The Meta-Mechanics 2009 Team Description Paper

Sven Olufs, Peter Einramhof, Georg Merzdovnik, Michael Weissbacher,
Gerald Wodni, Tom Gamper, and Markus Vincze

Automation and Control Insitute, Vienna University of Technology
olufs@acin.tuwien.ac.at
http://www.meta-mechanics.org

Abstract. This paper presents the Meta-Mechanics RoboCup@Home team, including two mobile robot platforms for the RoboCup@Home 2009 competitions. Some components that are of particular interest will be described in detail.

1 Introduction

The Meta-Mechanics RoboCup@Home team was established in late 2008 at the Vienna University of Technology. It is a mixed team of the Faculty of Electrical Engineering and Information Technology and the Department of Computer Science. In 2009 the Meta-Mechanics plan to participate in Graz[1] (World Cup) and German Open 2009 for the first time.
The team consists of a mixture of Bachelor, Master and PhD students, which are advised by professors from the University. We have two reasons to participate in RoboCup: First, to educate students in the field of mobile autonomous robotics. We believe that RoboCup provides the "best practise" for students in many fields such as how to build a robot or how to navigate in a cluttered environment. The second reason is to use RoboCup@Home as a field test for the outcome of EU-funded research projects that are carried out by our research group[2] (See [1–4] for some of our past research projects):

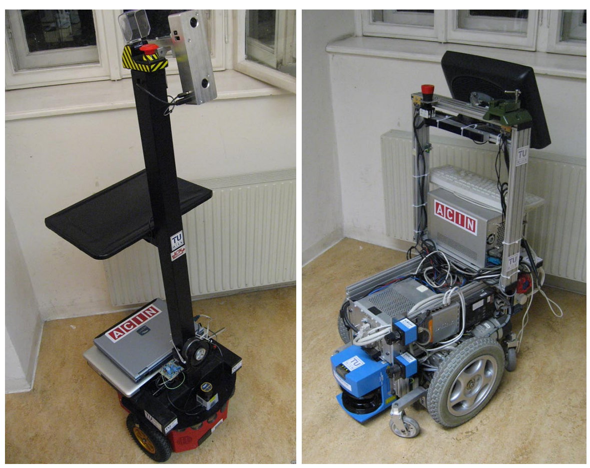
- Robots@Home[3]: The objective of robots@home is to provide an open mobile platform for the massive introduction of robots into the homes of everyone. The results of the robots@home project lead to an affordable, scaleable and versatile platform that is targeted for marketing in two to three years after project end open to any application that requires the safe navigation capability. We use an scenario-driven approach that is inspired by recent work in cognitive science, neuroscience and animal navigation: a hierarchical cognitive map incorporates topological, metric and semantic information.
- CogX[4]: The high level aim of this project is to develop a unified theory of self-understanding and self-extension with a convincing instantiation and implementation of this theory in a robot. By self-understanding we mean that the robot has representations of gaps in its knowledge or uncertainty in its beliefs. By self-extension

[1] Our Group is organizing the @Home League in Graz
[2] Vision for Robotics (V4R), Automation and Control Institute (ACIN)
[3] IST-FP6-2005-IST-6 http://www.robots-at-home.org/
[4] IST-FP7-ICT-2007-1 http://www.acin.tuwien.ac.at/?id=127

we mean the ability of the robot to extend its own abilities or knowledge by planning learning activities and carrying them out.
- Grasp[5]: The aim of GRASP is the design of a cognitive system capable of performing tasks in open-ended environments, dealing with uncertainty and novel situations. The design of such a system must take into account three important facts: i) it has to be based on solid theoretical basis, and ii) it has to be extensively evaluated on suitable and measurable basis, thus iii) allowing for self-understanding and self-extension.

Our main research interest is vision for robotics, i.e. visual navigation, vision-based grasping of unknown objects, object recognition and human-machine interaction. Some details of the work that is used and tested by our team will be shown in Chapter 5.
The remainder of this paper is organised as follows: Our robot platforms are described in Chapter 2 and the control software architecture is sketched in Chapter 3. In Chapters 4 and 5 we will provide some details about the components responsible for human-robot interaction and vision.

2 Hardware



(a) Mechs (b) James

Fig. 1. Our robots used for RoboCup

We use two robots for RoboCup that are built on different bases due to size and weight constraints of the three research projects. The first robot is "James", shown in Figure 1(b). It is a non-holonomic mobile robot manufactured by the Swiss company Bluebotics. It has a SICK LMS 200 laser range finder mounted to its front. The robot

[5] IST-FP7-IP-215821 `http://www.acin.tuwien.ac.at/?id=128`

features an onboard PC that controls the hardware (motors, sensors) and that runs the navigation software ANT (Autonomous Navigation Technology), which provides obstacle avoidance and self-localisation using the laser range finder's readings. An external PC can communicate with the onboard PC via ethernet or wireless and using a proprietary communication protocol. Using ANT and a (line-based 2D) map, the robot can move autonomously through the environment. It also provides a speed-control mode and access to its odometry as well as sensor data. The robot's maximum payload is around 100 kilos, the power supply consists of two 12 volts lead batteries with a capacity of 38 amperehours each that are connected in series (24 volts). We use an extra Quad-Core Barebone PC for our behaviour and vision modules mounted on top of the robot. The touch-screen as shown on figure 1(b) is only used for debug and is not your for RoboCup completions.
The secound robot is "Mechs", shown in Figure 1(a). The robot is based on the well known Pioneer 3-DX. First it was used for lab lectures on mobile robotics held by ACIN. Due to its reliability and ease of use we decided to use it in RoboCup. In its basic setup the Pioneer 3-DX only provides readings from eight sonar sensors, which is not appropriate for "serious" applications. So, we have augmented the robot with a Hokuyo URG-04LX laser range finder (mounted in front of the robot) as well as a stereo camera (mounted on top of a pole). In order to power the laser range finder (5 volts) from the robot's batteries, an additional DC-DC converter board had to be built. We use one Apple Pro Notebook for behaviours and three Dell Notebooks for Vision, Localisation and Speech recognition/synthesis.
Both robots are equipped with a custom build vertical stereo camera system from the MOVEMENT Project [1] with wide angle lens. We use a vertical stereo system due to space limitations on both robots. This setup allows us to detect humans within a range of 6m with an aperture angle of 120 degree.

3 Software

Our software-architecture is based on the component paradigm using Zero-C[6] Internet Communications Engine *ICE* that is also used by programs like Skype. It is a multi-platform object-oriented middleware with support for different programming languages such as C++ or Java. The functionality of each component, e.g. behavior or low-level driver is provided as a service that can be used by any other software module. That allows us to use the identical behavior software modules on both robots. This is possible due to both robots are using similar sensors and actuators. For instance the kinematics are modeled in the low-level differential driver and is hidden within the behavior.

3.1 Dual Dynamics

Our software (SW) methodology to robot behavior programming is based on Dual Dynamics (DD) [5–7], a mathematical model for robot behaviors developed by Fraunhofer

[6] http://www.zeroc.com

AiS. It integrates central aspects of a behavior-based approach with a dynamical systems representation of actions and goals. Robot behaviors are specified through differential equations, forming a global dynamical system made of behavioral subsystems that interact through specific coupling and bifurcation/induction mechanisms. Behaviors are organised in levels where higher levels have a larger time scale than lower levels. Since the activation of behaviors (activation dynamics) is separated from their actuator control laws (target dynamics), the approach is called Dual Dynamics. An important feature of DD is that it allows for robust and smooth changes between different behavior modes, which results in very reactive, fast and natural motions of the robots. Through the distribution of the overall control task on several simple controllers, we obtain a very robust system behavior.

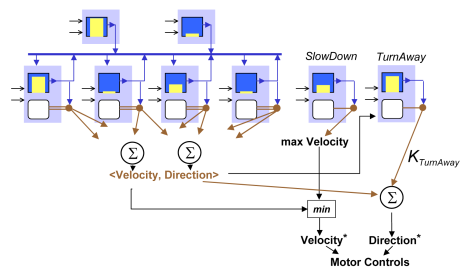


Fig. 2. Dual Dynamics, taken from [6]

3.2 Behavior

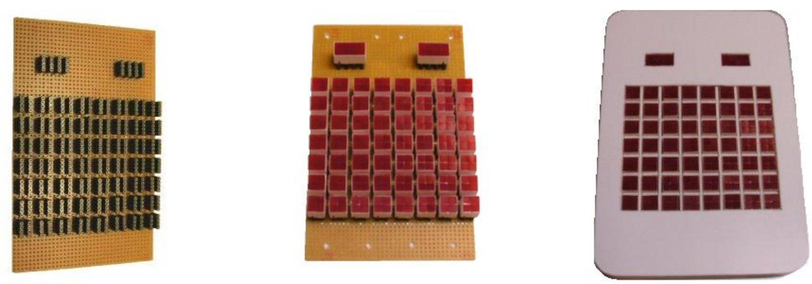
Behaviors are organised in three layers, where lower levels are always subsumed by the upper layers as compatible with the behavior based paradigm. The system has the properties of the well known Subsumption Architecture designed by Rodney A. Brooks. As many robotic researchers agree today, a favorable approach to solve real-life robotic problems, roboticists should follow a bottom-up approach rather than the usual top-down approaches which usually end up making the problems more complex than they actually are. In this sense, we achieved this goal by developing a bottom-up behavior system working with a minimum of sensor information. The main idea is to have a substitute behavior at the lowest useful level for urgent cases that usually occur in a typical RoboCup secnario.

4 Human Robot Interaction

Inspired by Crispin Jones' USB gadget "Tengu"[7], we decided to take a different approach to implementing a head for our robot. Instead of using mechanics for the head's

[7] http://www.tengutengutengu.com

face or displaying a rendered face on a monitor, we use an 8x8 LED matrix. As a result, the head has some kind of ”80s retro charm” and doesn’t look as disturbing as some of its antropomorphic counterparts.

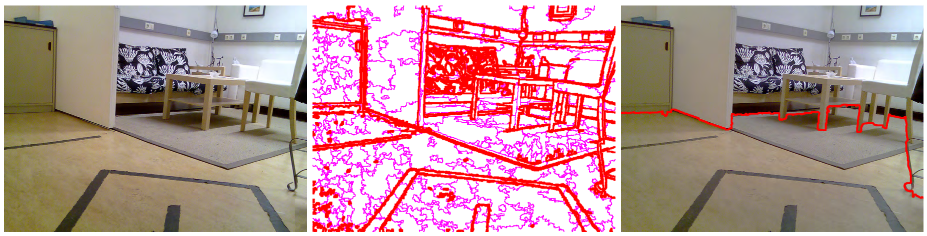


(a) Stage 1 (b) Stage 2 (c) Stage 3

Fig. 3. Various stages in constructing the head: (1) front of the board with the sockets for the LEDs, (2) all LEDs inserted into the sockets, (3) plastic body of the head without the white front cover over the LEDs.

Despite the low resolution, the LED matrix allows for a plethora of facial expressions, blinking of the eyes and motion of the ”lips” synchronous to speech output. Furthermore, the LED matrix can be used for displaying other contents than just the face such as scroll texts, a binary clock or a Pong game.
Figure 3 shows the first three stages of the construction of the head. The fourth stage - the implementation of the control electronics - is currently underway. The board with the electronics will be attached to the back of the head. Finally, a white translucent plastic cover over the front of the head will only let those LEDs be visible that are turned on.
The heart - or better the brain - of the head is an Atmel ATmega8515 microcontroller running at 8MHz. It receives the frames that are to be displayed via RS232 (or USB-to-RS232 converter) from a laptop. Using 38.4kbaud and 8 intensity levels per pixel (3 bits), up to 100 frames per second could be transmitted to the head.
The microcontroller generates frames for the LED matrix at a frequency of 700Hz. During such a frame patterns will be applied first to the even LED lines, then to the odd lines, one line at a time. For seven consecutive frames the same pattern will be applied to the LED matrix, but depending on each pixel’s 3bit intensity value (0..7) the respective LED will be either powered or not during such a ”subframe”. The head is powered by the robot’s 12 volt batteries and draws up to 250 milliamperes (if all LEDs were powered at maximum intensity).
We use the Microsoft Speech SDK for voice recognition and synthesis. In contrast to other speech synthesis solutions it generates facial expression for the LED face. This enables a more ”natural” interaction with the face. The necessary hardware for speech input and output - microphone and loudspeaker plus amplifier - are also mounted onto the superstructure that we have built on top of our mobile platform.

5 Vision



(a) Input Image (b) oversegmentation (c) Graph Cut result

Fig. 4. Ground Plane Segmentation with Superpixels, MRF and Graph-Cut

We use well-known standard techniques in computer vision like the Viola-Jones Face detector or Lowe SIFT detector as well as state-of-the art in vision techniques like boosted object detector [8] or Bag-of-Features [9, 10] approches. Figure 4 shows an example for the vision-based ground plane estimation using graph cuts. First we apply an oversegmentation using a superpixel [11] variant that has been proposed by Wildenauer et al. [12]. The main idea of the superpixel technique is to group similar pixels to "meaningful" clusters using colour, edges or texture information. Figure 4(b) shows the result of our oversegmentation at two different scales: The cyan lines corresponds to the low-scale, red corresponds to the high scale. One can see that the hi-scale oversegmentation groups the objects into almost meaningful regions, while some regions like the door and ground have been grouped together. This is due to the high gradient similarity of both regions. That is why we use an oversegmentation at a lower scaler to correct the region.

The oversegmentation of both scales is transformed into a graph-based representation

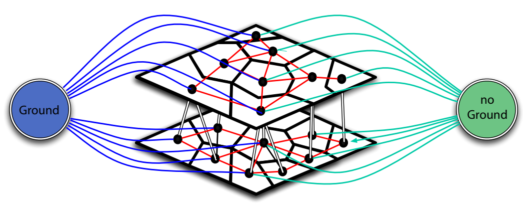


Fig. 5. Two-layer MRF with superpixels detected at two different scales

as shown in Figure 5. The main idea of graph-cut based image segmentation is to consider the image as a graph and to cut it into two independent pieces. This is also referred to as the Markov Random Field (MRF) segmentation method and is state of the art in computervision. A superpixel is used as node and is connected to its neighbours superpixels using links. A weight is assigned to each link that typically corresponds to the likelihood or similarity of the two nodes. This is usually referred to as the *smoothness term*. Each node is also connected to two ”super nodes”. The weight of these nodes represents the probability of the node of being a ground or no ground node. We use a learned model based on texture to calculate the weight of the two super-nodes. A learned model can be obtained from a few images and is invariant to light conditions or colour. Finally, the ground plane is obtained using a min-cut/max-flow graph cut. See [12] for more details.

Acknowledgements

We gratefully mention the continued support by the team of the Vienna University of Technology and the technical support of the European Community's Sixth Framework Programme (FP6/2003-2006) under grant agreement no FP6-2006-IST-6-045350 (robots@home)

References

1. P. Mayer, G. Edelmayer, G.J. Gelderblom, M. Vincze, P. Einramhof, M. Nuttin, T. Fuxreiter, and G. Kronreif. Movement -modular versatile mobility enhancement system. In *Procedings of IEEE International Conference on Robotics and Automation (ICRA), Rome*, 2007.
2. P. Einramhof H. Wildenauer M. Vincze, S. Olufs. Roboternavigation in büros und wohnungen. *Elektrotechnik und Informationstechnik (e&i)*, 1(2):25–32, Dec 2008.
3. M. Vincze P. Gemeiner, P. Einramhof. Simultaneous motion and structure estimation by fusion of inertial and vision data. *International Journal of Robotics Research*, 6(26):591 – 605, 2007.
4. W. Ponweiser M. Zillich M. Vincze, M. Ayromlou. Edge projected integration of image and model cues for robust model-based object tracking. *International Journal of Robotics Research*, 7(20):533 – 552, 2001.
5. Herbert Jäger and Thomas Christaller. Dual dynamics: Designing a behavior system for autonomous robots. *Artificial Life and Robotics*, 2:108–112, 1998.
6. Ansgar Bredenfeld and Hans–Ulrich Kobialka. Team cooperation using dual dynamics. 2001.
7. Hans-Ulrich Kobialka and Herbert Jaeger. Experiences using the dynamical system paradigm for programming robocup robots. In *AMiRE 2003 (2nd International Symposium on Autonomous Minirobots for Research and Edutainment)*, pages 193–202. Springer, 2003.
8. I. Laptev. Improvements of object detection using boosted histograms. In *BMVC'06*, pages 949–958, Edinburgh, UK.
9. J. Sivic, B. Russell, A. Efros, and booktitle =ICCV publisher =Beijing year =2005 A.] Zisserman and W. Freeman, title =Discovering object categories in image collections.
10. L. Fei-Fei and P Perona. A bayesian heirarcical model for learning natural scene categories. In *CVPR*, 2005.
11. X. Ren and J. Malik. Learning a classification model for segmentation. In *Tenth Intl. Conference On Computer Vision (ICCV-03)*, pages 10–16, Vancouver, British Columbia, Canada.
12. H. Wildenauer, B. Micusik, and M. Vincze. Efficient texture representation using multi-scale regions. In *ACCV07*, pages I: 65–74, 2007.