the representation of the probabilistic components of semantic object maps in KSOMas we envision them, since we need to consider, for example, spatial relations between objects in the environment, their attributes changing over time, their relevance to activities taking place and the effect these activities may have upon them.

#### 1.1.5 Spatio-Temporal Learning

Our intended research in learning spatio-temporal structures for semantic object maps in KSOMis inspired by research in visual analytics for the analysis of movement data. Andrienko et al. learn concepts like the working place, the living place, typical navigation routes and other spatio-temporal behavior patterns from GPS-data from their car Andrienko et al. (2007). While in their case, the data mining tasks are performed by human experts using visual analytics methods, other researchers such as Liao et al. Liao et al. (2007b,a) perform some of these learning tasks using probabilistic learning methods (hierarchical conditional random fields). Acquiring such models is also investigated in the pervasive systems community, where Philipose and his colleagues Pentney et al. (2007); Landwehr et al. (2007) and Intille and his research group Intille et al. (2006) learn models of daily activities from ubiquitous sensor networks.

#### 1.1.6 Entity Resolution

A key problem that has not yet received substantial attention in autonomous robotics is the inference about whether two different observations resulted from the same real world object. This type of inference is needed to recognize previously observed objects again, which in turn must be solved to build spatio-temporal environment models.

While early work has often phrased the problem as a classification problem where a pair of records would be independently classified as either “matching” or “non-matching” Fellegi and Sunter (1969), more recent approaches have phrased the problem as link prediction in statistical relational models, e.g. Singla and Domingos Singla and Domingos (2006): One defines an equivalence relation over entities and considers, in a probabilistic setting, any predicates that state the same things about two entities as evidence for the entities referring to one and the same real-world entity; the importance of certain predicates can be learned from statistical data. Most importantly, this relational approach considers the objects collectively and does not assume that the equivalence of a pair of objects is independent of other equivalences.

## 1.2 Appropriateness of research methodology and approach

In order to clarify the motivation from above, we identified the following three subgoals of the project:

**Subgoal 1: Representations in KSOM** will be hybrid by integrating geometric descriptions (such as primitives, meshes, voxels), first-order description-logic-based symbolic representations (e.g. symbolic object properties, relations between objects and object classes) and probabilistic first-order representations (including Markiv logic and Bayesian logic networks). Together with a representation language based on OWL (Web Ontology Language) and a concept ontology based on researchCyc (an encyclopedic knowledge base), these representations can be stored, extended and they can be queried by a Prolog-based inference mechanism.

**Subgoal 2: Perceptual mechanisms** for detection, categorization, recognition, localization and reconstruction of objects of daily use. Additionally, the robot needs capabilities to perceive and interpret arrangements of objects (spatially and/or temporally related). This includes methods for identity resolution and estimation of object state based on partial information and context to allow affordance-based manipulation.

**Subgoal 3: Learning of semantic object maps** for autonomous world state interpretation. Some of the identified learning problems are:

- Learning typical locations of objects such as storage or usage locations, and interpretation of containers based on a generalization of their contents (e.g. fridge);
- Learning action-related places are places where actions are performed or objects take on certain roles in actions. This allows a robot to learn e.g. the concept of a chair by its role to support people as opposed to purely geometric or appearance based recognition;
- Learning arrangements of objects is important to abstract from a continuous coordinate system of object locations to a symbolic, relational model of a scene which allows e.g. for context interpretation of a table arrangement;
- Knowledge-intensive learning from selected web sites such as researchCyc for encyclopedic knowledge, ehow.com or wikihow.com for task instructions and the OpenMind Indoor Common Sense (OMICS) knowledge base. This constitutes important background information for KSOM project.

These learning problems require autonomous and on-demand (partial) processing of sensor data to enable life-long learning.

All of the subgoals named above are to be implemented in a system that is to run on real robot platforms in real household environments.

The project's goals will be achieved in three work packages, reflecting the previously identified subgoals.

#### **WP 1: Representations of KSOM semantic object maps**

The first work package is concerned with achieving subgoal 1, the design and implementation of appropriate representational and reasoning mechanisms required by KSOM. It is subdivided into the following three tasks.

**Task 1.1: OWL Concept Taxonomy** The basic symbolic representation will be realized using SWI Prolog with an additional package for OWL (Web Ontology Language). Based on previous work of the return host, we will specify concepts for objects, object groups, places, object and scene states, and we will anchor these concepts in the data structures used for map representation and those generated by the perception routines to be developed in WP 2. As an example, objects of daily use will have locations attached to them, such as *home* places, state-dependant locations (dish washer when dirty), and locations where they belong (e.g. within arrangements of objects on a table).

**Task 1.2: Spatial Relations and Scene Representation and Reasoning** Representational primitives of key importance in KSOM project are qualitative as well as quantitative spatial relations. They are essential for parameterizing robot actions and communicating abstract information about the environment, respectively. While quantitative relations can be obtained in a straight-forward manner, qualitative will be realized following an approach inspired by Gapp (1995) and will be learned from observations of human activity.

**Task 1.3: Representations of Action-related Concepts** In order to relate objects, scenes, etc. to the actions and the activities they are involved in, we will develop *action-related concepts* as our key representational mechanisms. Preliminary investigations in this regard have been performed in the context of mobile pick-and-place tasks, where the return host group has developed the notion of *action-related places* (ARPLACES), i.e. the set of robot poses from which a given pick task is predicted to be successful. A second starting point is our previous work on *grounded action models* (GRAM), where we have introduced a new class of concepts in description-logic-based knowledge

representations, integrating actions and action models into the knowledge representation and inference mechanisms of intelligent systems.

**WP 2: Perception** The objective in WP 2 is to achieve subgoal 2, i.e. achieving the perceptual capabilities required for object and scene recognition and interpretation, object detection and reconstruction.

**Task 2.1: Object Detection and Perception** In this task, we will design, implement and analyze a perception system that enables the robot to infer *when which object is where* by accomplishing the following perceptual task: given (1) a set of perceptually distinct object instances (e.g. textured mugs, cereal boxes), (2) a set of perceptually indistinguishable object instances (e.g. plates, bowls) and (3) a set of regions of interest (e.g. tabletops, counters), detect and localize these objects of interest whenever they are present in those regions. The result is a set of time-stamped observations that include the object of interest, its pose, and the respective region of interest.

The research in this work package focuses mainly on object identity resolution and the inference about whether objects have been removed from regions of interest. Object identity resolution is the problem of deciding whether or not two partial views of some objects in the environment taken at different times  $t_1$  and  $t_2$  refer to the same object in the environment. The aspects of the problem that render this inference task hard are that the different views potentially have little overlap, that the views are corrupted by sensor noise and specular reflections and that views might be limited through occlusions. We will investigate how we can apply and adapt mechanisms from entity resolution originally developed in the area of data mining to better handle the spatial and temporal context of this perception task.

include paper by Blodow, Jain

**Task 2.2: Scene Perception, Interpretation and Analysis** This task will investigate the knowledge-based perception mechanisms for scenes and object arrangements. The inclusion of knowledge-based mechanisms will enable us to improve scene perception through the use of prior knowledge, the semantic object model of the environment, knowledge about everyday activities, and the effects of everyday activities on the situations in the environment. include WUP similarities

**Task 2.3: Perceiving Object States** In the project KSOM, the real-world perception of object states is beyond the scope of the project, but we still need to realize perception mechanisms for state estimation in order to realize a complete and integrated system.

We will simplify the state estimation in two ways. First, we will use — where-ever possible — sensor-equipped objects in distributed sensor networks that can estimate their own states, e.g. whether they are filled. Second, for those states that are needed but cannot be perceived otherwise, we will simplify the perception tasks by making the states perceptually distinctive, e.g. by using RFID tags to store state information or by color coding states.

**WP 3: Acquisition and Learning of KSOM semantic object maps** In this work package, we focus on the learning problems in KSOM which can be described as follows: *Given* (1) background knowledge needed for learning of semantic object maps, and (2) a stream of time-stamped partial observations of situations concerning what is on tables, on the counter, the states of objects, and object arrangements, *learn* a mapping for this particular environment.

**Task 3.1: Learning Grounded Representations** The type of information that we want to leverage here comes from web instructions. To make sense of web instructions, one first needs to apply natural language processing techniques. By using a parser based on probabilistic context-free grammars (PCFGs), return host group has showed how to obtain the structure of sentences and was able to assign to each word an appropriate synonym ring. They then used the WordNet lexical database in order to disambiguate word senses based on learned correlations between entity types, action verbs and prepositions, and to furthermore link the actions and entities appearing in instructions to concepts within the Cyc upper ontology. In this way, they can semantically interpret the set of instructions that constitutes a particular activity, which opens up new possibilities for detecting these activities within the environment.

**Task 3.2: Learning Important Locations** Key locations where certain actions take place need to be identified from observation data, both for the learning of action-related places and the understanding of human behavior. To this effect, we will need to develop extensions to existing methods such as hierarchical conditional random fields that also take related actions into account.

Additionally, the typical and storage locations for objects have to be inferred from the probabilistic models built by observations and web instructions. An important aspect to consider is that KSOM needs to learn the distributions of locations both for (1) specific objects and (2) objects of a specific type.

**Task 3.3: Learning Object Arrangements** We are planning to expand the perception and learning system we developed to consider spatial relations between different objects just as it considered relations between different parts of objects. We will be using *WUP* similarities. An example of such a feature is a qualitative spatial relation such as *left of*, which is represented as a pair of positions  $\langle o, refo \rangle$  where  $o$  is the position of an object and  $refo$  is the position of the reference object. So  $P(\text{left-of}(o, refo))$  denotes the probability that with respect to their coordinates the object  $o$  would be considered to be to the *left of* the object  $refo$ .

It is also important to note that qualitative spatial relations depend on the object types. Depending on whether the object pair would be a plate and a fork or a building and a gate, the expected positions satisfying the relation would be very different. We intend to learn appropriate probability distributions from observations of human activities.

**Task 3.4: Lifelong KSOM Learning** In this task, we will realize a *Passive Perception* system, i.e. a set of mechanisms that enable the robot to remember scenes that it encounters while performing its activities, storing them as memory structures in the knowledge base for incremental learning. In that sense we will implement fast logging mechanisms for raw and partially processed sensor data, derive sophisticated memory infrastructure formats (aka time-stamped objects store) and develop interfaces for continuous and on-demand processing of previously observed data.

### 1.3 Originality and Innovative nature of the project, and relationship to the state-of-the-art of research in the field

The ultimate goal of the project KSOM is the investigation of environment models that can be inferred from a combination of environment exploration and activity observation. To this end, we will design, implement and empirically analyze semantic model acquisition mechanisms for objects and everyday activities in order to generate semantic object maps of human living environment.

As an example, consider the following query that might arise within a household scenario: A human asks the robot to retrieve a glass he had been drinking from earlier. This simple request can spawn several queries to KSOM system, such as: Where is a glass that this person had contact with earlier? There might be multiple glasses, so the system has to select the one which was used for drinking and disregard other glasses that the human might have cleaned earlier. Also, for picking up the glass, grasp analysis might need a surface mesh of the object or a precomputed list of applicable grasps, and the motion planner needs an occupancy voxel map or surface mesh of the whole environment for collision avoidance or possibly additional manipulation constraints linked to the glass, such as maintaining a vertical orientation over the whole action. A system that is able to tackle above described challenges, to the best of our knowledge, does not exist.

The system will gain the maximum leverage from the algorithms and code available in Robot Operating System <sup>2</sup> as well as from the cooperation with the world renowned scientists involved in the PR2 Beta Program <sup>3</sup>. In particular we plan to collaborate with and search expertise from Rachid Alami from LAAS, Wolfram Burgard from University of Freiburg, Ben Pitzer from Bosch RTC, Brian Gerkey et al from Willow Garage Inc, Kei Okada from University of Tokyo, Herman Bruyninckx from KUL and Maxim Likhachev (previously UPENN, now CMU). All enlisted scientists have been consulted and confirmed their participation.

---

<sup>2</sup>[www.ros.org](http://www.ros.org)

<sup>3</sup><http://www.willowgarage.com/pages/pr2/pr2-community>

## 1.4 Timeliness and relevance of the project

With all the recent advances in the field of personal robotics and with first personal robots projected to be deployed in households within next 5 to 10 years, KSOM's relevance is enormous for robots to perform in highly dynamic and unpredicted situations. It will also enable researchers from other fields such as artificial intelligence, psychology, ergonomics, etc to get a grasp at the data and study human living patterns, social acceptance of robots and other high-level concepts. The KSOM project's relevance under the Marie Curie fellowship within a European setting is connected to the innovative ideas it introduces by combining in a novel way approaches to build semantic environment models to be used in robots' everyday operations. The applicant will need to be additionally trained within the host group, as will be discussed in Section 2, to pursue the proposed research, which is also a significant part of the Marie Curie program. Since the KSOM project is setup such to exploit the mutual leverage with the open source ROS community and the PR2 Beta community in terms of code re-usability, feedback from other experienced researchers and critical reviews, it will make a significant contribution in enhancing the continuously advancing scientific excellence within Europe. Both the applicant and the environment in which the project will be realized will benefit significantly from the realization of the proposed project, since also members, senior and junior ones, from the host group will have their opportunity to assist the KSOM research.

## 1.5 Host research expertise in the field and quality of the group/supervisors

Michael Beetz from Intelligent Autonomous Systems (in continuation IAS) group, Technische Universität München (TUM) will represent the return host organization in the project and Pieter Abbeel and Ken Goldberg, University of California at Berkeley (in continuation Berkeley) will represent the outgoing host organization. Prof. Beetz is a vice coordinator of the German national cluster of excellence COTESYS (Cognition for Technical Systems) where he is also co-coordinator of the research area "Knowledge and Learning". Prof Beetz was a member of the steering committee of the European network of excellence in AI planning (PLANET) and coordinating the research area "robot planning". He is associate editor of the AI Journal. His research interests include plan-based control of robotic agents, knowledge processing and representation for robots, integrated robot learning, and cognitive perception. Prof. Beetz has published over 200 articles in the first-class conferences and journals in the fields of robotics, computer vision and artificial intelligence<sup>4</sup>. Currently he is leading a group of 2 postdoctoral fellows, 25 grad students and a large number of undergrad students and visiting researchers. He is also serving as a principal investigator for TUM's PR2 Beta Program. Prof. Beetz has successfully coordinated and trained more than 15 post-doctoral fellows and PhD students, in addition to a larger number of master's and diploma students. Most of his group alumni currently hold faculty positions in Universities and Research Institutions in Europe and United States. The equipment that is required for the realization of the project at the host organization is already available. In particular, three robot manipulation platforms – a PR2<sup>5</sup>, an iCub<sup>6</sup> and a custom-built robot with KUKA lightweight arms and DLR/HIT hands – are available for real-world experiments on a daily basis. So is the testbed kitchen environment with an ample set of ubiquitous devices such as Kinect sensors, ceiling cameras, RFID readers, etc. Pieter, Ken: Above like description for you and Ken

Dejan: Shorten list of references

## References

- Andrienko, G., Andrienko, N., and Wrobel, S. (2007). Visual analytics tools for analysis of movement data. *ACM SIGKDD Explorations*, 9(2):38–46.
- Anguelov, D., Taskar, B., Chatalbashev, V., Koller, D., Gupta, D., Heitz, G., and Ng, A. (2005). Discriminative learning of Markov random fields for segmentation of 3d scan data. In *In Proc. of the Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 169–176.
- Bone, G., Lambert, A., and Edwards, M. (2008). Automated Modeling and Robotic Grasping of Unknown Three-Dimensional Objects. In *Proceedings of the 2008 IEEE International Conference on Robotics and Automation, Pasadena, USA*.
- Collet, A., Berenson, D., Srinivasa, S. S., and Ferguson, D. (2009). Object Recognition and Full Pose Registration from a Single Image for Robotic Manipulation. In *IEEE International Conference on Robotics and Automation (ICRA), Kobe, Japan*.
- De Raedt, L. (2008). *Logical and Relational Learning*. Cognitive Technologies. Springer.
- Fellegi, I. P. and Sunter, A. B. (1969). Ad//dx.doi.org/10.2307/2286061. *Journal of the American Statistical Association*, 64(328):1183–1210.

<sup>4</sup><http://ias.cs.tum.edu/publications>

<sup>5</sup><http://www.willowgarage.com/pages/pr2>

<sup>6</sup><http://www.robotcub.org>

- Galindo, C., Fernández-Madrigal, J.-A., González, J., and Saffiotti, A. (2008). Robot task planning using semantic maps. *Robot. Auton. Syst.*, 56(11):955–966.
- Gapp, K.-P. (1995). Object localization: Selection of optimal reference objects. In *COSIT*, pages 519–536.
- Getoor, L. and Taskar, B. (2007). *Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning)*. The MIT Press.
- Intille, S., Larson, K., Tapia, E. M., Beaudin, J., Kaushik, P., Nawyn, J., and Rockinson, R. (2006). Using a live-in laboratory for ubiquitous computing research. In Fishkin, K., Schiele, B., Nixon, P., and Quigley, A., editors, *Proceedings of PERVASIVE 2006*, volume LNCS 3968, Berlin Heidelberg. Springer-Verlag.
- Iocchi, L. and Pellegrini, S. (2007). Building 3D maps with semantic elements integrating 2D laser, stereo vision and INS on a mobile robot. In *2nd ISPRS International Workshop 3D-ARCH*.
- Lai, K. and Fox, D. (2009). 3D laser scan classification using web data and domain adaptation. In *Proceedings of Robotics: Science and Systems*, Seattle, USA.
- Landwehr, N., Gutmann, B., Thon, I., Philipose, M., and Raedt, L. D. (2007). Relational transformation-based tagging for human activity recognition. In *Proceedings of the 6th Workshop on Multi-Relational Data Mining (MRDM)*, Warsaw, Poland.
- Lepetit, V. and Fua, P. (2006). Keypoint recognition using randomized trees. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(9):1465–1479.
- Liao, L., Fox, D., and Kautz, H. (2007a). Extracting places and activities from gps traces using hierarchical conditional random fields. *International Journal of Robotics Research*.
- Liao, L., Patterson, D., Fox, D., and Kautz, H. (2007b). Learning and inferring transportation routines. *Artificial Intelligence*.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110.
- Mozos, O. M., Rottmann, A., Triebel, R., Jensfelt, P., and Burgard, W. (2006). Semantic Labeling of Places using Information Extracted from Laser and Vision Sensor Data. In *In Proceedings of the IEEE/RSJ IROS Workshop: From sensors to human spatial concepts*, Beijing, China.
- Mozos, O. M., Triebel, R., Jensfelt, P., Rottmann, A., and Burgard, W. (2007). Supervised Semantic Labeling of Places using Information Extracted from Laser and Vision Sensor Data. *Robotics and Autonomous Systems Journal*, 55(5):391–402.
- Neville, J. and Jensen, D. (2003). Collective classification with relational dependency networks. *Journal of Machine Learning Research*, 8:2007.
- Nüchter, A. and Hertzberg, J. (2008). Towards semantic maps for mobile robots. *Journal of Robotics and Autonomous Systems (JRAS), Special Issue on Semantic Knowledge in Robotics*, 56(11):915–926.
- Okada, K., Kojima, M., Tokutsu, S., Maki, T., Mori, Y., and Inaba, M. (2007). Multi-cue 3D object recognition in knowledge-based vision-guided humanoid robot system. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2007.*, pages 3217–3222.
- Pentney, W., Philipose, M., Bilmes, J., and Kautz, H. (2007). Learning large scale common sense models of everyday life. In *Proceedings of AAAI 2007*, Vancouver BC.
- Posner, I., Cummins, M., and Newman, P. (2008). Fast Probabilistic Labeling of City Maps. In *Proceedings of Robotics: Science and Systems*, Zurich.
- Quack, T., Ferrari, V., Leibe, B., Van Gool, L., and Zurich, E. (2007). Efficient mining of frequent and distinctive feature configurations. In *IEEE 11<sup>th</sup> International Conference on Computer Vision (ICCV)*, pages 1–8.
- Savarese, S. and Fei-Fei, L. (2007). 3D generic object categorization, localization and pose estimation. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8.
- Saxena, A., Driemeyer, J., and Ng, A. Y. (2008). Robotic Grasping of Novel Objects using Vision. *The International Journal of Robotics Research*, 27(2):157–173.
- Singla, P. and Domingos, P. (2006). Entity Resolution with Markov Logic. In *In ICDM*, pages 572–582. IEEE Computer Society Press.
- Taskar, B., Wong, M. F., Abbeel, P., and Koller, D. (2003). Link prediction in relational data. In *In Neural Information Processing Systems*.
- Thrun, S. and Wegbreit, B. (2005). Shape from symmetry. In *Proceedings of the International Conference on Computer Vision (ICCV)*, Beijing, China. IEEE.
- Triebel, R., Schmidt, R., Mozos, O. M., and Burgard, W. (2007). Instance-based AMN classification for improved object recognition in 2d and 3d laser range data. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), (Hyderabad, India)*.
- Ulrich, M., Wiedemann, C., and Steger, C. (2009). Cad-based recognition of 3d objects in monocular images. In *International Conference on Robotics and Automation*, pages 1191–1198.
- Vasudevan, S., Gächter, S., Nguyen, V., and Siegwart, R. (2007). Cognitive maps for mobile robots—an object based approach. *Robot. Auton. Syst.*, 55(5):359–371.
- Wuenstel, M. and Moratz, R. (2004). Automatic Object Recognition within an Office Environment. In *CRV '04: Proceedings of the 1st Canadian Conference on Computer and Robot Vision*, pages 104–109.
- Yan, P., Khan, S., and Shah, M. (2007). 3D Model based Object Class Detection in An Arbitrary View. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–6.

## 2 TRAINING

Dejan

### 2.1 Clarity and quality of the research training objectives for the researcher

How have I integrated and performed so far, how will that continue

### 2.2 Relevance and quality of additional research training as well as of transferable skills offered

How will I get better because of Pieter and the rest of the collaborators? Mention exchange with Bosch.

### 2.3 Host expertise in training experienced researchers in the field and capacity to provide mentoring/tutoring (outgoing and return host)

Continuation of tutoring, assistance through ROS, teaching, meetings, conference workshop, reports, reading groups, ....

### 3 RESEARCHER

#### 3.1 Research experience

Mr. Pangercic, the applicant of the KSOM project, received his B.Sc. degree with distinction from the University of Ljubljana in 2003. In his thesis he investigated a Wishbone Bus and its applicability to let the parts of an integrated circuit communicate with each other. In 2004 he joined the Laboratory for Integrated Circuit Design, Faculty of Electrical Engineering Ljubljana for 3 months to work on the system for Power and Data Transmission over one Line. His system was later on commercialized and deployed to the gas warehouse in west Slovenia. In 2005 Mr. Pangercic enrolled to the Technische Universität München (TUM) in to pursue Master's degree. In 2007 he successfully defended his Master's Thesis titled Monocular 3D SLAM for Indoor Environments which he carried out under the supervision of Prof. Michael Beetz of Intelligent Autonomous Systems Group. The thesis generated his first publication at the major international conference, namely Emerging Technologies on Factory Automation. Upon this he was accepted to the graduate school of the Computer Science Department of TUM where he is projected to graduate in early 2012. In his PhD studies Mr. Pangercic deepened the understanding and investigated the connection of the perception algorithms for personal robotics with the knowledge representation for robots acting in household environments. Hence the title of his thesis: Knowledge-enabled Scene Perception for Personal Robotics. During this time he published at the prime robotics conferences and four times appeared as first author, and contributed to seven publications as secondary author. He also co-authored two journal publications listed below as well. All publications are listed on pages XX and YY.

Mr. Pangercic actively participates in community activities and attends all major conferences, for which he also serves as a reviewer (ROMAN, ETFA, ICRA, IJRR, IROS, ICAR, CVPR, IJCAI). In 2009 he attended the Robot Learning Summer School where he gained a substantial knowledge on machine learning applied to the practical robotics problems. On the other hand he co-organized Player Summer School on Cognitive Robotics <sup>7</sup> and served as a main organizer for the CoTeSys-ROS Fall School on Cognition-enabled Mobile Manipulation <sup>8</sup> which hosted 65 researchers from Europe and USA. He also, together with Jürgen Sturm, organized a highly attended RGB-D Workshop on 3D Perception in Robotics <sup>9</sup> that took place the European Robotics Forum 2011.

Since May 2011 he is with Bosch Research and Technology Center in Palo Alto where he is working on the 3D Environment Reconstruction and Semantic Mapping system that will serve as a basis for the herein proposed work. The part of his work at Bosch will be demonstrated at the IROS 2011 conference which will take place between September 25th and 29th.

#### 3.2 Research results including patents, publications, teaching etc., taking into account the level of experience

The main research results of his PhD work at TUM are presented in publications, in peer-reviewed journals and his presentations in conferences and seminars, showing his potential in highly applied research. In most of these publications, Mr. Pangercic is the first author and primary researcher. Among the results are included: i) Global Radius-based Surface Descriptor which is a 3D feature useful to perform categorization of objects, ii) K-CoPMAN (Knowledge-enabled Cognitive Perception for Manipulation) which is a realization of a system that combines robotic perception algorithms with symbolic representations of thereof and thus enable e.g. life long learning of semantic object maps and ii) realization of a system for static semantic mapping of environments which will be of a high relevance to KSOM project. The results exhibit the integration of diverse skills of the applicant such as the ability to tackle a problem and apply the most efficient computational scheme to resolve it.

Mr. Pangercic has been involved into teaching and supervision of undergraduate student projects ever since he joined the PhD program at TUM. The list of classes that he has led alone include a seminar on Intelligent Systems (WS2010), a practical course on Sensor-enabled Intelligent Environments and a

<sup>7</sup><http://www9.cs.tum.edu/events/psscr07/>

<sup>8</sup><http://www.ros.org/wiki/Events/CoTeSys-ROS-School>

<sup>9</sup><http://ias.cs.tum.edu/events/rgbd2011>

practical course on the Introduction to Computer Science (WS2008). Additionally he has supervised 9 Bachelor or Master's theses as outlined in his CV below. Most of his students have graduated with distinction and their works have resulted in the publications at the international conferences. Furthermore, Rok Tavcar, Andreas Leha and Martin Schuster all moved on to their PhD studies in the field and remain in close collaboration with Mr. Pangercic. Several other students managed to get high-profile jobs (e.g. BMW) in world renowned companies.

Mr. Pangercic also served as a main organizer for the previously mentioned CoTeSys-ROS Fall School on Cognition-enabled Mobile Robotics where he among other matters was in charge for the technical program and selection of speakers. He also actively participates in discussions and roadmaps pertinent to the development of ROS and gives tutorials and educates researchers about ROS e.g. <sup>10</sup>. He is also a part of the expanded list of developers for Point Cloud Library <sup>11</sup>.

### 3.3 Independent thinking and leadership qualities

During his stay at TUM and after gaining experience in the field, the applicant researcher showed his independent thinking when he steered the goals of his research to the modelling of a system that will integrate the knowledge for robots representation concepts with the low-level concepts such as the ones from the perception for robots. His was one of the main architects behind the two demonstrations where IAS researchers taught their robots to prepare two simple meals: baking of pancakes and cooking of sausages that generated gross interest among the research as well as the public society. Both events are briefly summed up in the following article:

<http://ias.cs.tum.edu/news/robotic-roommates-shopping-for-and-preparing-bavarian-breakfast>.

Lack of the postdoc fellows in the IAS group forced Mr. Pangercic to take on senior research jobs such above mentioned supervising of undergraduate students and writing of research grants. For TUM he thus lead the writing of the PR2 Beta Program proposal titled CRAM (Cognitive Robot Abstract Machine) with which TUM received a \$400.000,00 worth robotic platform PR2 and a participation in a highly prestigious PR2 Beta Program that involves institutes such as Stanford, Berkeley, MIT, UCS, etc. Mr. Pangercic also contributed to the writing of a Large-scale integrating project proposal titled Web-enabled and Experience-based Cognitive Robots that Learn Complex Everyday Manipulation Tasks and was successfully awarded to the Principal Investigator Michael Beetz. Proposal's budget is estimated to €7.000.000,00.

Applicant's wide involvement into a research society and his recognition led to multiple collaborations between IAS group and researchers from other institutions. In 2010 Mr. Pangercic thus worked on the development of an RGB-D feature called VOSCH together with Asako Kanezaki <sup>12</sup> from University of Tokyo. With Jürgen Sturm from University of Freiburg he organized an earlier mentioned RGBD workshop and at the time of writing of this proposal he is staying with Bosch RTC in Palo Alto working on the environment reconstruction and semantic annotation of scenes using Mechanical Turk interface and WUP similarities.

### 3.4 Match between the fellow's profile and project

The project that we propose matches perfectly Mr. Pangercic's scientific profile as it combines his past research work in both robotic perception and knowledge representation for robots during his graduate years at TUM. His work on detection and recognition of everyday objects as described in *Multimodal Perception System for Novel Object Modeling and Re-detection* and work on acquisition of static maps as proposed in *Autonomous Semantic Mapping for Robots Performing Everyday Manipulation Tasks in Kitchen Environments* will constitute a solid base for the implementation of work proposed in WP1 above. Furthermore, his K-CoPMan system as described in *Combining Perception and Knowledge Processing for Everyday Manipulation* paper will provide foundations for the second and third task in WP3. His collaborative work with Martin Schuster and Dominik Jain on WUP similarities will enable an advancement of research on perception of scenes as proposed in task two of WP2.

Last but not least, his excellent network of fellow researchers, knowledge of implemented algorithms in

---

<sup>10</sup><http://ias.cs.tum.edu/teaching/ros-school-2011>

<sup>11</sup>[pointclouds.org](http://pointclouds.org)

<sup>12</sup><http://www.ros.org/wiki/vosch>

ROS and commitment to assistance from other lead scientists from PR2 Beta Program make matching between Mr. Pangercic profile and the project very promising.

### **3.5 Potential for reaching a position of professional maturity**

Mr. Pangercic is a skillful researcher, who has already had the chance to integrate in a highly competitive scientific environment other than the university at which he is conducting his studies. His participation in a European program like KSOM will enable him to expand his scientific knowledge and skills. Expanding his work in such a high-standard research and educational institutions, will enable her to further increase her research experience and will develop her collaboration skills, integrate and provide his ways to establish himself in the wide European scientific community. The gain of research experience in a competitive environment will be the step forward that will soon help him attain a position of professional maturity in a university or a research institution. The measures foreseen to help the applicant reach professional maturity are connected to the possibility of pursuing a novel challenging research project in a well established research institution and a group which has successfully trained junior scientists in the past. Specifically, within two more years of postdoctoral experience, new knowledge will be acquired together with new scientific and communicative/collaborative skills, additional teaching experience will be gained, as well as tutoring of graduate and undergraduate students in the partner group.

### **3.6 Potential to acquire new knowledge**

The applicant researcher has multiply demonstrated in the past the ability to acquire new knowledge after switching to a different research field and a new group. Although his studies were in the electrical engineering field, he very easily integrated and undertook a project in a new group conducting research in personal robotics. The researchers ability to acquire new knowledge was evident in terms of the different scientific fields and application domains he could easily adopt himself (e.g. 3D perception, machine learning and first-order logic). Assets, which show the high potential of the applicant to acquire new knowledge and use this knowledge for conducting high level research, modeling novel systems and guiding/proposing novel applications. In this respect, at the outgoing host institution, he will also have the potential to acquire new knowledge and skills. Specifically, he will have the possibility to be exposed to new tools and methods, such as apprentice learning schema, perception of deformable objects, Potential collaborations and interaction with the many excelling robotic groups in the area around Berkeley will also expand the applicants scientific horizons.

# Pangercic Dejan

---

CONTACT INFORMATION	Karlstr. 45 Department of Computer Science Technische Universität München 80333 Munich Last Update:	Voice: (+49) 1637194179 / (+1) 6508610951 Fax: (+49) 8928917757 E-mail: dejan.pangercic@cs.tum.edu <a href="http://ias.cs.tum.edu/people/pangercic">http://ias.cs.tum.edu/people/pangercic</a> August 9, 2011
RESEARCH INTERESTS	Personal Robotics, Knowledge-enabled Scene Perception, Computer Vision, Machine Learning	
EDUCATION	<b>Bosch RTC</b> , Palo Alto, California, USA Visiting Researcher, May 2011 - present Project: Semantic Environment Mapping and Object of Daily Use Indexing	
	<b>Technische Universität München</b> , Munich, Germany Ph.D. Candidate, February 2008 - present <ul style="list-style-type: none"><li>• Dissertation Topic: Knowledge-enabled Scene Perception for Personal Robotics</li><li>• Advisor: Michael Beetz, PhD</li></ul>	
	<b>Technische Universität München</b> , Munich, Germany M.S., Electrical Engineering/Computer Science, December, 2007 Thesis: Monocular 3D SLAM for Indoor Environments	
	<b>University of Ljubljana</b> , Ljubljana, Slovenia B.A., Electrical Engineering, May, 2003 Thesis: Wishbone Interface Standard	
REFERENCES	<ul style="list-style-type: none"><li>• Michael Beetz, PhD, Professor at Technische Universität München</li><li>• Dr. Radu B. Rusu, Research Scientist at Willow Garage</li><li>• Kei Okada, Assistant Professor in Creative Informatics at The University of Tokyo</li><li>• Andrej Trost, Assistant Professor at the Faculty for Electrical Engineering, UL</li></ul>	
EVENTS	<b>RGB-D Workshop on 3D Perception in Robotics</b> , April 2011, Västerås, Sweden - Co-organizer <b>ROS-CoTeSys Fall School 2010</b> , November 2010, Munich, Germany - Co-organizer	
HONORS AND AWARDS	Dr. Otto Likar and Karla Likar Foundation Fellowship, 2005-2008	
ACADEMIC EXPERIENCE	<b>Technische Universität München</b> , Munich, Germany <i>Supervision of student projects</i> <b>Julius Adorf</b> : Segmentation of Cluttered Scenes in Household Environments, Bachelor Thesis, March 2011-present <b>Monica Simona Opris</b> : Modeling Centers for Robotic Perception, Bachelor Thesis, February 2011-May 2011 <b>Shulei Zhu</b> : Contracting Curve Density Algorithm for 2D/3D Tracking, Master Thesis, August 2010-April 2011 <b>Florian Zacherl</b> : Detection of Household Objects Using Projected Light Patterns, Bachelor Thesis, May 2010-October 2010	<b>Aug, 2008 - present</b>

**Vladimir Haltakov:** Fast and Robust Object Detection in Household Environments using Vocabulary Trees with SIFT Descriptors, Practical Course, June 2010-present

**Hozefa Indorewala:** Online Semantic Mapping, Internship, June 2010-August 2010

**Andreas Leha:** Knowledge Representation and Machine Learning Techniques for Cognitive Factory, Diplom Thesis, July 2009

**Rok Tavcar:** (Co-supervised) Connecting High-Level Planning, Reasoning and Model-driven Vision into a Robotic System that Enables Everyday Manipulation Tasks, Diplom Thesis, February 2009

*Teaching Assistant*

**Oct, 2008 - present**

- Embedded Intelligent Systems, Lecture, WS 2010
- Intelligent Systems, Seminar, WS 2010
- Sensor Enabled Intelligent Environments, Practical Course, SS 2010
- Einführung in die Informatik 1, Praktikum (Eng: Introduction to Computer Science, Practical Course), WS 2008

**PROFESSIONAL  
EXPERIENCE**

**Bosch RTC, Palo Alto, California**

*Summer Internship*

**May, 2011 - present**

3D Reconstruction and Semantic Classification of Indoor Environments

**Laboratory for Integrated Circuit Design, Faculty of Electrical Engineering Ljubljana, Ljubljana, Slovenia**

*Summer researcher*

**January, 2005 - July, 2005**

Implementation of the system for open/close contact detection of the switches on one Serial Line – Power and Data Transmission over one Line

**Slovenian Institute of Quality and Metrology, Ljubljana, Slovenia**

*Laboratory Assistant*

**January, 2000 - December, 2003**

Testing and certification of the Ex-devices according through standards SIST EN 50014 to SIST EN 50020

**Reviewing**

ROMAN 2008, ETFA 2009, IROS 2009, ICRA 2010, ICRA 2011, IJRR 2010, IROS 2011, ICAR 2011

**PUBLICATIONS**

- Dejan Pangercic, Nico Blodow, Lucian Cosmin Goron, Zoltan-Csaba Marton, Thomas Rühr, Moritz Tenorth, and Michael Beetz, **Autonomous Semantic Mapping for Robots Performing Everyday Manipulation Tasks in Kitchen Environments**, 2011, IEEE/RSJ International Conference on Intelligent RObots and Systems
- Zoltan-Csaba Marton, Dejan Pangercic, Nico Blodow, Michael Beetz, **Multimodal Perception System for Novel Object Modeling and Re-detection**, 2011, International Journal of Robotics Research
- Nico Blodow, Zoltan-Csaba Marton, Dejan Pangercic, Thomas Rühr, Moritz Tenorth, Michael Beetz, **Inferring Generalized Pick-and-Place Tasks from Pointing Gestures**, 2011, IEEE International Conference on Robotics and Automation (ICRA), Workshop on Semantic Perception, Mapping and Exploration
- Michael Beetz, Ulrich Klank, Alexis Maldonado, Dejan Pangercic, Thomas Rühr, **Robotic Roommates Making Pancakes - Look Into Perception-Manipulation Loop**, 2011, IEEE International Conference on Robotics and Automation (ICRA), Workshop on Mobile Manipulation: Integrating Perception and Manipulation
- Zoltan-Csaba Marton, Dejan Pangercic, Radu Bogdan Rusu, Andreas Holzbach, Michael Beetz, **Hier-**

**archical Object Geometric Categorization and Appearance Classification for Mobile Manipulation**, 2010, Proceedings of 2010 IEEE-RAS International Conference on Humanoid Robots

- Dejan Pangercic, Moritz Tenorth, Dominik Jain, Michael Beetz, **Combining Perception and Knowledge Processing for Everyday Manipulation**, 2010, IEEE/RSJ International Conference on Intelligent RObots and Systems
- Zoltan-Csaba Marton, Dejan Pangercic, Nico Blodow, Jonathan Kleinehellefort, Michael Beetz, **General 3D Modelling of Novel Objects from a Single View**, 2010, 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
- Nico Blodow, Zoltan-Csaba Marton, Dejan Pangercic, Michael Beetz, **Making Sense of 3D Data**, 2010, Workshop on Strategies and Evaluation for Mobile Manipulation in Household Environments (RSS)
- Michael Beetz, Nico Blodow, Ulrich Klank, Zoltan-Csaba Marton, Dejan Pangercic, Radu Bogdan Rusu, **CoP-Man – Perception for Mobile Pick-and-Place in Human Living Environments**, 2009, Proceedings of the 22nd IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Workshop on Semantic Perception for Mobile Manipulation
- Ulrich Klank, Dejan Pangercic, Radu Bogdan Rusu, Michael Beetz, **Real-time CAD Model Matching for Mobile Manipulation and Grasping**, 2009, 9th IEEE-RAS International Conference on Humanoid Robots
- Andreas Leha, Dejan Pangercic, Thomas Ruehr, Michael Beetz, **Optimization of Simulated Production Process Performance using Machine Learning**, 2009, Proceedings of Emerging Technologies and Factory Automation (ETFA)
- Dejan Pangercic, Rok Tavcar, Moritz Tenorth, Michael Beetz, **Scene Detection and Interpretation using Knowledge-driven Description Logic**, 2009, International Conference on Advanced Robotics (ICAR 2009), Munich, Germany
- Christoph Ertelt, Thomas Ruehr, Dejan Pangercic, Kristina Shea, Michael Beetz, Integration of Perception, **Global Planning and Local Planning in the Manufacturing Domain**, 2009, Proceedings of Emerging Technologies and Factory Automation (ETFA)
- Florian Friesdorf, Dejan Pangercic, Heiner Bubb, Michael Beetz, **Mutually Augmented Cognition**, 2009, Proceedings of the International Conference on Social Robotics (ICSR)
- Thomas Ruehr, Dejan Pangercic, Michael Beetz, **Structured Reactive Controllers and Transformational Planning for Manufacturing**, 2008, Proceedings of the 13th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Hamburg, Germany
- Dejan Pangercic , Radu Bogdan Rusu , Michael Beetz , **3D-Based Monocular SLAM for Mobile Agents Navigating in Indoor Environments**, 2008, Proceedings of the 13th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Hamburg, Germany
- M. F. Zh, Michael Beetz , K. Shea, G. Reinhart, O. Stursberg, M. Ostgathe, C. Lau, C. Ertelt, Dejan Pangercic , Thomas Ruehr , H. Ding, T. Paschedag, **An Integrated Approach to Realize the Cognitive Machine Shop**, 2008, Proceedings of the 1st International Workshop on Cognition for Technical Systems, München, Germany

## 4 IMPLEMENTATION

### 4.1 Quality of infrastructure / facilities and international collaboration of host

Pieter, Ken: Below like for Berkeley

#### TUM

In recent years, the Technische Universität München (TUM) has been consistently ranked the top academic institution in Germany in several independent rankings. It provides an excellent environment (likely the best in Germany) by substantial funding from the Bavarian state government, the German government and many private companies alike. One of the missions of the university is to boost interdisciplinary research across engineering, natural sciences, medicine, and humanities. Today the TUM comprises 13 faculties with more than 23,000 students (about 20 percent of whom come from abroad shows the degree of internationalization), 420 professors, and roughly 6,500 academic and non-academic staff. The TUM is thus well positioned to create new knowledge and know-how in Europe and the world. In 2005 the federal and state governments started the so-called Excellence Initiative in Germany. Between 2006 and 2011 they will fund the expansion of top university research with up to 1.9 billion Euros in three funding categories: graduate schools, clusters of excellence and institutional strategies for universities. The TUM was recognized as one of the first three universities that succeeded in all three categories. The corporate concept "TUM The Entrepreneurial University" supports and advances the existing TUM strategy and promotes top-level research on multiple levels. In this context Entrepreneurial Spirit means to activate the diversity of human talent in a concerted, interactive way. In terms of top-level research, it entails combining a maximum of individual freedom with a supportive administration.

The proposed returning host is the TUM Intelligent Autonomous Systems Group (IAS<sup>13</sup>) from Professor Michael Beetz in the Munich city center, and its extension of the Munich Garching Campus. IAS consists of two full-time professors, 1 Junior Research Group Leader, 15 PhD students, 2 secretaries. This very dynamic group challenges successfully seven main research topics: Perception for Robots; Knowledge Processing; Plan-based Control; Cognitive Manipulation; Perception of Human Activities; Facial Expression Recognition. Its PhD students co-supervise student projects and nurture a solid number of workshops and seminars in German and English, ranging from ROS technologies, Knowledge Representation for Autonomous Robots (IROS 2011 Workshop for example), to 3D Perception in Robotics, Cognition-enabled Mobile Manipulation and up to Probabilistic Methods for Perceiving, Learning and Reasoning about Everyday Activities<sup>14</sup>. The group has extended contacts with the industry, as evidenced by the number of student placements in companies in Munich and elsewhere in Bavaria, and the ongoing research collaborations with major players in the domain. In particular in the robotics field, with companies such as Kuka GmbH, Aldebaran Robotics, Willow Garage, Robert Bosch LLC and others. The solid and excellent student basis allows this group to further develop its scientific path and will give the applicant upon his return to TUM the possibility of enriching his skills (in research, technology transfer, teaching and training). In this context, the applicant will benefit from the existing infrastructure to deliver thriving seminar courses to a talented student body. The recently established international CoTeSys Graduate School (2010) will extend its activities with the development of specific additional measures to qualify the next generation of PhD students, which the applicant will contribute to supervise and train. The returning host IAS has extended research academic collaborations abroad (USA, Japan, South-Korea, Europe), which is a key asset in the context of this IOF proposal, as the applicant will be called upon to contribute to the reinforcement of scientific and student exchanges with the USA and Berkeley. Prof. Beetz, IAS Director, is additionally a member and the vice-coordinator of the CoTeSys Cluster of Excellence<sup>15</sup> coordinated by TUM. In CoTeSys there is a close collaboration between scientists from various disciplines connecting neuro-cognitive and neuro-biological foundations to engineering sciences at leading research institutions in Munich: besides Technische Universität München, scientists from Ludwig-Maximilians-Universität München - LMU, Universität der Bundeswehr, Max-Planck Institute of Neurobiology and German Aerospace Agency DLR are involved. CoTeSys investigates cognition for technical systems such as vehicles, robots, and factories. Cognitive technical systems are equipped with artificial sensors and actuators, integrated and embedded into physical systems, and act in the physical world. They differ from other technical systems since they

---

<sup>13</sup><http://ias.cs.tum.edu/>

<sup>14</sup><http://ias.cs.tum.edu/events>

<sup>15</sup><http://www.cotesys.org/>

perform cognitive control and have cognitive capabilities. By cognitive capabilities they mean information processing that take into accounts: perception, attention, memory, action, learning, and planning. The aim of CoTeSys is to produce future innovations. This will represent another opportunity for applicant to diversify his activities and find additional partners for R&D projects (whether academic and local, or industrial and spread in Europe).

## 4.2 Practical arrangements for the implementation and management of the research project

Pieter, Ken: Below like for Berkeley

### TUM

The applicant will be inserted in the robust management structure of the IAS group. Applicant's direct connection with Prof. Michael Beetz, enriched by his interaction and supervision of undergraduate and graduate students will ensure the success of the project.

#### *Grant implementation:*

Eric Bourguignon (CoTeSys Proposal Manager), who has been associated to the non-scientific writing of this proposal, will deliver a set of Marie-Curie FP7 tools to manage the project, based on his long-years experience of EU grants management in the FP6 and FP7 programs. The tools to be developed specifically for the projects will include the basic synopsis of all contractual issues related to the project, the yearly reporting ready-to-use templates (science and finance) and regular follow-up meetings with the applicant and the group team to ensure a quality delivery of the expected results. The management style will aim at letting as much autonomy in the budget controlling to the applicant as possible. Both administrative staff members, Mrs. Walter and Mrs. Wagner, will take their load to ensure the smooth running of the project.

#### *IAS investments on this project:*

It is estimated that Prof. Beetz will spend 1,5 person-month on this project on the 36 months of the project duration. The IAS will lead regular reviewing of the project. Applicant's re-integration in year 3 will be facilitated by his current knowledge of the University structure, its people, its finance, its other support services.

A strong involvement from the applicant's side will be expected in the design and monitoring organization of workshops at major international workshops. It is planned to organize 3 workshops in next 3 years, 1 at each of the following conferences: ICRA, IROS and RSS. The applicant will coordinate teams of students to ensure that these events are successful. From the beginning, the applicant will fully dispose his travel budget to attend international high-level conferences as well as to develop his own scientific network and scientific partnerships. The applicant will supervise graduate and undergraduate students.

All details have been agreed in common with Prof. Goldberg and Prof. Abbeel at Berkeley and Prof. Beetz at TUM, during the preparation of this project proposal.

## 4.3 Feasibility and credibility of the project, including work plan

Eric: Check if this below is enough text

In Sections 1.1 and 1.2 we detailed the three tasks to be carried out in this project, including the research methodology and the expected results. In Table 1 we have represented the work plan including the three work packages, plus two additional extra tasks for the integration of the developed systems and the fellowship's training, and we assigned an approximate time schedule for each of them. The actual implementation steps are given in Section 1.2.

## 4.4 Practical and administrative arrangements and support for the hosting of the fellow

Pieter, Ken: check if below text is OK

A concrete help for all practical and administrative issues with local authorities related to the hosting of foreign researchers is given by the administration core of Berkeley and the administrative assistant of the host group. The administration of the partner group will provide the applicant with an office space in the host group area, as already mentioned, and a Desktop Computer at his disposal with internet access also

Table 1: Timeplan

to the computer resources mentioned above. Running costs and consumables which will be needed, like postage, copier and printer paper, printer cartridges, pens, scientific software, etc. will also be provided by the partner group assistant. In addition to the scientific issues mentioned in the previous sections, the administrative and financial management of the research fellowship will be provided by the institution's administration core and the secretary of the host group. Administration will also help in other important and minor issues, such as social security and health insurance, traveling, social and cultural activities, etc. Further help in all these issues will be provided directly by the supervisor and other members of the partner group.

At TUM, as mentioned earlier, Eric Bourguignon will assist with all the Marie Curie FP7 related matters. Both administrative staff members, Mrs. Walter and Mrs. Wagner, will take their administrative load to ensure the smooth running of the project.

## 5 IMPACT

### 5.1 Potential for acquiring competencies during the fellowship to improve the prospects of reaching and/or reinforcing a position of professional maturity, diversity and independence, in particular through exposure to transferable skills training

Eric: convert into text

- People Management training skills = supervision of people / students / PhD students.
- Budget Administration = knowing how to use a budget. Learning by doing with the project budget at Berkeley and at TUM.
- Intellectual property Rights training = follow courses in US and in Germany.
- Business acumen. How to create a company = 2 courses with UnternehmerTUM for scientists back at TUM in Year 3.
- Possibly another People Management course to follow at TUM in Year 3 with the Life-long learning support unit scheme (people management course for University staff).
- How to write a grant = Course to follow back at TUM.

### 5.2 Contribution to career development or re-establishment, where relevant

The benefits of the proposed training period at Berkeley from a mid/long-term perspective will be essential. It will provide for continuity in the career of the applicant, and will allow for the consolidation and widening of his career prospects. All skills obtained during the KSOM project, ranging from computational, methodological, communicative, and collaborative together with teaching and mentoring, will allow him to achieve professional maturity and to become an independent researcher in his area of interest. Not to mention the extremely beneficial outcomes of the applicant's stay in a highly advanced and excelling educational and research environment. A stay is expected to boost his academic skills and enhance his career development.

### 5.3 Potential for creating long term collaborations and mutually beneficial cooperation between Europe and the Other Third Country

Pieter, Ken: Check if OK

As a member of TUM and Berkeley, the applicant will endeavor to contribute to create long-term fruitful collaborations between Berkeley and TUM through a triple strategy agreed between the applicant, the IAS leader Prof. Beetz and the TUM Department of Computer Sciences:

*Exchanges of students up to Master level*

Study Projects will be planned in common between TUM and Berkeley (starting from applicant's former research group) mentors and mentored students. Short training periods will be planned on both sides to execute these projects, completed by a common **summer school** in the second year including all hands-on workshops, poster competition, invited internationally well-known speakers of the field. TUM students will benefit from travel bursaries to be applied from usual German sources (DAAD, Alexander von Humboldt Foundation, Bosch Foundation, etc). A duty of the newly recruited staff is to maintain a small database of potential grant possibilities to support this very relation. The applicant will be supported by the TUM-Forte office to identify the best funding sources, and will oversee the quality of the applications written by the students. This shall ensure a pipeline of regular financially supported student projects at the core of long-term exchanges with Berkeley. Berkeley students will equally apply to their usual funding bodies to finance study travel abroad. Both TUM and Berkeley students will be requested on their return to report on their experiences abroad to their mentors and to the new students (both technologically and

concerning the daily life). The existing TUM Master Program **Robotics, Cognition, Intelligence**<sup>16</sup>, shall be the pipeline of exchanges, and it will be steady enough after the conclusion of applicant's fellowship. Thanks to applicant's professional and personal experience, applicant will be a successful example and a role model for young TUM/Berkeley students / PhD students showing scientific and mobility achievements. One of applicant's main contributions will be to raise the awareness for the program in the US and to help in recruiting more US graduate students.

*Expected results:*

3 students exchanged between Berkeley and TUM during the period of applicant's fellowship; up to 5 students in the following five years after the completion of the IOF contract. A summer school during the second year (either in Munich or in Berkeley) with up to 50 participants.

*Exchanges of scientists from PhD student level to post-doc and senior scientists*

The German Research Foundation (DFG) offers a multitude of programs to sustain research exchanges and collaboration with the US, and complements nicely other programs from foundations for the same purpose (DAAD, Alexander von Humboldt Foundation, Bosch Foundation, etc). During the period at TUM and Berkeley, Mr. Pangercic will help to connect principal investigators of both institutions in order to win additional resources to sustain the flow of activities with University of California in Berkeley. In order to do that, he will exploit the potentials of the well-organized offices network at TUM, which will help him at three levels: a) by the TUM-Forte office concerning the non-scientific parts of the projects to be submitted; b) by the IAS Chair administration and the TUM Computer Sciences Financial Department concerning the financial details of such proposals; c) by the TUM Legal Office (in particular regarding Intellectual Property Rights - IPR). Applicant's role will then remain in the frame of scientific exchanges developments, while the administrative burden will be taken by these three supporting services.

*Expected results:*

- Up to 3 PhD students, 1 post-docs and 1 senior scientists from each side performing research training or experiments in the corresponding Berkeley or TUM laboratories.
- 1 common research projects developed in the first year; 1 common research project (eventually with other partners) financed in the second year. Clarity of IPR.
- Another possibility - however highly competitive - for projects developed in common will be to insert research teams from Berkeley into FP7 European Research proposals, with a request to the NSF, or the NIH to finance the US part.
- Ideally, high-level scientific publications can be acknowledged as part of this collaboration in internationally significant peer-review journals. A common patent with its exploitation plan would also show the success of this privileged scientific pipeline between University of California in Berkeley and TUM.

*Cultural and corporate understanding*

The socialization part is a key point of this plan and cultural awareness will be part of each proposed project. This cultural awareness will be acquired on both sides by ensuring that students and scientist have time during the exchanges to visit important historical areas, eventually be accommodated by locals (especially for the students), be proposed German and English courses. This cultural knowledge will be linked to the visit of main industrial sites to approach nearer the different industrial cultures (for some of the students it will be an additional opportunity to stay longer by identifying potential interested companies to work for). The TUM will contribute with enlisting the Berkeley visitors onto its IKOM (job fair) program of visits to industries.

## 5.4 Contribution to European excellence and European competitiveness

Applicant's research experience will bring to Berkeley and back to TUM a clear scientific competence. There are few competences of that kind in Europe that, at the same time, help to preserve from the "Brain Drain". With applicant's additional contacts, the applicant will help to solidify other projects ideas of the

---

<sup>16</sup><http://www.in.tum.de/fuer-studieninteressierte/master-studiengaenge/robotics-cognition-intelligence.html>

IAS group and of the Computer Science Faculty to answer the interests of European companies (main partners in this field are: Kuka Roboter GmbH, The Source Works, Aldebaran Robotics, Robert Bosch LLC Germany) and American ones such as Willow Garage and Robert Bosch LLC USA.

## 5.5 Benefit of the mobility to the European Research Area

As can be evinced from applicant's curriculum vitae, the applicant has a significant mobility ability and wishes not only to maintain his scientific collaborations with a recognized world-leading research institution (TUM, Berkeley, University of Tokyo, Stanford, Robert Bosch LLC in Palo Alto etc.), but also to use the scientific and technological links that he has in e.g. whole ROS community. For the applicant, it will be a thriving scientific time to work in the USA and back in Munich as the Munich R&D capital encompasses: Applied Science (i.e., 4 Institutes Fraunhofer Gesellschaft); Fundamental research (i.e., 12 Max-Planck Institutes); Aerospace & IT (i.e., 8 DLR Institutes including a new major one in Robotics; Health and Environment (i.e., 23 Helmholtz Center Institutes); Corporate R&D (about 20500 employees); IT infrastructure (i.e., Leibniz Supercomputing Centre); IT services for more than 100 000 university customers; and Competence Center for Networks (e.g., GEANT2, X-WIN, etc.).

The excellent quality of life in Munich, its dense and very dynamic industrial base and the excellent airport and train connections will support my mobility plans (professional duties as well as personal journeys) so that I will be in a best position. The dominant scientific language of the IAS group is English, including with administrative staff that can make the link to the external administrations (and with the support of TUM Dual Career office). The TUM has also strong life-long learning training offers (technology transfer entrepreneurship, business acumen). The applicant is convinced that this extraordinary environment will be beneficial to strengthen his knowledge, to reinforce his collaboration network, and will allow him to experience the excitement of working in a new country.

## 5.6 Impact of the proposed outreach activities

Dejan

- Demonstration at Berkeley and in Munich.
- Workshops at ICRA, IROS, RSS
- **Pieter: Day of the open doors?**
- Applicant's availability for promotion of Marie Curie programs
- TUM frequent appearance in media, e.g.  
<http://uk.reuters.com/video/2011/06/30/robot-roomies-collaborate-to-cook-breakfast?videoId=216643886&videoChannel=4000>
- summer school
- Youtube videos

## 6 ETHICS ISSUES

Eric: Could you provide this for me?