

Elements of Cognitive Robotics

Christian Goerick, Edgar Körner

2007

Preprint:

This is an accepted article published in it - Information Technology. The final authenticated version is available online at: [https://doi.org/\[DOI not available\]](https://doi.org/[DOI not available])



Elements of Cognitive Robotics

Elemente kognitiver Robotik

Christian Goerick, Edgar Körner, Honda Research Institute Europe GmbH

Summary In this paper we present our research aiming at creating a cognitive humanoid. We describe our understanding of the core elements of a processing architecture for such kind of an artifact. After these conceptual considerations we present our research results in the areas of task and body oriented motion control as well as visually and behaviorally oriented learning. ▶▶▶ **Zusammenfassung** In diesem Beitrag

präsentieren wir unsere Forschungsaktivitäten aus dem Bereich der kognitiven Robotik. Zunächst werden einige grundlegende Konzepte bezüglich der relevanten Elemente und deren Zusammenwirken in einer Architektur erörtert. Danach präsentieren wir konkrete Forschungsergebnisse aus dem Bereich der aufgaben- und körperorientierten Bewegungskontrolle sowie aus dem Bereich des visuellen und verhaltensorientierten Lernens.

KEYWORDS J.2 [Computer Applications: Physical Sciences and Engineering: Engineering] C.0 [General: System Architectures], cognitive robotics, humanoid robot, whole body control, visual object learning, behavior learning

1 Introduction

Service robotics is a well established field in the scientific research community. The possible applications are manifold, because the need for reliable and enduring service provision is continuously rising in the modern society. Especially humanoid robots are due to their “anatomy” in principle well suited for providing such kind of services in an anthropocentric environment. But based on this mechanical fundament especially cognitive abilities are necessary for providing the targeted services. A truly assisting service robot should be able to act autonomously and flexibly in a natural environment, learn about new things and situations in interaction or autonomously and apply the gathered knowledge under varying constraints. In this paper we report on our current research efforts towards cognitive robotics: the researched elements and the endeavors aiming at a brain-like control architecture for humanoid robots.

2 Towards an Architecture

The long term goal of our research is establishing a cognitive architecture for controlling humanoid robots. We are convinced that cognitive or intelligent performances of artifacts can only be achieved within an architecture orchestrating the individual elements in a phenomenologically coherent way. What those elements exactly are and how they are to be arranged is subject to current research. Nevertheless, it is understood that there is a minimal subset of those elements that have to be addressed. The analysis of biological systems teaches us about those principled elements and their possible role within animals’ brains [1]. We consider the following elements to belong to this subset:

- Sensory perception comprising extero- (vision, audition, tactile) and proprioception (measured internal states like posture).
- Presence or working memory, a short term representation of behaviorally meaningful per-

cepts or internally generated entities as the basis for external and internal actions. The content of this presence is modulated by top down attention processes.

- Plastic (i. e. learnable) long term memory for the storage and retrieval of consolidated persistent entities like objects, words and scenes.
- Elements for predictions and internal simulation for creating expectations about the world for selecting the relevant information from the sensory streams and choosing internally the most effective action from a set of possible alternatives without testing all alternatives externally.
- Basic motor control and coordination means for efficiently controlling systems with a large number of degrees of freedom.
- A basic behavior repertoire building on the motor control level for a more abstract and robust representation of actions.

- A more abstract behavior organization comprising the traditionally separated issues of communication and action, leading to behaviorally routed definitions of the semantics of language.
- A representation of goals and processes working on those representations for organizing a meaningful system behavior on mesoscopic time scales above purely reactive sensory actions and below strategically driven decision processes. Those elements will also influence the perception in a top down fashion for focusing on goal relevant entities in the external world.
- A set of internal drives and motivations for establishing internal forces towards a continuous self-development of the system and controlling the balance between explorative and exploitative actions.
- A value system (“emotions”) providing basic guidelines for the limits of autonomous behavior by unconditioned preferences concerning elements and states of the environment.

This subset is far from being complete, but we consider those elements as the most pressing research issues that will provide major progress in the field of cognitive robotics. A major issue for most of the elements stated above is learning and adaptation, always in interaction with the real world. We consider this as an important issue, because otherwise we may not ask the research questions stated above in the right way. Nevertheless, we take the freedom to ignore the issue of on-board computing resources for now, as long as we consider the property of scaling of the researched methods from the beginning.

The above stated elements are addressed as fundamental research projects at the Honda Research Institute Europe in Offenbach. In the following we will focus on selected issues, directly related to real robots.

3 Task and Body Oriented Motion Control

Before we can learn movements, actions and behaviors, we have to get to the point that the robot can behave in a non preprogrammed way, as it is the case currently. Controlling the physical motions of a biped humanoid robot is not a trivial task. The number of degrees of freedom that have to be coordinated is high, the balance of the system has to be maintained under all circumstances and it is currently possible for such kind of robots to mechanically destroy themselves by commanding position of the limbs that lead to self collisions. On the other hand, humans and animals effortlessly control their end-effectors for solving tasks without continuously reflecting on the level of joints about their current motions. Additionally, they have some kind of body image representing the anatomy and boundaries of the own body that helps acting in complex tasks without continuously having contacts with the own body or exter-

nal objects. From the constructive point of view, it seems desirable in a cognitive architecture to be able to cognitively control only the task relevant parameters and leave the “tedious” details to underlying levels of control. This should include the avoidance of self collisions, which is in technical systems much more disastrous than in natural ones.

As a result of current research we have established a motion interface that fulfills the requirements stated above and allows cognitive control processes to perform complex control tasks with the humanoid ASIMO with a minimum effort [2]. In contrast to classical joint level control, the robot is controlled by a task level description. The tasks can e. g. be defined by four separate targets for the two hands, the head gaze direction and a walking target for the robot. The respective coordinate systems are depicted in Fig. 1. The coupling between the tasks and the mapping to the actually controlled joints is performed by a whole body motion controller.

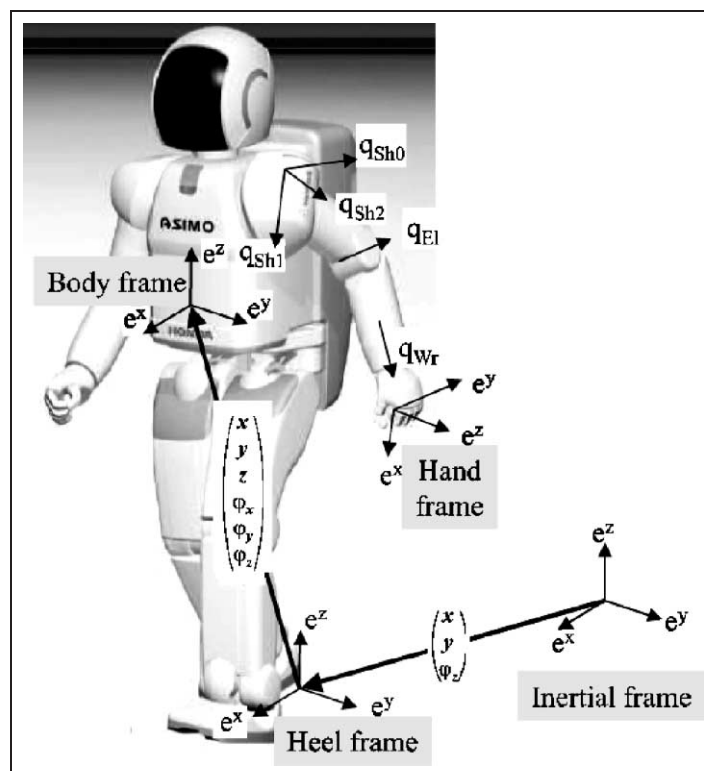


Figure 1 Kinematics model of ASIMO used for the whole body motion control.

This controller implements a redundant control scheme considering all degrees of freedom of the robot simultaneously. The commanded tasks usually do not determine all degrees of freedom uniquely. For the remaining degrees of freedom it is possible to state potential functions that model the task unspecific preferences of those. This could be closeness to rest positions and the avoidance of extreme joint positions close to the physical limits.

What do we gain by pursuing such kind of task description and whole body control in a cognitive architecture? First of all we have a description of tasks in a more natural way than in the joint space. For example, the right hand is commanded to a position in 3D space with a certain attitude. The necessary joint trajectories are computed automatically online on-board the robot. Therefore, higher level processes do not have to care about the details of the robot motion. Additionally, since the whole body including walking is employed for reaching the commanded target, the motion range is extended incrementally. Imagine a 3D position for the right hand is commanded that is not reachable by arm motions alone. The whole body motion controller first induces a leaning motion of the upper body in order to reach the target, and if this does not suffice ASIMO starts walking for finally reaching the commanded target. Again, higher level cognitive tasks still command only the 3D target position of the right hand. During those movements the appearance of the robot motion is naturally relaxed, because the redundant degrees of freedom are “softly” adapted to the requirements given by the hard tasks constraints. This fact can be envisioned by assuming springs between the segments of the robot, the task command would correspond to a force pulling the respective hand and the rest of the robot’s body adapting to the influence of this force. The walking is conceptually more advanced but

can also be treated within the same framework. Such kind of natural appearance can help solving acceptance problems with robots. Summarizing, the task space description and the whole body control approach gives more freedom to the motion control level and disburdens the higher levels of control. Further extensions in the same spirit are commanding task intervals instead of crisp tasks positions and including self collision avoidance on the level of motion control [3]. Current research is concerned with strategy selection based on internal simulation, allowing ASIMO to autonomously choose the hand to grasp the commanded target with. The proposed approach performs interactively with visually specified targets [4], which is in contrast to state of the art as described in [5; 6].

4 Visually and Behaviorally Oriented Learning

In the introduction we have stated that the long term goal of this work is the creation of a humanoid robot that is equipped with mechanisms for learning and development. The concrete goal here is to present an interactively behaving vision system for the humanoid that comprises already both kinds of mechanisms: autonomous developmental mechanisms governing the behavior generation and selection, and interactive learning mechanisms allowing for teaching the system new objects to be recognized online. Regarded separately, both mechanisms already represent a valuable step towards autonomous adaptive systems. But the emphasis is more on the principled combination of both. In contrast to statistical learning we are here less concerned about the representation of the variability of the input space but rather in learning behaviorally relevant external and internal entities.

For studying those issues we have created a biologically motivated interactive vision system with learnable basic behaviors being able to learn and recognize freely pre-

sented objects in interaction. The learning is governed by mechanisms building on internal needs dynamics based on unspecific and specific rewards governing and exploring the parameterization of the basic behaviors [7].

On a schematic level, the system can be described as follows: Based on the images from a stereo camera pair a set of features is computed. Those features comprise the general visual “interestingness” (saliency) S_v of image locations, the most prominent region in the image based on visual motion S_m , and the most prominent region in the image based on closeness S_d to the system. The information is represented as activation maps over the image location, with high values of the maps corresponding to interesting locations. Those maps are weighted with weights w_v , w_m , and w_d respectively, and added. Based on this combined saliency S and some memory about previous gaze directions the new gaze direction is determined by means of an integrative peak selection with hysteresis. This simple system exhibits some interesting behaviors. It is a homogeneous control loop that is constantly executed without any structural changes. The behavioral spectrum of the system is spanned by the weighting parameters w_v , w_m , and w_d of the maps. What kind of behaviors can such kind of a system show? Priming the saliency computation for a certain color like red and having a weight w_v greater than zero will yield a system that gazes at locations in its environment with red color. If there is no red color in the current view it will randomly look around and shifts its field of view until it finds a spot with red color. For an external observer it looks like the system is “searching for red color”, even though the system was never directly programmed to show a search behavior. Next, imagine having an interacting human in the scene approaching the system and trying to raise the attention of the system (see Fig. 2). Assigning a value to

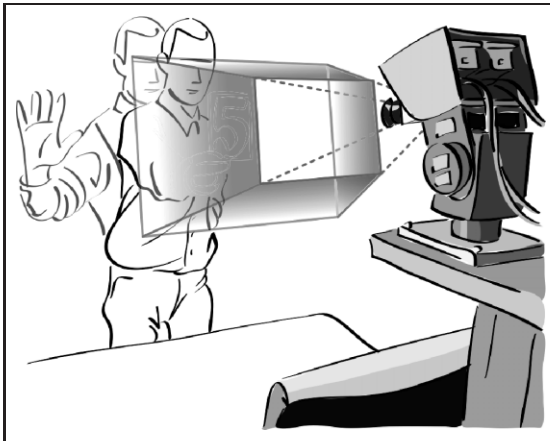


Figure 2 Interaction with the stereo camera head. Waving attracts attention, showing objects within the peripersonal space fixes attention to the presented object. The fixated object can be learned interactively and recognized immediately.

w_m greater than w_v will cause the system to look at a spot where the waving is located in the image, i. e., with such kind of parameter setting the system will look at regions containing visual motion rather than general visual saliency. Nevertheless, since the two channels are superimposed, they support each other for stabilizing the gaze selection. If the user continues to approach the system, the distance based map S_d will have an activation corresponding to the closest of the user's body to the system. If this contribution receives the highest weight w_d , the system will continuously focus on, e. g., the hand of the user. This behavior corresponds to the biological concept of a peripersonal space around the robot. The psychological concept of the peripersonal space is defined as the space wherein individuals manipulate objects, whereas extrapersonal space, which extends beyond the peripersonal space, is defined as the portion of space relevant for locomotion and orienting. Here, the peripersonal space establishes a very important concept for any further meaningful interaction: Sharing the attention between the system and the user. An interacting user can show something to the system and the system will focus on the shown object. Based on this capability we have addressed several new scientific concepts. The first one is the online learning of complex objects freely presented to the system. The object within the peripersonal

space is segmented from the image and processed by a biologically motivated visual object recognition and learning system [8]. Two different memory stages and speech interaction serve to continuously learn and label objects in real-time, allowing for online correction of errors during learning. Secondly we have introduced an internal homeostatic control system representing internal drives allowing for learning the weights of the different maps. This corresponds to learning the visual interaction behaviors instead of working with hard coded weights. The third one is the combination of the previous ones. The system learns to interact with the human user in such a way that its internal needs are in the temporal average equally satisfied including the curiosity for learning new objects. This represents a new quality in interactively learning systems.

The implementation is currently limited to controlling the gaze direction of a head, but the concepts are sufficiently general in order to allow for interactively learning behaviors including manipulators. This is currently being investigated. Research aiming in a similar direction can be found in e. g. [9].

We would like to point out that the system described above is internally not organized in terms and structures of externally visible behaviors. There is no "tracking" or "interaction" module within the system, even if those terms can be

attributed to externally observable behaviors. In our opinion it is crucial to make this distinction, because a system organized in terms of externally observable entities will by definition be confined to the predefined set of behaviors put into the system, and no self-driven cognitive self-development will be possible.

5 Summary

In this paper we have presented research concerned with elements of cognitive robotics. We consider them as steps towards a cognitive humanoid robot being able to truly provide services to humans in real anthropocentric environments. The elements stated above focus on the necessary elements of simple but powerful control of humanoid robots and learning human relevant objects and interaction behaviors based on internal mechanisms. It clearly points away from pre-designed motion and behaviors of most state-of-the-art robots.

Acknowledgments

The authors would like to thank Michael Gienger, Herbert Janßen, Hisashi Sugiura, Inna Mikhailova, Bram Bolder, Mark Dunn, Heiko Wersing, Stephan Kirstein, Julian Eggert, Antonello Ceravola, and Frank Joublin for their contributions and their support.

References

- [1] E. Kandel, J. Schwartz, and T. Jessell. *Principles of Neural Science, Fourth Edition*. McGraw-Hill, 2000.
- [2] M. Gienger, H. Janßen, and C. Goerick. Task-oriented whole body motion for humanoid robots. In *IEEE/RSJ Int'l Conf. on Humanoid Robots*, 2005.
- [3] H. Sugiura, M. Gienger, H. Janßen, and C. Goerick. Real-time self collision avoidance for humanoids by means of nullspace criteria and task intervals. In *IEEE-RAS Int'l Conf. on Humanoid Robots*. IEEE Press, 2006.
- [4] B. Bolder, M. Dunn, M. Gienger, H. Janßen, H. Sugiura, and C. Goerick. Visually guided whole body interaction. In *IEEE Int'l Conf. on Robotics and Automation*, 2007.

- [5] K. Nishiwaki, M. Kuga, S. Kagami, N. Inaba, and H. Inoue. Whole-body cooperative balanced motion generation for reaching. In *IEEE/RSJ Int'l Conf. on Humanoid Robots*, 2004.
- [6] N. Sian, K. Yokoi, S. Kajita, and K. Tanie. A framework for remote execution of whole body motions for humanoid robots. In *IEEE/RSJ Int'l Conf. on Humanoid Robots*, 2004.
- [7] C. Goerick, I. Mikhailova, H. Wersing, and S. KIRSTEIN. Biologically motivated visual behaviors for humanoids: Learning to interact and learning in interaction. In *IEEE/RSJ Int'l Conf. on Humanoid Robots*, 2006.
- [8] H. Wersing, S. KIRSTEIN, M. Götting, H. Brandl, M. Dunn, I. Mikhailova, C. Goerick, J.J. Steil, H. Ritter, and E. Körner. A biologically motivated system for unconstrained online learning of visual objects. In *Proc. Int'l Conf. Art. Neur. Netw. ICANN*, 2006.
- [9] A. Ude and G. Cheng. Object recognition on humanoids with foveated vision. In *IEEE/RSJ Int'l Conf. on Humanoid Robots*, 2004.



1



2

1 Dr.-Ing. Christian Goerick studied electrical engineering at the Ruhr-University, Bochum, Germany, and at the Purdue University, Indiana, USA. He holds a doctoral degree in electrical engineering and information processing from the Ruhr-University of Bochum. During his time in Bochum he was research assistant, doctoral worker, project leader and lecturer at the Institute for Neural Computation, Chair for Theoretical Biology. The research was concerned with biologically motivated computer vision for autonomous systems and learning theory of neural networks. Dr. Goerick is currently chief scientist at the Honda Research Institute Europe GmbH, with responsibility for embodied brain-like intelligence covering research on behavior based vision, audition, behavior generation, robotics, car assistant systems, systems architecture as well as hard- and software environments.

Address: Honda Research Institute Europe GmbH, Carl-Legien-Straße 30, 63073 Offenbach/Main, Tel.: +49-69-89011-742, Fax: +49-69-89011-759, E-Mail: christian.goerick@honda-ri.de

2 Prof. Dr.-Ing. habil. Edgar Körner studied electrical engineering, control engineering, and biomedical cybernetics at the Ilmenau Institute of Technology, Germany. From 1976 to 1984 he served as an assistant professor/senior staff researcher and established the Bionics Research Laboratory at the same university. The research activities included experimental work in neurophysiology and neural systems

modeling as well as in psychophysics and medical expert systems. He received his Dr.-Ing. in the field of biomedical cybernetics in 1977 and the Dr. Sci. (habilitation) in biocybernetics in 1984, both from Ilmenau Institute of Technology. From 1984 to 1987, he joined the Bioholonics Project of JRDC (Tokyo) as a research fellow dealing with brain-like vision systems. Back at the Ilmenau Institute of Technology, he continued research in biological vision and neurofuzzy control systems as an associate professor. In 1988, Dr. Körner was appointed full professor for biocybernetics and head of the Department of Neurocomputing and Cognitive Systems. In 1992 he moved to Japan to join Honda R&D's Wako Research Center near Tokyo, focusing as a chief scientist on the brain-like intelligence research. In 1997 he started research in computational neuroscience, evolutionary technology, and cognitive robotics at Honda R&D Europe, where he served as an executive vice president and head of the Future Technology Research Division. Since 2003, Dr. Körner serves as the president of the Honda Research Institute Europe GmbH. His research interest covers brain-like intelligence, with a special focus on self-referential control architectures, self organization of knowledge representation, and autonomous robots.

Address: Honda Research Institute Europe GmbH, Carl-Legien-Straße 30, 63073 Offenbach/Main, Tel.: +49-69-89011-750, Fax: +49-69-89011-759, E-Mail: edgar.koerner@honda-ri.de