# ARMAR-III: An Integrated Humanoid Platform for Sensory-Motor Control

T. Asfour, K. Regenstein, P. Azad, J. Schröder, A. Bierbaum, N. Vahrenkamp and R. Dillmann
University of Karlsruhe
Institute for Computer Science and Engineering (CSE/IAIM)
P.O. Box 6980, D-76128 Karlsruhe, Germany
Email: {asfour,regenst,azad,schroede,bierbaum,vahrenka,dillmann}@ira.uka.de

***Abstract*— In this paper, we present a new humanoid robot currently being developed for applications in human-centered environments. In order for humanoid robots to enter human-centered environments, it is indispensable to equip them with manipulative, perceptive and communicative skills necessary for real-time interaction with the environment and humans. The goal of our work is to provide reliable and highly integrated humanoid platforms which on the one hand allow the implementation and tests of various research activities and on the other hand the realization of service tasks in a household scenario. We introduce the different subsystems of the robot. We present the kinematics, sensors, and the hardware and software architecture. We propose a hierarchically organized architecture and introduce the mapping of the functional features in this architecture into hardware and software modules. We also describe different skills related to real-time object localization and motor control, which have been realized and integrated into the entire control architecture.**

## I. INTRODUCTION

The design of humanoid robots requires coordinated and integrated research efforts that span a wide range of disciplines such as learning theory, control theory, artificial intelligence, human-machine interaction, mechatronics, perception (both computational and psychological), and even biomechanics and computational neuroscience. These fields have usually been explored independently, leading to significant results within each discipline. The integration of these disciplines for the building of adaptive humanoid robots requires enormous collaborative resources that can be achieved only through long-term, multidisciplinary research projects.

Our current research interest is the development of humanoid robots which safely coexist with humans, interactively communicate with humans and usefully manipulate objects in built-for-human environments. In particular, we address the integration of motor, perception and cognition components such as multimodal human-humanoid interaction and human-humanoid cooperation in order to be able to demonstrate robot tasks in a kitchen environment as a prototypical human-centered one [1]. Recently, a considerable research work has been focused on the development of humanoid biped robots (see [2]–[5]). In order for humanoid robots to enter human-centered environments, it is indispensable to equip them with manipulative, perceptive and communicative skills necessary for real-time interaction with the environment and humans. The goal of our work is to provide reliable and highly integrated humanoid platforms which on the one hand allow the implementation and tests of various research and on the other hand the realization of service tasks in a household scenario.

The paper is organized as follows. In Section II, we describe the different components of the humanoid robot, its kinematics and sensor systems. Section III describes the control architecture including its hardware and software modules. The mapping of this architecture into a computer architecture is described in Section IV. The implemented features are presented in Section V. Finally, Section VI summarizes the results and concludes the paper.

## II. THE HUMANOID ROBOT ARMAR-III

In designing our robot, 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 robot must be designed under a comprehensive view so that a wide range of tasks (and not only a particular task) can be performed.

Fig. 1. The humanoid robot ARMAR-III.

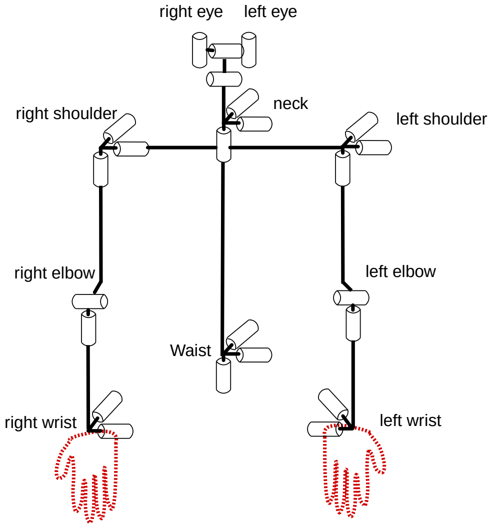

Fig. 2. Kinematics of ARMAR-III: The head has a total number of 7 DOFs. The Waist has 3 DOFs. Each arm has 7 DOFs. Each hand has 8 DOFs. The mobile platform has 3 DOFs.

The humanoid robot ARMAR-III (Fig. 1) has 43 degrees-of-freedom (DOF). From the kinematics control point of view, the robot consists of seven subsystems: head, left arm, right arm, left hand, right hand, torso, and a mobile platform. Figure 2 illustrates the kinematics structure of the upper body of the robot. The upper body has been designed to be modular and light-weight while retaining similar size and proportion as an average person. For the locomotion, we use a mobile platform which allows for holonomic movability in the application area.

### A. The Head/Neck System

The head has seven DOFs and is equipped with two eyes. The eyes have a common tilt and can pan independently. The visual system is mounted on a four DOF neck mechanism [6]. Each eye is equipped with two digital color cameras (wide and narrow angle) to allow simple visuo-motor behaviours such as tracking and saccadic motions towards salient regions, as well as more complex visual tasks such as hand-eye coordination. The head features human-like characteristics in motion and response, that is, the neck and the eyes have a human-like speed and range of motion.

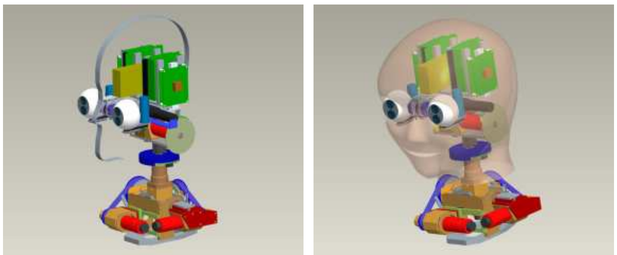

Fig. 3. Rendering of the head/neck system: Two cameras per eye. The eyes have a common tilt and can pan independently. The visual system is mounted on a 4 DOF neck which is realized as a Pitch-Roll-Yaw-Pitch mechanism.

We use the Point Grey Research Dragonfly camera in the extended version (`www.ptgrey.com`). The cameras can transmit color images with a resolution of $640 \times 480$ at 30 Hz. To reduce bandwidth it is possible to transmit the raw 8 bit Bayer pattern and perform RGB color conversion on the PC.

Furthermore, 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). This is necessary for a 3D acoustic localization.

### B. The Upper Body

The upper body of our robot provides 17 DOFs: 14 DOFs for the arms and 3 DOFs for the torso. The arms are designed in an anthropomorphic way: three DOFs in the shoulder, two DOFs in the elbow and two DOFs in the wrist. Each arm is equipped with a five-fingered hand with eight DOF ( [7], [8]). The main goal of our research is to build humanoid robots which can support people in their daily life. The main component of such a robot for handling objects is its manipulation system. The design of the arms is based on the observation of the motion range of a human arm. From the mechanical point of view, the human arm can be modelled by a first order approximation as a mechanical manipulator with seven DOFs. The links of the arm are connected by one DOF rotational joints, each specifying a selective motion.

The goal of performing tasks in human-centered environments generates a number of requirements for the sensor system, especially for that of the manipulation system. To achieve different control modalities, different sensors are integrated in the robot. Due to space restrictions and mechanical limitations we have to approach the sensor configuration in different ways. For example a sensor fitting into the elbow will most likely be too large for the wrist. In the current version of the arms we monitor motor revolution speed, position of axis and axis torque in each joint.

- For speed control, we use a persistent sensor concept. We deploy a motor configuration where the sensor is attached to the axis of the motor. Still, depending on the size of the motors, these sensors are optical or magnetical but have the same quadrature coded signal as output.
- To measure the position of all axes except the wrist we use an optical encoder in the axis itself. This encoder consists of an optical sensor scanning a reflective code-wheel. By reading the incremental and the coded track of this code-wheel an absolute position can be obtained after a marginal movement. Due to the space restrictions in the wrist a potentiometer is used to obtain an absolute position value.
- Joint torque sensors: The active joints of both arms are equipped with force sensors. For the three shoulder joints, torque can be measured separately for lifting, turning and swiveling. The torque for lifting the upper arm is measured via miniature load cells with a bi-directional measurement range of up to 1 kN (Novatech Measurements Ltd., `www.novatechuk.demon.co.uk`). The torque acting when turning the upper arm is

  determined with a sensor of the same type, but with a lower measurement range of up to 500 N, as this motion typically introduces less torque. For torque of the swiveling DOF a custom torque sensor utilizing strain gages has been developed [6]. The linkage system for moving the lower arm at the elbow joint has integrated load cells (FPG Sensors & Instrumentation, `www.fgp-instrumentation.com`) for measuring torque when turning and lifting the lower arm.

  The analogue sensor signals are acquired with local stand-alone CAN data acquisition modules. The sampling resolution is 10 bit with an adjustable sampling rate from 1000 Hz to 0.2 Hz. The measurement data is available to all connected CAN partners i.e. the PCs and motion control modules.

  This comprehensive system of torque sensors will be used for zero force control of the robot arms as described below. Furthermore, the sensor information may be used to control tactile contact initiated by the robot towards a human agent in a safe and careful way.

- Artificial skin: Advanced human-robot cooperation and interaction is made possible by the information provided by sensor pads made of artificial skin, as developed in ( [9], [10]). Four planar skin pads are mounted to the front and back side of each shoulder, thus also serving as a protective cover for the shoulder joints. Pressure applied to the sensor surface can be measured and localized with the shoulder skin pads. This tactile interface will be used for various purposes, e.g. the human operator may attract the attention of the robot by touching the shoulder or may guide tasks executed by the robot by varying force contact location on a pad. Similarly, cylindrical skin pads are mounted to the upper and lower arms respectively. These skin pads can measure the 3D torque vector that is externally applied to the skin, e.g. by a human grasping the upper arm for guiding the robot.

  The skin sensor information is processed by dedicated controllers and fed to the CAN network of the robot where the data is available to all CAN participants.
- Force/torque sensors in the wrist: For cooperative dual arm manipulating tasks, force/torque information in the wrist is very important. Therefore, dedicated 6D force/torque sensors (ATI Industrial Automation, `www.ati-ia.com`) are used in the wrist.

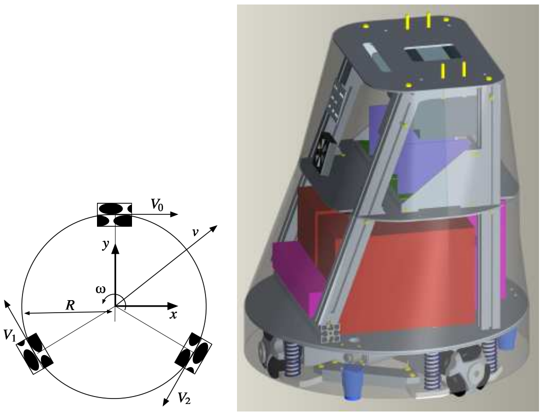

Fig. 4. Kinematics and rendering of the holonomic robot platform: three wheels with passive rolls at the circumference

### C. Platform Specifications and Kinematics

Due to the area of application like household, holonomic movability is a very important issue for flexible use in kitchens or other narrow environments. Since legged walking machines are another wide field of research, which is not to be considered, a wheel-based platform is going to serve for moving the upper body. One way to obtain holonomic flexibility is the use of wheels with passive rolls at the circumference. Such wheels are known as Mecanum wheels or Omniwheels. According to the type of wheel, the rolls are twisted upon 45 or 90 degrees to the wheel axis. To ensure that the platform moves according to the mathematical relations (see [11]), all wheels need to have the same normal force to avoid slackness effects. This needs also to be considered for the robot design and installation of heavy components. The use of only three active drives without any supporting rolls is the best way to guarantee this requirement. These main ideas, combined with other guidelines related to the upper body, result in the following platform specifications:

- Maximum height: 700 mm
- Weight of the upper body: 30 kg
- Translatory speed minimum: 1 m/s
- Holonomic drive
- Power supply (whole system): 8 h with 25% drive
- Spring-damper combination to reduce vibrations

Figure 4 shows the positions of the three wheels, arranged in angles of 120 degrees. For a desired robot movement, the necessary individual speeds are computed as follows: The input variables for the inverse kinematics formulas are translational velocity of the robot $v = \binom{V_x}{V_y}$ as well as angular velocity $\omega$ in respect to its center. The tangential velocities (velocities of the wheel mounting points at the base plate) $V_0, V_1, V_2$ consist of translational and rotational movements and are computed according to [11].

The sensor system of the platform consists of a combination of Laser-range-finders and optical encoders to localize the platform. Three Hokuyo scanners of type URG-X003S (Hokuyo Automatic Co.,Ltd. `www.hokuyo-aut.jp/products/`) are placed at the bottom of the base plate 120 degrees to each other. A scan range of 240 degrees per sensor allows complete observation of the environment. The maximum scan distance of 4 m is enough for use in a kitchen environment. A low scan plane of 60 mm was chosen due to safety reasons to detect small objects and foot tips. Optical encoders deliver a feedback about the actual wheel speeds to the speed control, and serve as a second input, together with the scanner data, to a Kalman-Filter which estimates the position of the platform. The platform hosts the power supply and the main part of the computer network for the entire robot.

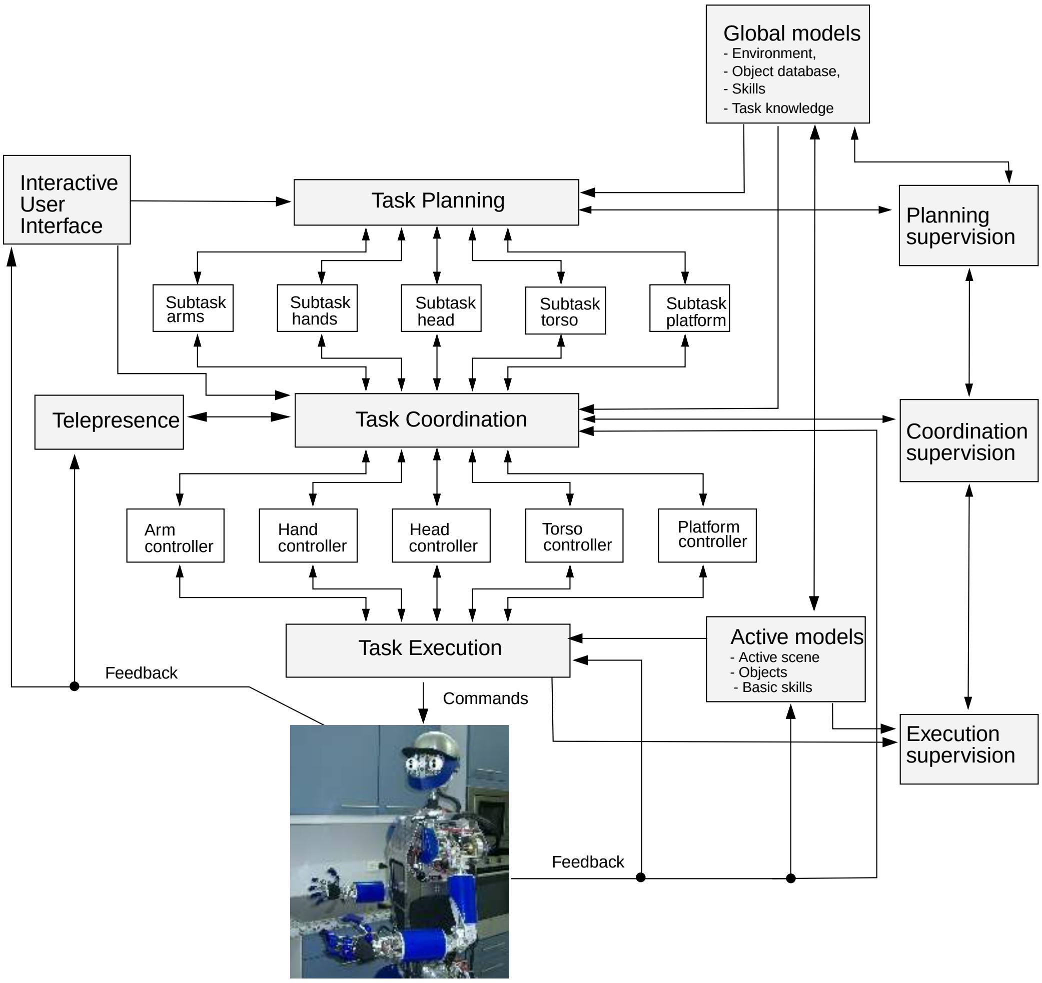

Fig. 5. Hierarchical control architecture for coordinated task execution in humanoid robots: planning, coordination, and execution level.

## III. CONTROL ARCHITECTURE

The control architecture is structured into the three following levels: task planning level, synchronization and coordination level, and sensor-actor level. A given task is decomposed into several subtasks. These represent sequences of actions the subsystems of the humanoid robot must carry out to accomplish the task goal. The coordinated execution of a task requires the scheduling of the subtasks and their synchronization with logical conditions, external and internal events. Figure 5 shows the block diagram of the control architecture with three levels, global and active models and a multimodal user interface [12]:

- The task planning level specifies the subtasks for the multiple subsystems of the robot. This level represents the highest level with functions of task representation and is responsible for the scheduling of tasks and management of resources and skills. It generates the subtasks for the different subsystems of the robot autonomously or interactively by a human operator. The generated subtasks for the lower level contain the whole information necessary for the task execution, e.g. parameters of objects to be manipulated in the task or the 3D information about the environment. According to the task description, the subsystem's controllers are selected here and activated to achieve the given task goal.
- The task coordination level activates sequential/parallel actions for the execution level in order to achieve the given task goal. The subtasks are provided by the task planning level. As it is the case on the planning level the execution of the subtasks in an appropriate schedule can be modified/reorganized by a teleoperator or user via an interactive user interface.
- The task execution level is characterized by control theory to execute specified sensory-motor control commands. This level uses task specific local models of the environment and objects. In the following we refer to those models as *active models*.
- The active models (*short-term memory*) play a central role in this architecture. They are first initialized by the global models (*long-term memory*) and can be updated mainly by the perception system. The novel idea of the active models, as they are suggested here, is the

ability for the independent actualization and reorganization. An active model consists of the internal knowledge representation, interfaces, inputs and outputs for information extraction and optionally active parts for actualization/reorganization (update strategies, correlation with other active models or global models, learning procedure, logical reasoning, etc.).

- Internal system events and execution errors are detected from local sensor data. These events/errors are used as feedback to the task coordination level in order to take appropriate measures. For example, a new alternative execution plan can be generated to react to internal events of the robot subsystems or to environmental stimuli.
- The user interface provides in addition to graphical user interfaces (GUIs) the possibility for interaction using natural language.
- Telepresence techniques allow the operator to supervise and teleoperate the robot and thus to solve exceptions which can arise from various reasons.

## IV. COMPUTER ARCHITECTURE

The computer architecture is built analogously to the control architecture proposed in Section III. This means we had to chose devices for the planning, coordination and execution level. For the first we could meet the requirements both with Industrial PCs and PC/104 systems. The requirements for the execution level could not be met with off-the-shelf products. Thus, we had to develop our own hardware: The Universal Controller Module (UCoM).

### A. *Universal Controller Module (UCoM)*

With the design of the UCoM we followed a modular concept, i.e. the UCoM is always used in combination with a plugon board. This can be a valve driver, a sensor acquisition board or like in ARMAR-III a motor driver board. In combination with the plugonboard 3-way-brushdriver, the UCoM is responsible for the sensory-motor control of the robot. In detail, the UCoM consists of a DSP and a FPGA on one board. By combining the 80 MHz DSP DSP56F803 from Motorola and the 30k gates EPF10k30a from Altera we achieve great reusability. Thus we developed a highly flexible and powerful device with the features given in Table I.

| Parameter | Description |
|---|---|
| Size: | 70 mm × 80 mm × 20 mm |
| Controller: | 80 MHz DSP, Motorola (DSP56F803) |
| Interfaces: | CAN, RS232, SPI, JTAG, 24 digital GPIO, 8 analog inputs |
| Power: | 3 motors at 24 V up to 5 A |
| Current sensing: | Differential measurement for each motor |
| Sensors: | 6 Quadrature Decoder (2 per driven axis) |
| Programming: | via JTAG or CAN-bus |

TABLE I
UNIVERSAL CONTROLLER MODULE (UCOM)

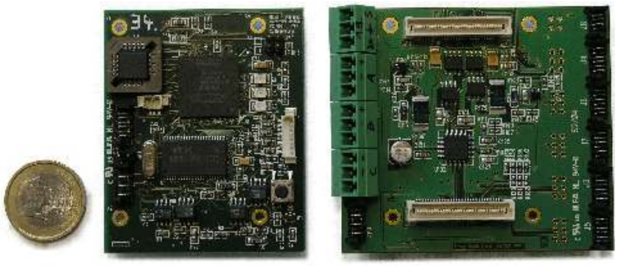

Fig. 6. The Universal Controller Module (UCoM) (left) and the 3-way-brushdriver (right).

On the UCoM, the DSP is connected to the FPGA via the memory interface. Via this interface the DSP can access the 3-way-brushdriver and read the encoder signals prepared by the FPGA. In other words, the distribution of the workload between DSP and FPGA is as follows: the DSP is responsible for calculations of current control variables. The FPGA is some kind of extended general purpose IO port with the ability to do some pre- and post-processing of values.

### B. *PC-Infrastructure and communication*

We use several industrial PCs and PC/104 systems. These PCs are connected via switched Gigabit Ethernet. The connection to the lab PC is established by wireless LAN on the master PC in the platform of the robot. To communicate between the UCoMs and the PC responsible for motion control we use four CAN buses to get real-time operation on the sensory-motor level. An overview over the structure of the computer architecture is given in Figure 7. According to the control architecture in Section III, we use the following components:

- Task planning level: One 1.6 GHz industrial PC. This PC establishes the connection to the lab PCs via wireless LAN and acts as a file server for the other PCs on the robot. Furthermore, it stores the global environment model.

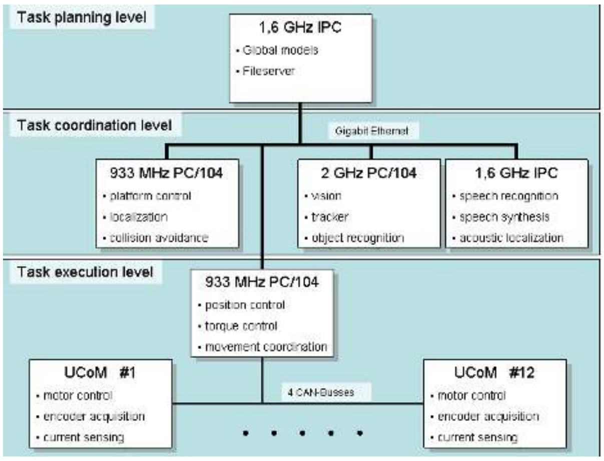

Fig. 7. Computer architecture: The used hardware is based on industrial standards and the developed Universal Controller Module (UCoM).

- Task coordination level: On this level we use one 933 MHz PC/104 system, one 2 GHz PC/104 system and one 1.6 GHz industrial PC. These PCs are responsible to gather sensor information such as camera signals, laser scanner data, force torque values, audio signals etc., and distribute them to the task planning and task execution level.
- Task execution level: On this level one 933 MHz PC/104 system and the UCoMs described above are used. Depending on the task goal issued by the task planning level and the sensor values gathered by the task coordination level the sensory-motor control is accomplished.

### C. *Software Environment*

The computers are running under Linux, kernel 2.6.8 with the Real Time Application Interface RTAI/LXRT-Linux. For the implementation of the control architecture we have used the framework MCA (`www.mca2.org`). It provides a standardized module framework with unified interfaces. The modules can be easily connected into groups to form more complex functionality. These modules and groups can be executed under Linux, RTAI/LXRT-Linux, Windows or Mac OS and communicate beyond operating system borders. Moreover, graphical debugging tools can be connected via TCP/IP to the MCA processes, which visualize the connection structure of the modules and groups. These tools provide access to the interfaces at runtime and a graphical user interface with various input and output entities.

## V. IMPLEMENTED SKILLS

In this section we will present first results related to real-time object localization, and motor control.

### A. *Perception Skills*

To allow the robot to perform the intended tasks in a household environment, it is crucial for the robot to perceive his environment visually. In particular, it must be able to recognize the objects of interest and localize them with a high enough accuracy for grasping. For the objects in the kitchen environment, which we use for testing the robot's skills, we have developed two object recognition and localization systems for two classes of objects: objects that can be segmented globally, and objects exhibiting a sufficient amount of texture, allowing the application of methods using local texture features.

Fig. 8. Typical result of a scene analysis. Left: input image of the left camera. Right: 3D visualization of the recognition and localization result.

Among the first class of objects are the colored plastic dishes, which we chose to simplify the problem of segmentation, in order to concentrate on complicated tasks such as the filling and emptying of the dishwasher. The approach we use is a combination of appearance-based and model-based methods; object models are used to generate a dense and highly accurate set of views by simulating the rotational space of interest. Throughout the recognition and localization process, potential colored regions are segmented and matched in the left and right camera image, rejecting regions outside the area of interest. Remaining regions are then matched based on their appearance in terms of gradients with the entries in the database. By combining stereo vision with the information about the orientation of the object that was stored with its view it is possible to determine the full 6D pose with respect to the object's 3D model at frame rate. An exemplary result of a scene analysis, which we have performed with ARMAR-III in our test environment, is illustrated in Figure 8. A detailed description of our approach is presented in [13]. An integrated grasp planning approach for ARMAR-III and its five-fingered hands, making use of our object recognition and localization system, is presented in [14].

Fig. 9. Scene analysis in a refrigerator: the traces visualize the correspondences of the local features between the learned view and the current view.

Among the second class of objects are textured objects such as tetra packs, boxes with any kind of food, or bottles, as can be found in any kitchen. For these objects, we have developed a system based on local texture features, combining stereo vision, Principal Component Analysis (PCA), a $kd$-tree with best-bin-first search, and a generalized Hough transform [15]. The correspondences between the learned view and the current view in a typical scene in a refrigerator are illustrated in Figure 9.

### B. *Motor Skills*

The execution of manipulation tasks is provided by different inverse kinematics algorithms [16]. This is necessary because most manipulation tasks are specified in terms of the object trajectories. Because of the kinematics redundancy of the arms an infinite number of joint angle trajectories leads to the same end-effector trajectory. We use the redundancy to avoid mechanical joint limits, to minimize the reconfiguration of the arm, and to generate human-like manipulation motions.

To avoid self-collision, the distances between joints of the robot are monitored by the collision avoidance module. The virtual representation of the environment is used to detect possible contacts with obstacles and agents. Joint configurations are only executed if they do not result in a collision.

### *C. Scenarios*

In the two German exhibitions CeBIT and Automatica we could present the currently available skills of ARMAR-III. In addition to the robot's abilities to perceive its environment visually, we also showed how we can communicate with the robot via natural speech. Speech recognition module for large vocabulary continuous speech recognition, 3D face and hand detection and tracking, developed in [17], were integrated and successfully demonstrated

Among the motor-skills we presented were the active tracking of objects with the head, combining neck and eye movements according to [18], basic arm reaching movements, early hand grasping tasks and force-based controlled motion of the platform. All skills were presented in an integrated demonstration.

## VI. CONCLUSION AND FURTHER WORK

We have presented a new 43 DOF humanoid robot consisting of an active head for foveated vision, two arms with five-fingered hands, a torso and a holonomic platform. The robot represents a highly integrated system suitable not only for research on manipulation, sensory-motor coordination and human-robot interaction, but also for real applications in human-centered environments.

The first results we obtained is an encouraging step in the effort toward the realization of different skills in human-centered environments. We believe that perception and action key components in ARMAR-III are advanced enough to define realistic benchmarks and test scenarios, which are representative for our target application area (kitchen).

One of our benchmarks is loading and unloading a dishwasher and a refrigerator with various "things" (tetra packs, bottles with tags, ketchup, beer, cola, etc.) This benchmark sets the highest requirements to perception and action abilities of the robot. Here, we will examine different scientific and technical problems, such as navigation, humanoid manipulation and grasping with a 5-finger hand, object recognition and localization, task coordination as well as multimodal interaction.

## ACKNOWLEDGMENT

The work described in this paper was partially conducted within the German Humanoid Research project SFB588 funded by the German Research Foundation (DFG: Deutsche Forschungsgemeinschaft) and the EU Cognitive Systems project PACO-PLUS (FP6-2004-IST-4-027657) and funded by the European Commission.

## REFERENCES

[1] R. Dillmann, "Teaching and Learning of Robot Tasks via Observation of Human Performance," *Robotics and Autonomous Systems*, vol. 47, no. 2-3, pp. 109–116, 2004.

[2] 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.

[3] J. L. I.W. Park, J.Y. Kim 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.

[4] 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*, 2002, pp. 2478–2483.

[5] K. Nishiwaki, T. Sugihara, S. Kagami, F. Kanehiro, M. Inaba, and H. Inoue, "Design and development of research platform for perception-action integration in humanoid robots: H6," in *IEEE/RSJ International. Conference on Intelligent Robots and Systems*, 2000, pp. 1559–1564.

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

[7] 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.

[8] S. Schulz, C. Pylatiuk, and G. Bretthauer, "A new ultralight anthropomorphic hand," in *IEEE/RSJ International. Conference on Intelligent Robots and Systems*, Seoul, Korea, 2001.

[9] O. Kerpa, D. Osswald, S. Yigit, C. Burghart, and H. Wörn, "Arm-hand-control by tactile sensing for human robot co-operation," *IEEE/RAS International Conference on Humanoid Robots*, 2003.

[10] O. Kerpa, K. Weiss, and H. Wörn, "Development of a flexible tactile sensor for a humanid robot," in *IEEE/RSJ International. Conference on Intelligent Robots and Systems*, Las Vegas, Nevada, Oct. 2003, pp. 1–6.

[11] J. Borenstein, H. R. Everett, and L. Feng, *Where am I? Sensors and Methods for Mobile Robot Positioning*. Ann Arbor, Michigan, USA: University of Michigan, Department of Mechanical Engineering and Applied Mechanics, 1996.

[12] T. Asfour, D. Ly, K. Regenstein, and R. Dillmann, "Coordinated task execution for humanoid robots," in *Experimental Robotics IX*, ser. STAR, Springer Tracts in Advanced Robotics. Springer, 2005.

[13] P. Azad, T. Asfour, and R. Dillmann, "Combining Appearance-based and Model-based Methods for Real-Time Object Recognition and 6D-Localization," in *International Conference on Intelligent Robots and Systems (IROS)*, Beijing, China, 2006.

[14] 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 *International Conference on Intelligent Robots and Systems (IROS)*, Beijing, China, 2006.

[15] K. Welke, P. Azad, and R. Dillmann, "Fast and Robust Feature-based Recognition of Multiple Objects," in *International Conference on Humanoid Robots (Humanoids)*, Genoa, Italy, 2006.

[16] T. Asfour and R. Dillmann, "Human-like Motion of a Humanoid Robot Arm Based on Closed-Form Solution of the Inverse Kinematics Problem." in *IEEE/RSJ International. Conference on Intelligent Robots and Systems*, 2003.

[17] R. Stiefelhagen, C. Fuegen, P. Gieselmann, H. Holzapfel, K. Nickel, and A. Waibel, "Natural human-robot interaction using speech, gaze and gestures," in *IEEE/RSJ International. Conference on Intelligent Robots and Systems*, 2004.

[18] A. Ude, C. Gaskett, and G. Cheng, "Support vector machines and gabor kernels for object recognition on a humanoid with active foveated vision," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2004, pp. 668–673.