RoboCup 2018 - homer@UniKoblenz (Germany)

Raphael Memmesheimer, Niklas Yann Wettengel, Lukas Debald, Anatoli Eckert, Thies Möhlenhof, Tobias Evers, Gregor Heuer, Nick Theisen, Lukas Buchhold, Jannis Eisenmenger, Simon Häring, Dietrich Paulus

Active Vision Group
University of Koblenz-Landau
Universitätsstr. 1
56070 Koblenz, Germany
`raphael@uni-koblenz.de`
`http://homer.uni-koblenz.de`

Abstract. This paper describes the robots *TIAGo* and *Lisa* used by team homer@UniKoblenz of the University of Koblenz-Landau, Germany, for the participation at the RoboCup@Home 2018 in Montreal, Canada. A special focus is put on novel system components and the open source contributions of our team. We have released packages for object recognition, a robot face including speech synthesis, mapping and navigation, speech recognition interface via android and a GUI. The packages are available (and new packages will be released) on `http://wiki.ros.org/agas-ros-pkg`.

1 Introduction

In 2017 our team homer@UniKoblenz was the winner of the RoboCup@Home Open Platform League in Nagoya, Japan and ended up first at the RoboCup German Open in Magdeburg, Germany. Furthermore we were awarded with four our of five prizes of the Europen Robotics League. In 2015 *Lisa* and her team won the 1st place at RoboCup World Championship in the RoboCup@Home league in Hefei, China.

Beside this success our team homer@UniKoblenz has already participated successfully as finalist in Suzhou, China (2008), Graz, Austria (2009) in Singapore (2010), where it was honored with the RoboCup@Home Innovation Award, in Mexico-City, Mexico (2012), where it was awarded the RoboCup@Home Technical Challenge Award and in Eindhoven, Netherlands (2013). Further, we participated in stage 2 at the RoboCup@Home World Championship in Instanbul, Turkey (2011). Our team achieved several times the 3rd place in the RoboCup GermanOpen (2008, 2009, 2010 and 2013) and participated in the GermanOpen finals (2011, 2012 and 2014).

Apart from RoboCup, team homer@UniKoblenz won the best demonstration award at RoCKIn Camp 2014 (Rome), 2015 (Peccioli), the 1st place in the overall rating, as well as the 2nd place in the Object Perception Challenge in the RoCKIn Competition (Toulouse, 2014). In the RoCKIn 2015 competition (Lisbon) team

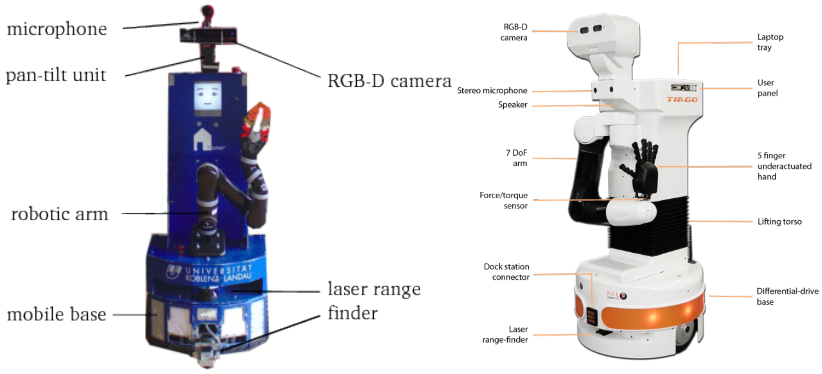

Fig. 1. Lisa (left) and TIAGo (right)

homer@UniKoblenz won the 1st overall rating together with SocRob, the Best Team Award, 1st place in the Navigation Challenge, 1st place in the Getting to Know my home task benchmark. Recently our Team won four of five possible prices of the European Robotics League.

In 2018 we plan to attend the RoboCup@Home in Montreal, Canada, with two robots: a sponsored PAL Robotics TIAGo[1] robot and a custom built Lisa (Fig. 1). Our team will be presented in the next Section. Section 3 describes the hardware used for Lisa. In Section 4 we present the software components that we contribute to the community. The following Section 5 presents our recently developed and improved software components. Finally, Section 6 will conclude this paper.

2 Team homer@UniKoblenz

The Active Vision Group (AGAS) offers practical courses for students where the abilities of Lisa are extended. In the scope of these courses the students design, develop and test new software components and try out new hardware setups. The practical courses are supervised by a research associate, who integrates his PhD research into the project. The current team is lead and supervised by Raphael Memmesheimer.

Each year new students participate in the practical courses and are engaged in the development of Lisa. These students form the team *homer@UniKoblenz* to participate in the RoboCup@Home. *Homer* is short for “home robots” and is one of the participating teams that entirely consist of students.

2.1 Focus of Research

The current focus of research is imitation learning by observation.

[1] `http://tiago.pal-robotics.com/`

Additionally, with large member fluctuations in the team, as is natural for a student project, comes a necessity for an architecture that is easy to learn, teach and use.

3 Hardware

In this year's competition we will use two robots (Fig. 1). The blue Lisa is built upon a CU-2WD-Center robotics platform[2]. Furthermore we will use a PAL Robotics TIAGo robot that is able to higher and lower it's torso has an wider working range. Currently, we are using a Workstation Notebook equipped with an Intel Core i7-6700HQ CPU @ 2.60GHz × 8, 16GB RAM with Ubuntu Linux 16.04 and ROS Kinetic.

Each robot is equipped with a laser range finder (LRF) for navigation and mapping.

The most important sensors of the blue Lisa are set up on top of a pan-tilt unit. Thus, they can be rotated to search the environment or take a better view of a specific position of interest. Apart from a RGB-D camera (Microsoft Kinect2) a directional microphone (Rode VideoMic Pro) is mounted on the pan-tilt unit.

A 6 DOF robotic arm (Kinova Mico) is used for mobile manipulation. The end effector is a custom setup and consists of 4 Festo Finray-fingers.

Finally, an Odroid C2 inside the casing of the blue Lisa handles the robot face and speech synthesis.

4 Software Contribution

We followed a recent call for chapters for a new book on ROS[3]. We want to share stable components of our software with the RoboCup and the ROS community to help advancing the research in robotics. All software components will are released on the Active Vision Group's ROS wiki page: `http://wiki.ros.org/agas-ros-pkg`. The contributions are described in the following paragraphs.

Mapping and Navigation

Simultaneous Localization and Mapping To know its environment, the robot has to be able to create a map. For this purpose, our robot continuously generates and updates a 2D map of its environment based on odomentry and laser scans. Figure 2 shows an example of such a map.

[2] Manufacturer of our robotic platform: `http://www.ulrichc.de`
[3] Call for chapters for a ROS book: `http://events.coins-lab.org/springer/springer_ros_book.html`

Navigation in Dynamic Environments An occupancy map that only changes slowly in time does not provide sufficient information for dynamic obstacles. Our navigation system, which is based on Zelinsky's path transform [8, 9], always merges the current laser range scans into the occupancy map. A calculated path is checked against obstacles in small intervals during navigation. If an object blocks the path for a given interval, the path is re-calculated.

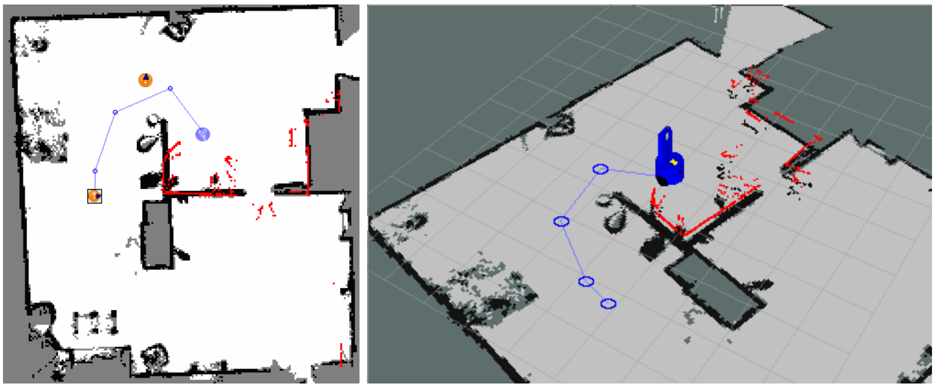

Fig. 2. 2D and 3D view of a map and a planned path (blue line). Red dots indicate the current laser scan, while orange points in the 2D map stand for navigation points.

Object Recognition

Object Recognition The object recognition algorithm we use is based on Speeded Up Robust Features (SURF) [1]. First, features are matched between the trained image and the current camera image based on their euclidean distance. A threshold on the ratio of the two nearest neighbors is used to filter unlikely matches. Then, matches are clustered in Hough-space using a four dimensional histogram using their position, scale and rotation. This way, sets of consistent matches are obtained. The result is further optimized by calculating a homography between the matched images and discarding outliers. Our system was evaluated in [3] and shown as suitable for fast training and robust object recognition. A detailed description of this approach is given in [5]. With this object recognition approach we won the Technical Challenge 2012 (Figure 3).

Human Robot Interaction

Robot Face We have designed a concept of a talking robot face that is synchronized to speech via mouth movements. The face is modeled with blender and Ogre3D is used for visualization. The robot face is able to show seven different

Fig. 3. Object recognition results during the Technical Challenge 2012.

face expressions (Figure 4). The colors, type and voice (female or male) can be changed without recompiling the application.

We conducted a broad user study to test how people perceive the shown emotions. The results as well as further details regarding the concept and implementation of our robot face are presented in [4]. The robot face is already available online on our ROS package website.

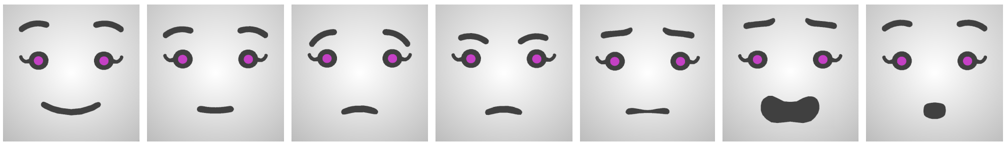

Fig. 4. Animated face of our service robot Lisa. The depicted face expressions are (from left to right): happy, neutral, sad, angry, disgusted, frightened, and surprised.

5 Technology and Scientific Contribution

5.1 General Purpose System Architecture

In the past years we have migrated step by step from our self developed architecture to ROS. Since 2014, our complete software is ROS compatible. To facilitate programming new behaviors, we created a architecture aiming at general purpose task executing. By encapsulating arbitrary functionalities (e.g. grasping, navigating) in self-contained state machines, we are able to start complex behaviors by calling a ROS action. The ROS action library allows for live monitoring of the behavior and reaction to different possible error cases. Additionally, a semantic knowledge base supports managing objects, locations, people, names and relations between these entities. With this design, new combined behaviors (as needed e.g. for the RoboCup@Home tests) are created easily and even students who are new to robotics can start developing after a short introduction.

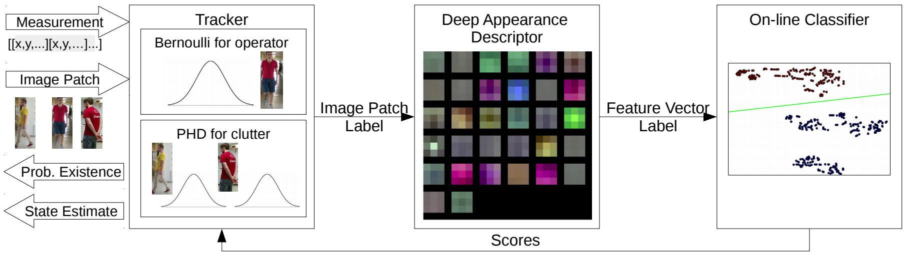

Fig. 5. Tracking Overview
A RFS Bernoulli single target tracker in cooperates with a deep appearance descriptor to re-identify and online classify the appearance of the tracked identity. Measurements, consisting of positional information and an additional image patch serve as input. The Bernoulli tracker estimates the existence probability and the likelihood of the measurement being the operator. Positive against negative appearances are contentiously trained. The online classifier returns scores of the patch being the operator.

5.2 People Detection and Tracking

We developed an integrated system to detect and track a single operator that can switch *off* and *on* when it leaves and (re-)enters the scene. Our method is based on a set-valued Bayes-optimal state estimator that integrates RGB-D detections and image-based classification to improve tracking results in severe clutter and under long-term occlusion. The classifier is trained in two stages: First, we train a deep convolutional neural network to obtain a feature representation for person re-identification. Then, we bootstrap a classifier that discriminates the operator online from remaining people on the output of the state-estimator. See Figure 5 for an visual overview. The approach is applicable for following and guiding tasks.

5.3 3D Object Recognition

For 3D object recognition we use a continuous Hough-space voting scheme related to Implicit Shape Models (ISM). In our approach [6], SHOT features [7] from segmented objects are learned. Contrary to the ISM formulation, we do not cluster the features. Instead, to generalize from learned shape descriptors, we match each detected feature with the k nearest learned features in the detection step. Each matched feature casts a vote into a continuous Hough-space. Maxima for object hypotheses are detected with the Mean Shift Mode Estimation algorithm [2].

5.4 Speech Recognition

For speech recognition we use a grammar based solution supported by a academic license for the VoCon speech recognition software by Nuance[4]. We combine continuous listening with a begin and end-of-speech detection to get good results even for complex commands. Recognition results below a certain threshold are rejected. The grammar generation is supported by the content o a semantic knowledge base that is also used for our general purpose architecture.

6 Conclusion

In this paper, we have given an overview of the approaches used by team homer@UniKoblenz for the RoboCup@Home competition. We presented a combination of out-of-the box hardware and sensors and a custom-built robot framework. Furthermore, we explained our system architecture, as well as approaches for 2D and 3D object recognition, human robot interaction and object manipulation with a 6 DOF robotic arm. This year we plan to use the TIAGo robot and blue *Lisa* for the main competition. Based on the existing system from last year's competition, effort was put into improving existing algorithms of our system (speech recognition, manipulation, people tracking) and adding new features (encapsulated tasks for general purpose task execution, 3D object recognition, affordance detection) to our robot's software framework. Finally, we explained which components of our software are currently being prepared for publication to support the RoboCup and ROS community.

References

1. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. SURF: Speeded up robust features. *ECCV*, pages 404–417, 2006.
2. Yizong Cheng. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):790–799, 1995.
3. Peter Decker, Susanne Thierfelder, Dietrich Paulus, and Marcin Grzegorzek. Dense Statistic Versus Sparse Feature-Based Approach for 3D Object Recognition. In *10th International Conference on Pattern Recognition and Image Analysis: New Information Technologies*, volume 1, pages 181–184, Moscow, 12 2010. Springer MAIK Nauka/Interperiodica.
4. Viktor Seib, Julian Giesen, Dominik Grüntjens, and Dietrich Paulus. Enhancing human-robot interaction by a robot face with facial expressions and synchronized lip movements. In Vaclav Skala, editor, *21st International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, 2013.
5. Viktor Seib, Michael Kusenbach, Susanne Thierfelder, and Dietrich Paulus. Object recognition using hough-transform clustering of surf features. In *Workshops on Electronical and Computer Engineering Subfields*, pages 169 – 176. Scientific Cooperations Publications, 2014.

[4] `http://www.nuance.com/for-business/speech-recognition-solutions/vocon-hybrid/index.htm`

6. Viktor Seib, Norman Link, and Dietrich Paulus. Implicit shape models for 3d shape classification with a continuous voting space. In *Proceedings of International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2015. to appear.
7. Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In *Proc. of the European conference on computer vision (ECCV)*, ECCV'10, pages 356–369, Berlin, Heidelberg, 2010. Springer-Verlag.
8. Alexander Zelinsky. Robot navigation with learning. *Australian Computer Journal*, 20(2):85–93, 5 1988.
9. Alexander Zelinsky. *Environment Exploration and Path Planning Algorithms for a Mobile Robot using Sonar*. PhD thesis, Wollongong University, Australia, 1991.

Name of team homer@UniKoblenz
Member Raphael Memmesheimer, Niklas Yann Wettengel, Lukas Debald, Anatoli Eckert, Thies Möhlenhof, Tobias Evers, Gregor Heuer, Nick Theisen, Lukas Buchhold, Jannis Eisenmenger, Simon Häring
Contact information raphael@uni-koblenz.de
Website `http://homer.uni-koblenz.de`
Hardware :

- PAL Robotics TIAGo Steel Edition
- Kinova Mico
- Lenovo P50 Notebook
- Directed Perception D46-17.5 PTU
- Rode VideoMic
- Microsoft Kinect 2
- CU2WD Robot platform
- Hokuyo Laser Scanner
- SICK LMS 100
- Odroid C2

Software :

- ROS
- OpenCV
- PCL
- Mary TTS
- Ogre3D
- Caffe
- Tensorflow
- Matlab
- Nuance VoCon
- strands
- Custom software for:
  - User Interface (homer_gui `https://gitlab.uni-koblenz.de/robbie/homer_gui`)
  - Object recognition (homer_object_recognition `https://gitlab.uni-koblenz.de/robbie/homer_object_recognition`)
  - Mapping / Navigation (homer_mapnav `https://gitlab.uni-koblenz.de/robbie/homer_mapnav`)
  - Robot face (homer_robot_face `https://gitlab.uni-koblenz.de/robbie/homer_robot_face`)
  - Speech Recognition / Speech synthesis (android_speech_pkg `https://gitlab.uni-koblenz.de/robbie/homer_android_speech`)