The Collaborative Research Center 588: “Humanoid Robots – Learning and Cooperating Multimodal Robots”

Regine Becher
Institut für Rechnerentwurf und Fehlertoleranz, Universität Karlsruhe (TH), Haid-und-Neu-Str. 7, Technologiefabrik, 76137 Karlsruhe, Germany becher@ira.uka.de

Peter Steinhaus
Institut für Rechnerentwurf und Fehlertoleranz, Universität Karlsruhe (TH), Haid-und-Neu-Str. 7, Technologiefabrik, 76137 Karlsruhe, Germany steinhaus@ira.uka.de

Rüdiger Dillmann
Institut für Rechnerentwurf und Fehlertoleranz, Universität Karlsruhe (TH), Haid-und-Neu-Str. 7, Technologiefabrik, 76137 Karlsruhe, Germany dillmann@ira.uka.de

Received (Day Month Year) Revised (Day Month Year) Accepted (Day Month Year)

This paper gives an overview of the current and forthcoming research projects of the Collaborative Research Center 588 “Humanoid Robots – Learning and Cooperating Multimodal Robots”. The activities can be divided into several areas: development of mechatronic components and construction of a demonstrator system, perception of user and environment, modelling and simulation of robots, environment and user, and finally cooperation and learning. The research activities in each of these areas will be described in detail. Finally, we will give an insight into the application scenario of our robot system, i.e. the training setup and the experimental setup “household”.

Keywords: Multimodality; Man-machine cooperation; Humanoid robots; Interaction; Mechatronics; Learning; Perception; Environment modelling; User modelling; Control.
1. Introduction

The Collaborative Research Center 588 “Humanoid Robots – Learning and Cooperating Multimodal Robots” (SFB 588) was established by the Deutsche Forschungsgemeinschaft (DFG) in July 2001 and is intended to run for 12 years. It is dedicated to investigate the various scientific challenges in the area of humanoid robotics. Goal of this project is the development of concepts, methods and concrete mechatronic components for a humanoid robot which shares its workspace with humans. By use of this specifically constructed partially anthropomorphic robot system, the step out of the “robot cage” and thus the direct contact to humans will be realised.

The evaluation of the concepts and prototypes developed in this project takes place in an everyday human scenario: Human user and robot act together in a kitchen environment. The integration of a robot system into human everyday life is one of the outstanding goals of this project.

In this interdisciplinary research center, 12 institutes from the faculties of computer science, mechanical engineering, electrical engineering and arts of the University of Karlsruhe, Germany, as well as other research institutions are involved. The more than 40 scientists belong to the following research organizations:

- University of Karlsruhe (TH)
- Forschungszentrum Karlsruhe
- Forschungszentrum Informatik
- Fraunhofer IITB

This paper gives an overview of the current and forthcoming activities of the Collaborative Research Center 588. In section 2, other humanoid robot projects are described. Section 3 discusses our main points of focus. Sections 4 to 7 explain our research in more detail. In section 8, we give an insight into our application scenario. The paper closes with a short summary (section 9).

2. Humanoid Robots in International Research

More than 30 years ago, first developments have taken place in the area of mechatronic sub-systems of humanoid robots. Since then, the technological environment (e.g. computing power, energy consumption, integration scale) has developed revolutionary. This has made acceptable solutions for sub-problems possible which have been unthinkable even 10 years ago. As a result of these advances, humanoid robots were rediscovered as a research topic. Completely built-up or in their development very advanced humanoid robots can mostly be found in the mother country of humanoid robotics, Japan, and in the United States.

In 1998, the Agency of Industrial Science and Technology (AIST) and the Ministry of International Trade and Industry (MITI) created a project in Japan which used a platform approach (HRP-1 = humanoid robot platform 1) for the construction of a humanoid robot. As a platform, the project used the humanoid robot P3 from Honda, which was designed as a successor of the P2 model. P2 was the
first humanoid robot world-wide which could walk very human-like and also climb stairs. Further information about the platform approach can be found in 1, about the robots P2 and P3 in 2. Further enhancements of the ability to move entered into the successive model Asimo.3 In competition to Honda, Sony developed the Dream Robot (SDR), which can be assigned to the field of entertainment.4

In an early stadium of development are the robots of the Waseda University in Japan:5 Hadaly-2 which is built onto a mobile platform and equipped with manipulators, and WABIAN which is a two-legged robot with a human-like walk. WABIAN can also carry objects. Main research goal of Hadaly-2 is the interaction with the human user on the levels of speech, gesture and image processing, whereas the abilities of WABIAN on this area are somewhat more restricted.

In 2000, the Humanoid Robotics Institute of Waseda University was founded. Current research includes among others Wendy, a robot which is supposed to interact with humans in their everyday environment. Wendy consists of a mobile platform, a torso, a head, two arms and hands. The robot Robita is used for investigating multimodal communication with humans.6 Other robot (sub-)systems are developed to investigate emotions of humans and robots, the use of sensor systems, human mimics etc. An integration of these very different functionalities on one robot system is currently not given.

A very well-known project from the United States is the robot Cog of the Massachusetts Institute of Technology.7 Cog was developed as a research platform for evaluating methods of multi-sensor use and robot-human interaction. Cog consists of a torso with two manipulators and a head which was equipped with various sensor systems. The complete system has 21 degrees of freedom for imitating human motions, and by use of its sensor systems, human senses (visual, vestibular, auditive and tactile) can be approximated. Another interesting project is the so-called Robonaut project of NASA’s Johnson Space Center which tries to develop a humanoid robot system that can function as an astronaut equivalent.8 The focus in this project is on manipulation, and the robot is still controlled by a human operator.

To do research on two-legged walking and running, the robot Johnnie was developed in Munich.9 Johnnie was built for the DFG Priority Program “Autonomous Walking”. The robot is able to walk independently. Goal is a human-like, dynamically stable walk.

Most of these projects have in common that they concentrate on developments in the field of mechatronics and control, but only peripherally care about the perception of environment and user.

Both at MIT and at Waseda University, interesting concepts of perception and cognition for human-machine interaction were developed, but mobility, flexibility or the ability to learn are investigated only peripherally.

The Collaborative Research Center 588 is outstanding in that it tries to integrate all abilities which are necessary for a humanoid robot system on one demonstrator system.
3. Main Research Topics of the SFB 588

For a robot system to be a helpful assistant in human everyday life, it must have various complex abilities and features:

- Humanoid shape
- Multimodality
- Ability to cooperate
- Ability to learn

To be accepted by humans and to be able to act together with humans, a human-like shape is at least advantageous. Thus, a mobile two-arm system with five-finger hands, a flexible torso and a sensor head with visual and acoustic sensors are constructed in our project. In addition, the motion system and thus the behaviour of the robot will be tailored to human-like motions.

In the background of this project, we use the term multimodality for the communication channels which are intuitive for humans like speech, gesture and haptics. They should be used for a direct commanding or instruction of the robot system.

Concerning cooperation between human and robot, e.g. for manipulating objects cooperatively, it is vital for the robot to recognize human intentions, to remember actions which have been carried out cooperatively and to use this knowledge correctly for each specific task. As the safety of the human plays a very important role, the architecture of the system is designed especially for this aspect of human-machine cooperation.

As an outstanding ability of our system, the ability to learn has to be emphasized, as it allows to make the robot carry out new, previously unknown tasks. New concepts and new objects, even new motions are learnable by help of a human, and can also be corrected interactively even by inexperienced users.

By analyzing these four features and abilities, the following problem fields and thus research areas were identified:

- Conception of mechatronic components and construction of the demonstrator system
- Perception of environment and user
- Modelling and simulation of robot, environment and user
- Cooperation and learning

4. The Demonstrator System

The SFB 588 demonstrator is developed and constructed by use of the experiences which have been made with the prototype system ARMAR. ARMAR is a robot system consisting of a mobile platform with a differential drive, route planning and collision avoidance, of a tiltable telescope torso, of two 7-axis light-weight arms with two-finger gripper, and of a sensor head which is equipped with a stereo camera system (cf. figure 1 left).
In the process of developing the demonstrator, concepts with different head forms, neck joints, arm constructions and tors and hip joints have been investigated. Below, the selected construction variant is discussed (cf. figure 1).

The conception as a whole consists of a hip joint with two degrees of freedom as well as a sensor head, a neck with four degrees of freedom, light-weight arms and a torso. The sensor head will be equipped with acoustic and visual sensors. In particular, a microphone array for a spacial analysis of the acoustic scene and a colour stereo camera system will be integrated.

The sensor head of this construction is moved by a gear rack neck (cf. figure 2 left). The 7-axis light-weight arms are driven by a combination of several drive
concepts.

The shoulder joint is equipped with three degrees of freedom, which are driven by the electric drive which is installed in the torso (cf. figure 2 right). Position sensors are realized by optical sensors in this case, and momentum sensors by power measurement, which is possible because of the extremely small-play drives with a very high efficiency.

The elbow joint with one degree of freedom is realized by a special linkage technique (cf. figure 3 left). Here, potentiometers are used for position measurement. The momentum sensors are realized by the installation of small force-torque sensors in the linkages, which are used in the torso area.

![Fig. 3. The elbow joint (left), the construction of the wrist (middle) and a prototype of the hand (right)](image)

The wrist with two degrees of freedom is driven hydraulically (cf. figure 3 middle), such that the actors are positioned in the torso here as well.

In addition to this system concept, an anthropomorphic five-finger light-weight robot hand\cite{Note1} with an appropriate control is built in the framework of the Collaborative Research Center. This hand, on basis of the flexible fluidic actor technology (cf. figure 3 right) has a human-like outer appearance as well as an appropriate motion sequence. The robot hand will be able to grasp objects of very different kinds with power or precision grasps, and offers possibilities for a force-based cooperation between human and robot by the integration of the sensor systems with an appropriate reactive control.

The local control and planning system on the hand level, for force- and precision-based control of the fingers as well as of a grasped object is built up hierarchically. On the uppermost level, there are grasping skills with a description semantics, below them is a real-time system for reactive force-position-based object position control on hand level. On the lowest level, the sensor data of the position sensors in the fluid actors in combination with the force and pressure sensors for the finger tips and finger joint surfaces are used for controlled grasps and the interaction with humans.
This “sensor skin” is not only attached to the hand itself, but also to other exposed places of the robot. As an example for this sensor system, parts of the human somatovisceral sensory system are used, which contains both the surface sensibility of the skin which records position and movement of body parts, and the forces which act on them. The haptic sensibility will technically be imitated by force sensors on the robot’s surface, force and angle sensors in the arm joints, force-momentum sensors between arm and hand, and by the integration of acceleration sensors on relevant positions on the robot.\textsuperscript{13} A classification of possible human-robot contacts was set up, whose formal description is used as a basis for later pre-classifications of sensor signals.

Apart from the control on hand level, each degree of freedom of the robot system must be controlled depending on the actual task context and state of the machine. Stability problems inside the robot control circuit which emerge during force-based human-robot cooperation and which can be potentially dangerous for the human are detected and assessed by an operative stability monitoring system consisting of model and fuzzy components.\textsuperscript{14} With the aid of a model- and fuzzy-based adaption algorithm, operative stability procedures are initiated in danger of instability.

Above the level of basic control units, the coordination of movements of the partially anthropomorphic two-arm system takes place.\textsuperscript{15} It provides the robot with a coordinated use of its extremities, its torso, and its platform, and thus allows for mobile two-arm manipulations (cf. figure 4). The coordination of movements will be optimized towards movements which are as human-like as possible. On the basis of movement coordination, the platform will make more and more challenging basic movement skills available.
5. Perception

To realize a multimodal communication interface which gives the human user an intuitive account of the robot system by means of the modalities speech and gesture, differentiated abilities of perception are necessary. These are, amongst others, speech recognition and synthesis, gesture recognition and generation, dialogue management, user recognition, user localisation and acquisition of information about the environment.

In the field of speech input, a system is developed which realizes a speaker-independent speech recognizer for spontaneous speech under contrarious acoustic conditions. In addition, a parser for multimodal semantic analysis of speech and gesture input is integrated into this system. Furthermore, a module for dialogue management is set up which allows for interaction with the user depending on the actual task and the actual context of interaction.

The gesture recognition takes place based on skin colour segmentation in combination with depth information which is extracted from stereo image analysis (cf. figure 5).

Another research goal is the development of an audio-visual attention control. It serves the robot to identify important objects in a static scene and to allow a closer analysis of the identified objects. In a dynamic environment, such an expensive analysis will be restricted to relevant changes or events, respectively. In particular, if a user does not react to a potentially dangerous situation, the robot should react adequately. For this purpose, already existing visual attention control systems will be combined and expanded. Among these expansions, the inclusion of the acoustic component, of the environment model and of the “focus of attention” of the user has to be emphasized particularly.

For multimodal recognition and localisation of the user, acoustic and optical data are fused. The acoustic component is used particularly for the initialisation and should help in situations in which the optical tracking does not work. A Kalman filter is used to fuse the data streams. Additionally, users will be identified on an
Preprocessed and pre-classified signals are the basis for each acoustic analysis. The signals result from an acoustic scene analysis. Beamforming, a method of spatial filtering, allows for a separation of the particular sources of noise, and is enhanced by the statistical method of blind-source separation. In this way, the speech signal for the speech recognition is enhanced in diffuse background noise. By separating the sources of noise, the robot system is also able to localize and track several sources of noise acoustically, even if they are active at the same time. Thus, not only the position of a speaker can be determined, but also the position of background noise can be pointed out. A classification of the separated sources of noise can be reached by neural network classifiers which extract time and frequency features from the signals.

6. Modelling and Simulation

Foundation of all features and abilities of a humanoid robot system are dedicated models to represent its knowledge about itself, about its environment and its user(s). Therefore, one of the main foci of this Collaborative Research Center is the conception of such models and the evaluation of their practical applicability.

The models which are used can be classified into the following categories:

- models of scene and objects
- models of the user
- models of motion (of human and robot)
- models of sensors and actors
- models of actions
- goals

Whereas both internal models for the kinematics and dynamics of robot systems and representations of objects and their features have been analyzed for quite a while
now, there is still a great need for research concerning the modelling of the user, his behaviour, his motions and his intentions.

One of the main concerns of the SFB 588 is thus the analysis and modelling of human basic motions. This suits several purposes: First, successive motions of the user can be predicted by a correct modelling of motions sequences via recording of user motions and context information. Secondly, the recognition of individually different patterns of motion of the user allows the robot to recognize communication partners. And finally, the robot which is developed in the course of this project should have a human-like shape to minimize fears of contact of the human user and to ease its operation for non-expert users. To this end, as human-like motions as possible are necessary on the part of the robot. They allow for a better estimation of the robot movements by the user and thus enhance safety and acceptance. At the same time, they ease the human-robot cooperation, as e.g. carrying objects together is significantly eased by similar motion trajectories.

The comprehension of human motions is thus the basis for an automatic recognition of human actions and intentions, for the execution of human-like motions by the robot and finally for a functional cooperation of human and robot. Therefore, an elementary goal is the kinematic analysis of human basic motions for simple object manipulations (e.g. grasping, turning, moving). In doing so, characteristic motion sequences in form of trajectories or joint angle coordinates are extracted (cf. figure 7). They are used as patterns for basic robot motions. To delimit the motion space of the robot concretely, motion studies are carried out which determine intra- and inter-individual variances in spatiotemporal motion executions with a series of subjects.

The motion data which is gathered by electromagnetical and optical motion recording is (under consideration of anatomical and geometric constraints) smoothed, registered and fused to determine the inter- and intra-individual characteristic features, i.e. motion trajectories must be deformed without loss of variability and degrees of freedom such that a comparison is possible and the data can be merged into a consistent framework. The data processing is achieved by fast signal transformations. The resulting prototypical evaluation data is analyzed statistically to get transition probabilities for successive motion patterns. These probabilities are
used to classify and predict user motions during robot operation. In addition, the mapping of human motions on the robot arm is investigated with different degrees of freedom.

Parallel to the modelling of the user, modelling the multimodal communication context is another very important task building up a multimodal interface for the robot system. Dialogue models, fusion methods for speech and gesture input and context models for correct interpretation of input data are part of this interface.

Apart from the models described above, other models are necessary for purposes of simulation, e.g. detailed sensor models for simulating the robot’s sensor systems. The tools which are developed and put into use in this Collaborative Research Center enable a simulation of kinematics and dynamics of the robot system and thereby help to construct new mechanical components. By an in-the-loop integration of robot software and the simulation of sensors and actors, the primary evaluation of new algorithms can be supported without putting humans at risk.

We aspire a real-time simulation and visualization of the application scenario; possibly constraints concerning the level of detail have to be taken for specific applications.
As the simulation can also reconstruct and predict the impact of manipulations by the participating agents on the environment, it is aspired to make it available as an offline optimization tool for the robot, i.e. the robot system will play through action sequences in VR space and be able to learn from these experiences. A fine optimization of motion trajectories by the simulation tool is planned.

Starting from detailed scene and object models (cf. figure 9) for the representation of current environment states as well as their actual and potential features, action and task models are necessary. Action models describe concrete, self-contained actions, which normally can be carried out context-free and are mapped on basic skills of the robot system. Task models are used to represent complex actions and represent the context for a series of actions. In the task models (cf. figure 10), e.g. temporal and spatial dependencies of actions and their distribution on the potential actors (distribution of competence between human and machine) are modelled. The application area of task models is laid out both for the recognition of the user’s intention from the observation of action sequences as well as for the appropriate activation of cooperation patterns.

Task models are represented by graph-based structures in which both possible sequence orders as well as possible parameterizations are modelled. The graph-based approach has the advantage of easier integration of further solutions which can be extracted either from additional demonstrations with different solutions or from a generalisation or recombination of already learned solutions. As far as generalisation or recombination are concerned, the consistency of the new solution hypothesis has always to be ensured. It is tested based on a precondition-postcondition system which is applied to the set of basic operators.

Both for the expansion of the known object knowledge as well as in a concrete application situation for a more precise specification of the environment, techniques for interactive environment modelling on basis of the multimodal communication interfaces will be developed, i.e. the user should be able to present new, previously unknown objects to the robot system and describe them via speech as well as via gestures (cf. figure 11). Additionally, a measuring setup is used to model these new objects with the necessary precision. This process happens in the training setup of the robot system (cf. section 8).
7. Cooperation and Learning

Fundamental features of a humanoid robot which is supposed to work together with humans in an environment that is adapted to humans are the robot’s ability
to cooperate and to learn. They arise from the claim that the human should ideally be able to behave towards the robot as if he or she would behave towards another human.

The field of learning spans most of the technical capabilities of the robot system. In the center of this Collaborative Research Center are the aspects of

- learning task knowledge and
- learning cooperation strategies and patterns.

The concepts which are used and enhanced for this purpose originate from the paradigm of Programming by Demonstration and enable a non-expert user, who has the necessary knowledge to solve a problem, to program the robot system by demonstrating the practical solution to the problem. The humanoid robot should be able to generate new, valid solutions from these demonstrated solutions simply by generalizing and recombining them. During the demonstration of these abilities, the user can assist the observation process by multimodal comments and commands.

Learning cooperation strategies is an enhancement of learning task knowledge, since problems of the allocation of competences between the cooperation partners in indirect cooperation and the specific cooperation skills in direct cooperation must be recognized and parameterized.

By the different learning procedures, the task memory of the robot system is built up, whereby the learning procedures are mainly conducted in a specific training setup (cf. section 8). At the moment, learning takes place as offline learning, i.e. the robot system does not learn these new actions during their application, but only via the training setup assisted by external sensor systems. For future stages of the Collaborative Research Center, it is planned to substitute the training setup more and more by robot-internal sensor systems and thus allow online learning in the experimental setup. The task memory which is such acquired now constitutes the central interface between the areas of learning and of cooperation. Results of the learning process are the knowledge about task solutions saved in task models and the cooperation patterns which are learned.

In the field of cooperation, the two aspects “implicit communication” and “cooperation” are more closely investigated starting from this task memory. In implicit communication, the robot is supposed to interpret the human intention based on sensor data and, if appropriate, conclude commands for itself. Additionally, the comprehension of human intentions is important to guarantee as much safety for humans as possible. In cooperation, the robot is supposed to solve problems together with the human user. In this mode, the human is acting in the working space of the robot – here, too, safety demands are extraordinarily high.

To reach as optimal a cooperation as possible with highest safety demands, the possibilities for cooperation between human and robot were classified, which led to 18 cooperation classes.

To ensure safety, we developed the concept of restricted reflexes, which means that the set of allowed reflexes in case of an unexpected situation is restricted by
the momentary action. Then, interaction primitives are defined for each cooperation class which allow the robot to distinguish communication classes by means of their sensor data. Such, it is possible for the robot to realize which type of cooperation begins, ends, or goes on to a certain point of time. To have as natural a communication between human and robot as possible, given domain restrictions should be lifted. For this end, a learning man-machine interface is developed by which unknown terms are interpreted by exploration, by interaction and by acquisition of all information which is to the disposal of the robot (speech, gesture and information from the environment model of the robot). The components speech recognizer, parser and dialogue manager are co-ordinated such that they build upon a shared linguistic knowledge (vocabulary, grammar). In the field of speech recognition, adding on new words is possible by detection of uncertain areas in a recording by confidence measures and successive phoneme recognition in these areas. Then, a clarification dialogue takes place, in which the user verifies candidates which were suggested by the system. The existing dialogue manager is extended such that it can react on emotions of the user. An optimization of the dialogue manager will e.g. use Reinforcement Learning.

Despite all means for correct and broad environment modelling, exceptions which could not a priori be expected – e.g. by falling objects during an object manipulation – can always happen. To ensure a sensible dealing with such problematic situations as well, it is important to analyse these situations constantly and, in the case of an exception, to draw conclusions on the causes and effects.
8. Training and Experimental Setup

Goal of the Collaborative Research Center is to develop a robot system which can be useful in everyday environments. Thus, specific elementary or higher actions or tasks must be laid down which the robot system should be able to execute in appropriate problem or task contexts. Table 1 shows part of the analysis result. With the kitchen environment, we found an application scenario which allows the execution of as many of these elementary actions as possible in an everyday context.

The application scenario of the demonstrator system can be divided into two parts, namely the training setup and the experimental setup.

In the training setup, conditions are controlled as it is necessary for many learning processes. For the demonstration of task solution knowledge, the external sensor systems of the training setup are available for observation, which are in particular two data gloves, magnetic field-based position and orientation sensor systems and visual sensors. These sensor systems assist in segmenting user demonstrations in elementary operations like e.g. approach movements, grasps (according to Cutkosky’s hierarchy), disapproach movements, free movements et cetera. They enable an easy and robust instruction via data glove commands to multimodally comment a user demonstration, and by the simple structure of the training setup, noisy and erroneous sensor data can be corrected mostly automatically and even parallel to the demonstration.

The same structure eases the modelling of new objects and will offer a scanning environment to interactively import detailed models e.g. in form of triangulated surface models.

All necessary tools to visualize the processes which are executed during learning are existing and facilitate operating for the user.

From the very first, the training setup was only designed as an addition to our robot system and with increasing abilities of the robot, the training setup should more and more take a back seat. To be able to adopt the methods developed for the training setup to the robot system without changes, we paid great attention to
Fig. 13. The training setup with data glove and stereo tracking system

providing both systems with identical sensor systems and computer architecture. The training setup offers, especially at the beginning of the Collaborative Research Center, the advantage of a relatively clear, heavily structured environment, and thus eases the development and evaluation of appropriate methods. Therefore, a complete data flow to all components of the robot system can be guaranteed from the very beginning.

The long-term objective of the SFB 588 is the development of a humanoid robot which can assist a human in real everyday-life environments like a kitchen. The experimental setup which is used in our project corresponds to a real kitchen environment with minor adaptations to the working space of our demonstrator prototype, comparable to a kitchen for a handicapped person who depends on a wheelchair, where the heights of the cupboards etc are appropriately adapted. In the course of our research, and with proceeding progressions in the area of hardware, the kitchen environment will more and more be adapted to off-the-shelf kitchens. At the very end stands a humanoid robot which acts in a completely normal kitchen. A picture of the current kitchen can be found in figure 14.

For safety reasons, the experimental setup is equipped with external sensor systems, too. A network of four colour camera units is currently installed, which will be expanded in future. The colour camera network can be used for fast 3d modelling of the environment, and thus to integrate the environment in the safety monitoring. In future phases of the Collaborative Research Center, this functionality will assist in online learning.

9. Summary

In this paper, we gave an overview over the current and forthcoming research activities of the Collaborative Research Center 588 “Humanoid Robots – Learning and cooperating multimodal robots”.
Long-term objective of the project is to develop a robot system which can be a helpful assistant of humans in everyday environments. The following important characteristics of a humanoid robot system were pointed out:

- (Partially) humanoid shape
- Multimodality
- Ability to cooperate
- Ability to learn

According to these demands, the following main foci of research were identified:

- Development of mechatronic components and construction of the demonstrator system
- Perception of environment and user
- Modelling and simulation of robot, environment and user
- Cooperation and Learning

These four key aspects were discussed in this overview paper. Finally, we described our application scenario kitchen and the appropriate training and experimental setup.

For more detailed information as we could give in this overview, we recommend to visit the web page of the Collaborative Research Center 588, where you can also find other current publications from the different research fields:

http://www.sfb588.uni-karlsruhe.de

Acknowledgements

The work of the Collaborative Research Center 588 “Humanoid robots – learning and cooperating multimodal robots” is supported by the Deutsche Forschungsgemeinschaft (DFG).

The authors would like to thank all members of the Collaborative Research Center for their support in the compilation of this overview.
References
**Regine Becher** received her M.A. degree from the University of Freiburg, Germany, in 2001. Since 2001, she is a Research Assistant at the Institut für Rechnerentwurf und Fehlertoleranz at University of Karlsruhe (TH), Germany, and works at the German Collaborative Research Center on Humanoid Robots.

**Peter Steinhaus** received his Dipl.-Inform. degree in Computer Science from the University of Karlsruhe (TH), Germany, in 1997. He currently is a Research Assistant at the Institut für Rechnerentwurf und Fehlertoleranz at University of Karlsruhe (TH), Germany, and coordinates the German Collaborative Research Center on Humanoid Robots.

**Rüdiger Dillmann** received his Dipl.-Ing. degree in Electrical Engineering and his Ph.D. degree from the University of Karlsruhe (TH), Germany, in 1976 and 1980, respectively. He is currently a Full Professor in the Department of Computer Science at the University of Karlsruhe (TH), Germany. His research interests include Humanoid Robotics, Programming by Demonstration and Medical Simulation systems. Rüdiger Dillmann is the head of the German Collaborative Research Center on Humanoid Robots in Karlsruhe.