NimbRo@Home Team Description for the RoboCup 2015 Competition

Max Schwarz and Sven Behnke

Rheinische Friedrich-Wilhelms-Universität Bonn
Computer Science Institute VI: Autonomous Intelligent Systems
Friedrich-Ebert-Allee 144, 53113 Bonn, Germany
nimbroathome @ ais.uni-bonn.de
http://www.NimbRo.net/@Home

Abstract. This document describes the RoboCup@Home league team NimbRo@Home of Rheinische Friedrich-Wilhelms-Universität Bonn, Germany, for the competition to be held in Hefei, China, in July 2015. Our team uses self-constructed humanoid robots for mobile manipulation and intuitive multimodal communication with humans. The paper describes the mechanical and electrical design of our robots Cosero and Dynamaid. It also covers our approaches to object and environment perception, manipulation and navigation control, and human-robot interaction.

1 Introduction

Our team NimbRo@Home competes with great success in the RoboCup@Home league since 2009, winning three international RoboCup@Home competitions in a row (2011 Istanbul [1], 2012 Mexico City [2], 2013 Eindhoven [3]). We also participate successfully in European competitions, winning RoboCup German Open for the last four years in a row (2011–2014).

Our robots Dynamaid and Cosero have been designed to balance requirements of indoor navigation, mobile manipulation, and intuitive human-robot interaction. These robots are equipped with omnidirectional drives for robust navigation in restricted environments, two anthropomorphic arms for human-like object manipulation, and a communication head for intuitive multimodal interaction with persons. In contrast to many other service robot systems, our robots are lightweight, relatively inexpensive, and easy to interface.

We developed advanced methods for real-time environment, person, and object perception using 3D sensors such as laser scanners and RGB-D cameras. Based on these percepts, we developed efficient planning methods for navigation, object manipulation, and the use of tools. Furthermore, the robots are equipped with a multimodal dialogue system.

2 Mechanical and Electrical Design

Our cognitive service robots Dynamaid and Cosero are shown in Fig. 1. They are equipped with omnidirectional drives for flexible locomotion in the restricted

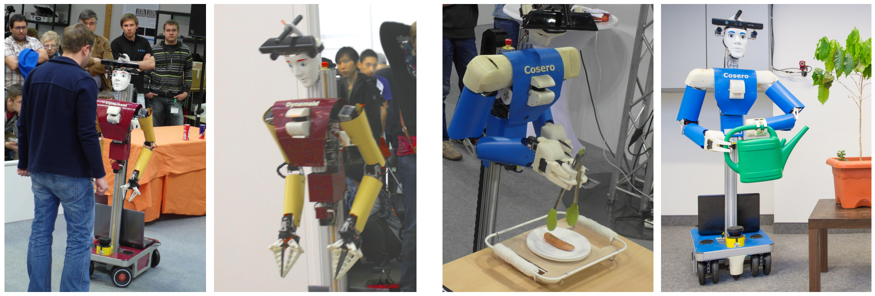

Fig. 1. Cognitive service robots Dynamaid (left) and Cosero (right), rightmost image by Volker Lannert.

spaces typical of domestic environments. Four pairs of directly driven, steerable wheels are mounted on the corners of the rectangular base.

The robots have an anthropomorphic upper body with two 7 DoF arms ending in two-finger grippers that comply to objects. The upper body can be moved up and down and twisted to extend the work space.

All joints are driven by Robotis Dynamixel actuators. Cosero has stronger motors and additional gears. It can lift objects up to 1.5 kg with a single hand.

The robots perceive their environment with multiple 3D sensors. A horizontally scanning SICK S300 laser range finder (LRF) is mainly used for 2D mapping and localization. For detection of small obstacles on the floor, a Hokuyo URG-04LX LRF is mounted between the front wheels. For 3D perception, a tilting Hokuyo UTM-30LX LRF is mounted in the chest. For measuring the height and distance of support surfaces, e.g. table tops, and detecting objects on these, a URG-04LX LRF is mounted on a roll joint on the hip. The head contains a Microsoft Kinect RGB-D camera and a directed microphone. A pan-tilt mechanism in the neck is used to directed the head sensors towards objects or persons of interest—and to indicate the driving direction. To detect the presence of objects, the grippers are equipped with IR distance sensors. Cosero has an additional camera in the belly for viewing manipulated objects from the side.

All computations are executed on a notebook with Intel Core-i7-Quad CPU. The robots are powered by Lithium-polymer rechargeable batteries.

3 Perception

3.1 Perception of Human Interaction Partners

For person detection and tracking, we combine complementary information from LRFs and the head camera. We also detect gestures like pointing towards and showing of objects in RGB-D images [4]. Our robots approach persons and enroll their faces for later recognition using a commercial library (VeriLook). For speech recognition and synthesis, we rely on Loquendo. The speech recognizer provides a parse tree according to a grammar, which we define for each test. Our

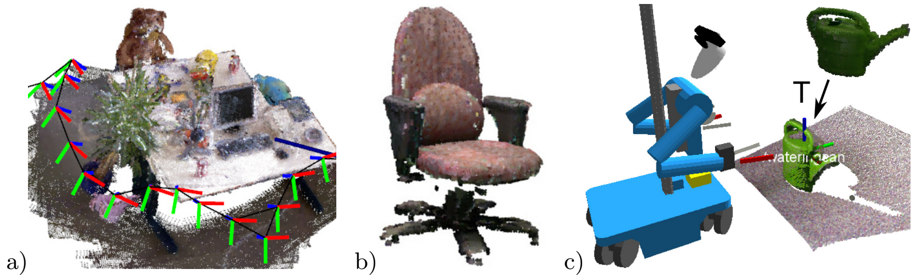

Fig. 2. Efficient RGB-D SLAM and model registration [5]. a) Model of a scene created by RGB-D-SLAM; b) Learned object model; c) Visual serving of an object for graping.

task execution module then interprets the resulting semantics and generates appropriate behavior.

3.2 RGB-D SLAM and Semantic Perception

In addition to state-of-the-art methods for simultaneous localization and mapping (SLAM) in 2D, we developed efficient RGB-D SLAM methods, based on Multi-resolution Surfel Maps (MRSMap) [5], which run in real time on a CPU. These can be used to model the environment and localize in these maps, or to obtain 3D object models from multiple views and track these in the camera images. Fig. 2 shows some examples. For difficult situations, we developed a particle filter-based method that detects and tracks MRSMap object models [6].

In addition to recognition of known object instances, we also developed methods for 3D semantic categorization, which are based on Hough forests for object detection [7], recognition of isolated objects based on features learned with convolutional neural networks [8], semantic mapping based on RGB-D SLAM and random forest object-class segmentation [9], object-class segmentation using learned conditional random field potentials [10], and object-class segmentation with convolutional neural networks [11]. Fig. 3 shows four examples.

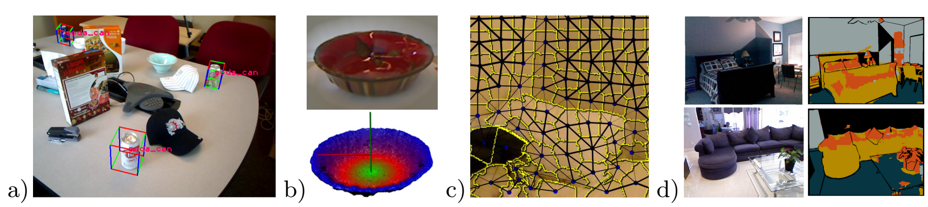

Fig. 3. Semantic RGB-D perception: a) Object class detection with Hough forests [7], b) Object recognition with object-centered coordinates [8]; c) Similarity between surface normals of superpixels [10]; b)Semantic segmentation of RGB-D scenes [11].

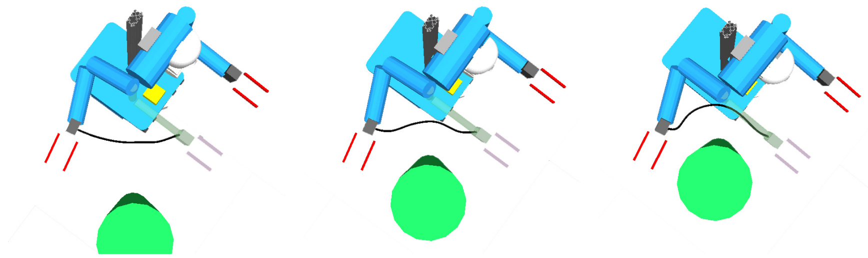

Fig. 4. Avoiding a dynamic obstacle by frequent trajectory optimization with multiresolution in time [12]

4 Behavior Control

The autonomous behavior of our robots is generated in a ROS-based modular control architecture. Hierarchical finite state machines are employed for high-level control. These configure individual mid-level modules, such as the perception of objects and persons, navigation planning, or the grasping and placement of objects. The motions are generated on the lowest layer of the hierarchy.

4.1 Robust Indoor Navigation

We developed an omnidirectional driving controller that coordinates the steering and rotation of the eight wheels to realize arbitrary combinations of 2D linear and rotational velocities. For navigation, we implemented path planning in occupancy grid maps and 3D obstacle avoidance using measurements from the LRFs and the depth camera.

4.2 Manipulation with One and Two Arms

We control the 7 DoF arms using differential inverse kinematics with redundancy resolution. The arms also support compliant control in task-space. For the avoidance of dynamic obstacles, we developed a method for fast trajectory optimization using multiresolution in time [12], which plans the initial parts of the trajectory with more detail than the later parts. This efficient approach allows for frequent replanning, as shown in Fig. 4.

Our robots can grasp objects on horizontal surfaces like tables and shelves efficiently using fast grasp planning [13]. We derive grasps from the top and the side directly from the raw object point clouds. The grasps are then executed using parametrized motion primitives, if the direct reach towards the object is not obstructed. In complex scenarios, such as in bin-picking, the robots plan collision-free grasps and reaching motions [14].

In addition, we developed solutions to pour-out containers, to place objects on horizontal surfaces, to dispose objects in containers, and to grasp objects

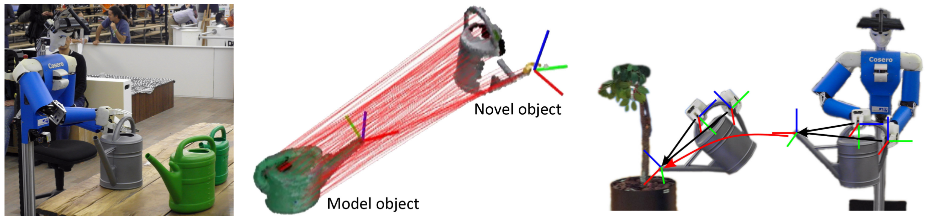

Fig. 5. Manipulation skill transfer: watering a plant with an unknown can based on non-rigid registration of a can model to the perceived novel can [16].

from the floor. Based on compliant control, we implemented mobile manipulation controllers to open and close doors, e.g. of fridges and closets [15]. We designed several task-specific motions, e.g. for switching on a cooking plate or for unscrewing the cap of a bottle.

Since our robots have two arms, they can grasp and manipulate larger objects. This has been used e.g. to push chairs to a specific location, to grasp a watering can and water a plant, or to carry a tray. In order to transfer manipulation skills from known object instances to novel ones, which differ in geometry and appearance, we developed efficient deformable registration methods [16]. As shown in Fig. 5, they determine a dense correspondence between the model and the novel object, which we use for transfer of grasp poses and end-effector frames to the novel object, which leads to adapted motion primitives. This has been demonstrated at RoboCup 2013 by the use of an unknown watering can.

We developed methods for using tools [17], some of which which have been equipped with a special handle that fits firmly into the gripper of our robot Cosero. Fig. 6 shows the opening of a bottle with a bottle opener, the use of a brush and a dust pan, and using a muddler to make Caipirinha.

Fig. 6. Tool use demonstrations at RoboCup 2014: Opening a bottle, using a brush and a dust pan, making Calpirinha [17].

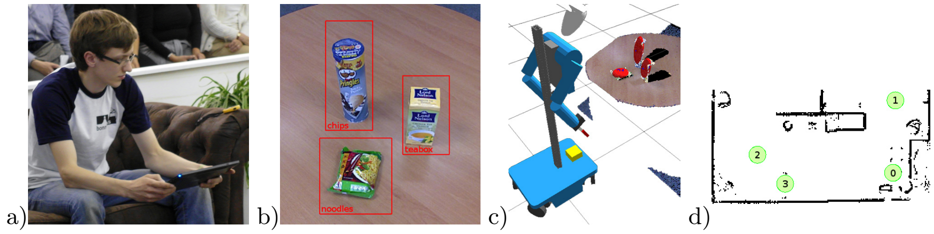

Fig. 7. Handheld teleoperation interface [20]. a) User with slate PC; b) Object selection for grasping; c); 3rd person view; d) Estimated locations of tagged objects [21].

4.3 Intuitive Human-Robot Interaction

For natural interaction with its users, we developed a multimodal dialogue system, which is based on the experiences made with our communication robot Robotinho [18]. This includes recognition and generation of speech (using Loquendo) and gestures.

The robots also perform physical human-robot interaction, e.g. by handing over objects from robot to human or vice versa and by cooperative carrying large objects, like a table [19].

For immobile users, it is important to have a possibility to remotely interact with the robot. To this end, we implemented a handheld teleoperation interface, based on a slate PC [20], see Fig. 7. Camera images, maps, localization, and object detections are displayed to the user who can control the robot on three levels of autonomy. These range from task level, where the user only specifies the goal of a complex task, over skill-level, where individual skills like grasping an object are triggered by the user, to body level, where the user directly controls base and arm motion. To avoid the search for frequently displaced objects, such as keys, we developed an object localization system, based on Bluetooth LE tags attached to objects and receivers distributed in the environment [21]. This allowed for keeping track of object positions in the teleoperation interface, such that the user can send the robot to fetch the objects without search (Fig. 7d).

5 Conclusion

Our cognitive robot system covers the basic functionalities required for the RoboCup@Home League standard tests well. Our system is very modular and contains some advanced functions, such as tool use, skill transfer, semantic perception, and teleoperation interfaces, which can be combined to design novel demonstrations for the open tests with limited effort. The system performed very well, winning seven competitions in a row (German Open 2011-2014 and international RoboCup 2011-2013).

We will continue to improve our system for RoboCup 2015 and to integrate new capabilities. The most recent information about our team (including videos) can be found on our web pages www.NimbRo.net/@Home.

Team Members

Currently, the NimbRo@Home team has the following members:

– Team leader: Max Schwarz, Sven Behnke
– Staff: David Droeschel, Matthias Nieuwenhuisen, Dirk Holz, and Michael Schreiber
– Students: Nikita Araslanov, David Schwarz, and Angeliki Topalidou-Kyniazopoulou

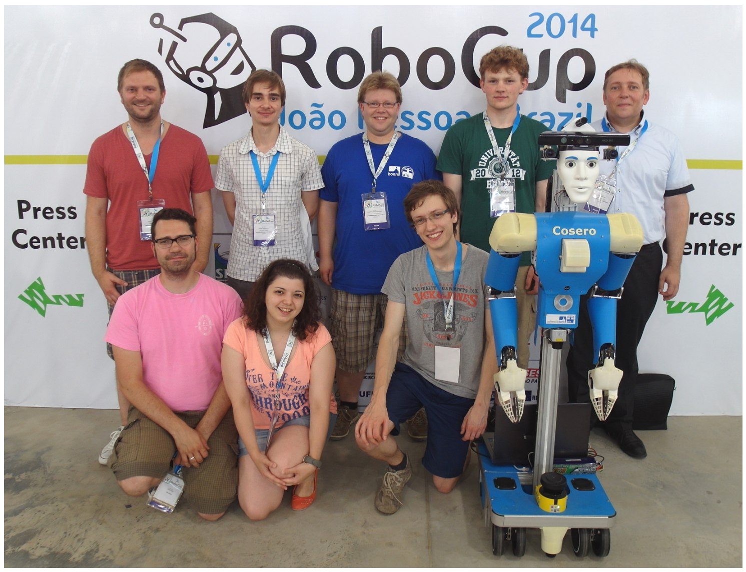

Fig. 8. Team NimbRo@Home at RoboCup 2014 in João Pessoa.

References

1. Stückler, J., Droeschel, D., Gräve, K., Holz, D., Kläß, J., Schreiber, M., Steffens, R., Behnke, S.: Towards robust mobility, flexible object manipulation, and intuitive multimodal interaction for domestic service robots. In: RoboCup 2011: Robot Soccer World Cup XV. Volume 7416 of LNCS., Springer (2012) 51–62
2. Stückler, J., Badami, I., Droeschel, D., Gräve, K., Holz, D., McElhone, M., Nieuwenhuisen, M., Schreiber, M., M.Schwarz, Behnke, S.: Nimbro@home: Winning team of the robocup@home competition 2012. In: RoboCup. Volume 7500 of LNCS., Springer (2013) 94–105
3. Stückler, J., Droeschel, D., Gräve, K., Holz, D., Schreiber, M., Topalidou-Kyniazopoulou, A., Schwarz, M., Behnke, S.: Increasing flexibility of mobile manipulation and intuitive human-robot interaction in robocup@home. In: RoboCup. Volume 8371 of LNCS., Springer (2014) 135–146
4. Droeschel, D., Stückler, J., Holz, D., Behnke, S.: Towards joint attention for a domestic service robot - person awareness and gesture recognition using time-of-flight cameras. In: Robotics and Automation (ICRA), IEEE International Conference on. (2011) 1205–1210

5. Stückler, J., Behnke, S.: Multi-resolution surfel maps for efficient dense 3D modeling and tracking. Visual Communication and Image Representation **25** (2014) 137–147
6. McElhone, M., Stückler, J., Behnke, S.: Joint detection and pose tracking of multi-resolution surfel models in rgb-d. In: Mobile Robots (ECMR), European Conference on. (2013) 131–137
7. Badami, I., Stückler, J., Behnke, S.: Depth-enhanced Hough forests for object-class detection and continuous pose estimation. In: 3rd Workshop on Semantic Perception, Mapping and Exploration (SPME) at ICRA. (2013)
8. Schwarz, M., Schulz, H., Behnke, S.: RGB-D object recognition and pose estimation based on pre-trained convolutional neural network features. In: Proc. of IEEE International Conference on Robotics and Automation (ICRA). (2015)
9. Stückler, J., Waldvogel, B., Schulz, H., Behnke, S.: Dense real-time mapping of object-class semantics from rgb-d video. Journal of Real-Time Image Processing (2014) available online.
10. Müller, A.C., Behnke, S.: Learning depth-sensitive conditional random fields for semantic segmentation of RGB-D images. In: Proc. of IEEE International Conference on Robotics and Automation (ICRA). (2014)
11. Schulz, H., Höft, N., Behnke, S.: Depth and height aware semantic RGB-D perception with convolutional neural networks. In: Proc. of European Symposium on Artificial Neural Networks (ESANN). (2015)
12. Steffens, R., Nieuwenhuisen, M., Behnke, S.: Continuous motion planning for service robots with multiresolution in time. In: 13th International Conference on Intelligent Autonomous Systems (IAS). (2014)
13. Stückler, J., Steffens, R., Holz, D., Behnke, S.: Efficient 3D object perception and grasp planning for mobile manipulation in domestic environments. Robotics and Autonomous Systems **61** (2013) 1106–1115
14. Nieuwenhuisen, M., Droeschel, D., Holz, D., Stückler, J., Berner, A., Li, J., Klein, R., Behnke, S.: Mobile bin picking with an anthropomorphic service robot. In: Robotics and Automation (ICRA), IEEE International Conference on. (2013)
15. Stückler, J., Behnke, S.: Perception of deformable objects and compliant manipulation for service robots. In: Soft Robotics: From Theory to Applications. Springer Vieweg (2015)
16. Stückler, J., Behnke, S.: Efficient deformable registration of multi-resolution surfel maps for object manipulation skill transfer. In: IEEE International Conference on Robotics and Automation (ICRA). (2014)
17. Stückler, J., Behnke, S.: Adaptive tool-use strategies for anthropomorphic service robots. In: 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids). (2014)
18. Nieuwenhuisen, M., Behnke, S.: Human-like interaction skills for the mobile communication robot Robotinho. International Journal of Social Robotics **5** (2013) 549–561
19. Stückler, J., Behnke, S.: Following human guidance to cooperatively carry a large object. In: IEEE-RAS International Conference on Humanoid Robots (Humanoids). (2011)
20. Schwarz, M., Stückler, J., Behnke, S.: Mobile teleoperation interfaces with adjustable autonomy for personal service robots. In: Human-Robot Interaction (HRI), ACM/IEEE International Conference on. (2014) 288–289
21. Schwarz, D., Schwarz, M., Stückler, J., Behnke, S.: Cosero, find my keys! Object localization and retrieval using Bluetooth Low Energy tags. In: Proc. of 18th RoboCup International Symposium, Joao Pessoa, Brasil. (2014)