

A Humanoid Robot Platform Suitable for Studying Embodied Interaction

Matthias Hackel and Stefan Schwöpe
MABOTIC – Robotics & Automation
Robotics Research Department
Hürth, Germany
{hackel, schwöpe}@mabotic.de

Jannik Fritsch, Britta Wrede and Gerhard Sagerer
Applied Computer Science
Faculty of Technology
Bielefeld University, Germany
{jannik, bwrede, sagerer}@TechFak.Uni-Bielefeld.de

Abstract—This paper presents the humanoid robot BARTHOC who has been developed to study human-robot interaction (HRI). The main focus of BARTHOC’s design was to realize the expression and behavior of the robot to be as human-like as possible. This allows to apply the platform to manifold research and demonstration areas. With his human-like look and mimic possibilities he differs from other platforms like ASIMO or QRIO, and enables experiments even close to Mori’s ‘uncanny valley’. The paper describes details of the mechanical and electrical design of BARTHOC together with its PC control interface. Through its humanoid appearance, it can imitate human behavior with its soft- and hardware. Currently, several components for HRI on a mobile robot platform are being ported to BARTHOC. Starting with these components, the robot’s human-like appearance will enable us to study embodied interaction and to explore theories of human intelligence.

Index Terms—Humanoid Robot, Human-Robot Interaction, Embodied Communication.

I. INTRODUCTION

Intelligent interaction between humans and machines is one of the main challenges in the future. The interaction between humans is supposed to be advanced; in that consequence a machine interface that reacts like a human would be the most universal interface. It will be even more effective if it represents a friendly and sympathetic human. This means that the core of this interface should also detect the actual mood of the user [1].

To learn more about human communication and human-robot interaction (HRI), many efforts in robotics research have already been achieved. But there is still much more work to do in future, especially in the field of humanoid robotics. A detailed knowledge of the way humans communicate via speech, gestures, and mimic [2]–[6] is required¹. The first task is to understand the meaning of human speech, mimic, and gestures by using sensors and algorithms. Second, a natural human-like reaction has to be generated based on this information. The current results of humanoid robotic research are far away from humans; the typical answer

¹Other human communication paths like the magic of direct eye contact or pheromones are interesting as well, but not discussed in here.

“Is this humanoid robot able to vacuum-clean my room or to wash my dishes?” has to be neglected while machines, designed for limited applications, are doing the job (e.g. autonomous vacuum cleaners).

The attraction of humanoid robots is their complexity which reflects our own (unconscious) complexity. Even if the interaction modules developed are “kind of poor” in correlation to human capabilities, their integration on a humanoid robot can lead to a surprisingly realistic human-like robot behavior. Moreover, only a robot with a humanoid appearance, having approximately the same actuators as a human body, will allow us to study embodied interaction. In human-human interaction, a variety of cues is applied by both communication partners during the interaction. Building a humanoid robot, therefore, is a prerequisite to study human-like embodied interaction and to develop algorithms suitable for implementing human-like capabilities on a robot. We present here the humanoid robot BARTHOC (Bielefeld Antropomorphic RoboT for Human-Oriented Communication) which has been developed in a tight cooperation between the manufacturer and the researchers who are going to use BARTHOC for studying embodied interaction.

The remainder of this paper is organized as follows: In the next section we will describe BARTHOC’s anatomy and its overall appearance. Section III describes details of the electronics and the low-level control software to operate the different degrees of freedom. A first sketch of the interaction architecture that will be used to control BARTHOC’s overall behavior is outlined in Section IV before the paper ends with a conclusion.

II. BARTHOC’S ANATOMY

In comparison to our former prototype H10 [7], BARTHOC’s hardware has benefited from a lot of improvements. Its mechanics consists of a mechatronic head and two arms including hands. These components are mounted on a steel-frame backbone. Though BARTHOC is a non-mobile humanoid robot, it could achieve mobility, e.g., by using a wheelchair [8] as basis. However, for the suggested fields of application the static design is sufficient. To give an

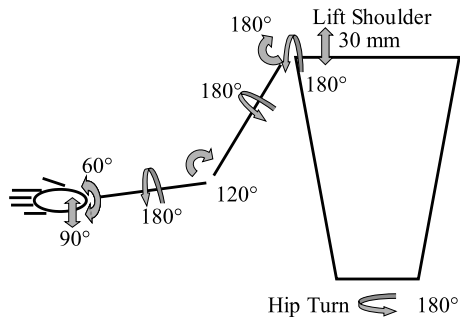


Fig. 1. Draft: Joints and angular paths of torso and right arm. Not drafted, each finger of the hand can be bend independently by servos.

overview about BARTHOC’s performance and possibilities, some mechanical details of its anatomy are explained in the following.

A. Torso, Arms, and Hands

The realized joints and their corresponding angular paths are drafted in Fig. 1. These degrees of freedom (DOF’s) supply BARTHOC with mechanical possibilities that comply with human ones. However, BARTHOC’s speed of motion and payload is dramatically reduced in that comparison. Additionally, it is not able to bend its back, which would be important for tracking a moving object with its whole body [9]. Nevertheless, the given DOF’s allow BARTHOC to perform human gestures and grasping actions in a satisfying way.

The joints of hip, shoulders, upper and lower arms are driven by planetary geared DC-motors with position feedback via precision potentiometers. Since no joint rotation is greater than 180 degrees, this method is reliable. To ensure a light construction of the arms, the DC-motors are used in a self-supporting way. The main axis of motor and gear always runs parallel to the corresponding “bone”. To redirect the turning moment, spiral bevel gears are used if needed. Some of these details are shown in Fig. 2. The shoulder lifting is simulated by only lifting the right and left part of the neck instead of the complete arms. As the mechanics of torso, arms, and hands will be covered by a usual shirt and gloves, this form of neck-lifting is sufficient for expressing the shoulder lifting.

The wrist has two joints to lift and turn the palm, driven by two servos via push-pull rods. The hand is constructed as an external actuator type [10]. Each finger is built from Polyoxymethylen (POM) with three spring pre-stressed joints driven by a flexible, high strain resistant nylon cable. If the cable is pulled, the corresponding finger will bend. To minimize friction and wear, the cable is running either above or inside a Teflon® or POM coating on its whole way. Five servos are mounted in the lower arm to pull the cables of each finger independently. To ensure that the fingers do not change their position if the pan is tilted, the nylon cables are

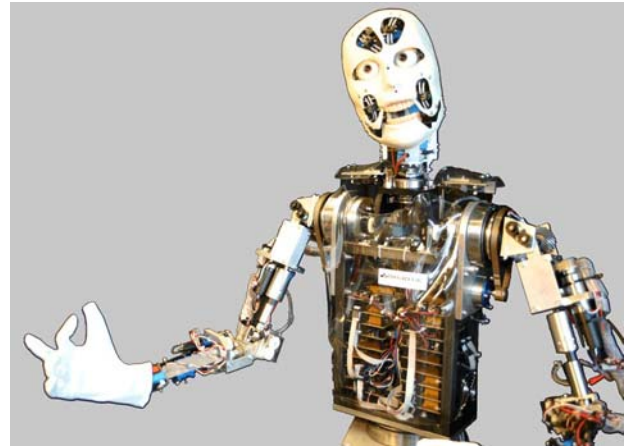


Fig. 2. Photo: BARTHOC. The self-supporting DC-motors in the upper arms and a toothed belt to drive the shoulder turning joint are observable. All the power and control electronics to drive the motors and servos fit inside the torso, see chapter III-A for more details.

laid exactly through the two rotation centers of the palm.

B. Head

While we used a TFT display for face representation in the past [7], now a complete mechatronic head has been developed to give BARTHOC a more human-like appearance and human-like features. To enable the implementation of cognitive processes, the hardware is equipped with sensors suitable for HRI. Therefore, a camera² is built in each eyeball, ear microphones are to follow. The eyes pan and tilt, opening and closing eyelids are provided, too. To permit changes in BARTHOC’s sex and character by affordable costs and effort, the head is constructed with the ability to exchange the removable mask. The most suitable material for a human-looking android mask seems to be foam latex or silicone³. Currently, we are experimenting with both materials to achieve good results. A picture of a first functioning prototype mask is shown in Fig. 3. The mask is driven by bowden cables and is designed to express basic human face expressions with movements of eyebrows and forehead, jowls and lower jaw. The lower jaw is rotated around the naturally given rotation point by a fast high-torque servo. This enables the simulation of lip movements by fast jaw action. A detailed photo of the naked head animatronics is shown in Fig. 4. The head is pivoted on a mechatronic neck. The neck allows to turn, nod and tilt the head, and a linear front and back movement to express interest or aversion.

Altogether, BARTHOC is equipped with 41 actuators that can be controlled by one USB interface (see next chapter). The highest voltage used is 12 V, so even non-contactless

²Synchronizable fire-wire cameras with 30 fps suitable for stereo-vision.

³An extensive overview is given at the the Android Head Projects Site at <http://www.androidworld.com>



Fig. 3. Foam latex mask over BARTHOC's mechatronic head. Due to its well defined folding lines (named wrinkles on humans) this mask shows good performance in dynamic face expressions. Only friendly expressions are hard to achieve with this character – maybe he has undergone too much trouble in his live.

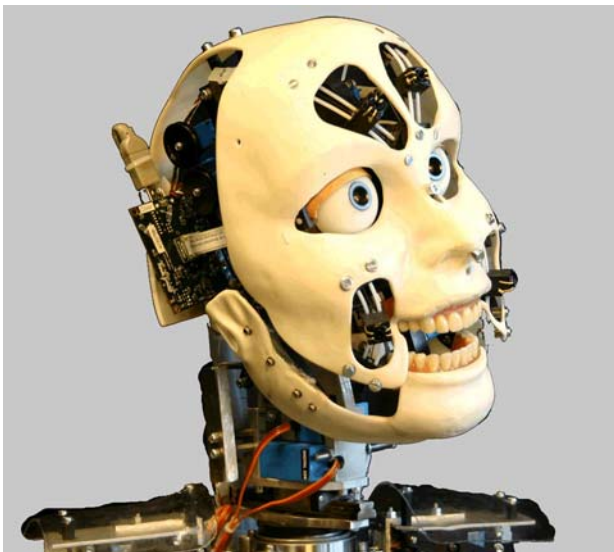


Fig. 4. The naked head animatronics with mask moving actuators (next to the upper lip and above the eyes) and eyeball cameras. The friendly expression shows its will to communicate.

experiments with BARTHOC are not dangerous from this point of view.

III. CONTROL ARCHITECTURE

The control architecture of BARTHOC is separated in three parts: mechanics, electronics and control software. As the mechanics have been described above, this chapter deals with some details of the developed electronics and low-level control software.

A. Control Hardware

The development of prototypes, especially in robotics, needs a large amount of time and money. To tackle this prob-

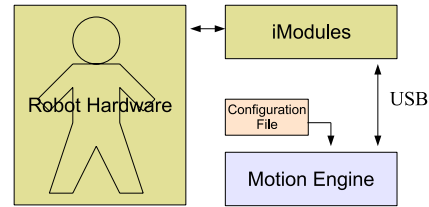


Fig. 5. The software/electronic/mechanic interface structure of BARTHOC.

lem we developed so-called iModules [7]. These components are reusable hardware modules with specialized functionality on each module (like object-oriented programming), so the user can concentrate on the core of development by using the iModules. To build a new machine, only the desired functionality has to be chosen, e.g., power control, motor control, or servo control. Each iModule has embedded control of the connected actuators or sensors, so no code for low-level control like interpreting sensor data and driving actuators has to be written in any control program on the PC. Instructions for the iModule system are send by an ASCII protocol via an USB interface. In Fig. 5 the interface between software, electronics and mechanics is drafted.

BARTHOC is using three different kinds of iModules: four iServo modules (44 servos can be connected); three iMotorControl modules with six 1×5 Amps power stage modules and three 2×2 Amps power stage modules connected (12 motors could be connected); and two independent iConnect modules for PC communication. One iConnect module is used for controlling the head and the other iConnect module controls the torso with the arms and hands. In this way, interaction components for controlling the facial expression and the gaze direction are separated from the componts controlling gestural expressions with the arms. This separation is especially important, as the arms have strong motors and are potentially dangerous for the surrounding humans.

The iServo modules generate a puls width modulation (PWM) signal with peaks between 0.8 ms to 2.2 ms at 50 Hz. The duration of one peak is converted to a bevel square/a position by the connected servo actuator.

One iMotorControl module generates PWM signals for four motors. Each signal is influenced by the current of the connected motor and the position of the controlled joint. To imitate human movements and to reduce the mechanical stress by high accelerations, the iMotorControl Module generates different ramp signals at different joint positions. Seven parameters can be set for each motor control, for example aging of the robot's mechanic can be balanced and also effects like jitter or slow movements can be imitated. The generated signals are send to power stages to provide enough power for the connected motors.

As the iConnect module is the interface to the PC, all commands from the iActuator system (see next chapter) are

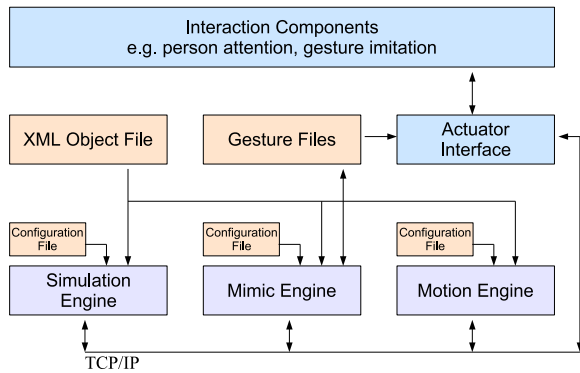


Fig. 6. The structure of the control software for driving BARTHOC's actuators in simulation and on the hardware via Motion Engine.

forwarded to the connected iModules. Also the temperature inside the robot can be interpreted and be used as “exhaustion” or “fatigue” indicator. The more strenuous a position is, the hotter the power stages of the motors become, the hotter the temperature inside the robot gets.

B. Control Software

The robot is controlled by the iActuator software architecture (see Fig. 6). Different programs – in the following called engines – are running at the same time and communicate via TCP/IP with each other. Each engine fulfills a special task concerning the robot's control, the simulation, or the behavior. Currently, three engines are implemented under Linux using Trolltech's Qt⁴: the Simulation Engine, the Motion Engine and the Mimic Engine. The iActuator Interface provides full communication to all engines and allows to build own engines. This open software structure ensures robustness, easy maintenance, and extension, which is of importance in this experimental stage. All mechanic, electric, and appearance properties of the robot are described in an XML object file. Configuration files for each engine allow flexible use on different computers. The delay from initiation of a command to beginning of the robot movement via TCP/IP and USB can be neglected due to the inertia of the mechanics.

The Simulation Engine (Fig. 7) visualizes the mechanic of the robot inside a 3D OpenGL environment. The simulated robot moves like its steel brother. Crash situations and gestures can be seen and tested before using the real robot. Background textures for sides, ground, and ceiling make the simulation more realistic. This engine helps to speed up development of own software modules. No laboratory time has to be shared and no precautions for humans and BARTHOC are required.

The Motion Engine controls the hardware of the robot. After receiving a command via TCP/IP from another engine,

⁴Qt is a cross-platform, C++ application development framework; <http://www.trolltech.com>.



Fig. 7. Screen shot: The robot Simulation Engine. This tool is indispensable for collision avoidance.

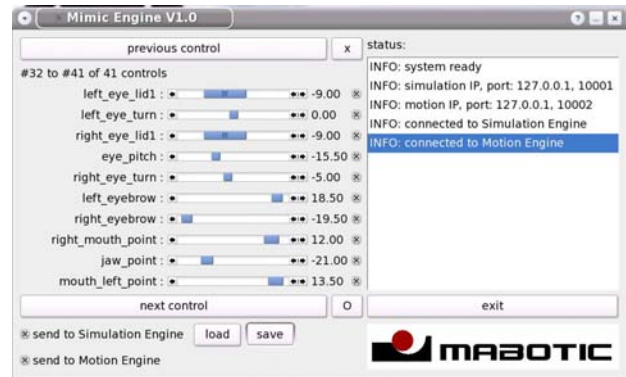


Fig. 8. Screen shot: The Mimic Engine. It is used to send movement commands to the robot and/or the simulation in an easy way. Positions defined can be stored and used with other engines.

it checks the command and all its parameters for validity. If the command and the parameters are valid, the Motion Engine sends appropriate movement or status commands to the iModule hardware via USB. No crash situations are checked at the moment. Collision avoidance and path planning for humanoid robots is an own field of research (see, e.g., [11]). It is a complete new topic in comparison to industrial robots, which not have to act social and human-like.

The Mimic Engine (Fig. 8) is a visual user interface for an application developer and allows to move every motor and servo of the robot by moving sliders. This way, manually generated commands can be forwarded to the Simulation Engine or/and the Motion Engine. All positions of the sliders – and so the position of the robot – can be saved in a gesture file. These files can also be used in other engines.

To build own engines for the iActuator system architecture, the iActuator Interface was developed. The C-language interface provides all functions to move the robot and/or the simulation and to replay gesture files. For example,

an Animation Engine for entertainment use of the robot is under development. Gesture files can be played after each other and the robot is able to play saved situations like an actress, an actor or a comedian would do. Other “more intelligent” iActuator engines are currently being realized (see next Section).

IV. ENABLING INTERACTION ON BARTHOC

Up to now, research on natural human-robot interaction has been performed using our mobile robot platform BIRON [12] (see Fig. 9). In order to enable BARTHOC to engage in a natural human-robot interaction, a variety of capabilities already developed for BIRON have been ported to BARTHOC. In the following we will briefly sketch these capabilities and their necessary modifications to apply them on BARTHOC.

BIRON’s basis is a Pioneer PeopleBot from ActivMedia™. The robot is equipped with an on-board PC (Pentium III, 850 MHz) for controlling the motors, on-board sensors, and for sound processing. An additional PC (Pentium III, 500 MHz) inside the robot is used for image processing and is equipped with a touch screen display. Both PCs each running Linux are linked by an 100 Mbit Ethernet LAN to a wireless LAN router enabling remote control of the mobile robot.

BIRON is equipped with a variety of sensors for locating communication partners and interacting with them. A pan-tilt color camera is mounted on top of the robot at a height of 141 cm for acquiring images of the upper body part of humans interacting with the robot. A VIDERE stereo camera at a height of 95 cm acquires stereo images of the humans in front of the robot. Two AKG far-field microphones are located right below the touch screen display. They enable speech processing and especially stereo-based speaker localization. Finally, distances within the scene and to humans are measured facilitating a SICK laser range finder mounted at the front at a height of 30 cm.



Fig. 9. BIRON.

The control of all components on BIRON and their communication is realized using the SIRCLE framework [13] which is based on a three-layer architecture using a central execution supervisor [14]. The framework itself enables easy reconfiguration and adaptation to new tasks. Therefore, SIRCLE can be applied easily for controlling the modules on BARTHOC. The overall component interaction that is being ported from the mobile BIRON to the static BARTHOC is depicted in Fig 10. However, some of the components require modifications in order to be used in BARTHOC’s architecture. In the following we briefly sketch the individual components and their modifications, if necessary.

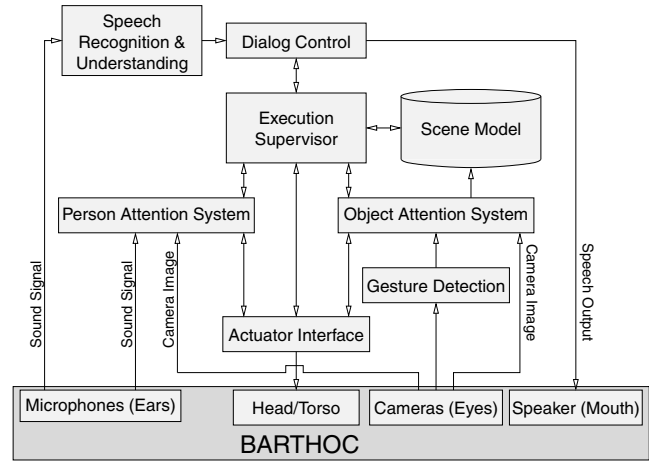


Fig. 10. The component interaction currently ported to BARTHOC.

A basic functionality of BIRON is its capability to track humans in its environment. This is achieved through a multi-modal anchoring approach using depth, vision, and sound data [15]. Based on the humans tracked in its surrounding and the multi-modal information associated with the individual humans, an attention mechanism allows to selectively pay attention to humans looking at the robot and speaking at the same time [16]. While BARTHOC does have vision and sound sensors (eyes and ears), it does not have a laser range finder for obtaining depth data like BIRON has. Consequently, the modification of BIRON’s person tracking and attention system was the first step in enabling BARTHOC to exhibit a natural attention behavior.

For BARTHOC’s attention behaviour the existence of a sound source is already sufficient to instantiate a person hypothesis. Subsequently, BARTHOC turns his head to the supposed person position to take an image of that area and search for a face at the expected position. Only if a face is successfully detected, the person hypothesis is retained. Different from BIRON where an animated face was displayed on the touch screen, BARTHOC can use mimic expressions to convey the state of the attention system to the user. In the inactive mode the mimics display a ‘sleeping’ face and if no person is around he has a ‘neutral’ face. On instantiating a person hypothesis, the mimics display an ‘interested’ face that changes to a ‘happy’ face if he is in an interaction with a human. During such an interaction (see also below), he displays a ‘thinking’ face if he processes speech input and a ‘searching’ face if he turns his head to search for an object referenced by the human.

However, a person attention behavior is not sufficient if a human wants to engage in a communicative interaction with the robot. In order to enable BIRON to understand natural language, we integrated components for speech recognition [16], speech understanding [12] and dialog [17]. These

components are all depending solely on a sound signal from the microphones and, therefore, do not require any modifications to be used on BARTHOC if the type of the instructions is kept unchanged.

In the most recent version of BIRON, the communication partner can not only get the robot's attention but can also control the robot's behavior by giving commands (for more details see [12]). For example, the command "Follow me" results in the robot following the human around. Obviously, this is an instruction that is not supported by BARTHOC and, consequently, the speech understanding and dialog components have to be modified to enable an adequate representation of what BARTHOC can understand and do.

The gesture detection and object attention components [18] do not require large modifications, as they are based only on camera images. Based on the static setup, their operation is even simplified, as BARTHOC's position in the room does not change and, therefore, the coordinate system is fixed. Moreover, the gesture detection could be made more robust by incorporating some kind of background subtraction.

Another important difference to BIRON is the humanoid body of BARTHOC. It has many more degrees of freedom that can be controlled during an interaction in order to exhibit a much more human-like behavior. On BIRON, the movements of the pan-tilt camera and of a face presented on its display reinforce the impression that the robot is paying attention and focusing on a communication partner, enabling humans to "read" the robot's current internal state. This somewhat human-like behavior is very much appreciated by users [19]. We expect that BARTHOC's human-like body will make this impression even stronger and will bring us further ahead in studying embodied interaction.

V. OUTLOOK

This paper has described our ongoing work for realizing a humanoid robot that supports research on embodied interaction. We have presented the technical details of the mechanics and the actuator control as well as the software tools for testing BARTHOC's actuators 'offline'. We are now porting the components developed for the mobile robot BIRON to the humanoid robot BARTHOC. This involves removing the dependency on the laser range finder from the person tracking module and modifying the speech processing components to cope with the different kind of instructions to enable natural interaction with it. At the end of the porting process, BARTHOC will be able to pay attention similar to a human by looking to the different communication partners around him. If BARTHOC is addressed by a human, it will focus its attention on this human and try to understand the instruction of the human. Through its humanoid appearance, BARTHOC will allow us to study a variety of aspects of embodied interaction that are becoming increasingly relevant for developing robot companions capable of natural human-robot interaction.

REFERENCES

- [1] K. Takekazu, M. Yasuhiro, S. Takeshi, "Cooperative distributed registration for robust face recognition," *Systems and Computers in Japan*, Vol. 33, No. 14, 2002, pp. 91-100
- [2] C. Breazeal, et al., "Humanoid robots as cooperative partners for people," *International Journal of Humanoid Robots*, 2(1), 2004
- [3] F. Tanaka, H. Suzuki, "Dance interaction with QRIO: a case study for non-boring interaction by using an entrainment ensemble model," 13th IEEE International Workshop on Robot and Human Interactive Communication, 2004
- [4] G. Sagerer, J. Fritsch, B. Wrede, "Bringing it all together: Integration to study embodied interaction with a robot companion," *AISB Symposium – Robot Companions: Hard Problems and Open Challenges in Human-Robot Interaction*, Hatfield, England, 2005
- [5] R. Stiefelhagen et al., "Natural Human-Robot Interaction using Speech, Gaze and Gestures," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Japan, 2004.
- [6] A. Ito, S. Hayakawa, T. Terada, "Why robots need body for mind communication – an attempt of eye-contact between human and robot," 13th IEEE International Workshop on Robot and Human Interactive Communication, 2004
- [7] M. Hackel, S. Schwope, "A Humanoid Interaction Robot for Information, Negotiation and Entertainment Use", *International Journal of Humanoid Robots*, Vol. 1, No. 3, 2004, pp. 551-563
- [8] A. Lankenau, T. Röfer, B. Krieg-Brückner, "Self-Localization in Large-Scale Environments for the Bremen Autonomous Wheelchair" in *Spatial Cognition III. Lecture Notes in Artificial Intelligence 2685*, C. Freksa, W. Brauer, C. Habel, K.F. Wender, Eds., Springer, 2003, pp. 34-61
- [9] I. Mizuuchi, et al., "The Design and Control of the Flexible Spine of a Fully Tendon-Driven Humanoid Kenta", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Lausanne, 2002
- [10] T. Hino, T. Maeno, "Development of a Miniature Robot Finger with a Variable Stiffness Mechanism using Shape Memory Alloy", *International Symposium on Robotics and Automation*, Querétaro, México, August 2004
- [11] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, B. Blumberg, "Learning from and about Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots," L. Rocha and F. Almedida e Costa (eds.), *Artificial Life*, 2004
- [12] A. Haasch et al., "BIRON – The Bielefeld Robot Companion," *Proc. Int. Workshop on Advances in Service Robotics*, E. Prassler and G. Lawitzky and P. Fiorini and M. Hägele, Eds., Fraunhofer IRB Verlag, Germany, 2004, pp 27-32
- [13] J. Fritsch, M. Kleinhagenbrock, A. Haasch, S. Wrede, G. Sagerer, "A Flexible Infrastructure for the Development of a Robot Companion with Extensible HRI-Capabilities," *Proc. IEEE Int. Conf. on Robotics and Automation*, Spain, 2005, pp. 3419–3425
- [14] M. Kleinhagenbrock, J. Fritsch, G. Sagerer, "Supporting Advanced Interaction Capabilities on a Mobile Robot with a Flexible Control System," *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Japan, 2004, pp. 3649–3655
- [15] J. Fritsch et al., "Multi-Modal Anchoring for Human-Robot-Interaction," *Robotics and Autonomous Systems*, Special issue on Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems, S. Coradeschi and A. Saffiotti, Eds., Elsevier Science, Vol. 43, No. 2-3, 2003, pp. 133-147
- [16] S. Lang et al., "Providing the Basis for Human-Robot-Interaction: A Multi-Modal Attention System for a Mobile Robot," *Proc. Int. Conf. on Multimodal Interfaces*, ACM, Canada, 2003
- [17] I. Topsis, S. Li, B. Wrede, G. A. Fink, "A Multi-modal Dialog System for a Mobile Robot," *ICSLP*, Korea, Vol. 1, 2004, pp. 273–276
- [18] A. Haasch, N. Hofemann, J. Fritsch, and G. Sagerer, "A multi-modal object attention system for a mobile robot," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Edmonton, 2005, to appear.
- [19] S. Li, M. Kleinhagenbrock, J. Fritsch, B. Wrede, G. Sagerer, "BIRON, let me show you something": Evaluating the Interaction with a Robot Companion," *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics*, Special Session on Human-Robot Interaction, W. Thissen, P. Wieringa, M. Pantic, and M. Ludema, Eds., 2004, pp. 2827-2834