Collaborative Research Center on Humanoid Robots (SFB 588)

Rüdiger Dillmann, Tamim Asfour

The Collaborative Research Center 588 "Humanoid Robots - Learning and Cooperating Multimodal Robots" was established in July 2001 by the Germany Research Foundation (Deutsche Forschungsgemeinschaft: DFG) and will run until June 2012. The goal of this interdisciplinary research project is the development of humanoid robots which resemble humans in their ways of acting in the world, of reasoning about the world and of communicating about the world. Apart from the development of mechatronics components for humanoid robots, the research efforts span a wide range of disciplines such as learning from observation, multi-modal human-humanoid interaction, human-humanoid cooperation, interpretation of human activities, as well as the execution of gasping and manipulation tasks in a household environment.

1 Introduction

Humanoid robotics is a new, challenging field of robotics and a good candidate to address the fundamental questions associated with personal, home robots. Humanoid robots are expected to exist and work together with human beings in the everyday world such as hospitals, offices and homes, and to serve the needs of elderly and disabled people. The term humanoid is commonly associated with the idea of robots whose physical appearance is similar to that of the human body. Beyond a physical resemblance, humanoid robots must resemble humans in their ways of acting in the world, of reasoning about the world and of communicating about the world. Most importantly, humanoid robots, like humans, should be provided with learning and adaptive capabilities to face a less predictable world. They should be able to interpret the actions of humans, in order both to react adequately to the humans' request and needs, as well as to learn from observing humans actions. The latter capability, often referred to as imitation learning or programming by demonstration, would facilitate human-humanoid interaction by providing humans with an intuitive way to teach humanoids.

Recently, considerable research work has been focused on the development of humanoid robots ([1, 9, 8, 5]). In contrast, our current research interest is the development of humanoid robots for applications in human-centered environments. In order for humanoid robots to enter such environments, it is indispensable to equip them with manipulative, perceptive and communicative skills necessary for real-time interaction with the environment and humans. In particular, we address the integration of perception, action and cognition in humanoid robots, which are rich in sensory and motor capabilities and hence provide on the one hand suitable framework for studying cognition and allow on the other hand the realization of service tasks in a household environment.

In this paper, we present the humanoid robots currently being developed within the SFB 588 and introduce the current state of work toward the realization of fully integrated humanoid robots in a household scenario. For a comprehensive overview the reader is referred to the project publications [1].

2 ARMAR-IIIa and ARMAR-IIIb

In designing our robots, we desire a humanoid that closely mimics the sensory and sensory-motor capabilities of the human. The robot should be able to deal with a household environment and the wide variety of objects and activities encountered in it. Therefore, the humanoid robots ARMAR-IIIa and ARMAR-IIIb (see Fig. 1) have been designed under a comprehensive view so that a wide range of tasks can be performed. The upper body of the robot has been designed to be modular while retaining similar size and proportion as an average person. For the locomotion, a holonomic mobile platform is used. From the kinematics control point of view, both robots consist of seven subsystems: head, left arm, right arm, left hand, right hand, torso, and a mobile platform. In the following the subsystems of the robot are briefly described. For detailed information the reader is referred to [3]. For a detailed description of the mechanics, the reader is referred to [2]. The head has seven DOF and is equipped with two eyes. The eyes have a common tilt and can pan independently. Each eye is equipped with two color cameras, one with a wide-angle lens for peripheral vision and one with a narrow-angle lens for foveal vision to allow simple visuomotor behaviours such as tracking and saccadic motions toward salient regions, as well as more complex visual tasks such as hand-eye coordination. The visual system is mounted on a four DOF neck mechanism (lower pitch, roll, yaw, upper pitch). For the acoustic localization, the head is equipped with a microphone array consisting of six microphones (two in the ears, two in the front and two in back of the head). Furthermore, an inertial sensor is installed in the head for stabilization control of the camera images.

The upper body of the robot provides 33 DOF: 14 DOF for the arms, 16 DOF for the hands and three DOF for the torso. The arms are designed in an anthropomorphic way: three DOF in the shoulder, two DOF in the elbow and two

---

[1] www.sfb588.uni-karlsruhe.de

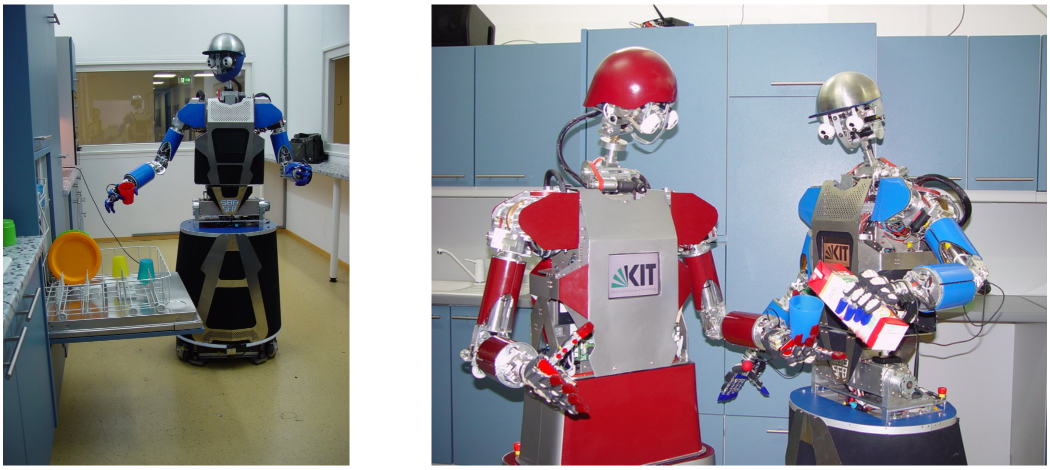

Figure 1: The humanoid robots ARMAR-IIIa (blue) and ARMAR-IIIb (red) in the kitchen. Each robot has an active head with foveated vision, two arms and two five-fingered hands and a holonomic mobile platform.

DOF in the wrist. Each arm is equipped with a five-fingered hand with eight DOF (see [10]). In order to achieve a high degree of mobility and to allow simple and direct cooperation with humans, the structure (size, shape and kinematics) of the arms has been designed to be similar to that of the human arm. The goal of performing manipulation tasks in human-centered environments generates a number of requirements for the sensor system, especially for that of the manipulation system. Each joint of the arms is equipped with motor encoder, axis sensor and joint torque sensor to allow position, velocity and torque control. In the wrists 6D force/torque sensors are used for hybrid position and force control. Four planar skin pads (see [6]) are mounted to the front and back side of each shoulder, thus also serving as a protective cover for the shoulder joints. Similarly, cylindrical skin pads are mounted to the upper and lower arms respectively.

The locomotion of the robot is realized using wheel-based holonomic platform, where the wheels are equipped with passive rolls at the circumference (Mecanum wheels or Omniwheels). In addition, a spring-damper combination is used to reduce vibrations. The sensor system of the platform consists of a combination of three Laser-range-finders (Laser-scanner) and optical encoders to localize the platform. The platform hosts the power supply of the robot and the main part of the robot computer system.

3 Programming by Demonstration

Programming by Demonstration facilitates human-humanoid interaction by providing humans with an intuitive way to teach humanoids. During the execution of a task in a so-called robot training-center, the user demonstration is observed using different sensors and analyzed to transfer the observed information for a specific task to robot-invariant knowledge, which can be used to generate the actions required to achieve the task goal. The sensor system used for the observation of human demonstrations include two data gloves for gathering the finger joint angles, magnetic trackers for obtaining the hands' position, and two active stereo cameras mounted on pan-tilt-units for object localization and tracking. From the sensor input data, so called elementary operators are extracted. Elementary operators (EO's) are actions that abstract a sensory motor coupling such as primitive movement types (linear moves etc.), grasping and ungrasping actions. These EO's abstract from the specific execution hardware and are robot-independent from a system point of view, although they have to be implemented on the specific target robot system. EO's are aggregated as primitives to so called macro-operators (MO's), containing the full task description. On a basic level, elementary move operators are aggregated to approach, depart and transport trajectories. Together with the grasp and ungrasp EO's, grabbing and releasing of objects can be constructed. Between those gripper operations, various basic manipulation operations are detected. On the highest level, a sequence of manipulation segments is subsumed under a whole task demonstration.

On each level in this hierarchical task representation, the changes to the environment are induced from lower levels of the hierarchy. The pre- and postconditions describing the environmental states before and after the execution of each macro-operator are propagated from the EO-level to the manipulation segment level and are computed from the environmental changes during the manipulation. The environment, its changes, and the pre- and postconditions are described in terms of first order predicate logics, based on a set of geometrical inter-object relations (on, under, to the right of, contained in, ...) and intra-object predicates such as object class (saucer, plate, table, ...) and internal state descriptors (opening angle, oven temperature, liquid level, ...), depending on the object class.

The system covers a broad range of manipulation classes drawn from the operations needed in a household environment: Transport operations (Pick&Place), device handling (operating household devices, opening a fridge) and tool handling (pouring liquids from a bottle to a glas). It depends on the condition that the user is doing all changes in the environment (closed world assumption). The result of the task acquisition process is a task description, containing a sequence of manipulation segments together with their pre- and post-conditions and the hierarchical decomposition into elementary operators.

4 Grasping and Manipulation Tasks

The central idea of our approach for the programming and execution of manipulation tasks is the existence of a database with 3D models of all the objects encountered in the robot workspace and a 3D model of the robot hand. This allows for an extensive offline analysis of the different possibilities to grasp an object, instead of focusing on fast online approaches. From this central fact we have developed an integrated grasp planning system, which incorporates a vision system for the localization and recognition of objects [4], a path planner for the generation of collision-free trajectories [12] and an offline grasp analyzer that provides the most feasible grasp configurations for each object [7]. The results provided by these modules are stored and used by the control system of the robot for the execution of a grasp of a particular object. A detailed description of the integrated grasping and manipulation system in given in [3].

5 Human-Robot Interaction

To enable multimodal human-humanoid interaction, a system for spontaneous speech recognition, multimodal dialogue processing, and visual perception of a user has been developed and integrated into the robot control architecture [11]. The system includes audio-visual localization, tracking, and identification of the user, recognition of pointing gestures, as well as the recognition of a person's head orientation. The building blocks of this system allow the interactive learning of dialog strategies as well as of new words, persons and objects. The dialog system allow the user to ask the robot questions related to the kitchen, to obtain recipes from the robot or to ask the robot to bring certain objects.

6 Summary

In this paper, we gave a short overview of some of the research activities of the Collaborative Research Center 588. Currently, our robots present highly integrated systems able to perform complex tasks in a kitchen environments such as "‘Bring me the apple juice from the fridge"’ or "‘Put this cup in the dishwasher"’. The emphasis of the research activities in the next period of the project is to study how higher-level cognitive capabilities emerge through exploration and interaction with the environments and on how developed skills and abilities can be transferred to varying situations and contexts.

References

[1] K. Akachi, K. Kaneko, N. Kanehira, S. Ota, G. Miyamori, M. Hirata, S. Kajita, and F. Kanehiro. Development of Humanoid Robot HRP-3. In *IEEE/RAS International Conference on Humanoid Robots*, 2005.

[2] A. Albers, S. Brudniok, and W. Burger. Design and development process of a humanoid robot upper body through experimentation. In *IEEE/RAS International Conference on Humanoid Robots*, pages 77–92, 2004.

[3] T. Asfour, P. Azad, N. Vahrenkamp, K. Regenstein, A. Bierbaum, K. Welke, J. Schröder, and R. Dillmann. Toward humanoid manipulation in human-centred environments. *Robot. Auton. Syst.*, 56(1):54–65, 2008.

[4] P. Azad, T. Asfour, and R. Dillmann. Stereo-based 6D Object Localization for Grasping with Humanoid Robot Systems. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, USA, 2007.

[5] G. Cheng, S. Hyon, J. Morimoto, A. Ude, and S. Jacobsen. CB: A Humanoid Research Platform for Exploring NeuroScience. In *IEEE/RAS International Conference on Humanoid Robots*, 2006.

[6] D. Göger, K. Weiß, K. Burghart, and H. Wörn. Sensistive skin for a humanoid robot. In *International Workshop on Human-Centered Robotic Systems (HCRS'06)*, Munich, 2006.

[7] A. Morales, T. Asfour, P. Azad, S. Knoop, and R. Dillmann. Integrated Grasp Planning and Visual Object Localization For a Humanoid Robot with Five-Fingered Hands. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, 2006.

[8] I. Park, J. Kim, J. Lee, and J. Oh. Mechanical Design of Humanoid Robot Platform KHR-3 (KAIST Humanoid Robot-3: HUBO. In *IEEE/RAS International Conference on Humanoid Robots*, 2005.

[9] S. Sakagami, T. Watanabe, C. Aoyama, S. Matsunage, N. Higaki, and K. Fujimura. The Intelligent ASIMO: System Overview and Integration. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2478–2483, 2002.

[10] S. Schulz, C. Pylatiuk, A. Kargov, R. Oberle, and G. Bretthauer. Progress in the development of anthropomorphic fluidic hands for a humanoid robot. In *IEEE/RAS International Conference on Humanoid Robots*, Los Angeles, Nov 2004.

[11] R. Stiefelhagen, H. Ekenel, C. Fügen, P. Gieselmann, H. Holzapfel, F. Kraft, K. Nickel, M. Voit, and A. Waibel. Enabling multimodal human-robot interaction fort the karlsruhe humanoid robot. *IEEE Transactions on Robotics, Special Issue on Human-Robot Interaction*, 23(5):840 – 851, 2007.

[12] N. Vahrenkamp, T. Asfour, and R. Dillmann. Efficient motion planning for humanoid robots using lazy collision checking and enlarged robot models. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 2007.

Contact
Prof. Dr. Rüdiger Dillmann
Universität Karlsruhe (TH)
Haid-und-Neu- Str. 7, 76131 Karlsruhe
Tel.: +49 (0)721-6083846
Fax: +49 (0)721-6084077
dillmann@ira.uka.de

Rüdiger Dillmann is Professor at the Computer Science Faculty, University of Karlsruhe (TH), Germany. His research interests include humanoid robotics, machine learning, computer vision, Programming by Demonstration, cognitive cars and medical applications of informatics.

Tamim Asfour is Sensior Research Scientist at the Institute for Computer Science and Engineering, University of Karlsruhe (TH), Germany. His research interests include humanoid robotics, grasping and manipulation, system integration, imitation learning and robot coaching.