FP6-004381-MACS

MACS

Multi-sensory Autonomous Cognitive Systems Interacting with Dynamic Environments for Perceiving and Using Affordances

www.macs-eu.org

Final Activity Report

Specifically Targeted Research Project (STReP)
FP6 IST Thematic Priority: 2.3.2.4 Cognitive Systems

Period covered: from Sep 1, 2004 to Nov 30, 2007
Date of preparation: 29.2.2008 Revision: Version 3

Co-ordinator: Erich Rome, Fraunhofer IAIS

Fraunhofer Gesellschaft zur Förderung der Angewandten Forschung e.V.
Fraunhofer Institut für Intelligente Analyse-und Informationssysteme (IAIS)

INDEX

MACS

Multi-sensory Autonomous Cognitive Systems Interacting with Dynamic Environments for Perceiving and Learning Affordances

www.macs-eu.org

1 INTRODUCTION

1.1 Abstract

This document is the final activity report of the EU research project MACS. Within the MACS project, six partners from Germany, Austria, Sweden and Turkey have investigated a new way for controlling a mobile robot with manipulation capabilities by adapting a concept from Cognitive Science.

In this final project report, we present the MACS approach to affordance-inspired robot control. An affordance, a concept from Ecological Psychology, denotes a specific relationship between an animal and its environment. Perceiving an affordance means perceiving an interaction possibility that is specific for the animal's perception and action capabilities. Perceiving an affordance does not include appearance-based object recognition, but rather feature-based perception of object functions. The central hypothesis of MACS is that an affordance-inspired control architecture enables a robot to perceive more interaction possibilities than a traditional architecture that relies on appearance-based object recognition alone. We describe how the concept of affordances can be exploited for controlling a mobile robot with manipulation capabilities. Particularly, we describe how affordance support can be built into robot perception, how learning mechanisms can generate affordance-like relations, how this affordance-related information is represented, and how it can be used by a planner for realizing goal-directed robot behaviour. We present both the MACS demonstrator and simulator, and summarise development and experiments that have been performed. By interfacing perception and goal-directed action in terms of affordances, we provided a new way for reasoning and learning to connect with reactive robot control. We show the potential of this new methodology by going beyond navigation-like tasks towards goal-directed autonomous manipulation in our project demonstrators.

1.2 Purpose

This document is the final activity report of the EU research project MACS. MACS is a specifically targeted research project that has been carried out between September 1, 2004 and November 30, 2007. MACS has been partly funded by the European Commission within their Sixth Framework Research Programme (FP6) under contract number FP6-004381. Within FP6, MACS contributed to the priority Information Society Technologies (IST) and its strategic objective "Cognitive Systems".

1.3 Contractors involved

The co-ordinator of the MACS project is the Fraunhofer Institute for Intelligent Analysis and Information Systems in Sankt Augustin, Germany (FhG/AIS). FhG/AIS provides expertise in mobile robotics, control architectures, AI, sensor systems and biologically motivated perception. Joanneum Research, from Graz, Austria (JR_DIB), combines know-how in cognitive visual perception systems and in machine learning. Linköpings Universitet, Sweden (LiU-IDA), contributes know-how in AI, knowledge representation and robotics. Middle East Technical University, Ankara, Turkey (METU-KOVAN), provides its know-how in the simulation of robotic systems and in control architectures. Cognitive science and learning methods are the research topics of OFAI, Österreichische Studiengesellschaft für Kybernetik, Vienna, Austria. The University of Osnabrück, Germany, (UOS) contributes expertise in AI planning, knowledge-based systems, and Robotics. A more detailed description of the consortium can be found in Annex B1.3.

1.4 Authors

Erich Rome	Fraunhofer IAIS	Sankt Augustin	DE
Lucas Paletta	Joanneum Research	Graz	AT
Erol Şahin	Middle East Technical University	Ankara	TR
Georg Dorffner	Österreichische Studiengesellschaft für Kybernetik	Vienna	AT
Joachim Hertzberg	Universität Osnabrück	Osnabrück	DE
Christopher Lörken	Universität Osnabrück	Osnabrück	DE
Ralph Breithaupt	Fraunhofer IAIS	Sankt Augustin	DE
Stefan May	Fraunhofer IAIS	Sankt Augustin	DE
Gerald Fritz	Joanneum Research	Graz	AT
Emre Uğur	Middle East Technical University	Ankara	TR

1.5 Document Structure

The remainder of this document is structured as follows. We start with motivating our approach and give a general introduction to the concept of affordances and how it can be utilized for robot control.

In the next section, we will describe the initial overall objectives of MACS. Subsequently, we will give an introduction to the central notion of ‘affordances’ and the main theses of the area of Ecological Psychology. In the main technical section of the document, we present our advances in formalising, perceiving, learning and representing affordances, describe the way our affordance-inspired robot control architecture works, and summarise the experiments that we have performed. We conclude with summarising our contribution to the covered research fields, with some general conclusions and recommendations for future work in the area of affordance-inspired robots. Appendices contain bibliographical references, basic facts of the project, a detailed presentation of the contractors, a list of publications, lists of other dissemination activities, and a glossary of terms and definitions.

2 MOTIVATION

Research and development in mobile robotics has made significant progress in the last decade. Some robots have entered mass production, like the Roomba (Fig. 1(a)), a floor cleaning robot, and some of its smarter colleagues. Some research prototypes expose impressive specialised skills, like the winner of the Grand Challenge contest in the USA. Stanley (Fig. 1(b)), a robotic car, travelled 150 miles in below 7 hours, controlled only by a computer program.

(a)

(b)

Fig. 1 (a) Roomba, floor-cleaning robot by iRobot. (b) Winner of Grand Challenge 2005: Stanley

But in general, mobile robots are still not suited for everyday use. The goal of deploying a sophisticated mobile robot to a human environment has not yet been accomplished. Particularly, robust mobile interaction and manipulation capabilities are yet to be developed. A mobile manipulating robot would clearly benefit from abilities to solve some real-world, everyday problems.

Let us give an example for the type of problem solving we address: Imagine yourself sitting at a table in a street café. On the table, there is a glass, a bottle of water, a coffee cup, a menu stand, an ashtray and some loose sheets of paper from a printed article. A light wind comes up and the sheets are in danger to be blown away. To prevent this, you need a paperweight. Unfortunately, you do not carry your nicely designed paperweight with you. Typically, you would use one of the other items on the table to function temporarily as a paperweight, for instance, the glass. Of course, you could as well use any of the other items on the table. A little while later the wind slows down, you notice that you are still a bit thirsty, and you drink from the glass, knowing that it is currently not needed as a paperweight.

This example shows that we can solve an everyday problem by improvising, namely by using an item in a way that it was not particularly created for. It also shows that we are able to identify interaction possibilities in our environment and that we can select and act upon such interaction possibilities depending on our current goals.

The cognitive ability to use objects in a variety of ways and to find alternative solutions for a given task would be a great benefit for a service robot. The described abilities are attributed to the practical *aspect* of intelligence, which is not yet well understood. So how can we incorporate an ability like the one described in a technical system? Or, to be more specific: How can we design a "cognitive" mobile robot system with manipulation capabilities that can, e.g.,

- find alternative solutions for a given task,
- interact with known and unknown objects in a meaningful and goal-directed way, and
- uses perception methods that are tailored for its tasks and its action capabilities, i.e. that are grounded in its actions?

A valid approach is to draw inspiration from cognitive science. For MACS, we have chosen Ecological Psychology as a starting point. One of the founders of Ecological Psychology, J.J. Gibson, has created the notion of 'affordances', an artificial noun. It denotes the specific interaction possibilities that the environment offers a particular animal. J.J. Gibson defined an affordance as a resource or support that the environment offers an animal for action, and required that the animal be able to directly perceive and employ it [1]. Gibson claimed that the animal is in a specific relation to its environment, which formed the basis of the concepts of 'situatedness' and 'embeddedness' that are used in several sub-areas of Cognitive Science, like embodied embedded cognition and situated cognition. The concept of affordances has, since its conception, proven to have a strong appeal in a wide range of fields, ranging from design [2] and neuroscience to robotics. In robotics and artificial intelligence, affordances offer an original perspective on coupling perception, action and reasoning.

Although J.J. Gibson did not provide a formalisation of the affordance concept, he characterized the concept in his works, and its most important properties shall be mentioned here.

1) Affordances are specific to animals and their action and perception capabilities, and to their body sizes as well. The same object can offer different affordances to different animals, as depicted in Fig. 2.

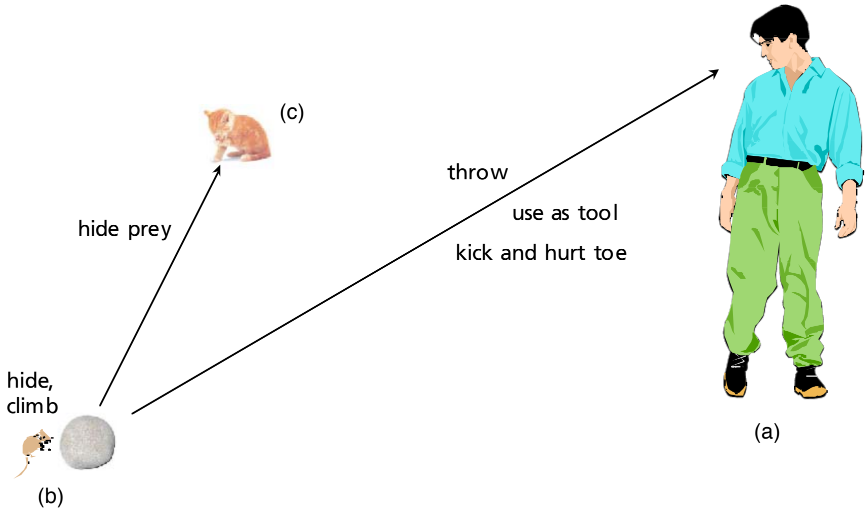


Fig. 2 The same object, here a fist-sized stone, offers different affordances to different animals. For an adult healthy human, it offers the affordance of throwing it, or to use it as a tool, to name just two. For a mouse, it offers the affordances to hide behind it or to climb on top of it, and for a cat it offers the affordance of hiding prey.

2) Affordances include utilities or functions of things in the world.

3) Affordances can be described by abstract features. Things in the environment that offer a human the possibility to sit upon them (affordance 'sitability') are typically horizontal, knee-high stable surfaces of a certain minimum size (cf. Fig. 3).

Fig. 3 Opportunities to sit offered by different entities in the environment (affordance of being "sitable").

4) Affordances need not necessarily be labelled. If a human knows what to do with a thing, it can use it, regardless of its name. However, it should be noted that many words in our languages are related to action and to haptic experience (like "to grasp something").

5) Affordances are perceived directly, no mental representation is required. This claim of Gibson's, called 'direct perception', is the mostly debated part of his theory of affordances. We will get back on this later.

How do affordances relate to 'objects' (which could be described as coherent entities or things)? Objects and affordances are complementary in the sense that one object class may offer a multitude of affordances, and one affordance may be offered by a multitude of object classes.

This complementarity of the object and the affordance notions can allow a robot a greater flexibility for performing tasks. A robot system that uses object-centred perception may need to abort a mission if objects of a certain class that are required to reach a (sub-)goal are not available. In those cases where an affordance (like a function) of this object is more important than its sensorial appearance, affordance-based perception may be more appropriate, since it allows the robot to perceive and use objects with the same function that belong to a completely different object class, that is, it helps finding alternatives for action.

So, intuitively, this concept seemed to be suited to form a basis for modelling the type of ability that we were looking for in MACS. However, before we could use it, there were three difficulties that needed to be overcome. First, Gibson elaborated much on perceiving affordances, but little on using or learning them. Second, Gibson did not provide a suited formalisation of affordances that could directly be employed for implementation in a technical system. And third, Gibson's claim of direct perception and lack of representations of affordances seemed to be insurmountable barriers for an implementation. But if affordance support cannot be explicitly built into a technical system X, then the behaviour of X can just merely be *described* in terms of affordances, and nothing is gained.

How did we overcome these difficulties? Let us start with some findings that we employed.

1) Although affordances are always there in the environment, we do not perceive them all at once. We have some filter mechanisms that prevent us from being flooded by affordances. The selection of the affordances that we act upon is, among other things, dependent on our current goals or motivations. For example, if we walk through a city and get hungry, then we are actively looking for restaurants and other opportunities to eat. If we walk through a city and get tired, we will be looking for opportunities to sit.
2) If we perceive an affordance, we do not always automatically act upon that affordance [14]. Instead, an act of will might be involved as well.
3) Affordances can be learned [23]. This holds for natural things, like learning to use a flint stone for making fire, and even more for technical systems. The affordances of an airplane cockpit cannot just be perceived, they must be taught, trained and memorized.

Our conclusion was: If we want to make use of affordances in a technical system in a goal-directed way, and if we want to reason about affordances, then we must represent them explicitly. And an explicit representation of affordances would benefit from a formalisation.

Since research on affordances has been continued until today, there were, of course, newer results that we inspected, too. Stoffregen [84] and Chemero [3] defined affordances as relations within the organism-environment system. We created an own formalisation that in general followed this idea. Our own 'agent affordances'—we chose this name to distinguish them from Gibson's original formulation of the concept—are represented as relations between features of the environment, the sequence of actions ('behaviour') that a mobile robot can perform on an affordance, and the outcome of these actions.

After we had resolved these difficulties, we were able to design an affordance-inspired control architecture for utilizing the concept of agent affordances. The architecture computes primitive features in a bottom up manner from the robot's sensory input. In an exploration phase, the robot can interact with its environment, and a learning module acquires knowledge about affordances and functions within the robot's environment. This knowledge is represented and stored for goal-directed usage.

In an application phase, the robot can goal-directedly act upon affordances. The robot can use knowledge about affordances for creating plans. These plans consist of a sequence of navigation and manipulation tasks descriptions, called *operators*. The plans are executed by invoking some basic robot control routines, the behaviours. The execution is monitored by the execution control module. The execution control module uses the abstract features of an affordance representation as cues for the presence of affordances in the environment. These cues function as a matched filter that prevents the robot from perceiving all affordances at once.

The architecture has been tested both in simulation and in a real demonstration test-bed including the mobile robot KURT3D with its basic manipulation capabilities. A physics-based simulator has been developed and employed for preliminary and for mass experiments. Some results achieved in simulation could be transferred to the real robot with only little changes required. Both facilities and some of the experiments are described in the main section 4.

3 MACS OBJECTIVES AND OVERALL METHODOLOGY

3.1 MACS Objectives

The main objective of the MACS project was to explore and exploit the concept of *affordances* for the design and implementation of autonomous mobile robots acting goal-directedly in a dynamic environment. The aim was to develop affordance-inspired control as a method for robotics. That involved making affordances a first-class concept in a robot control architecture. By interfacing perception and action in terms of affordances, the project aimed to provide a new way for reasoning and learning to connect with reactive robot control. The potential of this new methodology should be shown by going beyond navigation-like tasks towards goal-directed autonomous manipulation in the project demonstrators. All over, MACS aimed at embedding its technical results into cognitive science.

In MACS, there is explicit support for the affordance concept in the robot control architecture and the hypothesis was that the resulting performance of the robot will benefit in terms of robustness and generality. In fact, these are essentially the only criteria that can be used to evaluate empirically whether an affordance-based robotic system is better than a non-affordance-based one.

The main result of MACS should be a working, integrated robot system, based on the KURT3D robot, that serves as a proof of concept for the affordance-inspired robot control approach. Other results of the project should be a formal theory, a dedicated simulation environment, a specifically tailored learning approach for generating affordance representations, an affordance-based planner, feature extractors and other software for function-centred perception, plus dissemination of the results. In the conclusion, we will describe our achievements with respect to these initial objectives.

3.2 Affordances in Ecological Psychology

J.J. Gibson (1904–1979) is one of the most influential psychologists of the 20th century, who aimed to develop a 'theory of information pick-up' as a new theory of perception. He argued that an organism and its environment complement each other and that studies on the organism should be conducted in its natural environment rather than in isolation, ideas that later formed the basic elements of Ecological Psychology. The concept of affordance was conceived within this context.

Based on his studies of meaningful optical variables[1] and the Gestaltist conception of immediate perception of meanings of the things, J.J. Gibson built his own theory of perception and introduced the term affordance to refer to the action possibilities that objects offer to an organism, in an environment. The term affordances first appeared in his 1966 book [5], and is further refined in his later book [1]. In this book, the description of the affordance concept was discussed in a complete chapter, which generally laid out the fundamental aspects of affordances:

> "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, but the noun affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment." (J.J. Gibson, 1979/1986, p. 127)

J.J. Gibson believed that affordances were directly perceivable (a.k.a. *direct perception*) by the organism, thus the meaning of the objects in the environment were directly apparent to the agent acting in it. This was different from the contemporary view of the time that the meaning of objects were created internally with further 'mental calculation' of the otherwise meaningless perceptual data.

The discussions on the perception of object affordances naturally had some philosophical consequences on the much debated object concept.

> "The theory of affordances rescues us from the philosophical muddle of assuming fixed classes of objects, each defined by it common features and then given a name. ...You do not have to classify and label things in order to perceive what they afford." (J.J. Gibson, 1979/1986, p. 134)

Gibson goes on to state that

> "... But this does not mean you cannot learn how to use things and perceive their uses." ([1], p. 134).

And earlier:

> "... If you know what can be done with a graspable detached object, what it can be used for, you can call it whatever you please." ([1], p. 134).

Thus, objects and affordances are complementary in the sense that one object class may offer a multitude of affordances, and one affordance may be offered by a multitude of object classes.

J.J. Gibson's view of studying organism and environment together as a system (including the concept of affordance) has been one of founding pillars of Ecological Psychology. Following the formulation of the theory of affordances, the Ecological Psychology community started to conduct experiments in order to verify that people are able to perceive the affordances of the environment and to understand the mechanisms underlying this perception. These experiments ([6]–[11]) aimed at showing that organisms (mostly human) can perceive whether a specific action is *do-able* or *not-do-able* in an environment. This implies that what we perceive are not necessarily objects (e.g. stairs, doors, chairs), but the action possibilities (e.g. climbable, passable, sittable) in the world. Although the number of these experiments is quite high, the diversity in them is rather narrow. They constitute a class of experiments characterized by two main points: taking the ratio of an environmental measure and a bodily measure

[1] For example *optical centre of expansion* of the visual field was such an optical variable which was meaningful for a pilot trying to land a plane, indicating the direction of the glide, and helping him to adjust the landing behaviour.

of the human subject; and based on the value of this ratio, making a binary judgment of whether a specific action is possible or not.

The first point gives us a clue about how the experimenters interpreted affordances. Since affordances were roughly defined as the properties of the environment taken relative to the organism acting in it, the effort was to show that the ratio between an environmental measure and a bodily measure of the organism have consequences for behaviour. This ratio must also be perceivable, so that the organism is aware of this measure which, in a way, determines its behaviour's success.

Warren's stair-climbing experiments [6] have generally been accepted as a seminal work on the analysis of affordances, constituting a baseline for later experiments which seek to understand affordance-based perception. In these studies, Warren showed that organisms perceive their environment in terms of *intrinsic* or *body-scaled* metrics, not in absolute or global dimensions. He was able to calculate the constant, so called π proportions, that depend on specific properties of the organism-environment system. There exists one such ratio per each affordance, and they solely depend on the functionally relevant variables of corresponding actions. For instance, a humans judgment of whether he can climb a stair step is not determined by the global dimension of the height of the stair step, but by its ratio to his leg-length.

In [7], Warren and Whangs showed how the perception of geometrical dimensions such as size and distance is scaled relative to the "perceived eyeheight" [2] of the perceiver, in an environment where the subjects were to judge the affordance of walking through an aperture. Marks' surface sitting and climbing experiments [8] also incorporated a similar approach. Some of these studies ([9],[10]) criticized former studies because they limited themselves to only one perceptual source, namely visual information. Instead of limiting themselves to visual perception, they studied haptic perception in infant traversability of surfaces and critical slant judgment for walking on sloped surfaces. While in these experiments human subjects were asked to judge whether a certain affordance exists or not in a static environment, Chemero [11] conducted other experiments, in order to prove that changes in the layout of affordances are perceivable in dynamic environments, and found out that the results are compatible with *critical ratio* values. Another important work is Oudejans et. al.'s [12] study of *street-crossing behaviour* and perception of *critical time-gap* for safe crossing. This work is novel since it shows that not only static properties of the organism, but also its dynamic state is important when deciding on its actions.

An overview of the mentioned experiments shows that they are mostly focused on the perception aspect of affordances. Other cognitive processes such as learning, high level reasoning and inference mechanisms are not the subjects of these experiments, and the link between affordances and these higher level processes is not discussed.

4 WORK PERFORMED AND RESULTS

4.1 Affordance-inspired robot control architecture

a. Related Work

The concept of affordances is highly related to autonomous robot control and influenced studies in this field. We believe that for a proper discussion of the relationship of the affordance concept to robot control, the similarity of the arguments of J.J. Gibson's theory and reactive/behaviour-based robotics should be noted first.

[2] In [7], eyeheight is defined as the height at which a person's eyes would pass through the wall while walking and looking straight in a natural and comfortable position.

The concept of affordances and behaviour-based robotics emerged in very similar ways as opposing suggestions to the dominant paradigms in their fields. J.J. Gibson constructed his theory based on the criticism of the then dominant theory of perception and cognition, which favoured modelling and inference. Likewise, behaviour-based robotics was motivated by the criticism of the then dominant robotic architectures, which favoured modelling and inference. This parallelism between the two fields suggests that they are applications of the same line of thinking to different domains ([13], p. 244; [14]). Opposing to modelling and inference, J.J. Gibson defended a more direct relationship between the organism and the environment and suggested that a model of the environment and costly inferential processes were not needed. In a similar vein, behaviour-based robotics advocated a tight coupling between perception and action. Brooks, claiming that “the world is its own best model”, suggested an approach that eliminated all the modelling and internal representation [15]. As a result, one can see the underlying concepts of affordances in existence in robot control architectures such as subsumption architecture [16], the robot-schema architecture [17] and AuRA [18].

Some roboticists have already been explicitly using ideas on affordances in designing behaviour-based robots. For example, Murphy [19] suggested that robotic design can benefit from ideas in the theory of affordances such that complex perceptual modelling can be eliminated without loss in capabilities. She tried to prove her point with three case studies and drew attention to the importance of the ecological niche in the design of behaviours. Likewise, Duchon et al. [14] benefited from J.J. Gibson’s ideas on direct perception and optical flow in the design of behaviours and termed Ecological Robotics to be the practice of applying ecological principles to the design of mobile robots.

The use of affordances within Autonomous Robotics is mostly confined to behaviour-based control of the robots, and that its use in deliberation remains a rather unexplored area. This is not a coincidence, but indeed a consequence of the lacks in J.J. Gibson’s theory. The reactive approach could not scale up to complex tasks in robotics, in the same way that the theory of affordances in its original form was unable to explain some aspects of perception and cognition. The need to hybridize robotic control architectures can be considered similar to the attempts in Cognitive Psychology to view affordances as part of a complete cognitive model. While some cognitive models relate affordances only with low-level processes [20], others consider their role in cognitive processes as well ([21]–[23]). Similarly in robotics, some hybrid architectures inherit properties related to affordances only at their reactive layer ([18],[24]), while other studies exploit how affordances reflect to high-level processes such as learning ([25],[26],[23],[27],[28]), decision-making [29], and planning [30]. Recently a number of robotic studies focused on the learning of affordances in robots. These studies mainly tackled two major aspects. In one aspect, affordance learning is referred to as the learning of consequences of a certain action in a given situation ([27],[28],[30]). Stoytchev’s ([28], [30]) and Fitzpatrick et al.’s [27] work uses affordances as a higher-level concept, which a developing cognitive agent learns about by interacting with the objects in its environment. The robots in both studies execute certain actions on certain objects, and observe and learn the change in the environment as the consequence of the action. In other studies the focus lies on the learning of invariant properties of environments that afford a certain behaviour ([23],[26],[29]). In [23], MacDorman proposes an architecture, where the robot learns a sensory-motor mapping of its actions, and uses this learned model to make plans at the deliberative level. The learned model is then used to predict the affordances of objects in the environment. However, MacDorman defines affordances only in terms of internal values of the robot (like ‘tasty’ and ‘poisonous’ things), and not the physical changes it can create in the environment separating the process of predicting the outcome of actions, from the process of predicting affordances.

Some hybrid architectures inherit the properties of reactive architectures in their reactive components. For example, AuRA [18] is said to be influenced from J.J. Gibson’s theory of affordances for using action-oriented perception in the reactive component. In AuRA, each motor schema is associated with a perceptual schema, which extracts the sensory input relevant for the particular behaviour. Similarly, in the SSS [24] architecture, the communication of lower and upper layers is based on the idea of matched filters, which suggests that certain sensor states are equivalent if they call for the same motor response. Although not explicitly stated, we can further relate affordances to some deliberative processes in hybrid architectures. For instance, the AuRA [18] architecture can be said to perform deliberative modulation of perception, since plan execution occurs by activating motor schemas and the relevant perceptual schemas specified by the plan. Another example is the SFX [31] architecture in which the symbolic world model depends on the current behaviour, as a consequence of action-oriented sensor fusion.

We would like to note that the affordance theory of J.J. Gibson was mostly used as a source of inspiration in autonomous robotics. As a result, only certain aspects of the theory were used, and that no attempts to consider the implications of the whole theory towards autonomous robot control were made. In this sense, the development of an 'affordance-inspired robot control architecture' that is designed to learn, detect, and use the affordances in the environment [32] is an important contribution to the field.

b. The MACS approach to Affordance-inspired Robot Control

The vast majority of robot perception approaches are either close perception-action couplings for reactive behaviour or oriented towards object recognition on higher control levels. Also, object recognition is in many cases based on general computer vision methods that do not account for the specifics of the robot at hand, i.e. its sensory system and its actuator system. Only very few robot perception approaches deal with recognition of functions that the environment offers (cf. Sec. 4.3).

We can state that a function-centred perception approach realises a view of the environment that is orthogonal to object-centred perception. Such function-centred perception potentially allows a robot to find more alternatives for acting in its environment. A robot mission that requires to find—based on appearance only—and use certain objects in the environment will fail if one or more of these objects cannot be found. But often the identity or appearance of an object may not be relevant for completing a task. A task could, for instance, also be completed if the robot finds an alternative object that offers the same functions as the original one (in J.J. Gibson's terminology, one would say: it affords the same action possibility). As we initially hypothesized, an affordance-inspired robot control with a function-centred perception would allow a robot more flexibility in plan execution and thus increase the likelihood of successfully completing a mission. Thus, it would enhance a robot's abilities to perceive and utilize the potential for action that the environment offers, i.e. enable a robot to make use of affordances. This is the central hypothesis of MACS.

MACS aimed at realizing affordance-inspired control in a hybrid architecture that allows goal-directed behaviour based on function-centred perception, with functions related to and grounded in the robot's action capabilities. Affordance support in the sense sketched in the previous sections has been built into several levels of the architecture. In order to use affordance support for deliberate action, i.e. for planning, we needed an explicit representation of the potential for action or the functions that the environment offers, respectively. The formalisation that is the basis for such representations is described in [33]. In summary, a number of formalisations have been proposed to clarify the concept of affordance in the field of Ecological Psychology. To summarise briefly, Turvey [34] defined affordances as 'dispositions' in the environment that get actualized with the interaction of the organism and the environment. Different from Turvey's formalism, which attached affordances to the environment, Stoffregen [85] and Chemero [3] defined affordances as relations within the organism-environment system. Independent from these formalisations in Ecological Psychology, Steedman [36] formalised affordances in Linguistics by providing an explicit link to action possibilities offered by the environment, and by proposing the use of the concept in planning. The authors are not aware of other robot control methods that make use of explicit, symbolic affordance representations.

In order to distinguish our use of the term ‘affordance’ from the use in Ecological Psychology, we introduced the definition of an agent (or robot) affordance ([33],[32])[3]:

> **Definition 1 ((Agent) Affordance).** *An affordance is a relation between the agent and its environment as acquired from the interaction of the two.*

This definition states that the affordance is a perceivable relation between the subjective capabilities of an agent and the features of its surroundings. The agent affordance definition is used whenever we are referring to or describing the robot’s situation in its environment, e.g. in examples of the robot’s behaviour or in descriptions of experiments. For this purpose, we use the notions of entity, (observed) behaviour, and (observed) outcome. An example shall illustrate the meaning of these notions. Given our mobile robot KURT3D (cf. Sec. 4.6a) with its basic electromagnetic gripper as manipulation device, and given that there are magnetic cans in its environment, we could say: “The robot has successfully lifted the blue can.”, where some features of the blue can comprise the entity, lifting is the observed behaviour, and the successful execution resulting in the can attached to the robot’s electromagnet is the observed outcome. The entity can be represented by a set of features perceived prior to the lifting behaviour, the lifting behaviour can be represented as a sequence of basic actions, and the outcome by a set of features perceived after the lifting behaviour has been executed. This leads to a straightforward definition of an (agent) affordance representation [32]:

> **Definition 2 (Affordance Representation).** *An* affordance representation *or* affordance triple *is a data structure:*
>
> (cue descriptor, behaviour descriptor, outcome descriptor). (1)
>
> *Here, a cue descriptor or an outcome descriptor is specified as a list of attribute value pairs. A behaviour descriptor consists of one or more behaviour identifiers. Optionally, parameters for these behaviours can be specified.*

Such representations can either be handcrafted or learned during an extended initial learning phase as described in Sec. 4.4. The cue part of the representation can be used to hypothesize the presence of a certain affordance in the environment that the robot searches for achieving the planned outcome. The feature set comprising a cue needs only be sufficient for making such a hypothesis. It is neither required that the feature set is a sufficient representation of the manipulated object, nor that all the cue features belong to this object.

After a certain amount of affordance representations have been created, the robot shall make use of this information for deliberate action as described in Sec. 4.2. A mission defined by a human operator could be the task of searching ‘liftables’ and stack these in an arbitrary location. The planner would create operators that employ affordance representations, and an execution control would monitor, as usual, the progress of task completion.

In order to implement these concepts, the proposed affordance-inspired control architecture consists of two branches. A bottom-up branch goes from sensors via a perception module (cf. Sec. 4.3) to a learning module (cf. Sec. 4.4) that generates affordance representations. A top-down branch goes from a deliberation module via execution control down to a behaviour system that provides some basic robot skills, including but not limited to driving, braking, map-building and lifting, or moving and controlling the magnet.

The proposed and implemented affordance-inspired control architecture scheme is depicted in Fig. 4. In this diagram, a red, solid arrow between components *A* and *B* in the diagram is of

[3] A similar but alternative formalisation of affordances was also proposed in [37].

type control flow. The arrow indicates that the control is passed from *A* to *B*. The arrow does say nothing about the situations in which the control is passed, nor about the data that might be exchanged when passing control. The designations close to such an arrow indicate qualitatively the nature of the control flow, e.g. information request, configuration request etc. A blue, dashed arrow between components *A* and *B* in the diagram is of type data flow. The data flow arrow does not say anything about the circumstances, that is, the current control states, under which the data are transferred. The designations close to such an arrow indicate qualitatively the types of data that are passed from *A* to *B*. Bold arrows indicate flows between modules, thin arrows intra module flows. Data passed from module *A* to *B* are available to all components inside *B*. White boxes are specific affordance support oriented components that are usually not found in other control architectures.

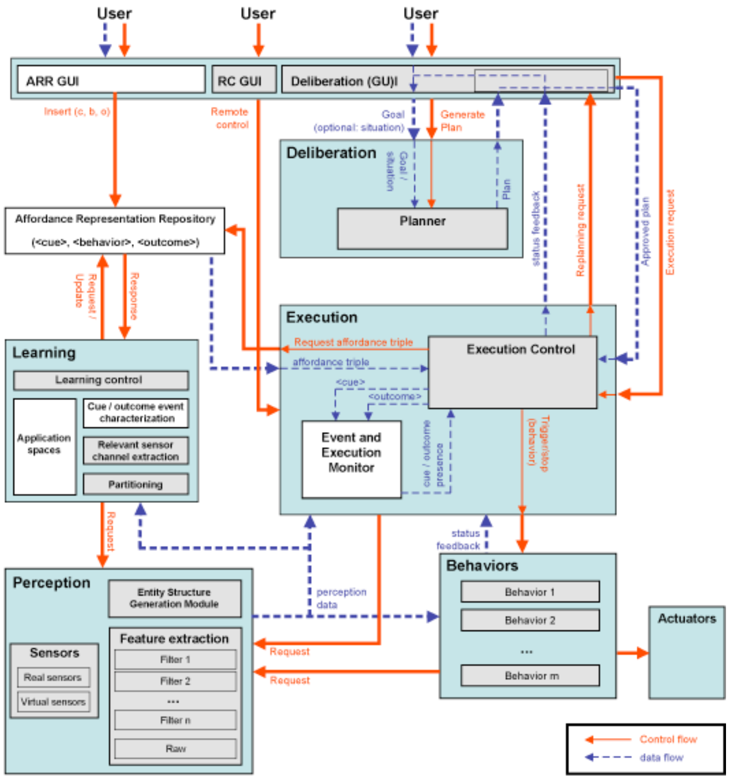


Fig. 4. Modules, data and control flow of the MACS control architecture. A solid arrow between components A and B indicates control flow, a dashed arrow data flow. White boxes are specific affordance support oriented components.

The main architectural building blocks in this diagram are:

User Interface displays status information and allows a user both to guide a robot manually through an action sequence and to just specify a mission goal for the robot. The

Deliberation module converts a mission goal into an executable affordance-based mission plan which is passed to the

Execution module. This module executes the mission plan, monitors its execution, including successful or unsuccessful acting upon affordances. The Execution module's new *Event and Execution monitor* checks the existence of affordance support cues and compares expected outcome with actual outcome of an executed behaviour control routine. The Execution control triggers behaviours of the

Behaviour System. This module provides a number of pre-programmed behaviour control routines that can be viewed as basic skills of the robot. Some behaviour control routines are parametrizable and can be configured by other modules, if necessary. The behaviours make use of

Actuators that enable the robot to move about and to interact with its environment. They include the drive motors, the sensor servos, and the crane arm motors. The

Sensors enable the robot to perceive its environment and its internal states, Virtual sensors provide software state information, real sensors yield data from the environment. All sensory data are first handled by the

Perception module. It relays sensory data, extracted features and status information (like active behaviours and their parameters) to the Learning module, Execution module, Behaviour System and Deliberation module. It can be configured to look just for certain features that relate to searched affordance support cues. Its *Entity Structure Generation Module* converts sensory data into appropriate data structures for architectural affordance support. The

Learning module takes input from the Perception module and generates affordance representations (affordance triples) to populate the new

Affordance Representation Repository. This repository is new and specific to our affordance-based approach. It provides affordance representations for use with the affordance-based Planner and Execution Module for goal-oriented mission planning.

This architecture is implemented in such a way that it can be connected both to the physical robot and the simulated robot via the same interface, just by pushing a button. This enables us to test the system both in simulated and in real environments.

In the next three sections, the affordance-based approaches of the main building blocks within this architecture, namely representation and planning (Sec. 4.2), perception (Sec. 4.3) and learning (Sec. 4.4), are explained. An elaborate description of the behaviour system and its basic skills can be found in [38]. In Sec. 4.5, our latest formalisation is presented. Before we conclude this report, we present the physics-based simulator MACSim and the experimentation environment, i.e. the demonstrator scenario and its elements in Sec. 4.6, and describe the experiments for the proof-of-concept in Sec. 4.7.

4.2 Representation for deliberation

Literature is rather sparse when it comes to more or less formal definitions for representations of affordances. This is not surprising, as representing them explicitly is actually against the ecological psychologist's interpretation of directly perceivable and usable affordances like it has always been postulated by Gibson. While such a view on affordance without representation and reason has as well been picked up in the area of robotic systems, e.g. by [19], we dissent from this view, arguing for the advantages of a formalised affordance concept throughout the MACS project.

For the benefit of affordances for robotics, we will instead follow the line of argument of, for instance, MacDorman who justifies learning and the explicit recognition, and thus implicit representation, of affordances by stating:

> "It is not surprising that Gibson underestimated the computational complexity of vision, since he wrote before researchers had begun to explore it seriously. [...] Thus, the brain may need to process sensorimotor data extensively and to spend time learning what kinds of invariance are useful in recognizing affordances." ([39], p. 1003)

We are furthermore convinced that it makes indeed sense to reason about affordances rather than acting directly upon an affordance percept. This point has been picked up by [14], too as they explicitly argue that an agent does not merely respond to a directly perceived stimulus by applying the action that is afforded in that situation. It is not controlled by the environment. It can rather use the information provided by the affordances of a situation and reason about them in a goal-directed manner selecting those afforded actions that will lead to its goal.

Now MACS has attempted to define an explicit, symbolic affordance representation on which the whole architecture and all its various components are based. Some of the following ideas were introduced in [33] and re-used in [38]. The overall idea, however, is primarily inspired by the work of Chemero [3] who first described an affordance as a perceivable relation between an agent and its environment or, as we interpret it, between the subjective capabilities of an agent and the features of its surroundings. We extended this idea by introducing the definitions of an (agent) affordance and an affordance representation (Def. 1 & 2, Sec. 4.1b).

Regarding Def. 2 of the affordance representation, one can understand its attributes as features of the environment or even internal states of the robot, while the values are not restricted to distinct values but can also represent value ranges.

The *cue descriptor* holds that piece of processed or raw sensor information, which supports the existence of the represented affordance, whereas the *outcome descriptor* contains the data as it was perceived by the robot while previously executing the behaviour referenced in the *behaviour descriptor*. That descriptor, on the other hand, refers to a robot behaviour and a set of parameters that were used with this robot behaviour when the respective cues and outcomes were monitored.

To summarise this definition, an affordance is represented by:

- The *cues* for an affordance that support it. These are the perceivable features or attributes of the environment for the agent and their values or value ranges. Attribute value pairs stored in a cue descriptor can thus be, for instance, the relative distance to a test object, its colour, or the different currents sent to the robot's motors.
- The *behaviour* descriptor refers to the behaviour or sequence of behaviours the robot has applied when this representation was created. To stick with the last example, this would be a lift action combined with the parameters like motor current or crane movement that were used for the particular action.
- The *outcome* of any action or behaviour executed upon the affordance. The outcome represents the changes of the agent and the environment, as far as they can be perceived by the agent. For example, a blue-coloured blob is being perceived at a higher position, relatively to the agent, if a lifting action has been performed.

The different affordance representation triples, which can both be hand-coded or learned (see Sec. 4.4), are then used during system runtime to build up and maintain a world model of the robot's surroundings that is represented as an affordance map. The different map regions hold the information whether a particular affordance type has been perceived in that area.

Given such a representation, it is left to describe how affordances can be exploited for robotics by reasoning about them. One approach has been attempted in the MACS project, namely, to ground plan operators by means of affordances; other types of reasoning with or about affordances are envisable. Here is the idea behind operator grounding: Assume the robot being in a situation where it has some goal to achieve, but one does not actually care about how or by which means to reach it. For example, it might want to weigh down a pile of paper; this could be done by putting a rock, a cup, or a book on that pile, achieving the same effect with each of these items. In other words, the robot may select just *any* item that affords the weighing action—the decision which concrete item to chose can be deferred to execution time, relying on affordance perception to re-identify the desired cue. On the other hand, if the plan currently under execution advises to watch out for a cue for the weighing-down affordance, it would lead to ignoring for the moment other affordances, which would not appear to lead to the desired goal as planned.

The MACS planning system is based on a complete domain and problem definition specified in the Planning Domain Definition Language (PDDL) [40]. The planner's world model contains knowledge of where what kind of affordance has been perceived . The planner uses the recorded availability of an affordance in a certain region of the environment to plan an action in that region (cf. Fig. 19). Consider the example that the robot has to open a door by putting some weight on a switch. The generated plan will be a sequence of operators to drive to a region where the liftability affordance has been perceived, to lift some liftable item there, to drive to the switch and put the item, whatever it may be, on the switch. The plan will thus contain a lift operator that gets implemented or grounded only during the execution phase of the plan. The robot simply has to select an affordance representation triple that belongs to the type of the liftability affordance and whose cues of its cue descriptor can currently be perceived. By acting as specified in that triple's behaviour descriptor, the robot implicitly chooses the next available liftable item and lifts it—be it a rock, a cup, or a book.

The deliberation part of the MACS architecture, as specified in [41], thus reasons about affordances in the sense that it selects goal-directedly the kind of affordance to act upon; i.e. it decides to use the liftability affordance rather than, for instance, pushability. So, the the action focus induced by some plan that is currently under execution helps the robot control selecting among the potentially very many affordances that may rush upon the robot at any time. Affordance representations may help plan-based robot control, on the other hand, by offering a direct way of grounding symbolic operators, as used in action plans, in physical execution.

It is an obvious topic for further research to explore more closely this interplay between affordance-based and plan-based robot control. In this context, it will also be useful to reintroduce objects (individuals as well as object classes) into the architecture, in addition to affordances. Imagine yourself dining with the Queen: you will desperately wish to distinguish between her glass and your glass rather than act upon drinkable-from affordance. So the problems of object detection and anchoring [89] will not go away, and it remains to be investigated how affordance representations and object representations interact. To start with, they may support each other, as, e.g., sitability is a property typically expected to be perceived of an object of the class chair. However, the connection in detail within a robot control architecture is unexplored.

Another line of future research is using more expressive plan representation languages in the planner. We have used a very simple propositional variant of PDDL here. A more expressive logical language, namely, TAL, has in fact been introduced into the project. It remains to be seen where and how this more expressive language could be brought to bear in affordance-based control—an obvious candidate for a planning system being TALPlan in this case.

4.3 Perception of affordances

In the context of ecological perception, as it was created by J.J. Gibson [1], visual perception would enable agents to experience in a direct way the opportunities for action. However, J.J. Gibson remained unclear about how this concept could be used in a technical system. Neisser [42] replied to this concept with the notion of a perception-action cycle that shows the reciprocal relationship of the knowledge (i.e., a schema) about the environment directing exploration of the environment (i.e., action), which samples the information available for pick up in the environment, which then modifies the knowledge, and so on. This cycle describes how knowledge, perception, action, and the environment all interact in order to achieve goals. Our work on affordance-like perception is in the context of technical systems based on a notion of affordances that 'fulfil the purpose of efficient prediction of interaction opportunities'.

In the project MACS we provided a refined concept of affordance perception by proposing two processing stages in terms of a predictive module, an interaction and an evaluation module (cf. Sec. 4.1b). Affordance-like perception aims at supporting control schemata for perception-action processing in the context of rapid and simplified access to agent-environment interactions. This framework is completely novel in the frame of affordance perception, in particular, concerning the learning of the selection of specific cues to predict opportunities for interaction.

a. *Related work*

Previous research on affordance based perception focused on heuristic definitions of simple feature-function relations to facilitate sensor-motor associations in robotic agents. Human cognition embodies visual stimuli and motor interactions in common neural circuitry (Faillenot et al. [43]). Accordingly, the affordance-based context in spatio-temporal observations and sensor-motor behaviours has been outlined in a model of cortical involvement in grasping by Fagg and Arbib [44], highlighting the relevance of vision for motor interaction [45]. Reaching and grasping involves visuo-motor coordination that benefits from an affordance-like mapping from visual to haptic perceptual categories (Wheeler et al. [46]). Within this context, the MIT humanoid robot Cog was involved in object poking and proding experiments that investigate the emergence of affordance categories to choose actions with the aim to make objects roll in a specific way (Fitzpatrick et al. [27]). The research of Stoytchev [30] analysed affordances on an object level, investigating new concepts of object-hood in a sense of how perceptions of objects are connected with visual events that arise from action consequences related to the object itself. Although this work innovatively demonstrated the relation between affordance triggers and meaningful robot behaviours, these experiments involve computer vision still on a low level, and do not consider complex sensor-motor representation of an agent interaction in less constrained, even natural environments.

Affordance based visual object representations are per se function based representations. In contrast to classical object representations, functional object representations (Stark and Bowyer [47], Rivlin et al. [48]) use a set of primitives (relative orientation, stability, proximity, etc.) that define specific functional properties, essentially containing face and vertex information. These primitives are subsumed to define surfaces and form the functional properties, such as ‘is sitable’ or ‘provides stable support’. Bogoni and Bajcsy [49] have extended this representation from an active perception perspective, relating observability to interaction with the object, understanding functionality as the applicability of an object for the fulfilment of some purpose. However, so far, function-based representations were basically defined by the engineer, and not learned from interaction. Fig. 5 depicts a schematic view about the embedding of affordance based perception into general frameworks of vision-based recognition, reflecting our view that it is determined by both function based and purpose vision.

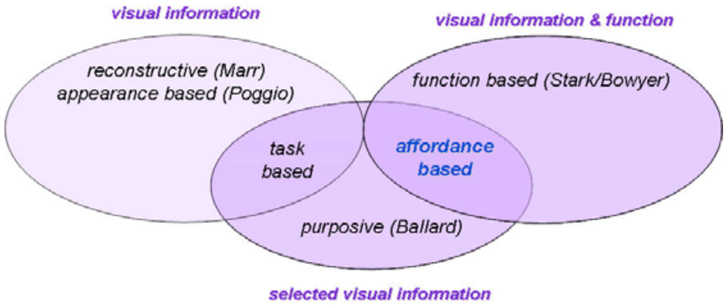


Fig. 5 Overview on the embedding of affordance based perception into various classical frameworks on vision based recognition. The function based approach relates to affordance based recognition in terms of its association of features with functions. However, in contrast to our approach it does not explicitly take the outcome of the function into account. Purposive vision in general relates goals with a selection of features, but as a global concept it does not consider the specific predictive structure including cues, behaviours and outcomes into account. Task based recognition simply does not explicitly relate parts of the environment with extracted features but can be understood to provide a global not local strategy to select features. Affordance based perception takes advantage of a predictive structure to relate the purposive selection of features with the association of functions, i.e., interactions.

b. *Stages in affordance perception*

We developed a refined concept on affordance perception [50] by proposing (i) an interaction component (affordance recognition: recognising relevant events in interaction via perceptual entities) and (ii) a predictive aspect (affordance cueing: predicting interaction via perceptual entities). This innovative conceptual step enables firstly to investigate the functional components of perception that make up affordance-based prediction, and secondly to lay a basis to identify the interrelation between predictive features and predicted event via machine learning technology ([50]–[53]).

Fig. 6 illustrates the various stages within the affordance based perception process for the example of the affordance *fill-ability* in the context of the opportunities for interaction with a coffee cup. Fig. 6(a) schematically illustrates the detection of perceptual entities that would provide affordance cues in terms of verifying the occurrence of a cup that is related to the prediction of being fill-able in general. Fig. 6(b) shows in analogy entities that would underlie the process of interaction of an agent with the cup by actually filling it up. Finally, Fig. 6(c) represents the entities corresponding to the final state of the interaction with the outcome of a successfully filled coffee cup. These figures illustrate that affordance cueing and affordance recognition must be conceptually separated and would involve different perceptual entities in general. While affordance recognition actually involves the recognition of the interaction process and its associated final state, affordance cueing will be solely determined by the capability to reliably predict this future event in a statistical sense.

(a) (b) (c)

Fig. 6 Affordance recognition in affordance based perception for the example of the affordance fill-able with respect to the impact of selecting appropriate features. The seemingly simple interaction of filling up a coffee cup can be partitioned into various stages in affordance based perception, such as, (a) affordance cueing by predictive features that refer to a fill-able object, (b) identifying perceptual entities that represent the process of the affordance related interaction (e.g., flow of coffee), and (c) recognizing the final state by detecting perceptual entities that represent the outcome of interaction (e.g., level of coffee in cup).

Fig. 7 depicts the innovative concept of feature based affordance perception worked out in the MACS project (cf. Sec. 4.1, Def. 2). We identify first the functional component of *affordance recognition*, i.e., the recognition of the affordance related visual event that characterizes a relevant interaction, e.g., the capability of lifting (lift-ability) an object using an appropriate robotic actuator. The recognition of this event should be performed by identifying a process of evaluating spatio-temporal information that leads to a final state. This final state should be unique in perceptual feature/state space, i.e., it should be characterized by the observation of specific feature attributes that are abstracted from the stream of sensory-motor information.

The second functional component of *affordance cueing* encompasses the key idea on affordance based perception, i.e., the prediction aspect on estimating the opportunity for interaction from the incoming sensory processing stream. In particular, this component is embedded in the perception-action cycle of the robotic agent. The agent is receiving sensory information in order to build upon arbitrary levels of feature abstractions, for the purpose of recognition of perceptual entities. In contrast to classical feature and object recognition, this kind of recognition is *purposive* in the sense of selecting exactly those features that efficiently support the evaluation of identifying an affordance, i.e., the perceptual entities that possess the capability to predict an event of affordance recognition in the feature time series that is immediately following the cueing stage of affordance based perception. The outcome of affordance cueing is in general a probability distribution on all possible affordances, providing evidence for a most confident affordance cue by delivering a hypothesis that favours the future occurrence of a particular affordance recognition event. This cue is functional in the sense of associating the related feature representation with a specific utility with respect to the capabilities of the agent and the opportunities provided by the environment, thus representing *predictive features* in the affordance-based perception system.

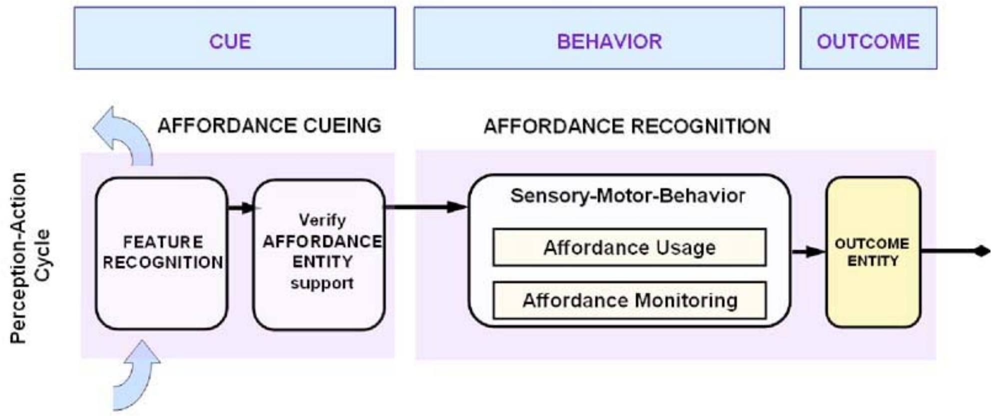


Fig. 7 Concept of affordance perception, depicting the key components of affordance cueing and recognition embedded within (left) an agents perception-action cycle (cf. Sec. 4.1b, Def. 2).

c. *Implementation: Perception Module*

The *perception module* includes an *Entity Structure Generation Module* (Fig. 4, Fig. 8, cf. also [33]) that generates appropriate data structures from sensory data in a framework of entity structures. Starting from simple structures (e.g. raw sensory data) these data structures are processed via transformation and/or combination into more abstract ones, describing the scene (e.g. regions of different colours and their relation) as well as affordances (e.g. regions with attributes like liftable, traversable, etc.).

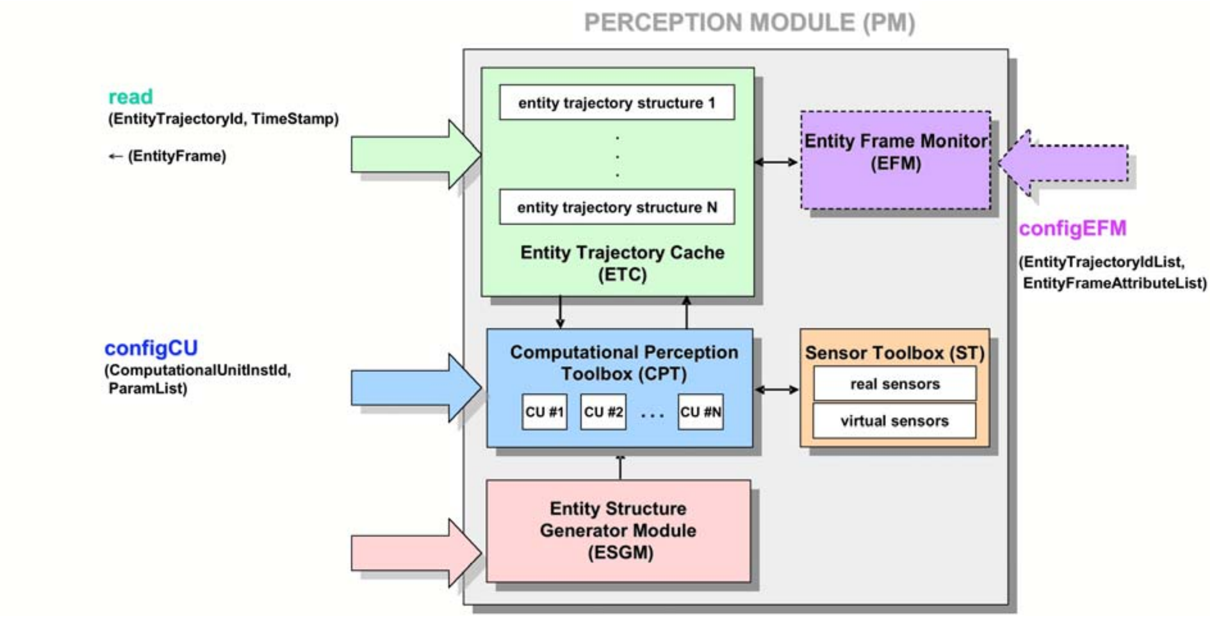


Fig. 8 System architecture of the Perception Module, depicting individual functional components and interfaces to other modules of the overall MACS system architecture (Fig. 4). Entity structures are generated in the ESGM. Each computational unit within the Computational Perception Toolbox (CPT) provides functions that output time stamped feature attributes which in turn are represented in the Entity Trajectory Cache (ETC) and can be read out from any other module.

The concept of *computational units* is employed to process these structures within an overall abstraction hierarchy. Computational units use *Entity Trajectory Streams* (i.e. series of entity structures over time) as input and produce entity trajectory streams as output. These entity trajectory streams provide input for the learning module, which learns suitable combinations of computational units for affordance cueing. For an example, several entity trajectory streams are combined in a final computational unit that classifies a particular region in the camera images into 'liftable' or 'non-liftable'. This classifier is encapsulated in the concept of the computational unit, with the benefit of a clear interface to other modules in the architecture of the system.

For finding salient locations that might be interesting during the robot's learning and mission phases, we employ a visual attention system called VOCUS [54]. VOCUS allows 'bottom up' detection of salient features in the environment as well as a 'top-down' search for certain features related to affordance cues [55]. The VOCUS system was also enhanced to work with two cameras in order to allow a triangulation of the position of salient regions relative to the robot. In order to accelerate VOCUS and to reduce CPU workload, it has been re-implemented such that it can run on a GPU. The latter variant can compute foci of attention at 60 fps, i.e. it can detect salient regions in both camera images at frame rate with little CPU usage [56]. This frees the CPU for other control tasks.

VOCUS is employed in terms of a *curiosity drive* for perceiving interesting locations. It computes foci of attention based on a saliency measure applied to elementary features like colour, brightness and orientation contrasts. The feature vector describing a salient location in an image is also provided as a computational unit in the Computational Perception Toolbox, i.e., to constrain the area in the image where affordance cueing is processed. The output of VOCUS can also be used by the learning module.

Fig. 9 shows sample results using the VOCUS attention system for the localization of test objects in the MACS scenario. VOCUS is here applied in bottom-up mode in the left camera image and, using the features from the attended region, in top-down mode in the right camera image in order to focus on a corresponding part of the environment. Then a colour region in the basin of attraction of the focus of attention is determined with a mean shift colour segmentation algorithm. Using triangulation the position of the coloured test object can be estimated with a distance error of ≈1cm within a maximum distance of 2 m in the demonstrator scenario. The estimated 3D position is then used to approach the object and proceed with affordance cueing and lifting behaviour to verify the affordance hypothesis if needed.

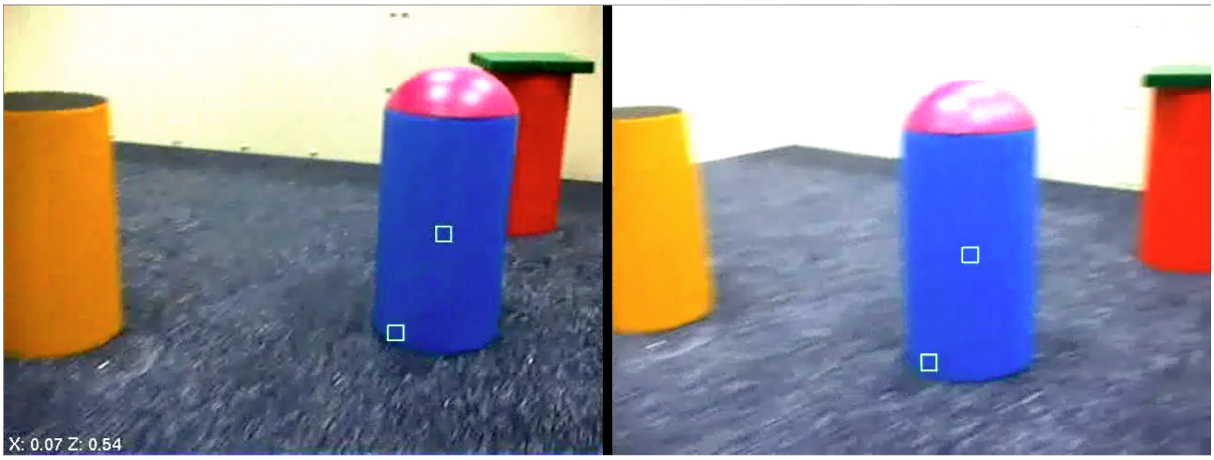


Fig. 9 Application of the VOCUS visual attention system for the estimation of 3D coordinates of objects in the MACS scenario. The left camera image shows the focus of attention (FOA; left square) as a result of the VOCUS bottom-up mode and a spatially associated colour region (right square). The system uses then the features of the FOA in order to find a corresponding FOA in the right camera image (left square) and the associated region. Triangulation using both FOA determines then the 3D coordinates for further execution of robot behaviours.

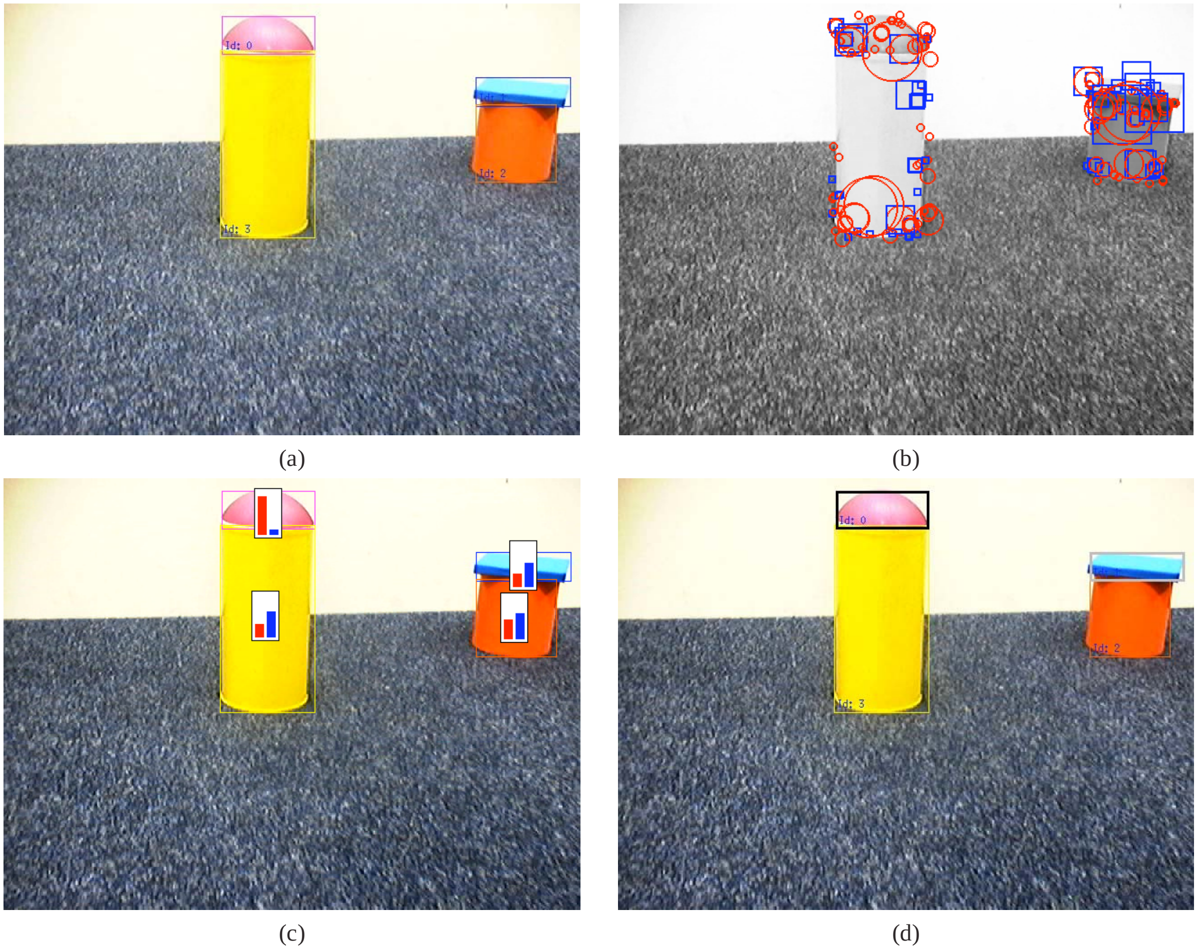


Fig. 10 Visual processing stages in affordance cueing. (a) Firstly, colour blobs are detected and the identity of the corresponding regions of interest are tracked using the KLT tracking methodology. (b) In each region of interest, local SIFT descriptors are extracted and classified into intermediate feature abstractions (circular and rectangular features or indicators). (c) For each region of interest, normalized voting provides then confidences (colour, region size, etc.) that will be classified by means of a decision tree. (d) Classification of single regions into affordance cues (bold and gray coloured bounding box) and no cues (bold and black bounding box; only top object regions are classified in this example).

Fig. 10 shows the application of local (SIFT) descriptors for the characterisation of regions of interest in the field of view. For this purpose, we first segment the colour based visual information within the image, and then associate integrated descriptor responses sampled within the regions to the region feature vector. The integration is performed via a histogram on local descriptors that are labeled with ‘rectangular’ and ‘circular’ attributes, respectively.

d. Learning of affordance cues

There are affordances that are explicitly innate to the agent through evolutionary development and others that have to be learned [1]. Learning the chains of affordance driven actions can lead to learning new, more complex affordances (cf. Sec. 4.4). In contrast to previous work on functional feature and object representations, we stress the fact that functional representations must necessarily contain purposive features, i.e., represent perceptual entities that refer to interaction patterns and thus must be selected from an existing pool of features by means of machine learning.

In this context we demonstrated the learning of causal relationships between visual cues and predictable interactions, using both 3D and 2D information ([51],[52]). We verified the concept with a concrete implementation applying state-of-the-art visual descriptors [57] and regions of interest that were extracted from a simulated robot scenario and prove that these features were successfully selected for their relevance in predicting opportunities of robot interaction by means of decision trees [58].

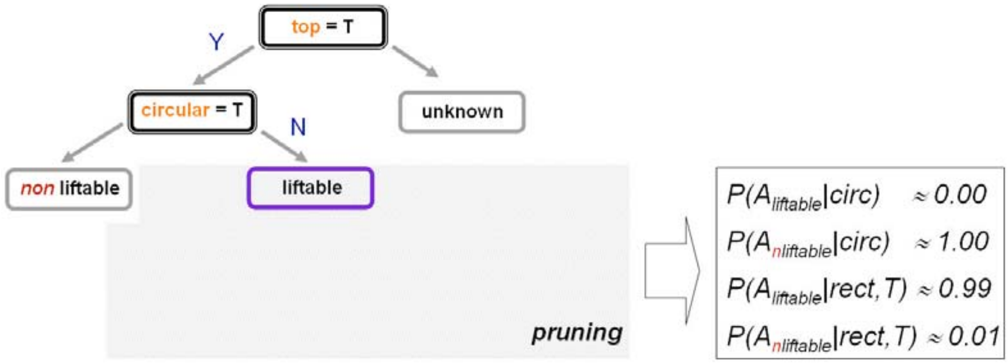


Fig. 11 Decision tree as classifier for affordance cueing. The outcome of the learning stage is the prioritizing of the attributes 'top region' (region geometrically on top of another region) and the focus on rectangular (i.e., not circular) features.

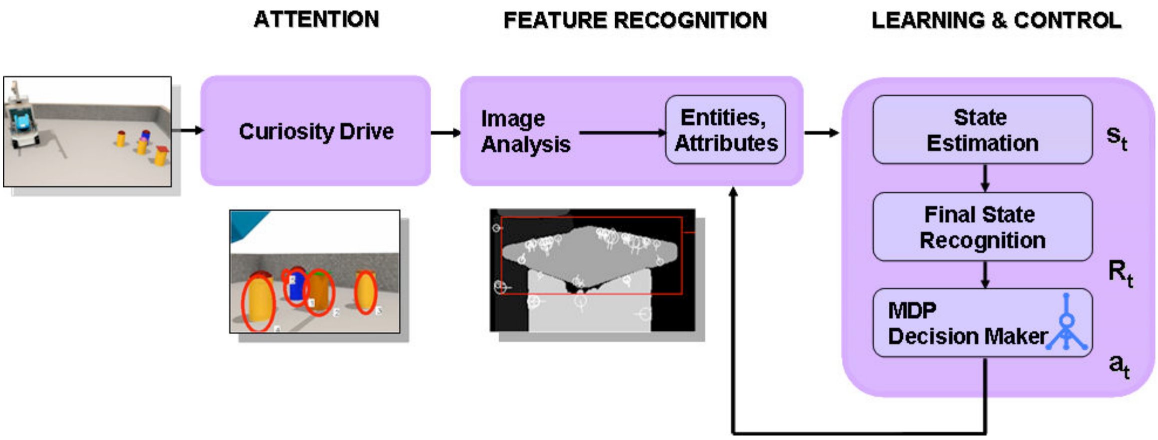
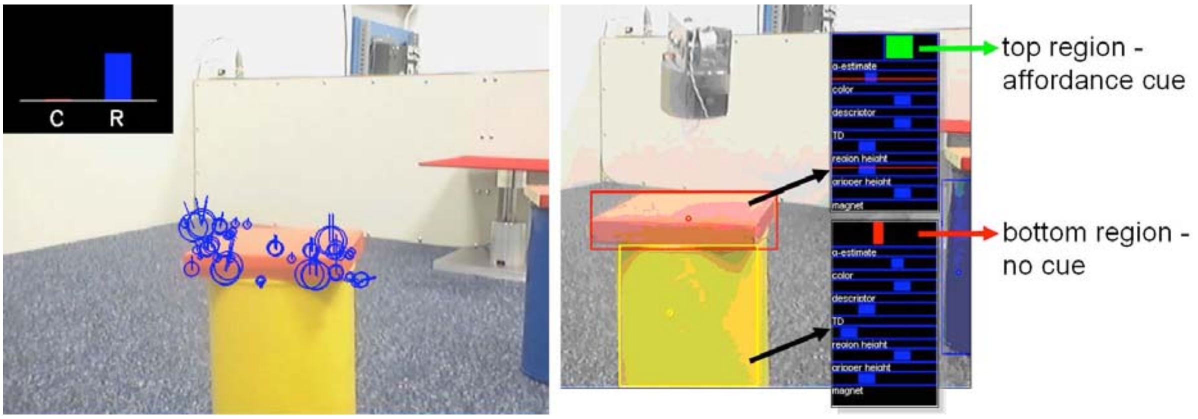


Fig. 12 Schematic diagram of closed-loop learning of affordance cues. The visual attention system VOCUS provides a focus of attention on objects in the environment, features are extracted and contribute to a multi-sensor state description. Based on the rewarding of approaching the goal state (e.g., lifting an object), the system can determine in an exploratory learning phase how to apply robotic actions (move, move magnetic gripper, etc.) in order to arrive at the goal. However, the affordance cue is associated with a perceptual state that is able to predict as early as possible the opportunity to interact and finally to arrive at the goal of the interaction, e.g., to end up with a lifted object.

Additional work was done in the direction of extending the scope of predictability via visual cueing using reinforcement learning [53]. Reinforcement learning ([59],[60]) as an on-line version of Markov decision processes (MDP) [61] is able to determine a specific perceptual state that owns the predictive characteristics for the representation of an affordance-like visual cue. The learning process is applied to bridging two basic components characterising the interaction component, i.e., affordance recognition, and the predictive aspect, i.e., affordance cueing, respectively. [62] presents the underlying theory and the experimental results from a robotic system scenario demonstrating how affordance recognition can provide the reinforcement signal to drive the propagation of reward information back in the affordance perception process. Upon convergence of the stochastic learning algorithm, we are able to identify an early perceptual state that enables to discriminate the capability to predict a future interaction opportunity with high confidence.



(a) (b)

Fig. 13 Affordance based cueing of region determined perceptual states from learned predictive cumulative rewards. (a) On the basis of a colour blob detector, local descriptors are classified into rectangular/circular (R/C) ones and the associated histogram feeds here into the recognition of an affordance cue (with respect to lift-ability). (b) Analysed top and bottom regions are correspondingly classified as cues for lift-ability or non-lift-ability, visualized in terms of green and red bars with bar sizes correlating to positive or negative reward, respectively (monitoring boxes, top), anticipating a lift-able event.

There is huge potential in research on affordance perception towards extending the feature based representations towards object driven affordance-based interaction, grounding the work on the visual descriptor information presented here. Furthermore, the learning of affordance cues can be viewed in the frame of developmental learning of meaningful sequences of affordance triplets [63], opening a broad avenue for future research.

4.4 Learning of affordances

The learning approach that was developed within the MACS project is an approach to acquire knowledge about relations that determine the interaction possibilities between an agent and its environment. Within this approach an artificial agent starts with basic interactions and uses more and more complex interactions over time and thus gathers experience about what happens before, during and after these interactions. These experiences are generalized by the agent, enabling it to act also in novel situations. Therefore the robot used starts with an initial set of reflex like actions and is designed to be able to deal with a growing set of (learned) actions. Thereby the approach is not limited to a special kind of actions.

The set of basic reflex-like actions shall enable the robot to stack building blocks. Whether two objects can be stacked or not dependents on the top region surface of the element that should provide the base, and depends on the bottom region surface of the element that should be stacked on this base. The two surfaces must be in a certain relation to each other for a successful stacking trial. In simple cases the necessary complementary shape is given over the entire top and bottom regions. More complicated objects may only share some of those complementary regions, but at least enough to keep a stacking element grounded on the base.

When objects are provided to an agent, the relevant surfaces cannot be perceived directly. Nevertheless humans are able to assume whether an object is stackable or not, without seeing this surface. They do it by using several cues based on their own experience to fulfil this task.

Consequently, in the MACS project, affordances within a robotic system are represented by relations between *cues*, *behaviours*,and *outcomes*.The space of learned affordances is thus a multi-relational repository from which *cue-behaviour-outcome* triples can be derived. To be able to extract these triples is not only crucial for learning by self-experience and for planning but also for learning by imitation to match observed *cue* and *outcomes* to previously made self-experience. That means that triples, which are *1:1:1* relations, are derived from that *o:m:n* relations database.

The *cues* and *outcomes*, their inter-relations as well as the relation to the causing actions are learnt from the incoming perceptual data stream. For a detailed description of the learning approach see paper *Learning of Interaction Possibilities* in this volume.

The schema in Fig. 14 depicts the key components of the developed *Learning Module* that is connected to and interacts with the overall affordances based architecture (see Sec. 4.1). The image shows which modules are required and how they are interconnected to realize the required data and control flow. The depicted modules, the used data structures, and the data / control flow are described in the following sub-sections.

a. Application Spaces Module

The agent applies actions to the environment. While doing that the agent permanently monitors its environment and the internal states before, during and after the application. The sets of resulting time series are stored within behaviour specific *Application Spaces* in the *Application Spaces Module* to be available for the learning processes. The begin and the end of the application of the actions must be marked within each stored time series. To be able to learn a cue for the existence of an affordance and the concerning outcomes (consequences of using an affordance), the recorded time series have to include data from a certain time interval before and after the application of the behaviour.

During the learning process, the *Application Spaces* are divided into partitions. This partitioning information, i.e. a *Partitioner Object*, as created by the *Partitioning Module* (see Sec. 4.4b), is stored in relation to its corresponding application space.

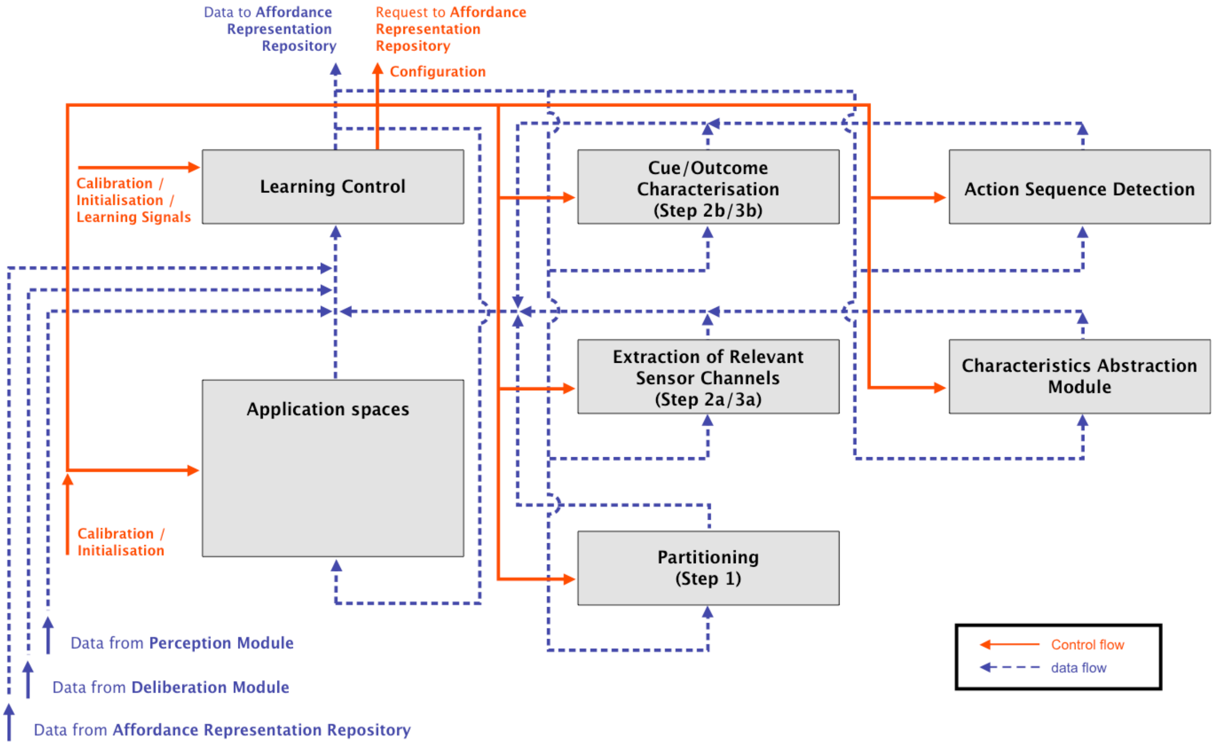


Fig. 14 The key components of the developed learning architecture, its modules, and the data (dashed lines) and control flows (solid lines) between the components.

The relevant sensor channel information that is extracted during the learning process (see Sec. 4.4c) as well as the characterization of these sensor channels (see Sec. 4.4d) are to be stored within the *Application Spaces* as well.

b. Partitioning

In the application space of an action sets of similar action application results should exist after a sufficient number of trials. For example in case of the application space of action “close gripper”, the following subsets could emerge:

- a set of results where the involved objects were gripped,
- a set of results where the objects slipped away,
- a set of results where the objects were not grippable at all.

As input the *Partitioning Module* receives the data of an Application Space. This *Application Space* contains a set of time series, resulting from several applications of the same action. The *Partitioning Module* provides a mechanism to discriminate these different types of action application results from each other.

The output of the *Partitioning Module* is a *Partitioner Object* that is stored and linked to the belonging application space. The *Partitioner Object* provides a function to decide for a given time series to which partition it belongs. A partition is thus defined as the set of all time series that are mapped to the same partition identifier.

When there is already a *Partitioner Object* assigned to an *Application Space* and the agent acquired new experiences concerning the related behaviour (new time series are stored within the *Application Space*) the partitioner must be adapted to these new experiences, if they deviate from the previous made ones. This re-learning process could change the *Partitioner Object* and thus could cause previously recorded time series to change their partition. It is also possible that new partitions emerge.

c. Relevant Sensor Channel Extraction

In the next step sensor channels are extracted from the time series of each partition of an application space that are representative for these partitions. Representative means, that these channels are in direct relation with the differing cues and outcomes, respectively. In case of performing an action that causes lifting an object, the partitions resulting from liftable and non-liftable objects will differ in the height (y-position time series) and (optionally) in the force sensor channel. For learning cues in this given example, the channel of the colour blob filter (y-position of the blobs) could be representative for the partitions. As input, the *Relevant Sensor Channel Extraction Module (RSCEM)* receives the data of an *Application Space* that is already partitioned.

To find characteristic channels concerning cues, the same process is applied to the pre-application part of the recorded time series of a partition. These partitions are sub-partitioned further.

As output, the *RSCEM* provides

- for each partition a set of channel identities of the relevant sensor channels concerning the cue events for the Application Space
- for each partition a set of channel identities of the relevant sensor channels concerning outcome events for the Application Space

These sets are stored in the relevant application space.

When there is already a *Partitioner Object* assigned to an Application Space and the two sets of relevant sensor channels, and the agent acquired new experiences in applying the related action (new data time series are stored within the *Application Space*) an adaptation of the set of relevant sensor channels for the partitions could be necessary in the case, that e.g.

- the previously gained knowledge was incomplete (new environmental configurations, e.g. new objects occurred)

- the configuration of the sensors or actuators has changed (e.g. broken or altered because of growth or enhancement)
- the partitioning has changed (e.g. partitions altered or new partitions emerged, see Sec. 4.4b).

d. Event Characterization

After the extraction of the relevant sensor channel(s), descriptions of what is characteristic for the relevant channel(s) of the partitions are to be derived, i.e. *cue characteristics* and *outcome characteristics*. These characteristics are used to enable the agent to recognize affordances (in case of characterizing cue related channels) or to monitor the outcome of the application of an action (in case of characterizing outcome related channels).

As input, the *Event Characterizer Module (ECM)* receives data of an *Application Space* in which the sets of relevant sensor channels for each partition (derived by the *Relevant Sensor Channel Extraction Module*, section 6.3) are stored.

As output the *Event Characterizer Module* provides

- for each partition a set of cue characteristics for the *Application Space*,
- for each partition a set of outcome characteristics for the Application Space.

Similar considerations to above with respect to new experiences apply.

e. Characteristics Abstraction

The task of the *Characteristics Abstraction Module (CAM)* is first to find similarities between the elements of a given set of outcome characteristics. Two or more characteristics could share a subset of characteristics, e.g. two different outcomes (different ball trajectories) of two behaviours applied to a ball (beating and kicking) share the characteristics, that the ball is moved and that the space in front of the agent is free after the behaviour application. On an abstract level of observation, looking at these examples regarding the 'change location' characteristic and neglecting the different time series that occur, both action applications and the corresponding outcomes are equal.

The described abstraction process, and the storage of this gathered abstracted outcome information in the *Application Space*s, enable an artificial agent to treat two or more actions as equal, with regard to the expected outcome of applying these actions in the context of the related cue characteristics. Regarding the above mentioned example where two different behaviours (beating and kicking) are applied on the same object (a ball), the two different behaviours are equal for reaching the outcome described by the derived abstracted 'change location' characteristic.

Additionally, objects or entities can be treated as equal, regarding the outcome that occurs by applying such *equal actions*. Even if objects or entities do not share visual features, they can be treated as equal under the context of applying those *equal actions* and gaining the abstracted outcome characteristics.

The outcome of the characterization process together with the extraction of abstracted characteristics by using similarity measurements on the level of characteristics (done by the *CAM*) enables the agent to measure similarities between entities on the abstract level of functionality. Thus the agent is enabled to achieve a level of semantic similarity measurement based on its perceptional similarity measurement abilities together with its behavioural experiences, which provides one method of solution to the complex problem of semantic similarity measurement in robotics [64].

4.5 Formalisation of affordances

In agreement with Chemero [3], we view affordances as relations within an ecology of acting, observing agents and the environment. Our starting point for formalising affordances is:

> **Definition 1 ((Agent) Affordance).**[4] *An affordance is a relation between the agent and its environment as acquired from the interaction of the two.*

Based on this definition, an affordance is said to be a relation that can be represented as

$$(\textit{environment}, \textit{agent}).$$

However, this formalism is too generic to be useful, and needs to be refined. As Chemero also asked in his formalisation, “which aspect of the environment is related to which aspect of the organism (agent), and in what way?” Therefore in this relationship, the environment and the agent should be replaced with “environmental relata” and “agent’s (organismal) relata” (as in Chemero’s terminology), to indicate the relevant aspects of the two. First, we use the term, *entity*, to denote the environmental relata of the affordance instead of *features* (as used by Chemero) or *object* (as generally used). In our formalism, *entity* represents the proprioceptive state of the environment (including the perceptual state of the agent) as perceived by the agent. The term *entity* is chosen since it has a generic meaning that is less restricting than the term *object*. Although for some affordances the term object perfectly encapsulates the environmental relata, for others, the relata may not be confined to an object and may be more complex. Second, the agent’s relata should represent the part of the agent that is generating the interaction with the environment that produced the affordance. Ideally, the agent’s relata should consist of the agent’s embodiment that generates the perception-action loop that can realize the affordance. We chose the term *behaviour* to denote this. In Autonomous Robotics, a *behaviour* is defined as a fundamental perception-action control unit to create a physical interaction with the environment. We argue that this term implicitly represents the physical embodiment of the interaction and can be used to represent the agent’s relata. Third, the interaction between the agent and the environment should produce a certain *effect*. More specifically, a certain *behaviour* applied on a certain *entity* should produce a certain *effect*, e.g. a certain perceivable change in the environment, or in the state of the agent. For instance, the *lift-ability* affordance implicitly assumes that, when the *lift behaviour* is applied to a *stone*, it produces the effect *lifted*, meaning that the *stone*’s position, as perceived by the agent, is elevated.

Based on these discussions, we refine our first definition as:

> **Definition 3.** *An affordance is an acquired relation between a certain effect and a (entity, behaviour) tuple, such that when the agent applies the behaviour on the entity, the effect is generated.*

and our formalisation as

$$(\textit{effect}, (\textit{entity}, \textit{behaviour})).$$ [5]

This formalisation explicitly states that an affordance is a relation which consists of an *(entity, behaviour)* pair and an *effect* such that there exists a potential to generate a certain *effect* when the *behaviour* is applied on the *entity* by the agent. In this formalism, we assume that this relation resides within the interacting agent. This means that all three components are assumed to be sensed by the agent. The *behaviour* denotes the executed perception-action routine that generated the interaction as sensed by the agent. The *entity* refers not to an abstract concept of

---

[4] Definition repeated from Sec. 4.1b.

[5] Within the MACS project, it was agreed to use this formalisation with a change of terminology in the form of: *(cue, behaviour, outcome)*, where *cue* and *outcome* denoted <*entity*> and <*effect*>.

an entity (such as a stone) but to its perceptual representation by the agent. Similarly, the *effect* refers to the change inflicted in the environment (including changes in the state of the agent) as a result of the *behaviour* acting on the *entity* as perceived by the agent.

The proposed formalisation, with its explicit inclusion of *effect*, can be seen as a deviation from J.J. Gibson's view at its outset. It is not. In J.J. Gibson's writings, the issue of effect had always remained implicit. For instance in the definition of the *lift-ability* affordance, the expected effect of *lifted* is implicitly present. Similarly, this has been implicitly included in Chemero's formalism where he named the relation as *Affords*-$\frown$ to exclude the instances that did not produce the affordance. On the other hand, in Turvey and Stoffregen's formalisations, the desired effect is represented as h and r respectively. The proposed formalisation is different from these, by not only making it explicit, also putting it on a par with the *entity* and the *behavior*.

Affordances should be relations with predictive abilities. We will propose four aspects through which multiple instances of interactions can be bound together towards discovering affordances.

Entity Equivalence

The class of *entities* which support the generation of the same *effect* upon the application of a certain *behaviour* is called an *entity equivalence class.* We would like to note that the concept of *entity equivalence* is related to the concept *invariance*, defined as "persistence under change" in broad terms by J.J. Gibson. He mentioned the concept in many contexts through his book and devoted one section in the Appendices to it. These invariants correspond to the properties which remain constant under various transformations, i.e. invariants of optical structure under changing illumination or under change of the point of observation. Although J.J. Gibson did not explicitly define these invariances, he gave some clues about the perception and usage of them.

> ". . . There must be invariants for perceiving the surfaces, their relative layout, and their relative reflectances. They are not yet known, but they certainly involve ratios of intensity and colour among parts of the array." (J.J. Gibson, 1979/1986, p. 310)

Behaviour Equivalence

The concept of affordance starts with equidistance to perception (through the entity in the environment) and action (through behaviour of the agent). Yet the role of action is often less pronounced than the role of perception, and most of the discussions concentrate on the perception aspect of affordances. We argue that, if we wish to maintain a fair treatment of the action aspect of affordances, then the same equivalence concept should be generalized to that aspect as well. For instance, our robot can lift a can using its lift-with-right-hand behaviour. However, if the same effect can be achieved with its lift-with-left-hand behaviour, then these two behaviours are said to be behaviourally equivalent. We would like to note that, similar to the entity equivalence, the use of behavioural equivalence will bring in a similar flexibility for the agent. Through discovery of the perceptual invariants of an entity equivalence class, the agent gains the competence to use a different entity to generate a desired effect, even if the entities that had generated the effect in the past are not present in its environment. Such a 'change of plan' is directly supported by the entity equivalence classes. A similar competence is gained through behavioural equivalence classes. For instance, a humanoid robot which lifted a can with one of its arms, loses its ability to lift another can. However, through behavioural equivalence it can immediately have a "change of plan" and accomplish lifting using its other hand.

Affordance Equivalence

Taking the discussion one step further, we come to the concept of *affordance equivalence*. Affordances like traversability, are obtainable by "walking across a road" or "swimming across a river"

Effect Equivalence

The concepts of entity, behaviour and affordance equivalence classes implicitly relied on the assumption that the agent, somehow, has effect equivalence. For instance, applying the lift-with-right-hand behaviour on a blue-can would generate the effect of "a blue blob rising in view". If the robot applies the same behaviour to a red-can, then the generated effect will be "a red blob rising in view". If the robot wants to join the two relation instances learned from these two experiments, then it has to know whether the two effects are equivalent or not. In this sense, all the three equivalences rely on the existence of effect equivalence classes.

We propose that an affordance can be formalised as:

$$(<\textit{effect}>,<(\textit{entity}, \textit{behaviour})>).$$

This formalism represents affordance from an agent's perspective. We will make this perspective explicit, and revise our definition as:

> **Definition 4 (Affordance (agent perspective))**. *An affordance is an acquired relation between a certain* $<$effect$>$ *and a certain* $<$(entity, behaviour)$>$ *tuple such that when the agent applies a* (entity, behaviour) *within* $<$(entity, behaviour)$>$, *an* effect *within* $<$effect$>$ *is generated.*

Different from the previous version of the definition, this one explicitly states that affordance is a *relation* between *equivalence classes*, rather than a *relation instance* between an *effect* and a *(entity, behaviour)*.

4.6 Simulator and Demonstrator

In this section, we describe the physical demonstrator and the physics-based simulator MACSim. The physical demonstrator consists of a demonstrator environment and a six-wheeled mobile robot with a simple crane arm manipulator, named KURT3D. The robot is available both in MACSim, and as a physical system (four units).

a. Physical demonstrator

The main elements of the physical demonstrator are a mobile robot, KURT3D, an experimentation arena called the demonstrator environment, and test objects for perception and manipulation experiments.

The MACS version of the KURT3D mobile robot platform consists of the KURT2 base platform, the KURT3D sensory enhancements, the MACS rack and a newly developed crane manipulator. The KURT2 base platform is a six-wheeled mobile robot platform of roughly one by one foot width and depth, and eight inches height. The robot has three wheels on each side, which are connected by a tooth-belt. Per side, a single DC motor drives all wheels via the tooth-belt. The drives and other low-level functions are controlled via a C167 and a TMC 200 controller board and special firmware. These microcontrollers are connected via CAN bus to an on-board notebook computer that runs the high-level control programs under Linux. The standard sensory equipment consists of tilt sensors and a number of distance transducers along the perimeter of the robot.

The KURT3D configuration consists of two additional enhanced sensor systems: a 3D Laser scanner and a stereo pan-tilt camera system, which both were developed at Fraunhofer IAIS.

An additional rack has been mounted on top of the robot in order to support a reversible notebook mount and the MACS crane arm (Fig. 15).

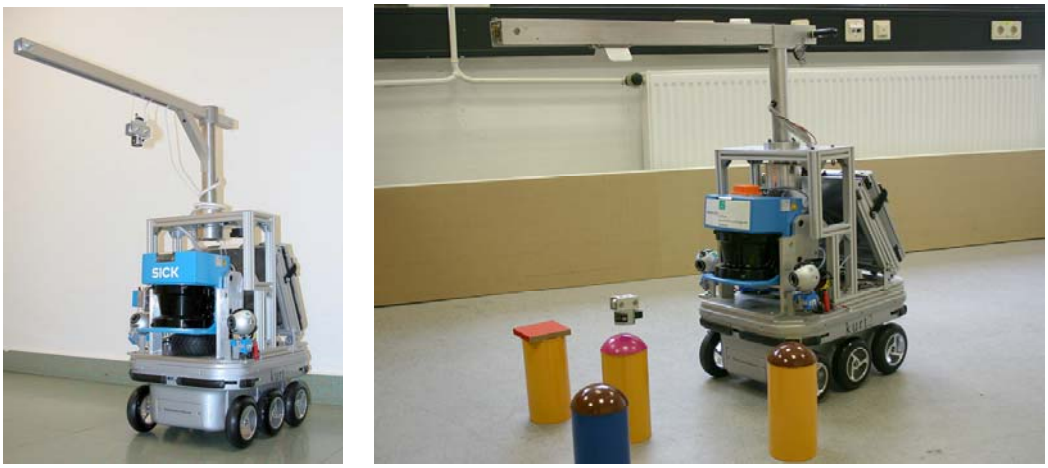


Fig. 15 The mobile robot KURT3D, equipped with a crane arm manipulator and a magnetic gripper.

The crane arm has three degrees of freedom. The arm itself can be rotated around a vertical axis. A small lorry can be moved horizontally along the crane's extension arm, and a magnetic griper can be raised and lowered along a rope that is hanging from the movable lorry. This construction allows most simple manipulation tasks, namely trying to 'grip' items in the environment with the electromagnetic gripper and lift them.

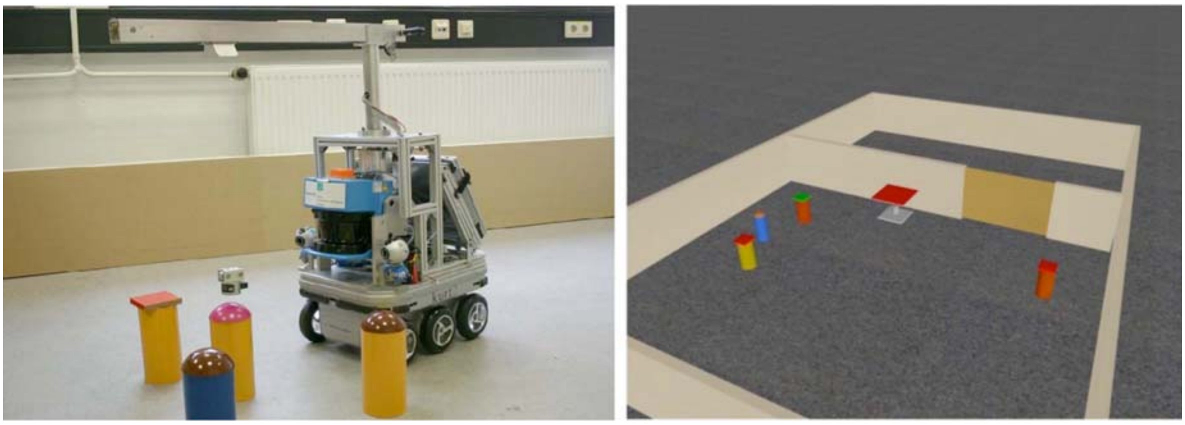
The demonstrator environment setup consists of a defined mission area with the dimensions of 2.5m x 3.5m. The area is surrounded by walls that are 40cm high and 5cm wide, made of heavy and robust wooden elements (Fig. 16(a)). As first passive elements to be manipulated by the crane arm we use tin cans with different colours, sizes and top designs. Some of them are magnetizable, some are not. Their weight can easily be altered by butting in heavy material at any time.

(a) (b)

Fig. 16 (a) The physical test arena, called demonstrator scenario, including a separating wall, a sliding door, and a (red) switch to open the door. (b) The switch can be operated by putting a weight on it, which will open the door. Removing the weight will close the door. This particular set-up has been chosen to train the robot to observe effects of its manipulation actions.

The demonstrator scenario contains also active components: a movable dividing wall with a motor driven sliding door that can divide the mission area into two separate rooms. The door can be opened and closed via a switch. The switch is operated depending on the weight put on it by the robot (Fig. 16(b)). The switch has a weighing area of 25cm x 25cm, is working on a high sensitive pressure sensor (strain gauge element) and is adjustable to trigger on weights between 15g and 7kg. This particular setup has been chosen to train the robot to observe effects of its manipulation actions.

The demonstrator environment and the test objects have been constructed both physically and in simulation (Fig. 17).



(a) (b)

Fig. 17 (a) Simple experimentation environment, showing robot KURT3D (FhG/AIS). (b) Total view of the demonstrator environment in MACSim.

b. Simulator MACSim

MACSim (Fig. 18) is a high fidelity simulation environment that models the KURT3D robotic platform and its environment. Built on top of a commercial quality open-source engine, ODE[6] (Open Dynamics Engine), MACSim accurately simulates the objects, robot parts, and their dynamics in a 3D world.

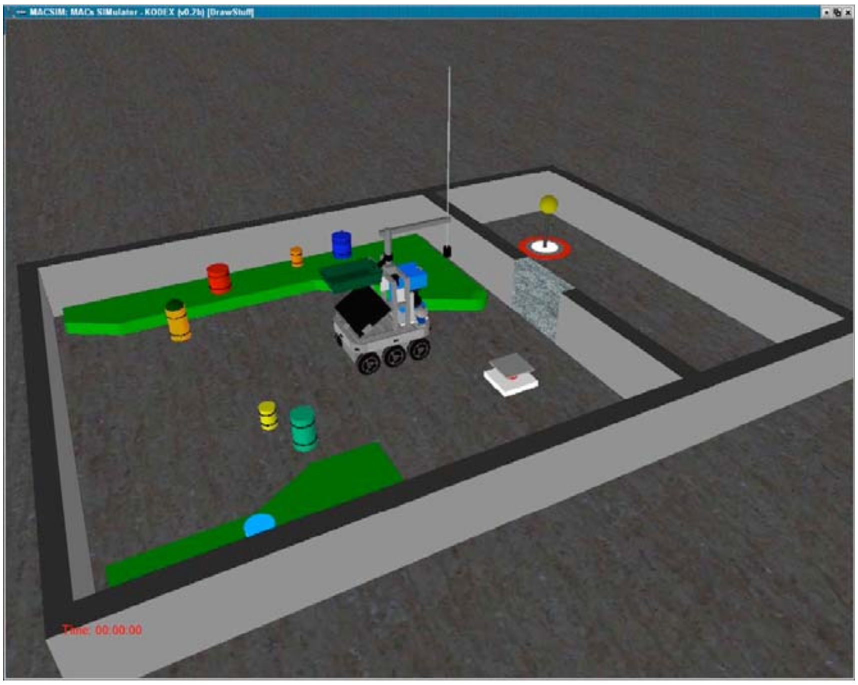


Fig. 18 A snapshot from MACSim where the KURT3D robot is modelled in an environment which is created for the demonstrator scenario.

The simulation model of our mobile robot provided in MACSim closely matches the real KURT3D robot in many aspects. Based on their physical properties, such as mass, size, and centre of mass, all parts that constitute the robot are modelled as rigid bodies. Later, junction locations of these components were measured, and they were assembled with appropriate joints to acquire the complete simulated robot. In order to simulate different actuators of the

[6] http://ode.org

robot, such as wheel systems or camera servo motors, the joints are virtually constrained and motorized with the parameters obtained from the real robot.

Realistic sensor modelling is also very crucial, since robot actions and control rely on the robot's perception of the world. While ODE provides excellent support for modelling rigid body dynamics based on laws of physics, similar to many low level engines, virtual sensors are not explicitly supported. For example, there is no ready-to-use acoustic signal or infra-red beam that could be sent or received. For laser scanner and infrared proximity sensors, ODE's ray geometry and collision detection routines are utilized, and ray intersection method is used. For colour cameras, OpenGL's back-buffer data is employed. Moreover, in order to close the gap between reality and simulation, sensor and actuator parameters are calibrated, based on the "same setup experiments" in virtual and real worlds.

The reality of the simulator is further verified in ([66]–[67]), where the robot controllers trained in the simulator are successfully transferred to the real robot. For example, in [66], a large set of training data (approximately 3000 samples) obtained from interactions of the robot with its environment is required to learn the perception of traversability affordances. MACSim is utilized in a training phase to decrease the time and cost of the learning process and to remove any risk of physical damage that might occur on the real robot. It is later shown that the robot is able to perceive the same affordances offered by the environment when encountered with same situations either in simulated or real world. Moreover, the physical effects created in the simulator and real world are compatible when the robot executes a certain action in that particular situation.

4.7 Proof-of-concept: Experimental results

a. Introduction

We have designed a demonstrator scenario and sketched a number of proof-of-concept experiments [65] that are suited to demonstrate the novelty of the MACS approach. The robot experimented with a variety of test objects and learned cues for the presence of certain functions or affordances. Cues in this sense may be invariants across a range of test objects with different appearances. The robot should use these test objects for different tasks, like operating a weight-sensitive switch for opening a door, or for freeing its way by pushing away pushable test objects. The final challenges were the use of new test objects that offer the same functions but have different appearances than the test objects that have been employed in the initial learning phase, and difficulties like taking away test objects that the robot needs for executing a plan.

Separate real-world experiments have been conducted with all modules of the architecture (perception of affordances, learning of affordances, planning using affordances, basic skills). Also, various experiments have been conducted in the simulator MACSim.

In this section, we provide an overview of the experiments we have conducted and which are described in several publications.

b. Experiments

The final demonstrator scenario has been employed for the proof-of-concept of the MACS approach to affordance-inspired robot control. In this scenario, our robot demonstrated the capabilities of its integrated affordance-inspired control architecture, where the planner creates affordance related operators using affordance maps (Fig. 19), and tasks related to learning and goal-directed use of affordances are performed based on these operators. However, for demonstrating and evaluating the benefits of our new robot control approach we do not rely on the final demonstrator alone. Instead, we decided to provide also proofs of concept for several significant steps during the development process. For this purpose, we defined a framework for selected experiments allowing us to analyse the performance of our approach in an isolated and well defined way, starting with simple tasks and increasing the complexity step by step towards the final demonstrator scenario ([65]–[68]). In the phase of the project where the basic skills and perceptual feature detectors necessary for performing affordance-related tasks were developed, specific experiments were performed to prove the explicit support for our affordance concept.

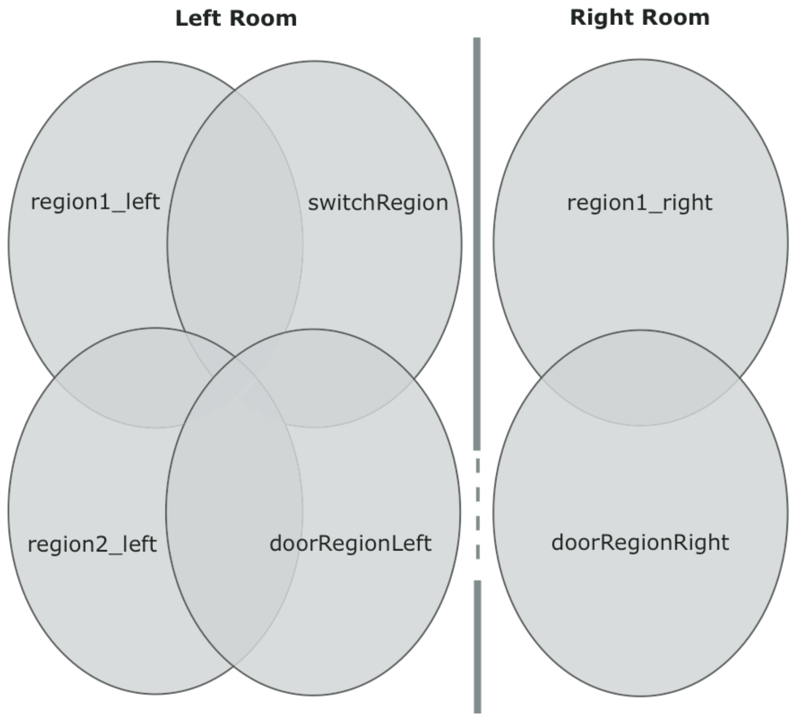


Fig. 19 The world (cf. Fig. 18) is separated into several regions. If the robot perceives an affordance within one of those regions the affordance type is added to the map. Plans are made based on the constructed affordance map.

For introducing affordance support into robot perception, we first examined the generation of a traversability map based on laser scanner input and a pre-programmed classification of traversable areas in the environment [33]. Next, we developed feature filters based on SIFT descriptors that enable ‘top down’ and ‘bottom up’ detection and classification of simple image features. This enables the robot to distinguish features in the top, body and bottom regions of test objects in the environment. These filters work equally well on simulated and real camera data ([69]–[72]).

A desired basic skill of our robot is the autonomous exploration of its environment. One question here was: Based on simple features alone, how can the robot find interesting spots in its environment that potentially contain items that it can manipulate, and thus enable it to learn from its actions. For this purpose, a special variant of the visual attention system VOCUS ([54], cf. Sec. 1.1) has been successfully employed. For driving towards the salient location, the robot’s Behaviour System provides basic navigational skills, namely driving through free space that is computed based on data of the 3D Laser scanner. The combination of basic navigational skills with salient region detection by VOCUS enables the robot to explore its environment by selecting potentially interesting areas and driving towards them until they are in range of the robot’s manipulator. We have informally named this combination ‘curiosity drive’ behaviour. For the creation of an appropriate Behaviour System, we also performed experiments in autonomous navigation and map generation with basic support for affordance information ([38],[73]).

Based on these concepts and results, the next development phase was dedicated to the integration of the architecture components and to the introduction of learning mechanisms. The latter ones are mainly employed to determine the descriptive feature sets that are either cues for a certain affordance, or descriptors of the outcomes of the robot's acting upon an affordance. The results of reinforcement learning experiments of predictive cues in affordance-based perception were presented in ([62],[53],[51],[52]). Learning mechanisms for environmental cues needed for perceiving the traversability affordance were demonstrated in ([74],[66]) and used in 'curiosity drive' experiments in [67]. Here, we tried to formally define the 'curiosity' notion beyond saliency measures by introducing an SVM based measure that helps the robot to decide whether an interaction possibility is worthwhile exploring (cf. also below). In this phase, we have used extensively both the MACSim simulator and synthetic data for our developments and experiments related to learning, which was beneficial.

In our recent experiments we accomplished the transition of primitive behaviours to goal-directed behaviour by using learnt behaviour-effect relations and situation awareness to achieve more complex behaviours ([66],[76]–[78]). In this study, the robot interacts with its environment by executing a set of primitive behaviours and collecting interaction samples. Based on these experiences, the robot discovers the different effects it can create in the environment, and associates an observed effect with the primitive behaviours and environmental situations that resulted in this effect. The robot then uses the learnt relations to achieve more complex behaviours. In our experiment, we used three primitive behaviours (turn-left, turn-right, and move-forward) and the learnt affordance relations of these behaviours to achieve three different goal-directed behaviours (traverse, approach, and avoid). Since the robot learns the affordance relations from its own experiences, it is not trivial for a human observer to specify a goal that 'makes sense' to the robot. One solution is to use as goal descriptors effect prototypes that can be learned from a range of similar observed effects. As an evaluation criterion, priorities can be assigned to learnt effect prototypes. This enables the robot to select and execute a primitive behaviour that would result in an effect similar to the goal, i.e. to the effect prototype having the highest priority. The results of this study are sketched in Fig. 20.

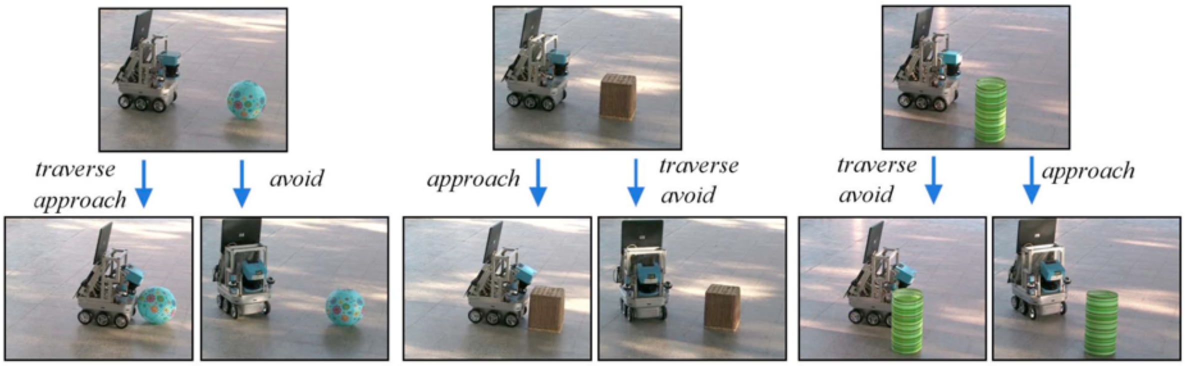


Fig. 20 Three cases in which different goal-directed behaviours (traverse, avoid, approach) make use of different primitive behaviours (move-forward, turn-right, turn-left) in the same setting of the environment (source: [75]).

The learned affordance relations can also be used as operators for planning, since they provide the capability to predict the effects of behaviours as discussed in the context of *cue-outcome based planning* [41]. In a recent study [79], we used these predictions to generate totally ordered plans which are composed of sequences of primitive behaviours. Forward chaining is used for this purpose. The robot starts with perceiving the present entity, and predicts the effects that each of its (five) primitive behaviours will create. Next, it estimates the five future entities that the robot will perceive after execution of corresponding behaviours by summing up the predicted effects and current entity. The robot then proceeds by predicting the effects of behaviours on those future entities and estimating next entities. This process can be viewed as the breath-first construction of a plan tree where the branching factor is the number of behaviours. Planning stops when any future entity or total predicted effect of the behaviour sequence satisfies the desired goal. Fig. 21 shows a number of sample plans generated using learned affordance relations for different goals in various environments.

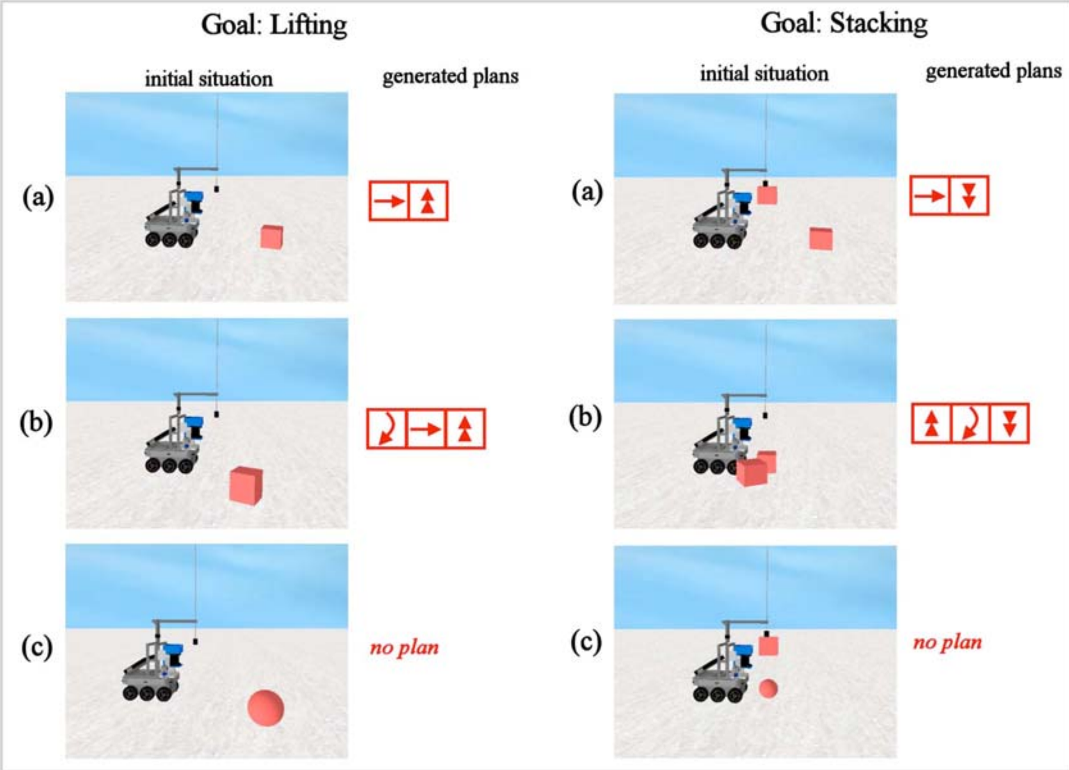


Fig. 21 The generated plans are shown as sequences of primitive behaviours. For lifting and stacking tasks the goals are defined as increase in the crane rope tension and decrease in the distance features on middle grids of the scan image, respectively. The primitive behaviours used in these experiments are turn-left, turn-right, and move-forward, lift, and release. (source: [80]).

In another series of experiments, we studied the learning of traversability affordances and investigated how the required number of interactions with the environment can be minimized with minimal degradation on the learning process. Specifically, we applied a two step learning process which consists of bootstrapping and curiosity-based learning phases. In the bootstrapping phase, a small set of initial interaction data were used to find the relevant perceptual features for the affordance, and were used to train a Support Vector Machine classifier. In the curiosity-driven learning phase it was determined whether a given interaction opportunity is worth exploring or not [67] (see Fig. 22 below).

Fig. 22 This sequence shows how the perception of traversability affordance are used to navigate in blocked situations. The initial position of the robot is shown in the left-most figure. The robot first goes forward, then turns left since trash-bin does not afford traversability. Third snapshot shows the robot driving over the spherical object. The path of the robot is shown in the last figure. (source: [67]).

The effects of two parameters of our learning system that serve as the curiosity threshold and the number of bootstrap samples were examined in systematic experiments. Selecting a small threshold keeps the system away from interacting with interesting situations, and selecting a large one slows down learning, since uninteresting samples are used for training. As for the number of bootstrap samples, small values degrade the performance of the system and large values beyond a certain threshold do not improve the performance. The affordance perception system, trained using optimized parameters, was tested in our scenario cluttered with objects of varying shapes. In this environment, the robot was able to predict the traversability affordances of the objects and wander around the room.

The trained controller was transferred to the real robot and was also successful in predicting the traversability affordance of real world objects. Further experiments will make use of the completely integrated robot control architecture until the full complexity of the final demonstrator scenario tasks will be achieved.

5 OVERALL CONCLUSIONS

5.1 Progress beyond State-of-the-Art

a. Formalisation

We believe that the proposed formalism [37] has laid out a good framework over which the concept of affordance can be utilized for autonomous robot control. The review of the affordance concept and its use in different fields, with special emphasis on Autonomous Robotics has provided a complete snapshot of the confusion surrounding the concept and through the formalisation has provided its implications at different aspects of robot control ranging from learning to planning. This is clearly a progress beyond state-of-the-art since the affordance theory has mostly been used as a source of inspiration in robotics. Most of the studies reviewed preferred to refer to J.J. Gibson’s original ideas as formulated in his books, ignoring modern discussions on the concept. As a result, only certain aspects of the theory have been used, and no attempts to consider the implications of the whole theory towards autonomous robot control have been made.

The affordance-based robot control architecture has provided a proof-of-concept implementation of how the formalism can be implemented and used in real-world. The experimental results obtained from partial implementations of the architecture has also provided convincing examples that the approach generates results that are beyond the current state-of-the-art in the field.

b. Representation for deliberation and planning

AI has envisaged for a long time robots acting as aided by symbolic plans, which are generated on the fly by state-of-the-art action planners. However, plan-based robot control has turned out to be difficult; one bag of problems associated with it is anchoring of objects in sensor data and of plan operators in physical action at execution. Using affordances as a first-class citizen in the control architecture has shed some new light upon the problem. Whenever the planner's symbolic domain representation can be made to include actions that correspond to affordances (like, e.g., a ‘lift’ action in the MACS scenario, derived from the ‘liftability’ affordance), then grounding its execution comes practically for free. The same is true for representations of affordances that appear in the preconditions of symbolic actions: Determining their validity at execution means to attend to the respective affordance.

As a result from our study, it turns out that affordance-based and plan-based control appear to be easy fellows in a joint robot control architecture. This is not too surprising, given that they both deal with action, under different perspectives. We are not aware, however, that this has been explored in robotics.[7]

[7] The planning literature includes some hints towards that, e.g., in work about situated action [90].

c. Perception

Affordance perception has been considered so far only in the frame of functional object recognition and in the context of action based perception. A key aspect that has been neglected in previous approaches is the novel type of generalisation that refers to the interaction of an agent with parts of its environment. Visual features should not be selected by the engineer but should be extracted for their relevance to predict opportunities for action. For this purpose we considered in MACS numerous types of 2D and 3D features and demonstrated that machine learning methodology can determine a meaningful selection of features that generalise across various appearances while they at the same time characterise a specific physical property which can be exploited in the predicted interaction. A particular contribution in MACS is not only to argue but also to verify that affordances are not associated to objects—a notion which is derived from an ontology that is grounded in human perception—but can be related to arbitrary chunks of information, such as visual regions, that possess a predictive nature in estimating future outcomes of an agent's interaction with the world and its physical properties. The predictive nature of affordance cues requires to perform feature selection with respect to the goal-directedness of behaviours, the relevance of the outcome: irrelevant outcomes of actions are negligible for the agent and therefore negligible for perception. Consequently we introduced a framework of feature selection and extraction in the context of the utility of actions through the association of behaviours to reward functions, rendering classification of features impacted by utility instead of identity as it is done in classical recognition frameworks.

The investigations performed in MACS contributed to pave the way for a new understanding of machine perception, moving away from manual labelling of image databases towards the autonomous emergence of symbols that are grounded in sensory-motor interactions with the real world and the associated relevance for the agent.

d. Learning

The learning approach that was developed within MACS and used in the Learning Module of the architecture is an approach to acquire knowledge about relations that determine the interaction possibilities between an agent and its environment. Affordances are thus represented in form of entity-action-outcome triples and thus refer to both, the agent (its capabilities) and its environment, as initially claimed by Gibson. In our work we describe how an artificial agent, starting with basic interactions and using more and more complex interactions, acquires this knowledge by using its gathered experience. These experiences are generalized by the agent, enabling it to act also in novel situations.

The approach is a general framework in which unsupervised learning methods as well as reinforcement learning methods are utilized. Hence the approach does not depend on a special learning mechanism and is thus open to future developments. Statistical approaches as well as traditional and novel self-organizing algorithms (as neural networks, Multi-SOMs [92], and ROLF [93]) are used.

The developed learning approach clearly addresses and defines how an artificial agent can gain the necessary knowledge required for the purposeful usage of affordances and the combination of different learning methods within this approach instead of using a monolithic learning method, which is novel compared to the state of the art. The approach starts with an initial set of reflex like actions and is designed to be able to deal with a growing set of (learned) actions. Thereby the approach is not limited to a special kind of actions.

Therefore the MACS affordance learning approach closes the developmental gap between an ‘infant agent’ and an agent that is able to use affordances purposefully. This provides a strong basis for using affordances in robotics to gain a more flexible and robust goal driven behaviour in artificial agents; and since the approach is a general concept that can be used to enable a wide range of artificial agents to learn affordances within its environment, it is not limited to a special scenario or agent. The final software implemented within MACS, demonstrated through a number of experiments, shows the viability of the approach in practice.

e. Architecture

We would like to note that before MACS, the affordance theory of J.J. Gibson was mostly used either as a source of inspiration for isolated features of control in autonomous robotics, or the concept of affordances was used to merely describe the behaviour of a robotic system. Consequently, only certain aspects of the theory were used, and no attempts to consider the implications of the whole theory towards autonomous robot control were made. In this sense, the development of a complete affordance-inspired robot control architecture that is designed to learn, detect, and use the affordances in the environment is an important contribution of MACS to the field.

5.2 MACS Objectives and Results

Although other researchers occasionally attempted to relate affordances to robotics ([19],[23],[27],[28],[30]), it is valid to state that MACS was the first dedicated investigation of using the concept of affordances in all stages of robot control, in perception, learning and planning. MACS has achieved its general goal "to explore and exploit the concept of affordances for the design and implementation of autonomous mobile robots acting goal-directedly in a dynamic environment." In this section, we summarise the MACS results and their relation to the project's objectives.

As far as perception of affordances is concerned, the concept of *affordance cueing*, i.e. applying a trained 'matched filter' when looking for affordances, is considered to be an important project result. Affordance cueing provides *perceptual economy* for robot perception, since it uses only a relevant fraction of the available perceptual input. Affordance cueing has been implemented in two different ways, namely first by using reinforcement learning for training, and second by using Support Vector Machines for training classifiers (the 'matched filters').

Although it has not been an initial goal of MACS, it turned out in the first year of the project that we need a formalisation of the affordance notion in order to make use of it. Such a formalisation provides clear definitions that aid defining representations and implementing the control architecture. A first formalisation has been drafted in our deliverable D4.2.1+4.3.1. It provided first definitions of agent affordances and their representations. Later in the project, a second formalisation has been developed that introduced a formalisation of the equivalence concepts (entity, affordance, and outcome/effect equivalence) that are the basis for the proposed flexibility. This second formalisation has been published in the *Adaptive Behavior* journal, and thus MACS has also fulfilled its aim to embed its results into Cognitive Science.

An integrated affordance-inspired robot control system including perception module, behaviour system, execution control module, planner, learning module and affordance representation repository has been implemented, tested and used. The proof-of-concept has been shown in various experiments with the simulator MACSim and the real robot KURT3D.

The dissemination of project results is a standard goal for all EU research projects. MACS has produced 16 peer-reviewed publications, two edited books, five book contributions, four theses, two online articles and twelve posters.

Research results have been presented at all major Robotics conferences (IROS, ICRA, ICDL, SAB), other relevant conferences and workshops, and at seminars and invited talks. MACS has organized a Dagstuhl seminar dedicated to its central topic, affordances for robot control.

The general public has been addressed by several press releases that have been echoed in print media and on more than two dozens web sites. An Industry Day for informing interested companies about the research being performed in the EU's Cognitive Systems objective has been organized and conducted by MACS, with the support of the Coordination Action eu-Cognition. At the Industry Day, six projects, including MACS, have been presented to 47 attendees, half of them from industry. A detailed list of the project's dissemination activities can be found in Annex B2. All in all, MACS has achieved its dissemination objective.

5.3 Conclusions and recommendations for future work

The concept of affordances has a strong appeal, since it seems to be intuitively understandable and applicable to a variety of areas. Several groups and researchers have been inspired by the concept of affordances. Affordances have been used in design of human-computer interfaces, in the development of new approaches for robot control, in linguistics, and in investigations of human way-finding strategies in large man-made infra-structures [81].

In all these areas, the major problem for utilization is to find a model that is suitable for the particular usage or implementation of the affordance concept. One major difficulty for finding operational models of the affordance concept is the vast generality of J.J. Gibson's affordance definition which he simply defined for all 'animals'. The questions arose whether it is really applicable to beings as different as crickets and humans, and whether it is applicable to animals at different levels of individual development.

In this final report of the MACS project, we have presented a comprehensive approach to affordance-inspired robot control. The approach is based on our own formalisation and operational model of affordances in the context of Robotics. It is built on a representational concept, the affordance representation triples, consisting of cue–behaviour–outcome descriptors. Such representations can be generated during an initial learning phase by analysing the streams of basic and complex perceptual features and applying a three stage learning approach. This comprises the 'bottom-up' part in the proposed architecture. The 'top-down' part of the architecture foresees the use of affordances for deliberate action, i.e. mission planning and execution. Thus, affordance support is built into all components of the presented robot control architecture.

The learned representations are grounded in the robot's actions and perceptions, thus they 'make sense' for the robot. However, for a human observer, the general comprehensibility of the learned affordance representations currently remains an open problem. Schönherr [82] presented a solution for a similar case by letting a human assign a symbol to a set of features designating a 'situation'. A similar solution could be applied to assign 'meaning' to affordance representations.

The presented approach to affordance-based planning foresees the goal-directed selection of the proper kind of affordance to act upon. This first version of the planner uses PDDL to define the plan operators. A future version of an affordance-based planner could introduce new operators that use the full range of cue-behaviour-outcome descriptors.

Current state-of-the-art appearance-based robot perception approaches typically can handle only about 100 everyday objects (German BMBF project DESIRE). This is an impressively high number of objects in terms of the state-of-the-art, but it is rather low compared to the number of objects that a human can recognize or that typically occur in a human environment like a flat. So for a human environment, a limited set of known everyday objects clearly limits a robot's interaction possibilities. The affordance-inspired approach helps overcoming these limitations by introducing a different ontological slice based on functional categories as provided by affordance perception. Here, a set of features that are characteristic for a certain interaction possibility are bundled as a 'matched filter' that works for a large number of objects that may belong to several different appearance categories. For example, the affordance or interaction possibility to drink from a drinking vessel for a human is offered by objects with quite different appearances like glasses, china cups, tin cups, paper cups, mugs, jars etc. Similarly, our robotic agent may traverse along a path by pushing differently looking things out of its way. These things must afford ‘pushability’, defined by spherical form features or cylindrical form features with a certain orientation (composing the matched filter). Sizes may vary in a certain range, and texture is irrelevant. In this way, the robot may interact with objects of quite different appearances.

However, there are also many situations and tasks where (object) recognition capabilities are indeed required. Some tasks require the ability to identify and recognise objects and to distinguish individual objects. For example, humans want to distinguish a coffee cup from a water glass, and ‘my coffee cup’ from ‘your coffee cup’. Neisser [83] proposed an approach that includes both affordance-related perception and object recognition. To date this approach has not been realized in robot control either. Thus, as a long term research question, the interaction between affordance perception and object recognition seems to be worthwhile to pursue. Investigating the little explored affordance-inspired perception and control is a prerequisite for a combined system along Neisser’s considerations.

One of the partners of MACS project (METU-KOVAN) is now involved in the ROSSI project. Funded within the first call of FP7 by the Cognitive Systems and Robotics unit, ROSSI aims at understanding the link between affordances and concepts in the form of language constructs such as nouns and verbs. The proposed affordance formalisation has already provided the framework for concept formation and will be used in the ROSSI project.

6 ACKNOWLEDGEMENTS

The authors would like to thank the following current and previous project members: Andreas Bartel[5], Maya Çakmak[3,11], Mehmet Doğar[3,13], Patrick Doherty[6], Selda Eren[3], Simone Frintrop[1,7], Fredrik Heintz[6], Dirk Holz[8], Martin Hülse[1,9], Maria Klodt[1,7], Manish Kumar[2], Torsten Merz[6,12], Frank Meyer[5], Bernd S. Müller[1,14], Fatih Nar[3], Tommy Persson[6], Jens Poppenborg[5], Piotr Rudol[6], Christin Seifert[2], Hartmut Surmann[1], Göktürk Üçoluk[3], Björn Wingman[6], Rainer Worst[1], Mariusz Wzoreck[6], M. Yavas[3] and Keyan Ghazi-Zahedi[1,10].

This work was partly funded by the European Commission’s 6th Framework Programme IST Project MACS under contract/grant number FP6-004381. The Commission’s support is gratefully acknowledged.

The authors would also like to thank Barbara Webb for connecting the concept of a ‘matched filter’ to affordance cueing at the Dagstuhl seminar 2006.

Affiliations:

[1]Fraunhofer IAIS, Sankt Augustin, Germany; [2]Joanneum Research, Graz, Austria; [3]Middle East Technical University, Ankara, Turkey; [4]Austrian Research Institute for Artificial Intelligence, Vienna, Austria; [5]University of Osnabrück, Germany; [6]University of Linköping, Sweden; [7]University of Bonn, Germany; [8]University of Applied Sciences Bonn-Rhein-Sieg, Sankt Augustin, Germany, [9]University of Wales, Aberystwyth, UK, [10]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [11]Georgia Institute of Technology, Atlanta, GA, USA, [12]CSIRO ICT Centre, Brisbane, Australia, [13]Italian Institute of Technology, Genova, Italy, [14]private/retired.

ANNEX A. REFERENCES

[1] J.J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA, 1979. also: 1986, Lawrence Erlbaum Associates, London.

[2] D.A. Norman. *The Psychology of Everyday Things*. Basic Books, New York, NY, 1988.

[3] A. Chemero. An outline of a theory of affordances. *Ecological Psychology*, 15(2):181–195, 2003.

[4] M. Steedman. Formalising affordance. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society*, pages 834–839, Washington D.C., 2002. Lawrence Erlbaum. Conference: Fairfax VA, August 2002.

[5] J.J. Gibson. *The senses considered as perceptual systems*. Houghton Mifflin, Boston, 1966.

[6] W.H. Warren. Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology*, 105(5):683–703, 1984.

[7] W.H. Warren and S. Whang. Visual guidance of walking through apertures: body-scaled information for affordances. *Journal of Experimental Psychology*, 13(3):371–383, 1987.

[8] L.S. Mark. Eyeheight-scaled information about affordances: A study of sitting and stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3):361–370, 1987.

[9] E.J. Gibson, G. Riccio, M.A. Schmuckler, T.A. Stoffregen, D. Rosenberg, and J. Taromina. Detection of the traversability of surfaces by crawling and walking infants. *Journal of Experimental Psychology*, 13(4):533–544, 1987.

[10] J.M. Kinsella-Shaw, B. Shaw, and M.T. Turvey. Perceiving walk-on-able slopes. *Ecological Psychology*, 4(4):223–239, 1992.

[11] A. Chemero. What events are. *Ecological Psychology*, 12(1):37–42, 2000.

[12] R. Oudejans, C. Michaels, B. van Dort, and E. Frissen. To cross or not to cross: The effect of locomotion on street-crossing behaviour. *Ecological Psychology*, 8(3):259– 267, 1996.

[13] R.C. Arkin. *Behaviour Based Robotics*. MIT Press, Cambridge, MA, 1998.

[14] A.P. Duchon, W.H. Warren, and L.P. Kaelbling. Ecological robotics. *Adaptive Behaviour*, 6(3):473–507, 1998.

[15] R.A. Brooks. Intelligence without representation. *Artificial Intelligence*, (47):139–159, 1991.

[16] R.A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2(1):14–23, 3 1986.

[17] D. Lyons and M. Arbib. A formal model of computation for sensory-based robotics. *IEEE Transactions on Robotics and Automation*, 5(3):280–293, 1989.

[18] R.C. Arkin and T. Balch. AuRA: Principles and practice in review. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2):175–189, 4 1997.

[19] R.R. Murphy. Case studies of applying Gibson's ecological approach to mobile robots. *IEEE Transactions on Systems, Man, and Cybernetics*, 29(1):105–111, 1999.

[20] J. Norman. Ecological psychology and the two visual systems: Not to worry! *Ecological Psychology*, 13(2):135–145, 2001.

[21] E.J. Gibson. Perceptual learning in development: Some basic concepts. *Ecological Psychology*, 12(4):295–302, 2000.

[22] U. Neisser. Multiple systems: A new approach to cognitive theory. *The European Journal of Cognitive Psychology*, 6:225–241, 1994.

[23] K.F. MacDorman. Responding to affordances: Learning and projecting a sensorimotor mapping. In *Proc. of 2000 IEEE Int. Conf. on Robotics and Automation*, pages 3253–3259, San Fransisco, California, USA, 2000.

[24] J.H. Connell. SSS: a hybrid architecture applied to robot navigation. In *Proceedings of the 1992 IEEE International Conference on Robotics and Automation (ICRA '92)*, volume 3, pages 2719–2724, Los Alamitos, CA, 1992. IEEE Computer Society. Conference, Nice, France, May 12–14, 1992.

[25] R. Cooper and D.W. Glasspool. Learning action affordances and action schemas. In R.M.French and J.P. Sougne, editors, *Connectionist Models of Learning, Development and Evolution*, Perspectives in Neural Computing, pages 133–142, London, 2001. Sixth Neural Computation and Psychology Workshop, Springer-Verlag.

[26] I. Cos-Aguilera, L. Canamero, and G.M. Hayes. Using a SOFM to learn object affordances. In *In Proceedings of the 5th Workshop of Physical Agents*, Girona, Catalonia, Spain, 3 2004.

[27] P. Fitzpatrick, G. Metta, L. Natale, S. Rao, and G. Sandini. Learning about objects through action – initial steps towards artificial cognition. In *Proceedings of the 2003 IEEE International Conference on Robotics and Automation, ICRA*, pages 3140–3145, 2003.

[28] A. Stoytchev. Toward learning the binding affordances of objects: A behaviour-grounded approach. In *Proceedings of AAAI Symposium on Developmental Robotics*. Stanford University, 2005. March.

[29] I. Cos-Aguilera, L. Canamero, and G.M. Hayes. Motivation-driven learning of object affordances: First experiments using a simulated Khepera robot. In *In Proceedings of the 9th International Conference in Cognitive Modelling (ICCM'03)*, pages 57–62, Bamberg, Germany, April 2003.

[30] A. Stoytchev. Behaviour-grounded representation of tool affordances. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pages 3071–3076, 2005. Conference: Barcelona, Spain, April 18–22, 2005.

[31] R.R. Murphy and R.C. Arkin. SFX: An architecture for action-oriented sensor fusion. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 1992*, pages 1079–1086, Raleigh, NC, July 1992.

[32] E. Rome, E. Şahin, R. Breithaupt, J. Irran, F. Kintzler, L. Paletta, M. Çakmak, E. Uğur, G. Üçoluk, M.R. Doğar, P. Rudol, G. Fritz, G. Dorffner, P. Doherty, M. Wzoreck, H. Surmann, and C. Lörken. Evaluation of existing control architectures for using affordances. Technical Report MACS/2/2.2 v1, Fraunhofer Institut für Intelligente Analyse- und Informationssysteme (IAIS), Sankt Augustin, Germany, 2006.

[33] P. Doherty, T. Merz, P. Rudol, and M. Wzorek. Tentative proposal for a formal theory of affordances; Tentative proposal for an affordance support architecture; Prototype: Affordance-based motion planner. Technical Report MACS/4/2.1 v1, Linköpings Universitet, IDA Group, Linköping, Sweden, 2005.

[34] M. Turvey. Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, 4:173–187, 1992.

[35] T.A. Stoffregen. Affordances are enough: Reply to Chemero et al. *Ecological Psychology*, 15(1):29–36, 2003.

[36] M. Steedman. Plans, affordances, and combinatory grammar. *Linguistics and Philosophy*, 25(5–6):723–753, December 2002.

[37] E. Şahin, M. Çakmak, M.R. Doğar, E. Uğur, and G. Üçoluk. To afford or not to afford: A new formalisation of affordances towards affordance-based robot control. *Adaptive Behaviour*, 15(4):447–472, 2007.

[38] C. Lörken. Introducing affordances into robot task execution. In K.-U. Kühnberger, P. König, and P. Ludewig, editors, *Publications of the Institute of Cognitive Science (PICS)*, volume 2, 2007. University of Osnabrück, Osnabrück, Germany, May 2007. ISSN 1610-5389.

[39] K.F. MacDorman. Grounding symbols through sensorimotor integration. *Journal of the RSJ (The Robotics Society of Japan)*, 17(1):5, January 1999.

[40] D. McDermott. PDDL – The planning domain definition language. Technical report, Yale University, 1998.

[41] C. Lörken and J. Hertzberg. A specification for a propositional planner and its interface to the MACS execution control module. Deliverable MACS/2/3.2, University of Osnabrück, Institute of Computer Science, Osnabrück, Germany, 2007.

[42] U. Neisser. *Cognition and Reality: Principles and Implications of Cognitive Psychology*. W.H. Freeman and Co, 1976.

[43] I. Faillenot, I. Toni, J. Decety, M.-C. Grégoire, and M. Jeannerod. Visual pathways for object-oriented action and object recognition: functional anatomy with pet. *Cerebral Cortex*, 7:77–85(9), January 1997.

[44] A.H. Fagg and M.A. Arbib. Modelling parietal–premotor interactions in primate control of grasping. *Neural Networks*, 11(7-8):1277–1303, 1998.

[45] M.G. Edwards, G.W. Humphreys, and U. Castiello. Motor facilitation following action observation: a behavioural study in prehensile action. In *Brain Cognition*, volume 53, pages 495–502, 2003.

[46] D.S. Wheeler, A.H. Fagg, and R.A. Grupen. Learning prospective pick and place behaviour. In *Proc. 2nd International Conference on Development and Learning*, pages 197–202, Cambridge, MA, June 2002. IEEE Computer Society.

[47] L. Stark and K.W. Bowyer. Function-based recognition for multiple object categories. *Image Understanding*, 59(10):1–21, 1994.

[48] E. Rivlin, S.J. Dickinson, and A. Rosenfeld. Recognition by functional parts. *Computer Vision and Image Understanding: CVIU*, 62(2):164–176, 1995.

[49] L. Bogoni and R. Bajcsy. Interactive recognition and representation of functionality. *Computer Vision and Image Understanding: CVIU*, 62(2):194–214, 1995.

[50] L. Paletta, G. Fritz, E. Rome, and G. Dorffner. A computational model for visual learning of affordance-like cues. In *Proc. 29th European Conference on Visual Perception, ECVP06*, St. Petersburg, Russia, August 2006.

[51] G. Fritz, L. Paletta, R. Breithaupt, E. Rome, and G. Dorffner. Learning predictive features in affordance-based robotic systems. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2006*, pages 3642–3647, Beijing, China, October 2006. Springer-Verlag.

[52] G. Fritz, L. Paletta, M. Kumar, G. Dorffner, R. Breithaupt, and E. Rome. Visual learning of affordance based cues. In *Proc. International Conference on the Simulation of Adaptive Behaviour, SAB*, LNAI 4095, pages 52–64, Berlin, Germany, September 2006. Springer-Verlag. Conference: Rome, Italy.

[53] G. Fritz and L. Paletta. Reinforcement learning for the selection of predictive cues in affordance-based perception. In *Proc. 29th European Conference on Visual Perception, ECVP06*, St. Petersburg, Russia, August 2006.

[54] S. Frintrop. *VOCUS: A Visual Attention System for Object Detection and Goal-directed Search*. LNAI 3899. Springer-Verlag, Berlin, Heidelberg, 2005. also: PhD thesis, University of Bonn.

[55] S. Frintrop, M. Hülse, E. Rome, and L. Paletta. Saliency detection with visual attention. Technical Report MACS/3/1.3 v1, Fraunhofer Institut für Autonome Intelligente Systeme, Sankt Augustin, Germany, 2005.

[56] S. May, M. Klodt, E. Rome, and R. Breithaupt. GPU-accelerated affordance cueing based on visual attention. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2007*, San Diego, CA, 2007.

[57] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.

[58] J.R. Quinlan. *C4.5 Programs for Machine Learning*. Morgan Kaufmann, San Mateo, CA, 1993.

[59] R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. MIT Press, Cambridge, MA, 1998.

[60] S.D. Whitehead and D.H. Ballard. Learning to perceive and act by trial and error. *Machine Learning*, 7(1):45–83, 1991.

[61] M.L. Puterman. *Markov Decision Processes*. John Wiley & Sons, New York, NY, 1994.

[62] L. Paletta and G. Fritz. Reinforcement learning of predictive features in affordance perception. In E. Rome, J. Hertzberg, and G. Dorffner, editors, *to appear in: Towards Affordance-based Robot Control – Proceedings of Dagstuhl Seminar 06231*, volume 4760 of *LNAI*, Berlin, Germany, 2007. Springer-Verlag.

[63] L. Paletta, G. Fritz, F. Kintzler, J. Irran, and G. Dorffner. Learning to perceive affordances in a framework of developmental embodied cognition. In *Proc. 6th International Conference on Development and Learning, ICDL 07*, London, UK, July 2007. IEEE CD-ROM ISBN: 1-4244-1116-5.

[64] K. Janowicz. Extending semantic similarity measurement by thematic roles. In *First International Conference on GeoSpatial Semantics (GeoS 2005)*, Raleigh, NC, 2005. Conference: Mexiko City, Mexiko, 2005.

[65] R. Breithaupt, S. Frintrop, J. Hertzberg, E. Rome, and B.S. Müller. Specification of final demonstrator. Technical Report MACS/6/1.1 v2, Fraunhofer Institut für Autonome Intelligente Systeme, Sankt Augustin, Germany, 2004.

[66] E. Uğur, M.R. Doğar, M. Çakmak, and E. Şahin. The learning and use of traversability affordance using range images on a mobile robot. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA 07)*, pages 1721–1726. IEEE, 2007. Conference: Rome, Italy, April 10–14, 2007.

[67] E. Uğur, M.R. Doğar, M. Çakmak, and E. Şahin. The curiosity-driven learning of traversability affordance on a mobile robot. In *Proc. of the IEEE International Conference on Development and Learning (ICDL 07)*, 2007. Conference: London, UK, July 11–13, 2007.

[68] R. Breithaupt, S. Frintrop, E. Şahin, J. Hertzberg, P. Pölz, P. Rudol, E. Uğur, P. Doherty, E. Rome, and B.S. Müller. Report on experiment design. Technical Report MACS/6/4.1 v1, Fraunhofer Institut für Autonome Intelligente Systeme, Sankt Augustin, Germany, 2004.

[69] L. Paletta, G. Fritz, M. Kumar, J. Hertzberg, and F. Schönherr. Top-down and bottom-up symbol grounding. Technical Report MACS/3/1.1 v3, Joanneum Research Institute of Digital Image Processing Computational Perception (CAPE), Graz, Austria, 2005.

[70] L. Paletta, G. Fritz, E. Şahin, and M. Kumar. Affordance recognition from visual cues. Technical Report MACS/3/1.2 v1, Joanneum Research Institute of Digital Image Processing Computational Perception (CAPE), Graz, Austria, 2005.

[71] L. Paletta, G. Fritz, E. Rome, S. Frintrop, J. Hertzberg, and M. Kumar. Prototypical affordance based object detection for MACS scenario. Technical Report MACS/3/1.4 v1, Joanneum Research Institute of Digital Image Processing Computational Perception (CAPE), Graz, Austria, 2005.

[72] E. Rome, L. Paletta, G. Fritz, H. Surmann, S. May, and C. Lörken. Multi-sensor affordance recognition. Technical Report MACS/3/2.1 v2, Fraunhofer Institut für Autonome Intelligente Systeme, Sankt Augustin, Germany, 2006.

[73] D. Holz and C. Lörken. Continuous 3D environment sensing for autonomous robot navigation and mapping. In *Proc. of the 9. Fachwissenschaftlicher Informatik-Kongress, Lecture Notes in Informatics (LNI)*, pages 39–42, Bonn, Germany, March 2007. Gesellschaft für Informatik (GI).

[74] E. Uğur. Direct perception of traversability affordance on range images through learning on a mobile robot. M.Sc. thesis, Middle East Technical University, Kovan Laboratory, Ankara, Turkey, 2006.

[75] E. Uğur, M.R. Doğar, M. Çakmak, and E. Şahin. From primitive behaviours to goal-directed behaviour using affordances. Technical Report METU-CENG-TR-2007-02, Middle East Technical University, Kovan Laboratory, Ankara, Turkey, 2007. Short version published for IROS 2007.

[76] E. Uğur, M.R. Doğar, M. Çakmak, and E. Şahin. From primitive behaviours to goal-directed behaviour using affordances. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2007*, San Diego, CA, USA, 2007.

[77] E. Uğur, M.R. Doğar, M. Çakmak, and E. Şahin. Report on experimental results in simulator. Technical Report MACS/6/4.2 v1, Middle East Technical University Dept. of Computer Engineering, Ankara, Turkey, 2007.

[78] M.R. Doğar, M. Çakmak, E. Uğur, and E. Şahin. Report on experimental results in demonstrator. Technical Report MACS/6/4.3 v1, Middle East Technical University Dept. of Computer Engineering, Ankara, Turkey, 2007.

[79] M. Çakmak, M.R. Doğar, E. Uğur, and E. Şahin. Affordances as a framework for robot control. In *Proc. of the International Conference on Epigenetic Robotics, EpiRob 2007*, 2007. Conference: Piscataway, NJ, USA, November 5–7, 2007.

[80] M. Çakmak. Robot planning based on learned affordances. M.Sc. thesis, Middle East Technical University, Kovan Laboratory, Ankara, Turkey, 2007.

[81] M. Raubal. *Agent-based Simulation of Human Wayfinding: A Perceptual Model for Unfamiliar Buildings*. PhD thesis, Institute for Geoinformation, Vienna University of Technology, Vienna, Austria, 2001.

[82] F. Schönherr. *Verankerung der Semantik veränderlicher Situations-Fakten und symbolischer Aktionen in der hybriden Roboterkontrollarchitekur DD&P*. Fraunhofer series in information and communication technology 2004, 8. Shaker, Aachen, 2004.

[83] U. Neisser. *Cognitive Psychology*. Prentice-Hall, Englewood Cliffs, NJ, 1967.

[84] T.A. Stoffregen. Affordances as Properties of the Animal-Environment System. *Ecological Psychology* 15(2): 115–134, 2003.

[85] T.A. Stoffregen. Affordances and Events. *Ecological Psychology* 12(1): 1–28, 2000.

[86] K.B. Bærentsen and J. Trettvik. An activity theory approach to affordance. In *NordiCHI '02: Proc. 2nd Nordic conf. on Human-computer interaction*, p. 51–60. ACM Press, 2002.

[87] C.A.A. Calderon and H. Hu. Goal and actions: Learning by imitation. In *Proceedings of the Second International Symposium on Imitation in Animals and Artifacts*, 2003.

[88] J.J. Gibson. The theory of affordances. In R. Shaw and J. Bransford, editors, *Perceiving, Acting and Knowing*, pages 67–82. Wiley, New York, 1977.

[89] S. Coradeschi and A. Saffiotti. An introduction to the anchoring problem. *Robotics and Autonomous Systems* 43(2-3):85–96, 2003. Special issue on perceptual anchoring.

[90] P. Agre and D. Chapman. Pengi: An implementation of a theory of action. In *Proc. AAAI-87*, pages 268–272. Morgan Kaufmann, 1987.

[91] E. Uğur, M.R. Doğar, and E. Şahin. Planning using learned object affordances. Submitted to *AAAI 2008*.

[92] N. Goerke, F. Kintzler, and R. Eckmiller: Multi-SOMs: A new approach to self organised classification. In: S. Singh et al (eds.): Pattern Recognition and Data Mining: Third Int. Conf. on Advances in Pattern Recognition, ICAPR 2005, Bath, UK, Proc., Part I. Lecture Notes in Computer Science (LNCS) 3686. Berlin, Heidelberg: Springer, pp. 469–478, 2005.

[93] R. Schatten, N. Goerke, and R. Eckmiller: Regional and online learnable fields. In: Proc. of the Int. Joint Conf. on Neural Networks (IJCNN), Budapest, 2004.

ANNEX B. PROJECT DATA

B1. General Project Data

B1.1 Project key data

EC contract number:	FP6-004381
EC affiliation:	FP6, Strategic Objective 2.3.2.4 “Cognitive Systems”
EC project type:	Specifically Targeted Research Project (STReP)
Start date:	September 1, 2004
End date:	November 30, 2007
Maximum project costs:	2,597,113 €
Maximum EC contribution:	1,894,000 €
Total work force effort:	353 person months: 318 funded, 35 unfunded
Project web site:	http://www.macs-eu.org/

B1.2 Project Consortium Overview

IAIS Fraunhofer Institut Intelligente Analyse- und Informationssysteme	Fraunhofer IAIS, Sankt Augustin, Germany Acronym: FhG/AIS ART Department (Co-ordinator) Co-ordinator: Dr.-Ing. Erich Rome
JOANNEUM RESEARCH	Joanneum Research, Graz, Austria Acronym: JR_DIB CAPE Group Local Manager: DI Dr. Lucas Paletta
LINKÖPINGS UNIVERSITET	Linköpings Universitet, Linköping, Sweden Acronym: LiU-IDA IDA Department Local Manager: Prof. Dr. Patrick Doherty
Middle East Technical University	Middle East Technical University, Ankara, Turkey Acronym: METU-KOVAN KOVAN Department Local Manager: Asst. Prof. Dr. Erol Şahin
Ö F A I	Österreichische Studiengesellschaft für Kybernetik, Vienna, Austria Acronym: OFAI Neural Computation and Robotics Local Manager: Prof. Dr. Georg Dorffner
UNIVERSITÄT OSNABRÜCK	Universität Osnabrück, Osnabrück, Germany Acronym: UOS Department of Mathematics/Computer Science Local Manager: Prof. Dr. Joachim Hertzberg

B1.3 Project Consortium Description

B1.3.1 Fraunhofer IAIS, Sankt Augustin, Germany

Main researcher:

MACS is coordinated by Dr.-Ing. Erich Rome, a senior scientist in the ART Department of Fraunhofer IAIS.

Organisation description:

The *Fraunhofer Institut für Intelligente Analyse-und Informationssysteme* (IAIS) has been founded in July 2006 as a result of a merger of the original MACS co-ordinating Fraunhofer Institut für Autonome Intelligente Systeme (AIS) and the Fraunhofer Institut für Medienkommunikation (IMK), both located in Sankt Augustin, Germany. IAIS has now about 250 employees, including a research staff of about 140 scientists. Its main business areas are process intelligence, digital media asset management, M3 – marketing, market research and media analysis, preventive security, and high-tech experimental environments. In these areas, it performs pre-market research leading to application-oriented concepts and individual solutions for industrial, scientific, and governmental clients.

Within the Fraunhofer Society, FhG/AIS is a member of Fraunhofer Group Information and Communication Technology and a member of the Fraunhofer Alliance Vision. FhG/AIS is cooperating with universities from its home region, particularly with the University of Bonn and the University for Applied Sciences Bonn-Rhein-Sieg. The institute is headed by Prof. Dr. Stefan Wrobel.

On a European level, FhG/AIS has participated in the TMR network VIRGO (Vision-based robot navigation network), in the NoEs EURON (European Robotics Research Network) and PLANET, PLANET-II (European Network of Excellence in AI Planning), where FhG/AIS members have actively participated in the *Robot Planning* unit, in the SSA "Roberta goes EU", and the CA Ethicbots. MACS has been conducted in the Adaptive Reflective Teams Department (ART). ART is pursuing R&D projects in the areas of Process Intelligence and Preventive Security. In the latter area, ART is currently coordinating the Integrated Project IRRIIS (FP6) and the Collaborative Project DIESIS (FP7).

B1.3.2 Joanneum Research, Graz, Austria

Main researcher:

DI Dr. Lucas Paletta was the MACS local manager of JOANNEUM RESEARCH. He is head of the Computational Perception (CAPE) Group at the Institute of Digital Image Processing of Joanneum Research Forschungsgesellschaft mbH.

Organisation description:

JOANNEUM RESEARCH is one of the most important focal points of scientific research and economy in Austria. As a "partner in innovation" for business enterprises and political decision-makers, the non-profit institution owned by the Province of Styria focuses on applied research and development in the key technologies of the future. Their highly qualified staff of more than 370 employees working in 20 research institutes implements their know-how in all sectors of innovation, both at national and international levels. The Institute of Digital Image Processing is internationally recognised as a centre of excellence in the fields of industrial machine vision, cognitive vision, mobile mapping, and remote sensing, having a staff of over 50 scientists.

The Computational Perception Group researches and develops in cognitive vision and machine learning with the aim to provide efficient, robust and fast image interpretation in active perception and mobile mapping being inspired from cognitive science. Related projects investigate attentive image analysis, in particular on mobile visual object detection and positioning in urban environments, with the aim to achieve autonomous and robust interpretation of visual context in dynamic scenes. The Group consists of 4 permanent and 4 associated academic members that were and are being involved in EU-funded projects on Cognitive Vision (DETECT, MACS, ECVision), multimodal interfaces (coordinator of MOBVIS) and Remote Image Understanding (FIREGUARD, GMOSS), as well as national projects (JRP-ASF “Cognitive Vision”, ASF “Semantic Mapping of Pedestrian Visual Attention”, “Mobile City Explorer”), and has been involved in the organisation of related symposia on cognitive vision (ICVS2003, WAPCV '03–WAPCV '08).

B1.3.3 Linköpings Universitet, Linköping, Sweden

Main researcher:

Prof. Dr. Patrick Doherty was the MACS local manager of Linköpings Universitat. He is head of the Artificial Intelligence and Integrated Computer Systems Division / Knowledge Processing Laboratory (AIICS) of the Department of Computer and Information Science (IDA).

Organisation description:

The Department of Computer and Informatin Science at Linköping University, Sweden, is one of the largest departments for computer and information science in northern Europe. It has more than 270 employees, of which over 50 faculty members (Ph.Ds), including 17 full professors. There are approximately 175 Ph.D students enrolled in Ph.D programs, including 80 doctoral student positions. The department is divided into five divisions of which AIICS – The Artificial Intelligence and Integrated Computer Systems Division is one.

AIICS is led by Patrick Doherty and has 34 members with 2 research labs and 2 research groups. The main focus of interest for the AIICS division is intelligent artifacts, that is, man-made physical systems containing computational equipment and software that provide them with capabilities for receiving and comprehending sensory data, for reasoning, and for rational action in their environment. Such artifacts range from PDAs to UAVs and mobile robots.

B1.3.4 Middle East Technical University, Ankara, Turkey

Main researcher:

Asst. Prof. Dr. Erol Şahin is the MACS Local Manager of the Middle East Technical University. He is also head of the KOVAN Research Lab.

Organisation description:

The Middle East Technical University (METU) (http://www.metu.edu.tr) is the leading technical university in Turkey. METU is a point of reference for Turkey’s industry, engineering, production and new e-based activities. The fast developing and expanding Turkish economy and industry has traditionally looked to renowned institutions like METU for attracting engineers for leadership, for consulting and problem solving collaboration.

KOVAN research group (http://www.kovan.ceng.metu.edu.tr) consists of two staff members, two Ph.D. and three M.Sc. students. The team is founded and headed by Dr. Erol Şahin to study the design of artificial autonomous systems through understanding natural systems. We have four active research tracks. The design and understanding of swarm robotic systems is the main active research track. Current research topics include; a) development of a parallized evolutionary system for evaluating fitness values of robot systems in parallel over a group of networked computers, b) physics-based modelling and simulation of swarm robotics systems, c) self-organized pattern formation in swarm robotic systems, d) study the use of evolutionary methods in developing behaviours for a swarm of robots. The second track focuses on the measurement of large closed spaces by mobile robots. Inspired by ant scouts which can successfully assess the area of potential nest sites, we investigate how we can use the same strategies on mobile robots as an engineering method. Our third track is on the use of evolutionary methods for solving problems. We have successfully developed DARWIN, a programming language that enables easy programming in Evolutionary Computing. Our fourth track explores the use of adaptation models (evolution and learning) to be used on mobile robots. We have made some preliminary studies on how a neural model of conditioning can be applied to on-line learning of simple behaviours on simple robots.

KOVAN is participating to the Swarm-bots Project (IST-2000-31010) as a sub-contractor. KOVAN is also awarded a METU grant for the "Virtual Robot Colony" project.

B1.3.5 Österreichische Studiengesellschaft für Kybernetik, Vienna, Austria

Main researcher:

Prof. Dr. Georg Dorffner was the MACS Local Manager of the Austrian Research Institute for Artificial Intelligence (OFAI). He was also the head of OFAI's **Neural Computation Group.**

Organisation description:

The Austrian Research Institute for Artificial Intelligence (OFAI) was founded in 1984 with support from the Austrian Federal Ministry for Science and Research. Within the framework of the Federal Development Program of the Austrian Government Microelectronics and Information Processing the Institute was assigned key institute for research area S7 Artificial Intelligence. According to this development program, OFAI maintains close links to the Department of Medical Cybernetics and Artificial Intelligence of the University of Vienna (IMKAI). At OFAI basic and applied research is performed in several areas of AI, such as neural computation, machine learning, knowledge based systems, software agents and natural language processing.

OFAI has, among others, been partner in around 25 multinational projects in the EU programs ESPRIT, BIOMED, BRITE-EURAM and IST (FP V). Recent projects include Biomed-2 project SIESTA (Development of a new standard for computerised sleep analysis), of which OFAI is the co-ordinator, the ESPRIT project METAL (A meta-learning assistant), and the FP V projects SOL-EU-NET, MOSART and SAFIRA (Supporting Affective Interactions for Real-Time Applications). Furthermore, OFAI is a member of several EU Networks of Excellence, such as the network on Intelligent Technologies for Smart Adaptive Systems (EUNITE) and the Network of Excellence for Agent-Based Computing (AgentLink).

The Neural Computation group was founded in 1988. The group has been involved in major European projects, such as the ESPRIT-II project NEUFODI (Neural Networks for Forecasting and Diagnosis Applications), the BIOMED 1 and 2 projects ANNDEE and SIESTA on EEG signal analysis and the ongoing IST-FET project SIGNAL (Systemic Architectures for Growing Up Artefacts that Live) on cognitive robotics. Today, the group is one of the largest and most renowned neural computation research groups in Austria. The main foci of work at the group include time series and signal processing, cognitive neuroscience and cognitive modelling.

B1.3.6 Universität Osnabrück, Osnabrück, Germany

Main researcher:

Prof. Dr. Joachim Hertzberg was the MACS Local Manager of University of Osnabrück's (UOS) Department of Mathematics/Computer Science. He is also the head of the Knowledge-based Systems Research Group and currently the dean of the Department of Mathematics/Computer Science.

Organisation description:

The *University of Osnabrück* (UOS, www.uni-osnabrueck.de) is among the younger universities in Germany, starting operation in 1974. With about 10,000 students 600 PhD students and a staff of 1,350, it is also one of the smaller universities. Students can choose from over 80 degree programmes in the Humanities, Social Sciences, Natural Sciences, Law, and Business Studies. Most study programmes follow the Bachelors/Masters scheme, using the ECTS credit point rating. Complementary to its classical faculty structure, the university features a high number of interdisciplinary units (such as the Institute for Cognitive Science) and study programmes, exploiting the flexibility of a young and small university.

The Institute for Computer Science (www.inf.uos.de) is part of UOS's Department of Mathematics/Computer Science. It currently comprises seven working groups, four of which have been founded (3) or re-founded (1) within the last three years. It is planned to further enlarge the institute, together with its move from offering only secondary-subject studies on Diploma, Bachelors, and Masters levels to additionally offering a consecutive Master in Informatics (starting in Winter Semester 2006/07). Three of the four recently founded working groups (Neuroinformatics, Knowledge-Based Systems, Technical Informatics) use mobile robots as a focus of their work in teaching and research.

The Knowledge-Based Systems (KBS) Group (www.inf.uos.de/kbs) has been founded in September 2004, when Prof. Joachim Hertzberg joined UOS. In addition to him, the group currently consists of four full staff researchers (two on teaching positions, one in MACS, one on the BMBF-funded project LISA) plus secretariat and infrastructure staff. In 2006, three PhDs (two of them by externals), four Masters, and four Bachelors have graduated under primary supervision by KBS. The current research topics are grouped around robot navigation, SLAM, semantic mapping, plan-based robot control, and rescue robotics. Current equipment of the KBS group includes four KURT2 robot platforms and one KURT3D platform. Partners in formalised cooperations include Fraunhofer IAIS and IFF; the companies Fox GmbH, Schunk and Jenoptik. Currently informal cooperation is close with the Universities Bonn, Hannover and TU Munich; with AASS (Örebro, Sweden) and LAAS (Toulouse, France); and with Telekom Labs (Berlin).

B2. MACS Dissemination Activities

B2.1 Publications

Peer-reviewed journal articles

E. Şahin, M. Çakmak, M.R. Doğar, E. Uğur, and G. Ücoluk (2007): To afford or not to afford: A new formalisation of affordances towards affordance-based robot control. *Adaptive Behaviour*, volume 15, number 4, pp. 447–472, 2007.

Peer-reviewed conference papers

J. Irran, F. Kintzler and P.M. Pölz (2006): Grounding Affordances, In: Trappl R. (ed.): *Cybernetics and Systems 2006*, Proc. Of 18th European Meeting on Cybernetics and Systems Research (EMCSR), Vienna: Austrian Society for Cybernetic Studies, ISBN 3 85206 172 5.

L. Paletta, G. Fritz, E. Rome and G. Dorffner (2006): A computational model for visual learning of affordance-like cues. *ECVP06 European Conference on Visual Perception*, St. Petersburg, Russia, August 20–25, 2006, Perception, Vol. 35, abstracts, page 18, Pion Ltd., London, UK.

G. Fritz, L. Paletta, R. Breithaupt, E. Rome, and G. Dorffner (2006): Learning predictive features in affordance-based robotic systems. In: *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2006*, pp. 3642–3647. Conference: Beijing, China, October 9–15, 2006.

G. Fritz, L. Paletta, M. Kumar, G. Dorffner, R. Breithaupt, and E. Rome (2006): Visual Learning of Affordance based Cues. In: *Proc. Ninth International Conference on the Simulation of Adaptive Behaviour, SAB 2006*, pp. 52–64, Springer-Verlag, Berlin, LNAI 4095. Conference: Rome, Italy, September 25–29, 2006.

D. Holz and C. Lörken (2007): Continuous 3D Environment Sensing for Autonomous Robot Navigation and Mapping. In: *Proc. of the 9. Fachwissenschaftlicher Informatik-Kongress*, Lecture Notes in Informatics (LNI), Series of the Gesellschaft für Informatik (GI), ISBN 978-3-88579-439-4, pp. 39–42, Bonn, Germany, March 2007.

E. Uğur, M.R. Doğar, M. Çakmak, and E. Şahin (2007): The learning and use of traversability affordance using range images on a mobile robot. In: *Proc. of IEEE International Conference on Robotics and Automation, ICRA 2007*, pp. 1721–1726, IEEE, 2007. Conference: Rome, Italy, 2007.

L. Paletta and G. Fritz (2007): Reinforcement Learning of Predictive Features. In: *Proc. 31st Workshop of the Austrian Association for Pattern Recognition, AAPR / ÖAGM 2007*, pp. 105–112. Conference: Krumbach, Austria, May 3–4, 2007.

L. Paletta, G. Fritz, F. Kintzler, J. Irran, and G. Dorffner (2007): Learning to Perceive Affordances in a Framework of Developmental Embodied Cognition. In: *Proc. IEEE 6th International Conference on Development and Learning, ICDL 2007*, pp. 110–115, IEEE. Conference: London, UK, July 11–13, 2007.

E. Uğur, M.R. Doğar, M. Çakmak, and E. Şahin (2007): Curiosity-driven Learning of Traversability Affordance on a Mobile Robot. In: *Proc. IEEE 6th International Conference on Development and Learning, ICDL 2007*, pp. 13–18, IEEE. Conference: London, UK, July 11–13, 2007.

A. Bartel, F. Meyer, C. Sinke, T. Wiemann, A. Nüchter, K. Lingemann, and J. Hertzberg: Real-Time Outdoor Trail Detection on a Mobile Robot. In: *Proc. of the 13th IASTED International Conference on Robotics and Applications*, pp. 477–482, Würzburg, Germany, August 2007.

L. Paletta, G. Fritz, F. Kintzler, J. Irran, and G. Dorffner (2007): Perception and Developmental Learning of Affordances in Autonomous Robots. In: *Proc. 30th German Conference on Artificial Intelligence, KI 2007*, pp. 235–250, Springer-Verlag, Berlin, LNCS 4667. Conference: Osnabrück, Germany, September 10–13, 2007.

S. May, M. Klodt, E. Rome, and R. Breithaupt (2007): GPU-accelerated Affordance Cueing based on Visual Attention. In: *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2007*, pp. 3385-3390, IEEE. Conference: San Diego, CA, USA, October 29 – November 2, 2007.

M.R. Doğar, M. Çakmak, E. Uğur, and E. Şahin (2007): From Primitive Behaviours to Goal-Directed Behaviour Using Affordances. In: *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2007*, pp. 729–734, IEEE. Conference: San Diego, CA, USA, October 29 – November 2, 2007.

M. Çakmak. M.R. Doğar, E. Uğur, and E. Şahin (2007): Affordances as a Framework for Robot Control. In: *Proc. of International Conference on Epigenetic Robotics (EpiRob 2007),* Conference: Piscataway, NJ, USA, November 5–7, 2007.

M.R. Doğar, E. Şahin, E. Uğur, and M. Çakmak (2008): Using Learned Affordances for Robotic Behaviour Development. Accepted for the *IEEE International Conference on Robotics and Automation, ICRA 2008*.

Ch. Lörken and J. Hertzberg (2008): Grounding planning operators by affordance. Accepted for the *International Conference on Cognitive Systems 2008 (CogSys 2008)*.

Edited books

E. Rome, J. Hertzberg, and G. Dorffner (eds): *Towards Affordance-based Robot Control. Proceedings of Dagstuhl Seminar 06231*. Springer-Verlag, Lecture Notes in Computer Science – LNAI 4760, Springer-Verlag, Berlin, Germany, February 2008.

L. Paletta and E. Rome (eds): *Attention in Cognitive Systems.* Lecture Notes in Computer Science – LNAI 4840, Springer-Verlag, Berlin, Germany, January 2008.

Book chapters and online articles

L. Paletta and G. Fritz (2007): Reinforcement Learning of Predictive Features in Affordance Perception. In: *Towards Affordance-based Robot Control, Proceedings of Dagstuhl Seminar 06231*, Springer-Verlag, LNAI 4760, E. Rome, J. Hertzberg and G. Dorffner (eds.), February 2008.

L. Paletta and G. Fritz (2007): Reinforcement Learning for Decision Making in Sequential Attention, in Paletta, L., and Rome, E., Eds., *Attention in Cognitive Systems,* Lecture Notes in Computer Science – LNAI 4840, pp. 293-306, Springer-Verlag, Berlin, Germany.

J. Hertzberg, K. Lingemann, C. Lörken, A. Nüchter, and S. Stiene (2007): Does it Help a Robot Navigate to Call Navigability an Affordance? In: *Towards Affordance-based Robot Control, Proceedings of Dagstuhl Seminar 06231*, Springer-Verlag, LNAI 4760, E. Rome, J. Hertzberg and G. Dorffner (eds.), February 2008.

J. Irran, F. Kintzler, and G. Dorffner (2007): Learning of Interaction Possibilities. In: *Towards Affordance-based Robot Control, Proceedings of Dagstuhl Seminar 06231*, Springer-Verlag, LNAI 4760, E. Rome, J. Hertzberg and G. Dorffner (eds.), February 2008.

E. Rome, L. Paletta, E. Şahin, G. Dorffner, J. Hertzberg, R. Breithaupt, G. Fritz, J. Irran, F. Kintzler, C. Lörken, S. May, and E. Uğur (2007): The MACS project: An approach to affordance-based robot control. In: *Towards Affordance-based Robot Control, Proceedings of*

Dagstuhl Seminar 06231, Springer-Verlag, LNAI 4760, E. Rome, J. Hertzberg, and G. Dorffner (eds.), February 2008.

E. Rome, J. Hertzberg, G. Dorffner and P. Doherty (2006): Towards Affordance-based Robot Control – Executive Summary of the Dagstuhl Seminar 06231. In: *Online proceedings of Dagstuhl Seminar 06231* http://drops.dagstuhl.de/opus/volltexte/2006/725/pdf/06231.SWM.Paper.725.pdf

E. Rome, J. Hertzberg, G. Dorffner, P. Doherty et al. (2006): Towards Affordance-based Robot Control – Abstracts Collection of the Dagstuhl Seminar 06231. In: *Online proceedings of Dagstuhl Seminar 06231* http://drops.dagstuhl.de/opus/volltexte/2006/722/pdf/06231_abstracts_collection.722.pdf

Posters

M.R. Dogar and S. Eren (2005): MACS project overview. First EURON Summer School on Perception and Sensor Fusion in Mobile Robotics Conference: Ancona, Italy, September 1–7, 2005.

L. Paletta (2006): poster presentation, euCognition Inaugural Meeting, Conference: Nice, France, February 16–17, 2006.

F. Kintzler and J. Irran (2006): poster presentation, euCognition Inaugural Meeting, Conference: Nice, France, February 16–17, 2006.

F. Kintzler and J. Irran (2006): *Affordance-based Learning*. Robot Challenge 2006. Conference: Vienna, Austria, March 11, 2006.

F. Kintzler and J. Irran (2006): *Affordance-based Learning*. Brain Awareness Week 2006. Conference: Vienna, Austria, March 13–18, 2006.

F. Kintzler and J. Irran (2006): *Affordance-based Learning Approach*. CogSys II – 2nd Cognitive Systems Networking Meeting. Conference: Nijmegen, The Netherlands, April 12–13, 2006.

G. Fritz, L. Paletta, R. Breithaupt, E. Rome, and G. Dorffner (2006): *Affordance perception as a basis for cognitive development*. poster for COGROB06 (The Fifth International Cognitive Robotics Workshop / AAAI), Boston, MA, USA, July 16–17, 2006.

L. Paletta and G. Fritz (2006): Reinforcement Learning for the Selection of Predictive Cues in Affordance-based Perception, *ECVP06 European Conference on Visual Perception*, St. Petersburg, Russia, August 20–25, 2006, Perception, Vol. 35, abstracts, page 48, Pion Ltd., London, UK.

L. Paletta and G. Fritz (2006): Local Descriptor Groupings in Reinforcement Learning of Sensory-Motor Attention, *ECVP06 European Conference on Visual Perception*, St. Petersburg, Russia, August 20–25, 2006, Perception, Vol. 35, abstracts, page 48, Pion Ltd., London, UK.

L. Paletta and G. Fritz (2008): poster presentation, Cognitive Systems Industry Day, Conference: Sankt Augustin, Germany, January 29, 2008.

R. Breithaupt, E. Rome, S. May (2008): poster presentation, Cognitive Systems Industry Day, Conference: Sankt Augustin, Germany, January 29, 2008.

J. Hertzberg, C. Lörken, F. Meyer, A. Bartel (2008): poster presentation, Cognitive Systems Industry Day, Conference: Sankt Augustin, Germany, January 29, 2008.

Theses

Frank Meyer:
Surface detection in 3D range data for mobile crane manipulation. Bachelor Thesis, University of Osnabrück, Institute for Cognitive Science, March 2008.

Christopher Lörken:
Introducing Affordances into Robot Task Execution. Masters Thesis, University of Osnabrück, Institute for Cognitive Science, 2006. Available online as PICS Volume 2-2007, URL: http://www.cogsci.uni-osnabrueck.de/cogsci/de/m.ikwPublications.php.

Emre Uğur:
Direct Perception of Traversability Affordance on Range Images Through Learning on a Mobile Robot. M.Sc. Thesis, Middle East Technical University, Kovan Laboratory, 2006.

Maya Çakmak:
Robot Planning based on Learned Affordances. M.Sc. Thesis, Middle East Technical University, Kovan Laboratory, July 2007.

Mehmet Remzi Doğar:
Using Learned Affordances for Robot Behaviour Development. M.Sc. Thesis, Middle East Technical University, Kovan Laboratory, September 2007.

B2.2 Conferences, Workshops, Seminars, Networking

Conferences and Workshops organized by MACS staff

June 5–9, 2006 Wadern, Germany	**Dagstuhl Seminar "Towards Affordance-Based Robot Control"** Erich Rome (FhG/AIS), Georg Dorffner (OFAI), Joachim Hertzberg (UOS): Seminar organisation
Jan 8, 2007 Hyderabad, India	**WAPCV 07 –** 4th International Workshop on Attention in Cognitive Systems Lucas Paletta (JR_DIB), Erich Rome (FhG/AIS): Workshop organization
Sep 10–13, 2007 Osnabrück, Germany	**KI 2007 –** 30th German Conference on Artificial Intelligence Joachim Hertzberg (UOS): Organisation
Jan 29, 2008 Sankt Augustin, Germany	**Cognitive Systems Industry Day** Erich Rome (FhG/AIS): Organization Erich Rome (FhG/AIS): Presentation on MACS

Presentations on Conferences and Workshops

Mar 11, 2006 Vienna, Austria	**Robot Challenge 2006** Florian Kintzler and Jörg Irran (OFAI): poster presentation
Mar 13–18, 2006 Vienna, Austria	**Brain Awareness Week 2006** Florian Kintzler and Jörg Irran (OFAI): poster presentation
April 18–21 Vienna, Austria	**EMCSR 2006 –** 18th European Meeting on Cybernetics and Systems Research 2006 Florian Kintzler and Jörg Irran (OFAI): Grounding Affordances
Jul 16–17, 2006 Boston, MA, USA	**COGROB 06 –** The Fifth International Cognitive Robotics Workshop (in conjunction with AAAI) Gerald Fritz (JR_DIB): Affordance perception as a basis for cognitive development

Aug 20–25, 2006 St. Petersburg, Russia	**ECVP 2006 –** European Conference on Visual Perception Lucas Paletta (JR_DIB): 1. A computational model for visual learning of affordance-like cues, 2. Reinforcement Learning for the Selection of Predictive Cues in Affordance-based Perception
Sep 25–29, 2006 Rome, Italy	**SAB '06** – FROM ANIMALS TO ANIMATS 9: The Ninth International Conference on the Simulation of Adaptive Behaviour Lucas Paletta (JR_DIB): Visual Learning of Affordance based Cues
Oct 9–15, 2006 Beijing, PR China	**IROS 2006 –** IEEE/RSJ International Conference on Intelligent Robots and Systems Stefan May (FhG/AIS): Learning predictive features in affordance-based robotic systems
Mar 30–31, 2007 Bonn, Germany	**Informatiktage 2007** – 9. Fachwissenschaftlicher Informatik-Kongress Dirk Holz (FhG/AIS): Continuous 3D Environment Sensing for Autonomous Robot Navigation and Mapping
April 10–14, 2007 Rome, Italy	**ICRA 2007** – 2007 IEEE International Conference on Robotics and Automation Emre Uğur (METU-KOVAN): The learning and use of traversability affordance using range images on a mobile robot
July 11–13, 2007 London, UK	**ICDL 2007** – 2007 IEEE International Conference on Development and Learning (ICRA 2007) Lucas Paletta (JR_DIB): Learning to Perceive Affordances in a Framework of Developmental Embodied Cognition Emre Uğur (METU-KOVAN): Curiosity-driven Learning of Traversability Affordance on a Mobile Robot
Sep 10–13, 2007 Osnabrück, Germany	**KI 2007** – 30th German Conference on Artificial Intelligence Joachim Hertzberg (UOS): Organisation Lucas Paletta et al. (JR_DIB): Developmental Learning of Affordances in Autonomous Robots
Oct 29 – Nov 2, 2007 San Diego, CA, USA	**IROS 2007** – 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems Stefan May (FhG/IAIS): GPU-accelerated Affordance Cueing based on Visual Attention Mehmet Doğar (METU-KOVAN): From Primitive Behaviours to Goal-Directed Behaviour Using Affordances
Nov 5–7, 2007 Piscataway, NJ, USA	**EpiRob 2007 –** International Conference on Epigenetic Robotics Maya Çakmak (METU-KOVAN): Affordances as a Framework for Robot Control
April 2–4, 2008 Karlsruhe, Germany	**CogSys 2008 –** International Conference on Cognitive Systems Joachim Hertzberg (UOS): Affordances as a Framework for Robot Control
May 19–23, 2008 Pasadena, CA, USA	**ICRA 2008** – 2008 IEEE International Conference on Robotics and Automation Mehmet Doğar (METU-KOVAN): Using Learned Affordances for Robotic Behaviour Development

Presentations on Seminars

June 5–9, 2006 Wadern, Germany	**Dagstuhl Seminar "Towards Affordance-Based Robot Control"** Erol Şahin (METU-KOVAN): Oral presentation on formalising affordances Lucas Paletta (JR_DIB): Oral presentation on perception of affordances Florian Kintzler (OFAI): Oral presentation on learning of affordances Erich Rome (FhG/AIS): Oral presentation on MACS
July 16, 2007 Edinburgh, UK	**Seminar at University of Edinburgh, Institute for Perception, Action and Behaviour** Emre Uğur (METU-KOVAN): The learning and use of traversability affordance on a mobile robot

Networking and other Dissemination

Oct 28–30, 2004 Bled, Slovenia	**Cognitive Systems Networking Meeting** Erich Rome (FhG/AIS): oral presentation, introduction to MACS
Nov 15–17 The Hague, The Netherlands	**IST Congress 2004** Lucas Paletta (JR_DIB): Cognition networking session, Nov 16
Feb 16–17, 2006 Nice, France	**euCognition Inaugural Meeting** Lucas Paletta (JR_DIB): poster presentation Florian Kintzler and Jörg Irran (OFAI): poster presentation
April 12–13, 2006 Nijmegen, The Netherlands	**CogSys II –** 2nd Cognitive Systems Networking Meeting Erich Rome (FhG/AIS): oral presentation on MACS Florian Kintzler and Jörg Irran (OFAI): poster presentation learning approach
Jan 24, 2007 Karlsruhe, Germany	**VDI-GMA Fachausschusssitzung 4.13 Robotersysteme** Erich Rome (FhG/AIS): Invited MACS Overview Talk
Jan 29, 2008 Sankt Augustin, Germany	**Cognitive Systems Industry Day** Erich Rome (FhG/AIS): Presentation on MACS Lucas Paletta and Gerald Fritz (JR_DIB): poster presentation J. Hertzberg, C. Lörken, F. Meyer, A. Bartel (UOS): poster presentation R. Breithaupt, E. Rome, S. May (FhG/AIS): poster presentation

B2.3 MACS Website

The MACS project is documented on its project web site at URL http://www.macs-eu.org.

B2.4 Education

During the execution of MACS, several students completed their theses on topics directly related to and contributing to MACS.

Christopher Lörken:
Introducing Affordances into Robot Task Execution. Masters Thesis, University of Osnabrück, Institute for Cognitive Science, 2006. Available online as PICS Volume 2-2007, URL: http://www.cogsci.uni-osnabrueck.de/cogsci/de/m.ikwPublications.php.

Emre Uğur:
Direct Perception of Traversability Affordance on Range Images Through Learning on a Mobile Robot. M.Sc. Thesis, Middle East Technical University, Kovan Laboratory, 2006.

Maya Çakmak:
Robot Planning based on Learned Affordances. M.Sc. Thesis, Middle East Technical University, Kovan Laboratory, July 2007.

Mehmet Remzi Doğar:
Using Learned Affordances for Robot Behaviour Development. M.Sc. Thesis, Middle East Technical University, Kovan Laboratory, September 2007.

B3. List of Deliverables

Del. no.	MACS Deliverable title	Nature
D0.1.1	Project handbook and quality management plan	R
D1.1.1	Specification of SW development environment	R
D2.1.1	Identification of architectural requirements of an affordance-based control	R
D4.1.1	A conference or journal article summarising the results of task 4.1.	P
D5.1.1	Overview of existing affordance learning approaches	R
D6.5.1	Project web site and Project Presentation	S, W
D1.1.2	Specification of module interfaces	R
D2.2.1	Evaluation of existing control architectures for using affordances	S, R
D4.2.1	A conference or journal article summarising the results of task 4.2.	P
D6.1.1	Specification of final demonstrator	R
D6.4.1	Report on experiment design	R
D1.1.3	Implementation of SW development environment	S, DO
D3.1.1	Top-down and bottom-up symbol grounding	S, R
D3.1.2	Affordance recognition from visual cues	S, R
D4.3.1	A specificationof a software module for affordance representation	R
D5.2.1	Implementation of unsupervised and reinforcement learning algorithms	R, S
D3.1.4	Prototypical affordance based object detection for MACS scenario	R, S
D5.3.1	Robotic learning architecture that can be taught by manually putting the robot through action sequences	S, R
D3.1.3	Saliency detection with visual attention	R, S
D1.2.1	Simulation of KURT2 Platform	S, R
D6.2.1	Simulation model of final demonstrator scenario	S, R
D1.3.1	Integrated implementation of reference control system	S, R
D2.2.2	Development of an affordance-based control architecture	S, R
D3.2.1	Multi-sensor affordance recognition	DO, S, R
D3.3.1	Prototypical sensormotor based affordance recognition	R, S
D4.3.2	A software prototype for affordance support	S
D5.3.2	Robotic learning architecture capable of autonomously segment action sequences into affordances	S, R
D5.4.2	Prototypical software for representing and learning visual affordance support	R, S
D6.3.1	Physical robot demonstrator and scenario	R, DO, S
D2.3.1	Implementation of the affordance-based control architecture	S, R
D4.4.1	A software prototype for an affordance monitoring module with empirical testing using various MACS robotics platforms	S, R
D5.3.3	Robot protoype learning affordances through self-experience	S, R
D6.4.2	Report on experimental results in simulator	R
D3.3.2	Sensorimotor decision making and affordance recognition	S, R
D2.3.2	A specification for a propositional planner and its interface to the MACS Execution Control Module	R
D5.4.5	Outlook towards affordance usage observation and imitation	R
D5.3.3	Robot prototype learning affordances through self-experience V2	S, R
D6.4.3	Report on experimental results in demonstrator	R
D4.3.3	A software prototype of the propositional MACS planning module	S, R
D3.3.2	Sensorimotor decision making and affordance recognition	S, R

Legend: R: Report, S: Software, DO: Documentation, M: Meeting/Workshop

Del. no.	MACS Deliverable title	Nature
D4.4.3	An evaluation of the MACS planning module in the context of the MACS architecture	R
D6.5.2	Industry Day	M
D4.4.4	Submission of a conference or journal article describing the results of D4.3.4 and D4.4.3	P
D0.1.7	Publishable Final Activity Report	R

Legend: R: Report, S: Software, DO: Documentation, M: Meeting/Workshop

ANNEX C. Definitions, Terms & Acronyms

C1. Definitions

Initial Definition 1 ((Agent) Affordance). An *affordance* is a relation between the agent and its environment as acquired from the interaction of the two.

Definition 2 (Affordance Representation). An affordance representation or affordance triple is a data structure:

(*cue descriptor, behaviour descriptor, outcome descriptor*). (1)

Here, a *cue descriptor* or an *outcome descriptor* is specified as a list of attribute value pairs. A *behaviour descriptor* consists of one or more behaviour identifiers. Optionally, parameters for these behaviours can be specified.

Refined Definition 3 ((Agent) Affordance). An *affordance* is an acquired relation between a certain effect and a *(entity, behavior)* tuple, such that when the agent applies the behavior on the entity, the effect is generated.

Definition 4 (Affordance (agent perspective)). An *affordance* is an acquired relation between a certain <*effect*> and a certain <(*entity*, *behavior*)> tuple such that when the agent applies a (*entity*, *behavior*) within <(*entity*, *behavior*)>, an effect within <*effect*> is generated.

C2. A Glossary of Terms

Affordance Representation Repository: Stores *affordance representations* (→def.) generated through manual design or through learning.

Action: An action is a distinguished subset of events deemed under an agent's control. In artificial intelligence, actions generally have a structure consisting of a precondition representing the constraints required to successfully invoke the action, durative conditions (if the action has duration) representing the causal dependencies and durative effects the action has during its execution and a post-condition representing the direct effects the action has. In the context of a causal theory, one often distinguishes direct effects from indirect effects or ramifications of the action. The term →*outcome* will be used as another term for the direct effect of an action, or in the case of causal theories, for the combination of direct, durative and indirect effects.

Actuators: "a mechanical device for moving or controlling something" (definition from Merriam-Webster, [1]).

Affordance-based behaviours: An affordance-based behaviour is an enhancement of a robot behaviour by enriching the sensor space and knowledge base of an artificial agent by affordance information. Robot behaviours that are especially prepared for such enhancements are called →*affordance-related behaviours.*

Affordance-related behaviours: A group of →*high-level behaviours* that are directly related to *agent affordances* (→def.) in terms of manipulating objects or trying to manipulate objects. Examples would be behaviours that lift, push or carry items.

Affordance hypothesis verification: Suppose the robot perceives a cue c that supports the presence of an affordance A represented by at least triple *(c,b,o)*. A offers the potential of executing a behaviour *b*. The robot hypothesizes that the execution of the behaviour *b* leads to the outcome described by *o*. During the execution of *b* (and eventually a certain time thereafter) it observes the outcome *o'*. During the whole time period, the outcome *o'* must match to the hypothesized outcome *o*. In this case the affordace hypothesis is verified. The step of comparing *o'* and *o* is then called Affordance Hypothesis Verification.

Behaviour system: The collection of →*robot behaviours* for a given robot.

Behaviour: See →*robot behaviour*.

Components: Refers to modules and other parts of the robot control architecture. Typically, a part of a module would be named component.

Cue: "Something serving as a signal or suggestion ... <hint> ... a suggestion for action given briefly or in an indirect manner ..." [2]. Here: the perceptual data part of an →*affordance representation* that supports the existence of the associated agent affordance.

Deliberation module: →*Components* of the control architecture that enable deliberate or planned acts.

Entity schema: An entity structure (→def.) that consists of pairs of attributes and value ranges, e.g.—arbitrary notation used here—((form, circular), (sizes, (10:20, 30:50, 80:120))).

External sensor: →*Robot sensor* for perceiving the world. An external or exteroperceptive sensor is a device that senses the surrounding environment, i.e. a sensor that takes measurements of the surrounding environment and translates these measurements into useful data for the robot. Including, but not limited to: cameras, ranging sensors, odometry sensors, gyroscopes.

Feature extraction: "Feature extraction is a special form of dimensionality reduction and is in the area of image processing also connected with shape recognition. ... It can be used in the area of image processing which involves using algorithms to detect and isolate various desired portions or shapes (features) of a digitized image or video stream." (definition from Wikipedia, [3]). Feature extraction may employ specialized →*filters* for implementing the extraction of particular features from an image.

Filter: A computer program to transform or selectively reduce the amount of information in a set of data or a data stream. Also: analog or digital image processing (IP) operations to enhance or modify an image. Here, the data stream is a stream of sensory data. Filters for a stream of image data are called image filters. Image filters take an image as input, perform certain operations on it, and return the result image. An example is a Gaussian blur filter that smoothes a small image portion around a centre pixel in a specified radius.

Goal: "Usually specified as one or a set of world states. There are three kinds of goals: maintaining some condition, preventing some condition from occurring, or sequencing activities." [4]

High-Level / Complex behaviours: Goal-oriented robot behaviours that fulfil more complex tasks than merely reacting to perceptual stimuli. They often may use other high-level or →*low-level behaviours* to reach their →*goal*. For example, a behaviour that explores the environment by wandering around may use reactive behaviours for driving and for avoiding obstacles. All →*affordance-related behaviours* are high-level behaviours.

Internal sensor: An internal or proprioceptive sensor is a →*robot sensor* device that monitors and senses conditions of the robot and its hardware. Including, but not limited to: temperature sensor for monitoring the internal temperature, inclinometers for sensing the robot pose, voltage sensors for monitoring the battery status.

Learning module: Uses →*machine learning* techniques for acquiring and structuring information and for realizing adaptivity.

Low-Level / Simple behaviours: See →*reactive behaviours*.

Machine learning: "Machine learning refers to the ability of computers to automatically acquire new knowledge, learning from, for example, past cases or experience, from the computer's own experiences, or from exploration. ... Machine learning enables computer software to adapt to changing circumstances, enabling it to make better decisions than non-AI software." (definition from Stottler Henke Associates, Inc. Glossary of AI terms, [5]).

Outcome: Here: the perceptual data part of an →*affordance representation* that is employed to verify the robot's hypothesis of the expected results of its behaviours (→*affordance hypothesis verification*). Outcome refers to a situation and thus may have a temporal component.

Perception module: "... 3a: awareness of the elements of environment through physical sensation ..." (definition from Merriam-Webster online [6]). Also: "In perception the environment is scanned by means of various sensory organs, real or artificial, and the scene is decomposed into separate objects in various spatial relationships." (definition from Encyclopedia Britannica online, article on Artificial Intelligence [7]). In the context of robotic systems, this module should generate an interpretation, using the raw spatiotemporal multimodal data incoming from extero- and proprioceptive →*robot sensors*, for the benefit of the →*execution module*, the →*behaviour system*, and the →*Learning module*.

Plan: "A plan is a representation of a course of action." [4]

Planning: "Deciding upon a course of action before acting. A →*plan* is a representation of a course of action. A finished plan is a linear or partially ordered sequence of operators. Planning is reasoning about future events in order to verify the existence of a reasonable series of actions to take in order to accomplish a →*goal*. " [4]

Planner: "Artificial Intelligence (AI) Planning is areas of study concerned with the automatic generation of a →*plan* to solve a problem within a particular domain. Given an initial state, the planner tries to find the actions required to achieve some →*goal conditions*." [8]

Raw sensor data: →*Robot sensor* data that are provided by the →*Perception module* without prior →*feature extraction*. For the MACS architecture, raw sensor data might have passed synchronisation services and/or transformation into a suitable data structure, like an entity frame (→def.). Such services are provided, for instance, by the →*Entity Structure Generation Module*.

Reactive behaviours: →*Robot behaviours* with a very tight perception-action coupling that have the need to react fast to sensory input. For example, a reactive brake behaviour will stop the robot directly if an obstacle is detected in close proximity in heading direction.

(Robot) behaviour: A robot control routine using a perception-action cycle for performing certain limited actions. For example, a "wall following behaviour" may use range sensors to keep the robot at a specified distance from a wall while driving along that wall at a specified speed.

Robot sensors: A sensor for a mobile robot is a physical device that detects or senses a signal or physical condition of the robot or the robot's environment. A robot sensor yields sensory data that can be processed by a robot's perception software, e.g. the →*perception module* in general and its →*feature* extraction part, for instance. The robot's hardware sensors are coarsely divided into →*external* and →*internal sensors*. →*Virtual (software) sensors* can be added.

Scheduler: “Scheduling is deciding how to allocate one or more resources to accomplish particular activities over time so that input demands are met in a timely and cost-effective manner. Most typically, this involves determining a set of activity start and end times, together with resource assignments, which satisfy all temporal constraints on activity execution, satisfy resource capacity constraints and optimize some set of performance objectives to the extent possible.” [8]

Virtual sensor: Virtual sensors are software modules that filter or combine sensory data in order to generate new information. →*Feature extraction* may be viewed as a virtual sensor. Monitors for control program state information may be implemented as virtual sensors, too.

References for Glossary of Terms

[1] Merriam-Webster online. Definition: Actuator, 2006. http://www.m-w.com/dictionary/actuator.

[2] Editors of Merriam-Webster, editor. Webster’s New Encyclopedic Dictionary. Federal Street Press, Jan 2002.

[3] Wikipedia. Definition: Feature extraction, 2006. http://en.wikipedia.org/wiki/Feature extraction.

[4] Patrick Doyle. Definitions: Planner, plan, scheduler, 2006. http://www.cs.dartmouth.edu/brd/Teaching/AI/Lectures/Summaries/planning.html.

[5] Inc. Stottler Henke Associates. Definition: Machine learning (glossary of AI terms, 2006). http://www.stottlerhenke.com/ai general/glossary.htm.

[6] Merriam-Webster online. Definition: Perception, July 2006. http://www.m-w.com/dictionary/perception.

[7] Encyclopedia Britannica online. Definition: Artificial Intelligence, July 2006. http://www.britannica.com/eb/article-219082?query=perception&ct=.

[8] EU NoE Planet. Definitions: Planner, plan, scheduler, 2006. http://scom.hud.ac.uk/planet/repository/.