NICO – Neuro-Inspired COmpanion: A Developmental Humanoid Robot Platform for Multimodal Interaction

Matthias Kerzel[1], Erik Strahl[1], Sven Magg[1], Nicolás Navarro-Guerrero[1], Stefan Heinrich[1], and Stefan Wermter[1]

Abstract— **Interdisciplinary research, drawing from robotics, artificial intelligence, neuroscience, psychology, and cognitive science, is a cornerstone to advance the state-of-the-art in multimodal human-robot interaction and neuro-cognitive modeling. Research on neuro-cognitive models benefits from the embodiment of these models into physical, humanoid agents that possess complex, human-like sensorimotor capabilities for multimodal interaction with the real world. For this purpose, we develop and introduce NICO (Neuro-Inspired COmpanion), a humanoid developmental robot that fills a gap between necessary sensing and interaction capabilities and flexible design. This combination makes it a novel neuro-cognitive research platform for embodied sensorimotor computational and cognitive models in the context of multimodal interaction as shown in our results.**

I. Introduction

Currently, most robots are still used in industry, where they carry out predefined tasks in specialized environments, often secured behind physical barriers to avoid (potentially harmful) interaction with humans. In contrast, humanoid companions have to be designed to integrate into dynamic domestic environments and thus need to be capable of adapting to novel environments and tasks as well as of learning from experience and human instruction.

These functions all depend on multimodal processing as can be seen in human studies and are actively investigated in research fields such as developmental robotics and embodied cognition. This research requires robotic platforms with rich sensory and motor capabilities. However, these human-like capabilities come at the cost of an increased complexity of hardware and control framework, which often leads to monolithic systems that are difficult to modify and to maintain. Thus, robots often only offer the necessary complexity for one specific task, prohibiting general multimodal experiments in the field of human-robot interaction.

To overcome this challenge, we introduce the NICO (Neuro-Inspired COmpanion)[2], see Figure 1, a novel developmental humanoid that features a flexible, modular design, which is adaptable to individual experimental set-ups. The NICO robot fills an existing gap for developmental humanoid robot platforms that either have the necessary equipment for multimodal research or are optimized for intuitive human-robot interaction.

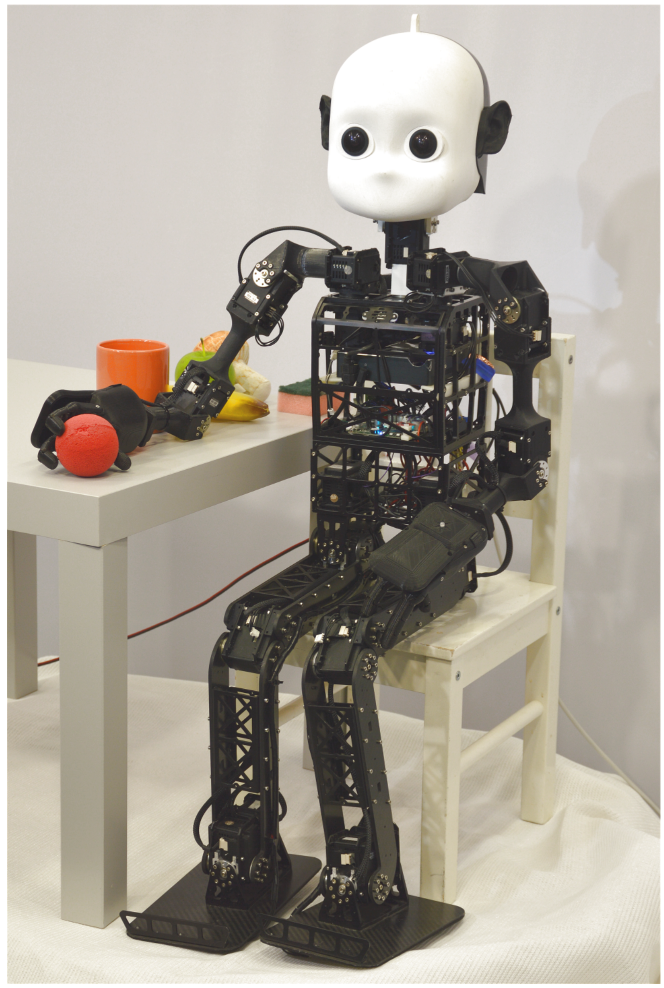

Fig. 1. NICO (Neuro-Inspired COmpanion), a multimodal robot platform.

NICO is not designed to be human-like in every aspect, but to fulfill the requirements as an affordable research platform for multimodal human-robot interaction and embodied neuro-cognitive models. At the same time, its open and modular hardware and software framework ensure adaptability to novel research directions focusing on different modalities. This strategy maximizes the long-term value of the robot platform and ultimately leads to a co-development of robot hardware and neuro-cognitive models.

[1]All authors are with Knowledge Technology, Department of Informatics, Universität Hamburg, Vogt-Kölln-Str. 30, 22527 Hamburg, Germany. `kerzel/strahl/magg/navarro/heinrich/wermter@informatik.uni-hamburg.de`

[2]Visit `http://nico.knowledge-technology.info` for further information and video material.

II. Requirements of Multimodal Research for Developmental Humanoid Robots

Robotic platforms with human-like sensorimotor capabilities are essential for research on neuro-cognitive models and multimodal human-robot interaction. Human-like sensorimotor capabilities allow for the adaptation of neuro-cognitive theories for the bio-inspired development of computational models. Conversely, only a robot with such capabilities can be used to evaluate theories and models about human sensorimotor abilities.

Current research on developmental robotics focuses on instructing or teaching robots to perform tasks instead of programming abilities explicitly [3]. Since humans are experts on how to perform tasks with a human body, it is advantageous that robots have anthropomorphic bodies to teach them in a natural and intuitive manner. Furthermore, a human-like appearance is essential for research on human-robot interaction where the robot displays non-verbal social communication cues like gestures, facial expressions, or body postures. The anthropomorphism of developmental robots also plays an important role in safe human-robot interaction, simply by having human-like dynamics that make their movements more predictable. Finally, service robots benefit from an anthropomorphic form because they are supposed to operate in an environment designed by and for humans.

We developed a list of requirements that a small humanoid developmental robot platform for multimodal interaction should fulfill:

1) Anthropomorphism: For research in bio-inspired multimodal human-robot interaction, it is desired for the robot to have an anatomy that is human-like without eliciting the uncanny-valley effect.

2) Scaled to realistic environment: The robot needs to be sized large enough to perform tasks like grasping, manipulating, and transporting objects, particularly fetching items from standard furniture in realistic environments while still not being too big to impede movement through the environment or create safety problems.

3) Human-like perception: Research on bio-inspired models for vision requires RGB cameras with high resolution and overlapping fields of view for depth estimation with disparity maps. For auditory perception, stereo microphones located at both sides of the head and placed inside realistic pinnae are necessary for sound source localization (SSL). Haptic perception requires proprioception and tactile sensing. Depending on the experimental design different distributions of tactile sensors should be possible.

4) Bimanual manipulation: The robot needs to be able to perform bimanual manipulation and haptic exploration of objects which could be handled by a child in terms of size and weight. It needs to be able to touch, grasp, lift, move around, and turn these objects.

5) Locomotion: Walking (and sometimes crawling) are preferred ways of locomotion for humanoid robots. As walking locomotion is often slow and unstable, wheeled locomotion can be an efficient complementary alternative for moving a robot, in cases where walking and a humanoid lower body is not in the focus of the research.

6) Modular open source and hardware design: The robot design has to be highly flexible to adapt to changing experimental constraints. Thus, a modular design in both software and hardware is essential, to be able to replace and upgrade components without compromising the overall functionality of the robot and hence ensuring its long-term value for a research group.

7) Initial and maintenance cost: Finally, the initial and expected maintenance costs of a robot platform are an important factor for wider dissemination to other research, private, or industrial settings.

III. Available Humanoid Robots

Due to the number of purchasable robots and prototypes, only examples are listed to show the limitations and advantages of different robot classes. Overall, humanoid robots can be grouped roughly into three categories. 1) Semi-humanoids or service robots often have a human-like upper body and are mostly endowed with wheeled locomotion. 2) Small humanoids have a size of about 50cm. 3) Medium-size humanoids have a size of up to approx. 150cm height.

A. Semi-Humanoid Robots

Owing to the inherent stability issues of bipedal locomotion, semi-humanoids are usually developed for stationary use or offered with a wheeled base. Besides stability, wheeled platforms can also offer a higher payload compared to bipedal robots and are suited for a large range of indoor research and service applications.

PR2 from Willow Garage is a wheeled robot with an adjustable height between 133cm and 164cm. It has two arms with seven degrees of freedom (DoF) endowed with a gripper with tactile sensors. The Care-O-bot-4 and its predecessors from the Fraunhofer Institute are wheeled robot platforms that are equipped with up to two manipulators but no discernible head. Instead, the Care-O-bot-4 offers a tablet screen for interaction. To provide adaptability for different scenarios ranging from domestic assistance to hospitals, the platform is configurable from a set of standardized components [20]. Both robots are designed to work with humans but without mimicking human appearance. They offer a broad range of sensors including high-resolution cameras, depth sensors, and microphones and can be controlled via ROS [18]. All in all, these features make them attractive platforms for neuro-cognitive research. However, their high price, 250,000–400,000 USD, have hindered a wide distribution.

Baxter from Rethink Robotics is originally a non-mobile robot torso with two industrial arms and a tablet for displaying its face. It is designed for industrial tasks involving manipulation of objects but has also gained interest in the research community due to its safety enhancing compliance and low cost [10]. It is also compatible with ROS and has a price of about 25,000 USD. In comparison, the Pepper robot from Aldebaran was particularly designed for human-robot interaction tasks with current applications focusing largely on emotion recognition and display of prototypical body postures associated with human emotions. It has a height of 120cm and a human-like torso that is fitted onto a wheeled platform. Pepper has overall 20 DoF and human-like arms with five-finger hands with fingers designed mainly as means for physical expression rather than for grasping. The Pepper has a price of ca. 15,000 USD including a monthly fee. The largest advantage of Baxter and Pepper are their affordability which can be explained by two main factors: The limited number of sensory and behavioral capabilities, and the fact that they run to a great extent on proprietary software making extensions challenging to integrate.

B. Small Humanoids

NAO from Aldebaran is one of the most commonly used small humanoid robots in developmental and multimodal research, e.g. [17], [8]. NAO features a large number of sensors and degrees of freedom for its size of ca. 57cm [6]. Its main drawbacks are the lack of stereo vision, the poor quality of the microphones, and the high ego-noise, which complicate its applicability to tasks that require e.g. speech recognition. NAO has a current price of about 10,000 USD and, like Pepper runs on proprietary software.

The DARwIn-OP from Robotis is a small humanoid robot that was primarily developed for robot soccer [7]. The DARwIn-OP has 20 DoF and is powered by Robotis Dynamixel[3] servomotors that can be controlled by an open and easy-to-access framework, overcoming the limitations of Baxter, Pepper, and NAO. DARwIn-OP also has an open hardware design thus it can be more easily maintained and extended. DARwIn-OP comes for a price similar to the NAO.

C. Medium-size Humanoids

iCub is a developmental humanoid that resembles a 3.5-year-old child with its body height of 104cm [15], [16]. It features 53 DoF, including five-finger hands and actuated eyes for human-like gaze shifts with its two cameras. Optional sensory additions for the iCub such as tactile sensors are available. Although having many features that make the robot interesting for multimodal research, the hardware of the iCub is optimized for its purpose and has a holistic design, which does not allow for an easy extension with alternative multimodal sensors [16]. The basic version of the iCub has a price of about 250,000 EUR.

Addressing the issues of modularity and cost-efficiency in medium-sized humanoids, the NimbRo-OP and the Poppy robots have been suggested. Both robots are powered by Robotis Dynamixel servomotors and thus share a similar mechanical design. On the one hand, the NimbRo-OP robot from the AIS group of the University of Bonn is designed as an open robot for the RoboCup TeenSize league [21]. Thus, many features useful for human-robot interaction (HRI) research are left out in its design, such as arms and hands with high numbers of DoF. The open design of the robot, however, allows the development and integration of extensions for HRI research. A major limitation of NimbRo-OP is its custom made software framework optimized for robot soccer which, to our knowledge, has not been used outside the original group. Currently, the NimbRo-OP has an overall cost of ca. 25,000 EUR. On the other hand, the Poppy robot is a humanoid focused on bipedal locomotion and full-body physical human-robot interaction [13]. Like NimbRo-OP, Poppy has non-actuated hands and the amount, as well as the type of available sensors in the current design, are not sufficient for multimodal interaction research. The Poppy robot project has developed a general purpose library (PyPot library[4]) to control Robotis Dynamixel servomotors, which facilitates the development of any project including such motors.

D. A Gap in the State-of-the-Art for Humanoids

Comparing the available robot platforms against the requirements established above, it becomes apparent that there is a gap for flexible and modular humanoids with rich sensorimotor capabilities for multimodal human-robot interaction. The requirements that we have defined in Section II are distributed over the three robot categories, with no platform being able to fulfill more than a few at once.

Among the semi-humanoids, PR2, Care-O-bot-4 and Baxter (with mobile base) are large enough and offer robust locomotion, but their design is not anthropomorphic, and they rely on non-human perception for most tasks. Due to their large size, especially PR2 and Baxter have an intimidating effect which may negatively affect natural human interaction. Although being able to manipulate objects with haptic feedback, the manipulators are based on industrial designs and do not allow for fine-grained, dexterous manipulation that can be compared to a human-like hand, making it difficult to compare robotic solutions to human manipulation. The Pepper, in contrast, has a human-like appearance but is not designed for object manipulation. All of these platforms are also difficult to modify due to their complex or closed source designs.

Small humanoids ($\sim$ 50cm) in comparison are not intimidating and invite natural interaction due to their anthropomorphic, child-like appearance. However, due to their size and the simplicity of their grippers, they have limited object manipulation capabilities. Even with the optional gripper in the case of the DARwIn-OP, the robots are too small and light to lift standard-sized objects and easily fell over in our trials when walking while carrying objects, thus violating the requirement for a sensible scale, locomotion, and bimanual manipulation. Additionally, they offer insufficient sensors for many multimodal tasks and cannot be easily extended due to their size or holistic design.

Teen-sized robots offer a good trade-off between advantageous child-like appearance and necessary size for realistic tasks. iCub offers many of the required qualities like an anthropomorphic design, dexterous bimanual manipulation, and human-like perception, but is complex, not easily modifiable, and expensive. NimbRo-OP and Poppy allow for this

[3] http://en.robotis.com/index/

[4] https://github.com/poppy-project/pypot

adaptability with respect to the robotic hardware and the control software, and are available for a much lower price, but they had to forgo sensor capabilities, focus on walking, and arms with grippers or hands, losing again most features that make the iCub so interesting for multimodal manipulation research.

There is a clear gap between necessary sensor and manipulation capabilities, on the one hand, that allow for experiments involving human-like perception and manipulation, and an affordable anthropomorphic and flexible design that, on the other hand, allows for intuitive and non-intimidating human-robot interaction. Exploiting the advantages of the affordable teen-sized robots in terms of modular design and open control framework, we propose the NICO platform that can bridge the gap and combine the advantages of both sides to offer a teen-sized humanoid with advanced sensorimotor capabilities, especially focussed on object manipulation and haptic perception.

IV. NICO, A DEVELOPMENTAL ROBOT FOR MULTIMODAL INTERACTION

The NICO robot is designed as a highly flexible and modular platform for multimodal human-robot interaction and neuro-cognitive modeling.

NICO is developed on the basis of the first version of NimbRo-OP (see Section III-C) and likewise, features a humanoid anatomy resembling a child.

However, the focus for the NICO is on human-like perception and interaction as well as object grasping and manipulation, instead of bipedal walking research.

With a height of 101cm and a weight of 7kg, NICO is constructed mainly from Robotis Dynamixel servomotors and 3d-printed parts, which adds to its high flexibility and simple maintenance.

Redesign mainly took place in the upper body including a) the neck and the head, with a strong emphasis on the audio and visual sensory modalities and a display for facial expressions as well as b) the shoulders, the arms, and the hands, with a focus on advanced manipulation. In total, the robot has 30 DoF which are distributed as follows. Two DoF can perform yaw and pitch movements of the head. The shoulders feature a tight cluster of motors that offer three DoF, mimicking the physiology of the human shoulder ball joint, while an additional degree of freedom allows bending at the elbows.

Seed Robotics SR-DH4D articulated hands[5] add four DoFs for wrist rotation, wrist flexion, controlling the two index fingers, and controlling the opposable thumb. Each leg includes three DoF in the hip joint, one DoF in the knee and two DoF in the foot. Due to its modular design, the legs can be removed at the hip, if not required, and the torso can be mounted on a wheeled platform.

A. Head, Facial Expressions, Visual and Auditory Modalities

The outer shell of the robot's head is based on the open-source design of iCub because the form and facial design recreate a child-like appearance with pleasant and symmetrical features as shown in Figure 2 (left). This design avoids the uncanny-valley effect while still facilitating intuitive human-robot interaction, especially for scenarios where a human adopts the role of a parent-like teacher.

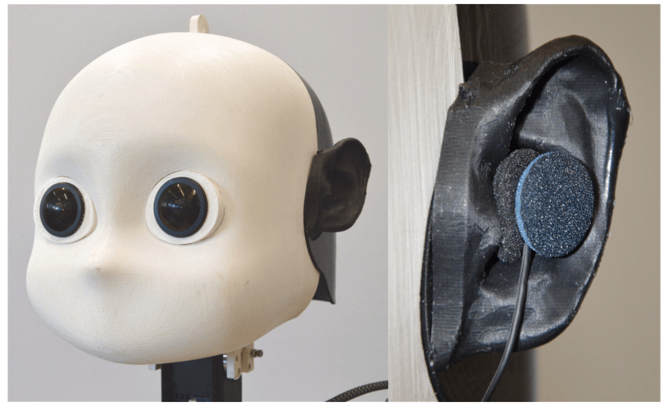

Fig. 2. Left: NICO head based on the iCub head. It includes two cameras with an overlapping field of view and two stereo microphones mounted on realistic 3D-printed pinnae. Right: close up of one pinna.

For the visual modality, NICO's head houses two parallel camera mounts with exchangeable lenses and can be equipped with a variety of cost-effective or state-of-the-art electronics such as the 2 Megapixel sensor of a Logitech C905[6]. Additionally, the mounts include holders for the optics that enable to adapt the optics to different scenarios. This way the NICO can e.g. include wide-angle optics with a field of view of 180 degrees for research in RoboCup football, or lenses that mimic a human-like field of view (70°) for research in bio-inspired perception models.

For the auditory modality, NICO's head is endowed with two Soundman OKM II binaural microphones embedded in realistically shaped and 3D-printed pinnae as shown in Figure 2 (right). This design is particularly convenient to study human-like binaural hearing for sound source localization (SSL) and speech recognition. Specifically for SSL, the interaural time and level differences for sounds perceived by both ears are essential. Thus, the location of the microphones, the dampening factor of the head, and also the form of the pinnae (necessary for vertical SSL) have been carefully designed close to human-child anatomy for providing a realistic distortion of the sounds. In contrast to other humanoids such as iCub, NICO's head does not include any moving part, computing unit, or cooling device. This minimizes the ego-noise that can superimpose to frequency components that are important for speech recognition.

NICO's head is also fitted with three LED arrays in the mouth and eye areas. These arrays enable the NICO robot to show stylized facial expressions during human-robot interaction, see Figure 7 for examples.

[5]http://www.seedrobotics.com/

[6]http://www.logitech.com/video/webcams/

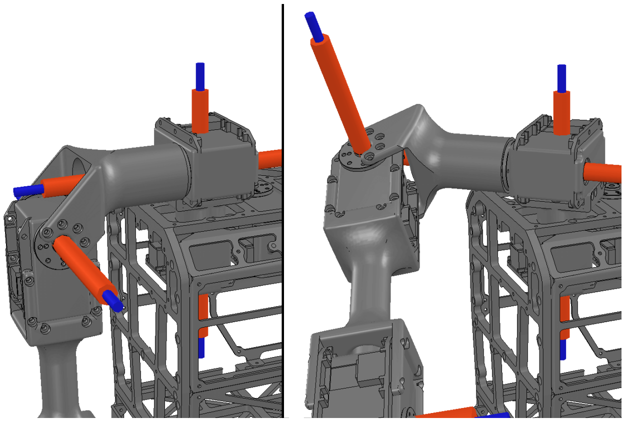

Fig. 3. Shoulder and upper arm of NICO. Three degrees of freedom in the shoulder area mimic the function of a human shoulder ball joint.

B. Arms, Hands, and Haptic Sensors

The open design of NimbRo-OP allowed redesigning the torso and shoulder area including a modified placement of motors inside the torso, increasing the robot's working space to introduce bimanual manipulation in front of the body. Figure 3 shows our new design of the upper arm and shoulder area.

For the hands, we integrated the three-fingered SR-DH4D hands from Seed Robotics into the NICO, see Figure 4 (left). Their tendon-based mechanics with force feedback allow for segmented, bendable fingers with three joints per finger and thus for seamless handling of objects with different shapes. The two index fingers are controlled together as a single DoF as is the opposable thumb, which reduces the complexity of the grasping or object manipulation while still preserving the advantages of a fingered hand over a rigid gripper. The hands use the Dynamixel bus system and protocol and can, therefore, be seamlessly integrated into the control framework of the NICO robot. For multimodal human-robot interaction, the human-like kinematics of the arm and hand provides the added advantage of enabling the robot to learn grasping and gestures from human demonstration.

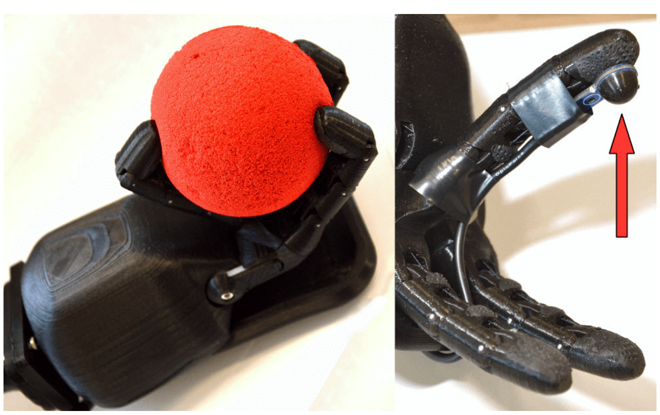

Fig. 4. Left: Three-fingered Seed Robotics SR-DH4D articulated hand on the NICO robot, grasping a foam ball. Right: OPTOFORCE OMD-10-S-10N 3D force sensor mounted on a finger.

In humans, the haptic sensory modality subsumes two distinct sensory subsystems [14]: 1) The tactile modality is perceived via sensory cells in the skin that react to deformation, vibration, and temperature. 2) Proprioception provides information about the body posture, movement and exerted forces. The integration of haptic submodalities is required for haptic exploratory procedures such as lateral motions for texture information, unsupported holding for weight information, or contour following for shape information.

For tactile sensing, we integrated OPTOFORCE OMD-10-SE-10N[7] force sensors, see Figure 4 (right). This sensor consists of a rubber dome that can measure pressure and sheer forces of up to 10N in three dimensions and has a temporal resolution of up to 1,000Hz. Overall this sensor enables deformation and vibration sensing at fingertips that are vital for gaining information about handled materials. Deformation forces normal to the material surface give information about the compliance of a material while forces along the movement direction during lateral motions reveal static friction. Once the static friction is overcome, both low and high-frequency vibrations can be observed, giving information about regularities in the texture, e.g. a corrugated surface and sliding friction respectively. Combining information about these properties can accomplish discrimination of various materials and adds a novel modality to the NICO which can, for instance, be used to learn an embodied understanding of haptic adjectives [4].

Proprioceptive information such as torque from the motors can be combined with tactile information from the OPTOFORCE OMD-10-SE-10N sensors on the fingertips to provide a haptic sensory modality for the NICO robot.

C. Modular Open Source Software Framework

Similar to our approach of utilizing existing hardware designs, we build upon existing open source and modular software frameworks to create the NICO API to support multimodal interaction research. To synergy effects in open-source development over several platforms, the NICO API acts as a wrapper of different frameworks. It standardizes function calls and provides high-level functionalities which are internally built from different libraries. This function call standardization is inspired by the closed source NAOqi API.

Our software framework is developed in Python due to its ideal trade-off between ease of use and computational efficiency. Further, Python allows for an easy integration with and from other existing frameworks. The low-level motor control of the NICO API is based on PyPot [13] which we extended to support the Seed Robotics hands described earlier. Perception capabilities are based on existing Python libraries such as OpenCV for the cameras and pyaudio for microphones. Facial expressions are provided by a Python-wrapper for the Arduino-controlled LED array. This API also makes it easy to control the real and simulated robot by simply switching a flag. To further increase the flexibility and modularity of the NICO API we also created a ROS [18]

[7] `http://optoforce.com/3dsensor/`

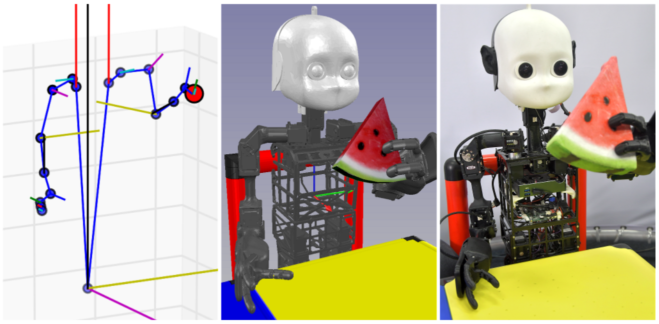

Fig. 5. NICO as a kinematic model, in V-REP, and in real.

interface which offers the same functionality and enables streaming of sensory data such as stereo images and stereo audio not available in the standard Python API.

The use of Python and the integration of libraries that are actively developed in the robotic community has been crucial to expedite development of the NICO robot and to connect it to efforts of a wider community.

D. NICO in Virtual Robotic Environments

Virtual robotic environments enable extended and controlled experiments and learning phases for developmental robots. Repetitive tasks can be trained with reduced effort and strain to the robotic hardware before they are embedded in physical robots. In some cases, certain movements can be dangerous for the human or for the robot itself and thus such behavior should be tested in a safe virtual environment. Simplifications of the robot model, like a reduced model of the kinematic chain, can drastically reduce the learning time for the algorithms as the calculations are much cheaper than a full physics simulation. To enable this, we use three modes of operation:

- Calculating only the kinematic chain of arms and legs
- 3D-Physics simulation of the full robot including involved masses, resulting forces, gravity, and friction
- Operation on the real robot

The virtual realization of the NICO is based on the Unified Robot Description Format (URDF)[8] (see Figure 5). This specification contains information about the kinematics of the robot as well as its collision model and visual representation. The URDF file contains all information to build up the kinematic chains of the robot for lowest abstraction level and in the next step to build the full physics model in the robot simulator.

The NICO-API comes with integrated support for the robot simulator V-REP [19]. This is inherited from the PyPot framework, which the NICO-API is based on, and leads to a comfortable switching capability from simulated to real robot, with nearly no effort.

To have a high interchangeability between the real world and the simulator, not only the robot, but also the laboratory environment has been re-created in the simulator. To simplify this process, we used standard building blocks for children environments. For working in a ROS environment, the Gazebo [12] simulator can be used as well utilizing the same URDF model of NICO.

[8] `http://www.ros.org/wiki/urdf`

V. Results

To show the capability of NICO, we present three studies that are exemplary for our ongoing evaluation. We evaluated the ability of the haptic sensors to mimic human haptic perception, the ability of NICO to display and recognize emotion expressions, and the effect of the emotional display on how humans subjectively perceive NICO.

A. Haptic Material Classification

First, we evaluated the material classification performance of the novel haptic sensors of NICO [11]. In this study, the sensor was used to perform two haptic exploratory procedures, pressing down and lateral motion. The haptic sensor was moved by a robotic arm employing the same motors and low-level control framework as NICO, see Figure 6. Samples were collected from 32 different materials including metal, glass, different types of paper, different fabrics, wood, leather, and plastic. The data was used to train a multi-channel convolutional neural network for classification of two general haptic properties, compliance and texture, as well as classification of specific materials. Using the combined information of both exploratory procedures the classification accuracy for the two material properties reached 99.5% and 99.4% while the accuracy for classifying specific materials reached 98.8%. Though the high accuracy was partly attributed to the controlled laboratory conditions during sampling, these results demonstrate that the optical force sensor of NICO is suitable for human-level haptic perception with regard to material classification. The study also showed the successful integration of NICO's haptic sensors and low-level motor control with a neural network architecture.

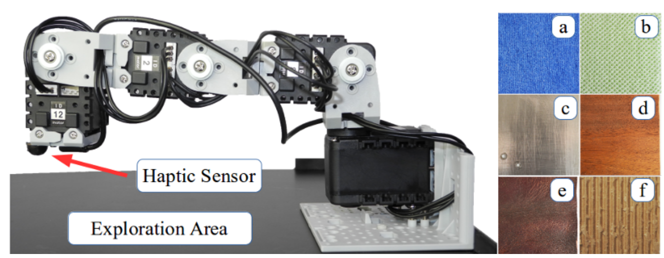

Fig. 6. Left: Robot arm with haptic sensor. Right: Material samples: carpet (a), cloth (b), steel (c), wood (d), leather (e), and cardboard (f).

B. Learning Emotion Expressions

We also conducted a study on teaching emotion expressions to NICO using deep neural architectures [5]. First, we evaluated how well humans can recognize the seven emotional expressions of NICO, see Figure 7. Results showed that all 20 participants (7 female, 13 male) could identify a subset of five expressions (neutral, happiness, sadness, surprise, and anger) with an accuracy $\geq 75\%$. In the second experiment,

Fig. 7. NICO is displaying facial expression for happiness, anger, and sadness using LED arrays in the mouth and eye areas.

10 participants trained NICO to recognize and replicate their emotional expressions. For this, the participants expressed one of the five emotions mentioned above to NICO. Recorded images from the robot camera were analyzed by a neural network and linked with an action representation, i.e. the robot would mirror back an emotion expression. This, in turn, was rewarded by affirmative or negative emotion expression by the participants to facilitate learning. Depending on the experimental condition, the neural network learned to mirror 4 out of 5 emotion expressions correctly.

This study demonstrates the effectiveness of NICO's facial expression display and also shows that NICO is suitable for HRI teacher-learner scenarios.

C. *Effect of Emotion Display on Subjective User Rating*

To evaluate how NICO is perceived by human interaction partners, we conducted a user study using the established Godspeed test [1]. The test encompasses 24 items of semantic differential scales, where participants are asked to state their opinion on a five-point scale between two bipolar adjectives, e.g. *fake* and *natural*. The items are organized into five categories: anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety.

All 20 participants (7 female, 13 male) of the study were between 19 and 49 years old and reported English speaking abilities on a conversational level or better. The participants came from eleven different countries, including countries in Europe, Asia, South America and the Middle East. In the fist part of the experiment, participants were seated at a table with a non-moving NICO. They were asked to report their initial impression of the robot using the Godspeed test. In the second part of the experiment, users saw NICO using its facial display to express different emotions, which the participants were asked to recognize. Participants were informed that the emotional display is controlled by a human operator. Afterward, the participants were asked to fill the Godspeed questionnaire again. During this time, a happy face was displayed on NICO to prevent priming effects due to the last displayed emotion.

Figure 8 shows the aggregated results of both questionnaires. In the first condition, NICO was rated positively with regard to likeability (3.6), perceived safety (3.6) and slightly positive about its perceived intelligence (3.1). It was, however, only rated neutral concerning its anthropomorphism (2.6) and slightly negative with regard to its animacy (2.2), which was expected in the static experimental condition. Our

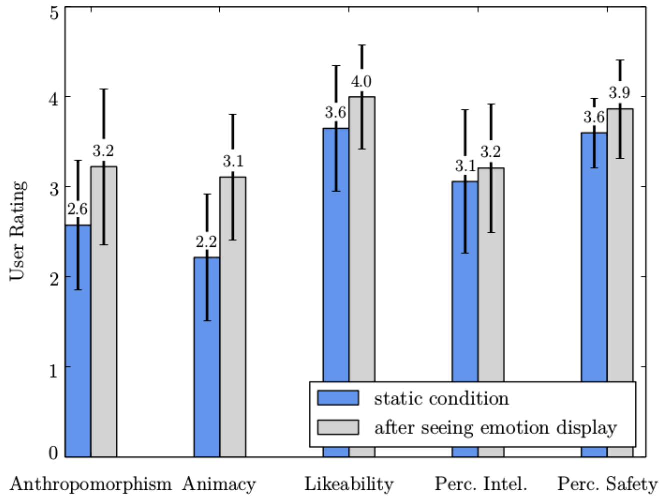

Fig. 8. Results of the Godspeed questionnaire before and after seeing the emotion display of NICO. High ratings indicate a positive subjective impression of the participants. Seeing the emotion display results in overall higher ratings in all five categories of the test: anthropomorphism, animacy, likeability, perceived intelligence and perceived safety.

hypothesis was that NICO would be rated more positive after users have seen it display emotions; p-values were determined with a one-tailed, paired t-test. The results of the second experiment not only show a significantly more positive evaluation regarding anthropomorphism ($+0.6, p < 0.001$), animacy ($+0.9, p < 0.001$), and likeability ($+0.4, p < 0.001$) but also a slightly more positive assessment of perceived safety ($+0.3, p = 0.018$). Only the rating of perceived intelligence ($+0.1, p = 0.105$) was mostly unaffected.

VI. Conclusion

NICO is a novel multimodal humanoid robot platform with rich sensorimotor capabilities. It was designed for research in the areas of neuro-cognitive modeling and multimodal human-robot interaction, aiming to fill a gap between advanced, multimodal sensor capabilities and affordable humanoid appearance for intuitive human-robot interaction. The result is an open, flexible, and affordable humanoid that has the necessary equipment for state-of-the-art research in multimodal perception and natural interaction in realistic environments.

Our current research focuses on a thorough end-to-end evaluation of the sensorimotor and multimodal interaction capabilities of the NICO. We employ different evaluation environments, ranging from virtual simulations over virtual-reality environments with 180-degree image projection and loudspeaker arrays [2] to special laboratories that recreate realistic home environments. NICO will be used to research crossmodal integration of sensory information in neuro-cognitive models. A typical scenario is learning embodied representations of word meanings from the multimodal interaction with small objects while receiving spoken instructions from a human teacher [9]. In such a scenario aspects of natural language, for example, words like *soft*, *heavy*, *blue*, or *rattling*, are grounded in sensory experiences of the robot, thus fostering crossmodal human-robot interaction based on a common embodied understanding of the world.

In this paper, we have outlined the necessary robot requirements for such experiments in section II: NICO has an anthropomorphic design, with a child-like appearance that aids natural interaction between robot and humans. To enhance human-robot interaction the robot can display stylized facial expressions and, despite its child-like size, is still large enough to operate in realistic domestic environments, particularly comparable to that of developing children. NICO features cameras for stereo vision and microphones embedded in realistic pinnae for SSL mounted on a humanoid head at biologically correct positions. 3D force sensors in the fingertips for tactile perception and motor feedback currents from all main joints and fingers provide high-quality proprioceptive feedback. The kinematics of the arms is human-like, and with its three-fingered hands it can grasp, manipulate, and haptically explore objects. NICO is capable of legged locomotion, but can easily be modified for torso-only setups or wheeled locomotion. The hardware and the control framework are highly modular and flexible, making it easy to integrate new sensors and other hardware. Due to the use of URDF models and their integration in the control framework, the switch between kinematic calculations, 3D physics simulation, and the real robot is easy and fully supported in the software. The adaptability and parallel use of different levels of abstraction increase the long-term value of the robot, especially for neurocognitive approaches where long simulated training phases are followed by testing on the real hardware.

Individual modules of the robot are evaluated on different benchmarks. The motor capabilities are assessed on object manipulation and fetch-and-follow tasks. Vision sensors are tested on object identification and localization tasks. Auditory sensors are evaluated on their ability for binaural sound source localization in a virtual-reality environment. The haptic sensors are tested on object and material classification task. Finally, the facial expression display is evaluated in HRI user studies. Initial results have been reported and show very promising results for both haptic and multimodal experiments as well as human acceptance of the robot as interaction partner. Overall, NICO can be used in different experimental setups and research areas, providing state-of-the-art capabilities. NICO provides a platform that can be continuously improved and evaluated and thus is fit for long-term research in multimodal human-robot interaction.

Acknowledgment

This work was partially funded by the German Research Foundation (DFG) in project Crossmodal Learning (TRR-169) and the Hamburg Landesforschungsförderungsprojekt CROSS.

We want to thank the NimbRo-OP team from the AIS group of the University of Bonn as well as the iCub team from the Istituto Italiano di Tecnologia (Genova, Italy) for the fruitful discussions.

We also like to thank our student assistants for their valuable contributions to NICO.

References

[1] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, 2009.

[2] J. Bauer, J. Dávila-Chacón, E. Strahl, and S. Wermter. Smoke and mirrors–virtual realities for sensor fusion experiments in biomimetic robotics. In *Proc. IEEE Conf. on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 114–119. IEEE, 2012.

[3] A. Cangelosi and M. Schlesinger. *Developmental robotics: From babies to robots*. The MIT Press, Cambridge, US, 2015.

[4] V. Chu, I. McMahon, L. Riano, C. G. McDonald, Q. He, J. M. Perez-Tejada, M. Arrigo, N. Fitter, J. C. Nappo, T. Darrell, and K. J. Kuchenbecker. Using robotic exploratory procedures to learn the meaning of haptic adjectives. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 3048–3055, 2013.

[5] N. Churamani, M. Kerzel, E. Strahl, P. Barros, and S. Wermter. Teaching emotion expressions to a human companion robot using deep neural architectures. In *Proc. Int. Joint Conf. on Neural Networks (IJCNN)*, pages 627–634. IEEE, 2017.

[6] D. Gouaillier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourcade, B. Marnier, J. Serre, and B. Maisonnier. Mechatronic design of NAO humanoid. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 769–774, 2009.

[7] I. Ha, Y. Tamura, H. Asama, J. Han, and D. W. Hong. Development of open humanoid platform DARwIn-OP. In *Proc. SICE Annual Conf. (SICE)*, pages 2178–2181. IEEE, 2011.

[8] S. Heinrich. *Natural Language Acquisition in Recurrent Neural Architectures*. Dissertation, Department of Informatics, Universität Hamburg, 2016.

[9] S. Heinrich, C. Weber, S. Wermter, R. Xie, Y. Lin, and Z. Liu. Crossmodal language grounding, learning, and teaching. In *Proc. NIPS2016 Workshop on Cognitive Computation (CoCo@NIPS2016)*, pages 62–68, Dec 2016.

[10] Z. Ju, C. Yang, and H. Ma. Kinematics Modeling and Experimental Verification of Baxter Robot. In *Proc. Chinese Control Conf. (CCC)*, pages 8518–8523. IEEE, 2014.

[11] M. Kerzel, M. Ali, H. G. Ng, and S. Wermter. Haptic material classification with a multi-channel neural network. In *Proc. Int. Joint Conf. on Neural Networks (IJCNN)*, pages 439–446. IEEE, 2017.

[12] N. Koenig and A. Howard. Design and Use Paradigms for Gazebo, An Open-Source Multi-Robot Simulator. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 2149–2154, 2004.

[13] M. Lapeyre, P. Rouanet, J. Grizou, S. N'Guyen, A. Le Falher, F. Depraetre, and P.-Y. Oudeyer. Poppy: Open Source 3D Printed Robot for Experiments in Developmental Robotics. In *Proc. Joint IEEE Int. Conf. on Development and Learning and Epigenetic Robotics (ICDL-Epirob)*, pages 173–174, 2014.

[14] S. J. Lederman and R. L. Klatzky. Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19(3):342–368, 1987.

[15] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, A. Bernardino, and L. Montesano. The iCub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks*, 23(8–9):1125–1134, 2010.

[16] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori. The iCub humanoid robot: An open platform for research in embodied cognition. In *Proc. Workshop on Performance Metrics for Intelligent Systems*, PerMIS, pages 50–56. ACM, 2008.

[17] N. Navarro-Guerrero. *Neurocomputational Mechanisms for Adaptive Self-Preservative Robot Behaviour*. Dissertation, Department of Informatics, Universität Hamburg, 2016.

[18] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng. ROS: An open-source Robot Operating System. In *Proc. ICRA Workshop on Open Source Software*, 2009.

[19] E. Rohmer, S. P. Singh, and M. Freese. V-REP: A versatile and scalable robot simulation framework. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 1321–1326, 2013.

[20] J. Schäfer, S. Feustel, and R. Kittmann. Ein Bild von einem Roboter: Der Care-O-bot 4 als Gentleman. In *Mensch und Computer – Usability Professionals*, pages 144–151. De Gruyter, 2015.

[21] M. Schwarz, J. Pastrana, P. Allgeuer, M. Schreiber, S. Schueller, M. Missura, and S. Behnke. Humanoid TeenSize Open Platform NimbRo-OP. In *Proc. Workshop on Humanoid Soccer Robots at IEEE-RAS Humanoids*, pages 568–575, 2014.