Computational Models of Affordance in Robotics: A Taxonomy and Systematic Classification

Adaptive Behavior
2017, Vol. 25(5)

DOI: 10.1177/1059712317726357
www.sagepub.com/
SAGE

Philipp Zech[1], Simon Haller[1], Safoura Rezapour Lakani[1], Barry Ridge[3], Emre Ugur[2], and Justus Piater[1]

Abstract

J. J. Gibson's concept of affordance, one of the central pillars of ecological psychology, is a truly remarkable idea that provides a concise theory of animal perception predicated on environmental interaction. It is thus not surprising that this idea has also found its way into robotics research as one of the underlying theories for action perception. The success of the theory in this regard has meant that existing research is both abundant and diffuse by virtue of the pursuit of multiple different paths and techniques with the common goal of enabling robots to learn, perceive and act upon affordances. Up until now there has existed no systematic investigation of existing work in this field. Motivated by this circumstance, in this article we begin by defining a taxonomy for computational models of affordances rooted in a comprehensive analysis of the most prominent theoretical ideas of import in the field. Subsequently, after performing a systematic literature review, we provide a classification of existing research within our proposed taxonomy. Finally, by both quantitatively and qualitatively assessing the data resulting from the classification process, we highlight gaps in the research terrain and outline open questions for the investigation of affordances in robotics that we believe will help inform future work, prioritize research goals, and potentially advance the field towards greater robot autonomy.

1 Introduction

Over the last two decades, affordances have gradually occupied an increasingly important role in robotics (Jamone et al., 2016). The emerging subfield of developmental robotics, in particular, substantially builds on the idea of affordances (Cangelosi & Schlesinger, 2015). Affordances comprise one of the key concepts when it comes to formalizing and coding elements of exploratory learning where robots autonomously interact with the environment to develop an understanding of the world. In this work, we study the application of the concept of affordances in robotics by collating and categorizing its core aspects, and look to forming a structured approach to further investigation by analyzing the state of the field under this paradigm.

Horton, Chakraborty and Amant (Horton et al., 2012) provided a first brief survey on computational models of affordance in robotics. Recently Thill *et al.* (2013) and Jamone *et al.* (2016) published two thorough reviews on both theories and computational models of affordance (Thill et al., 2013) as well as the use of affordances in psychology, neuroscience and robotics (Jamone et al., 2016). Their work shows that the field of affordance related research is truly both vast and diverse, and has been growing rapidly since the early 2000's. However, these studies also show that, in practice, no common consensus about what really comprises a computational model of affordance exists. We see this as an unfortunate deficit in the field, as it is presently extremely difficult to quantitatively compare different computational models of affordance to study their competitiveness.

Ugur *et al.* (2011) published an early, coarse classification of various applications of the concept of affordance in robotics. In total, their classification considered 16 published works by studying their applied learning schemes and internal representations. Their classification assumes an internal representation of affordances in an agent's memory thus—by definition—contradicting Gibson's original formulation of the theory of affordances in which they are directly perceived (Gibson, 1966; Gibson, 1977; Gibson, 1979). Indeed, the question of whether or not computational models of affordances can ever hope to fully satisfy Gibson's conception of direct perception is fraught with difficulty and is touched upon throughout the remainder of this paper. Further to the emphasis on internal representation, the underlying taxonomy of Ugur *et al.* fails to capture the concept of affordances in full detail (see Section 3) by neglecting to study a number of important facets of their theoretical analysis, such as their inherent hierarchy, competitiveness and dynamics.

Jones (2003) claims that the inception and need for such an internal representation (Chemero & Turvey, 2007) stems from Gibson's evolving discussion

[1]University of Innsbruck, Department of Computer Science, Innsbruck, Austria.
[2]Bogazici University, Department of Computer Engineering, Istanbul, Turkey.
[3]Humanoid and Cognitive Robotics Laboratory, Department of Automatics, Biocybernetics and Robotics at the Jožef Stefan Institute, Ljubljana, Slovenia.

Corresponding author:
Philipp Zech, Intelligent and Interactive Systems Group. University of Innsbruck, Technikerstrasse 21a, Innsbruck, Tyrol Austria
Email: philipp.zech@uibk.ac.at

of his theory of affordances where Gibson used different definitions of them as being either properties of objects or, broadly speaking, emergent properties of the complementarity of animals and their environment (see Section 2). As this obviously exacerbates the study of affordances by constraining their existence to at least an environment, in this work, we motivate a perceptual shift on affordances. Instead of studying them just as emergent properties in the context of animal-environment systems, we ascribe them a constrained fundamental nature independent of perception (Gibson, 1966; Gibson, 1977; Gibson, 1979; Chemero, 2003; Michaels, 2003) (see Sections 2 and 3). This abstracted interpretation attributes to affordances a primary imperative role commensurate with those of real physical entities, which they actually are (Chemero, 2003). Using this interpretation as a basis, we study affordances in full detail to construct our taxonomy (see Section 4).

Min *et al.* (2016) recently published a survey on the use of affordances in developmental robotics. Despite citing close to 200 papers, their study only classifies about 30 papers on this research topic that have so far been published. Further, the classification of selected works that is provided by Min *et al.* similar to Ugur *et al.* fails to capture affordances in full detail (see Section 3) by only classifying them as either deterministic or probabilistic approaches. Aside from that, and as this work shall thoroughly demonstrate, the volume of publications resulting from the establishment of developmental robotics as a research discipline unto itself has been more expansive than the 30 papers by Min *et al.* by at least an order of magnitude.

The goal of our article is to define classification criteria for computational models of affordance in robotics that aim at capturing affordances over a sufficiently broad range of analytical viewpoints that have emerged from their theoretical interrogation. The defined classification criteria are crucial in providing a formal definition of the aspects of computational models of affordance for robotics. We further give a thorough overview of existing affordance related research in robotics by classifying relevant publications according to these criteria in a systematic way (see Section 5). As a result of this classification we then provide a comprehensive and qualitative discussion on existing research to identify both promising and potentially futile directions as well as open problems and research questions to be addressed in future (see Section 6). Though existing studies also give similar discussions (Ugur et al., 2011; Horton et al., 2012; Thill et al., 2013; Jamone et al., 2016), our discussion is distinct in being motivated by a quantitative study.

Contribution The core contribution of this article is the introduction of a comprehensive taxonomy for classifying computational models of affordance in robotics. A systematic search of the keywords *affordance* and *robot* resulted in 1980 hits, which were methodically (see Section 5) reduced to 398 considered papers. Out of those, we identified and classified 146 major contributions. For each publication it was possible to classify the employed computational models of affordance. Given the resulting classification we then discuss the current state of the art of affordances in robotics. Finally, on the basis of this discussion, we identify promising directions for future research.

Intentional Limitations In this work, only computational models of affordance that have an application in the field of robotics will be considered. Thus we will not discuss any research related to the concept of affordance from any other scientific fields such as psychology or neuroscience. Apart from that, we avoid classifying papers that just build on existing models (see Section 5).

Structure Section 2 reviews relevant theory on affordances and briefly discusses the role of affordances in robotics. Section 3 elaborates our refined definition of affordance in terms of the fundamentality of their existence independent of perception. Section 4 introduces our taxonomy for computational models of affordance in robotics by defining relevant classification criteria. Next, in Section 5, relevant publications are selected and classified according to our taxonomy. In Section 6 we then present and discuss our results followed by Section 7 which outlines open research challenges on the grounds of our classification results. Finally, Section 8 draws conclusions.

2 Theories of Affordance

Since Gibson's seminal definition of the term *affordances* as directly perceived action possibilities, much controversy has abounded regarding their exact nature. As Jones discussed (2003), the difficulty in understanding what exactly an affordance is, might be due to the "evolved" and "unfinished" formulation of the concept in Gibson's own writings. In the following we provide an overview of varying, and sometimes antithetic, definitions on the nature of affordances that have been formulated over the last number of decades.

A relevant summary and discussion of the history of affordance is given by Dotov *et al.* (2012). Contrary to their review however, we aim at discussing these theories from the perspective of roboticists. Towards the end of this section we link the discussed theories of affordance to robotics research and how it is currently applied.

Affordances

Gibson's original definition of the notion of affordances as set down in *The Senses Considered as Perceptual Systems* (Gibson, 1966) essentially defines them as action opportunities in an animal-environment system, that is, encapsulations of what objects in the environment *afford* an animal,

> When the constant properties of constant objects are perceived (the shape, size, color, texture, composition, motion, animation, and position relative to other objects), the observer can go on to detect their affordances. I have coined this word as a substitute for values, a term which carries an old burden of philosophical meaning. I mean simply what things furnish, for good or ill. What they afford the observer, after all, depends on their properties (Gibson, 1966, p. 285).

Obviously this vague and early definition allows for contentious discussions as to the nature of affordances. In his later work *An Ecological Approach to Visual Perception* (Gibson, 1979), Gibson aimed at complementing his earlier definition,

> 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 (Gibson, 1979, p. 127). . . . [Objects] can all be said to have properties or qualities: color, texture, composition, size, shape and feature of shape, mass elasticity, rigidity, and mobility. Orthodox psychology asserts that we perceive these objects insofar as we discriminate their properties or qualities. Psychologists carry out elegant experiments in the laboratory to find out how and how well these qualities are discriminated. The psychologists assume that objects are composed of their qualities. But I now suggest that what we perceive when we look at objects are their affordances, not their qualities. We can discriminate the dimensions of difference if required to do so in an experiment, but what the object affords us is what we normally pay attention to. The special combination of qualities into which an object can be analyzed is ordinarily not noticed (Gibson, 1979, p. 134).

One can notice how the concept evolved: the first quote (Gibson, 1966, p. 127) explicitly states that the perception of the properties of objects such as size and shape are required for (or at least precede) affordance detection, whereas the second quote (Gibson, 1979, p. 134) suggests that that the discrimination of such qualities is not the basis for or the precursor to their detection. According to Jones (2003), this latter description of the perception of affordances as being quite distinct from the perception of object properties (Gibson, 1979, p. 134) differs markedly from his earlier description where it is asserted that once an object's properties are initially perceived, it is then possible to subsequently detect its affordances (Gibson, 1966, p. 127). Although Gibson had aimed at clarifying what is actually required from the environment such that affordances can emerge, these differing definitions led to a series of decade-lasting contentious discussions on their true nature. It is worth emphasizing that, being one of the founders of the field of ecological psychology, Gibson naturally assumed that affordances (and other perceivable entities) are directly perceived. We do not seek to question this, but—in an effort to both dissect and illuminate Gibson's argument—are sympathetic to this view, and attempt to analyze ways in which direct perception might be both interpretable and implementable in robotics research in the remainder.

Barwise (1989) was one of the first to comment on Gibson's idea of affordances by speculating about their meaning. According to his view, their meaning resides in the interaction of real, living things and their actual environment (Barwise, 1989), thus providing an important description of affordances as being emergent relational properties of the mutuality of both an animal and its environment. This mutuality or reciprocity of the relations between animals and the environment was already emphasized by Gibson (1977). Broadly speaking, it is the complementarity of an object's functionality and an animal's ability that allows the animal to perceive this functionality as an affordance that can be acted upon by exploiting the ability. Section 3 further elaborates on this homeomorphism between functionality and affordance.

Contrary to Barwise, Turvey argues that affordances are properties of objects in the environment that are of a dispositional nature. More precisely, Turvey considers affordances to be dispositional properties of the environment, i.e., the means by which it affects the agent's behavior and its constituent objects, complemented by dispositional properties of an animal (Turvey, 1992). This, however, immediately implies the presence of an animal such that affordances become apparent, i.e., that they require such a context to *become manifest* or to be *actualized*. Thus, in Turvey's view, affordances essentially only exist under specific circumstances. Apart from that, Turvey considers affordances central to prospective control, ultimately conceding them a leading role in shaping an animal's behavior.

Vera and Simon (1993) argue that affordances comprise internal symbolic representations that arise from semantic mappings between symbolic perception and action representation. According to them, such an internal representation is vital for affordances to be part of a cognitive system. This, however, seems to break with Gibson's classical view of affordances as being directly perceived. If a mental representation has to be built in the first place, then it can be argued that direct perception is not in fact occurring. Ultimately they also nest affordances in the environment instead of treating them as emergent properties of the animal-environment system.

In his book *Encountering the World* (Reed, 1996), Reed defines affordances as resources for an animal at the scale of behavior. Similarly to Barwise, Reed also sees affordances as emergent properties. They are characteristics attributable to the intrinsic properties that features, objects, and events possess, by virtue of their structure, that are specified relative to a specific perceiver-actor (Reed, 1996). Reed thus provides an interpretation of affordances that does not rely on either the environment or an agent, but—as Gibson already suggested—on their mutuality. A key characteristic of Reed's interpretation of affordances is that they exert selection pressure, thus—similar to Turvey—pointing out their relevance for regulating behavior.

Half a decade later, Shaw (2001) argued that affordances are intentional— a harness for directing causes but not causal by themselves. That is, the true purpose of affordances is to satisfy the intentions of an animal. They comprise a source of information for an animal for taking behavioral decisions. In his argument, intentionality is satisfied by effectivity, i.e., the specific combination of the functions of tissues and organs taken with reference to an environment (Shaw, 1982). Further, he argues that affordances and effectivities are complementary to each other by defining the boundary conditions of the perceiving–acting cycle (Shaw, 2001; Heft, 2003). In this vein, affordances remain dormant until completed by an effectivity of an agent's intention. He therefore, again, ultimately defines the reference of affordances to be the environment.

Steedman (2002) presents a definition of affordances in terms of a computational representation. In this spirit, his fundamental idea is to define mappings between actions (functions) and affordances. That is, by perceiving the object and having knowledge about its functionality, affordances are transparently perceived. Further, he requires specific preconditions to be fulfilled in order for corresponding actions to be applicable. Thus, Steedman ultimately—similarly to Turvey—treats affordances as properties of objects and their related functionality.

Heft (2003) treats affordances as percepts, as multidimensional, intrinsic properties that objects and events possess as part of their make-up; their value however is extrinsic (Heft, 2003). This follows from his view of affordances as being action-related qualities that are context dependent. Ultimately, he defines affordances to be dynamical functional relations of environmental processes and animal-related factors. Hence, they are not static properties of the environment but rather features embedded in a confluence of ongoing dynamic processes in a continuously changing world. Heft, in this way, similarly to Barwise and Reed, also takes the view that affordances are emerging properties of the animal-environment system.

Stoffregen argues in a similar vein as Barwise, Reed, and Heft. However, he explicitly treats affordances as emergent *relational* properties of the animal-environment system that do not inhere in either an animal or the environment but in their mutuality (Stoffregen, 2003). Stoffregen also considers affordances as determining what can be done, again thereby—similarly to Turvey and Reed—pointing out their relevance for behavior sculpting. Central to his argument is that affordances are—as Gibson already suggested—direct percepts, that is, they are directly perceived by an animal without any further cognitive processing.

Chemero argues in agreement with Stoffregen's treatment of affordances by stating that they are emergent relational properties between the abilities of animals and features of the environment (Chemero, 2003). He describes the process of perceiving affordances as one involving the placement of features in the environment, while pointing out that affordances are directly perceived as the activities a situation allows for. Another important aspect of Chemero's argument is related to the existence of affordances. Chemero argues that affordances essentially exist without the presence of an animal, however, for describing a real physical entity that can be studied, an affordance requires at least the *potential* existence of an animal that could perceive it. Chemero thus, similarly to Reed, provides an interpretation of affordances that allows them to exist outside the presence of an animal. Gibson however, though arguing that affordances are always perceivable, does not discuss the constraints that permit the existence of an affordance after all (Gibson, 1977) thereby—as we believe—attributing them a rather fundamental, independent nature,

> The concept of affordance is derived from these concepts of valence, invitation, and demand but with a crucial difference. The affordance of something does not change as the need of the observer changes. The observer may or may not perceive or attend to the affordance, according to his needs. But the affordance, being invariant, is always there to be perceived. An affordance is not bestowed upon an object by a need of an observer and his act of perceiving it. The object offers what it does because it is what it is. To be sure, we define what it is in terms of ecological physics instead of physical physics, and therefore it possesses meaning and value to begin with. But this is meaning and value of a new sort (pp. 138-139).

Describing affordances as being *invariant* supports Chemero's existence criterion in terms of being bound to the perceivability by at least one animal. If not perceivable, we resort to just dealing with a functionality of an object. Turning a functionality into an affordance is subjective. It is based on—as pointed out by Gibson—an animal inherent value system, some means of intrinsic or extrinsic motivation that goads the animal. Section 3 further elaborates on this homeomorphism between functionality and affordance.

Michaels (2003) finally gives an interpretation on the notion of affordances and their origin which is in line with both Gibson's and Chemero's views as being emergent properties embodied in relations of the animal-environment system. In Michaels' view, affordances refer to a specific action level, that is, they are not to be equated with arbitrary action components or action aggregation. Additionally, according to Michaels, affordances exist independently of their perception, but require effectivities for their actualization, and are subjective. In other words, an affordance exists in the absence of an animal with the effectivities necessary to perceive it, but cannot be made manifest in terms its perception in the absence of those effectivities, and moreover, those effectivities are particular to a given animal. In this sense, Michaels contradicts Gibson's theory of affordances (Gibson, 1979),

> [...] These are the invariants that enable two children to perceive the common affordance of the solid shape despite the different perspectives, the affordance for a toy, for example. Only when each child perceives the values of things for others as well as for herself does she begin to be socialized (p. 141).

Biological evidence for such *mirrored* perception is evident by the well-adapted theory of mirror neurons (di Pellegrino et al., 1992; Rizzolatti et al., 1996). Further, Thill *et al.* (2013) in their survey give an ample discussion of the tight link between mirror neurons and affordances thereby reinforcing Gibson's argument.

To summarize our discussion from this section Table 1 provides an overview of the different concepts various authors used for discussing their view on affordances.

The Role of Affordances in Robotics

The concept of affordances as imagined by Gibson provides a powerful notion for making sense of, and finding meaning in, the environment or in a given situation. As such, affordances comprise a powerful tool for learning, reasoning and behavioral decision making for artificial systems. Duchon *et al.* (1998) were probably the first roboticists to recognize the relevance of certain ideas of ecological psychology for building autonomous robots. In their work they investigated the notion of direct perception by capitalizing on optical flow as a source of information for decision making using Warren's law of control (1988) to avoid obstacles in order to survive in a game of tag. Nearly a decade later, Sahin *et al.* (2007) set forth a novel formalization of the concept of an affordance and its subsequent application in mobile robots for autonomous navigation. Though in their interpretation, affordances are also relations, they break with an important aspect of Gibson's theory of affordances by defining

Table 1. Summary of the various definitions of affordances as given by the previously discussed authors as well as the concepts on which their definitions build.

Author	*Definition*	*Concepts*
Gibson (1966, 1977, 1979)	emergent object properties	valence, invitation, demand, mutuality
Barwise (1989)	emergent relational properties of the animal-environment system	functionality, ability, mutuality
Turvey (1992)	dispositional properties of the environment	prospective control
Vera & Simon (1993)	semantic mapping of symbolic perception and action representation	mental representation
Reed (1996)	resources for an animal for decision making	mutuality
Shaw (2001)	source of information for decision making	intentionality, effectivity
Steedman (2002)	mapping between action (function) and affordances	computational entities
Heft (2003)	dynamical functional relations of environmental processes and animal-related factors	intrinsic properties, extrinsic value
Stoffregen (2003)	emergent relational properties of animal-environment system	mutuality
Chemero (2003)	emergent relation properties of the animal-environment system	activities
Michaels (2003)	emergent properties embodied in relations of the animal-environment system	independence, subjectivity

these relations to be mental representations. This immediately requires some kind of cognitive processing such that these representations can emerge. Thus, according to Sahin *et al.*, affordances are neither properties of an object or an animal, nor properties that emerge from the animal-environment system, but rather properties that emerge from cognitive processing. This, however, contradicts Gibson's assumption in his theory of affordances that states that they are directly perceivable. According to Chemero and Turvey, such a representationalist view as taken by Sahin *et al.* is unnecessary. Direct perception along with dynamical modeling already comprise the necessary conceptual tools for equipping artificial systems with the ability to perceive Gibsonian affordances (Chemero & Turvey, 2007).

The discussions of this section clearly reveal that over the past few decades there has been much controversy regarding the nature of affordances, that is, regarding how affordances come about and where they ultimately come from. We claim that the definitions of both Gibson and Chemero are complementary. Whereas Gibson defined the original meaning of affordances as being directly perceivable action opportunities, Chemero described the nature of their constrained existence as to be perceivable by at least one animal (Chemero, 2003). Such a definition seems quite rational as it defines affordances to be properties that are subjective to an animal given its embodiment and the features of objects in the animal's environment. This subjectivity was also pointed out by Michaels, who additionally emphasized that affordances exist independently of perception (Michaels, 2003). We treat affordances—as Chemero suggested—as real physical entities of fundamental nature that exist independently of perception— as highlighted by Michaels (Section 3). This is strongly in line with Gibson's fundamental argument of his theory of affordances that they are directly perceivable. It is this interpretation of affordances that nourishes our extended definition of affordances and motivates our taxonomy as elucidated further in the two subsequent sections.

One open question that remains unanswered however is *why is robotics research related to affordances apparently unaffected by these controversies* as they exist in ecological psychology. We guess that this is due to the circumstance that generally, roboticists view the idea of affordance as an interesting concept that can be helpful in the development of autonomous agents, yet they neglect to stick to its original Gibsonian definition with regard to direct perception. As our classification shows (see Section 5), it is generally the case that most research circumvents this problem of direct perception by drawing on an internal representation of affordances. Though this clearly contradicts Gibson's original formulation, it immediately allows for modern machine learning techniques to build computational models and structures that can easily be processed and manipulated for learning. This paradigm is also evident from the results of our classification (see Section 6), which clearly support our argument that roboticists generally extract features as a basis for affordance detection and learning, thereby implicitly building an internal representation.

3 An Extended Definition of Affordances

As argued earlier, we attribute affordances a fundamental, independent nature. Informally speaking, this defines affordances as properties of the animal-environment system that essentially are always present. This has already been argued by Gibson (1977),

> An important fact about the affordances of the environment is that they are in a sense objective, real, and physical, unlike values and meanings, which are often supposed to be subjective, phenomenal, and mental. But, actually, an affordance is neither an objective property nor a subjective property; or it is both if you like. An affordance cuts across the dichotomy of subjective-objective and helps us to understand its inadequacy. It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither. An affordance points both ways, to the environment and to the observer (p. 129).

The crucial question then is how can such an interpretation be in line with the fact that affordances are also emergent properties of the animal-environment system? To answer this question, we first want to elaborate on the difference between an affordance and a *functionality* which is crucial for our argumentation. Obviously there is a tight relationship between these two terms, not least because an affordance eventually possesses a functionality. Softening this argument, one essentially arrives at the equality of these two terms, that is, an affordance *is* a functionality. In favor of Gibson, we argue that this equality is sound. However, we also see a clear difference between functionality and affordance. To become affordable by an animal an object's functionality must relate to an animal's embodiment. Only then is it a germane affordance for an animal. Otherwise, it only comprises another source of information that—by slightly abusing terminology—in the context of our study at least, is also treated as an affordance. Consider for example a stone used for hammering. Contrary to a hammer as such—remember, it had to be invented to fulfill its purpose—the functionality of hammering is fundamental, or arguably primordial, to stones (of course proper size and shape are required for effectivity). However, perceiving this functionality as an affordance requires (i) knowing about hammering and (ii) being in need of a hammer. It is now clear that functionalities of tools essentially are there in nature. However, to be perceivable we first must learn about them. At the end of the day, this boils down to learning new affordances. Section 4 further elaborates on this issue.

According to this separation of functionality and affordance, it now becomes evident how affordances actually emerge in an animal-environment system. No matter whether perceived or not, an object's or the environment's functionality is fundamental or possibly even primordial. We argue that these fundamental facets of object functionality already comprise affordances that essentially can always be perceived but not yet necessarily acted upon. However, upon relating an affordance to an animal's skills and its homeostasis it immediately emerges as a reasonable action opportunity from the mutuality of the animal and the object or the environment by becoming meaningful. According to Chemero, this is expressed by fully grounding[*] the relation

$$\text{Affords–}\phi(\text{feature, ability}),$$

[*]Observe that partial grounding by a feature according to our argument then just entails a functionality.

where ϕ denotes a specific affordance, i.e., support or danger. Clearly, this describes an emergent property that becomes real and physical upon an animal possessing a specific ability it can relate to a feature in the environment in order to apply an affordance. Despite thus providing an elementary and pivotal formalism for the nature of affordances, we however argue that this relation does not capture the various aspects of affordances in full detail as discussed next.

Recent research (Ellis & Tucker, 2000; Wells, 2002; Cisek, 2007; Borghi & Riggio, 2015) has shown that affordances—outside of their fundamentality and independence—are, further to that, of a hierarchical, competitive and dynamic nature. This requires an extension of the relational scheme as introduced by Chemero (2003).

The Hierarchy of Affordances

The hierarchical nature of affordances is fundamental. If not so, mammalian cognition could only reason in terms of atomic actions instead of being able to develop solutions for complex higher-level action goals. Upon this notion, Ellis and Tucker (2000) introduce their concept of micro-affordances. They argue that micro-affordances are the potentiated components of a perceived affordance. For example, in the case of grasping, perceiving graspability of an object triggers specific preparation actions, execution actions and possibly cleanup actions. These are referred to as the micro-affordances, the effects of affordance perception on action selection. Following their line of reasoning, affordances are a rather abstract entity that trigger multiple micro-affordances in sequence to be acted upon. Consequently, micro-affordances are low-level actions that need to be taken to achieve some higher-level action goal, i.e., applying a higher-level affordance subsumes these micro-affordances. Observe that Chemero's classical interpretation of the structure of affordances, i.e., Affords–ϕ(feature,ability), fails to capture such hierarchical relations (see Section 2). This is as a result of defining them as constrained emergent relational properties of the animal-environment system, i.e., features of both the animal and the environment. Yet, the relation does not capture connections between affordances such that their hierarchy can be studied.

Turvey also discussed the hierarchical nature of affordances by attributing them the foundation on which other affordances might be based (Turvey, 1992). From a behavioral point of view, these interrelations and dependencies between affordances play a crucial role for action selection and behavior sculpting that goes beyond the present moment and allows planning into the future by chaining different affordances together. This ultimately results from the coordination of perception and the possibilities for action that a situation affords at a given time and place (Wells, 2002). For capturing this idea of a hierarchy among affordances coupled with the paradigm of behavior affording behavior (Gibson, 1977) in a relational manner, we introduce the relation

$$\text{Affords–}\Phi(\text{Affords–}\phi_1, \text{Affords–}\phi_2, \ldots, \text{Affords–}\phi_n,\ \text{ability})$$

We define Affords–Φ to be an $(n+1)-$ary relation that subsumes an arbitrarily large but countable number of affordances. Here, Φ denotes the higher-level action goal in terms of an affordance. Informally speaking, Affords–Φ collects low-level actions to be taken to achieve a higher-level action goal relative to an animal. Observe that this relation also includes an ability. However, in this context, it does not refer to some low-level motor skill but rather to cognitive reasoning abilities that allow for thinking in terms of higher-level action goals.

The Competitiveness of Affordances

The competitive nature of affordances is obvious. Without competitiveness, there could be no concept of deliberate choice as, essentially, at every instant of our lives we would not have the option of choosing between different action possibilities, but would be forced to take a preordained next step. Cisek studied the idea of affordance competition by developing a computational model that mimics action selection in animals (Cisek, 2007). His research is motivated by the single argument that at every moment, the natural environment presents animals with many opportunities and demands for action (Cisek, 2007). This obviously requires some mechanism that decides which affordance is ultimately acted upon in an environment. Cisek's model presumes such a mechanism to be present in the cognitive system of an animal and grounds its workings on an internal representation. This, however, contradicts Gibson's argument that affordances are directly perceived. Contrary to Cisek, we argue that affordance competition does not occur in the cognitive system of an animal but rather is the result of the needs and embodiment of an animal in relation to the stimulus of an affordance to cause selective attention, i.e., by focusing on what provides for good rather than for ill. This competition ultimately happens at the affordance level in an animal-environment system, and we formalize it by introducing the relation

$$\text{Competes–}\Upsilon(\text{Affords–}\Upsilon_1, \text{Affords–}\Upsilon_2, \ldots, \text{Affords–}\Upsilon_n)$$

where $\Upsilon \in \{\phi, \Phi\}$. That is, we allow affordance competition among both affordances relating to low-level action goals as well as among those relating to higher-level action goals where Υ denotes the affordance under question. The competitiveness is reflected in forcing all usages of the Afford–Υ predicate to ground (i.e., to bind to a specific affordance instance) either to Φ or ϕ. The arity of Competes–Υ is ultimately restricted to an arbitrarily large but countable number of affordances that may compete with each other. However, we do not constrain the ability necessary for applying an affordance as (i) it may differ for distinct affordances or (ii) may even be ungrounded (which would relate to some plain functionality). This latter treatment is vital to also allow functionalities to compete with each other, e.g., hammering with a stone or an actual hammer.

The Dynamics of Affordances

Attributing affordances a dynamic nature is crucial for explaining effects in a changing environment. Without having an effect that reveals novel affordances or expunges previous ones, the state of an animal-environment system would not change. Such an existent nature was already suggested by Chemero in constraining affordances to be perceivable by at least one animal (Chemero, 2003). Borghi and Riccio further elaborate on this idea of the dynamics of affordances (Borghi & Riggio, 2015). They argue that affordances can be both stable and variable. Stable affordances relate to static object properties[†], e.g., graspability at the handle of a mug, whereas variable affordances relate to temporary affordance of an object or the environment, i.e., fillability in the case of an empty mug, or *jumpability* over a stream in the case of taking a run-up instead to attempting to make a standing jump. This dynamic nature was already anticipated by Gibson when he suggested that acting upon a given affordance causes new affordances to become available, that is, behavior affords behavior (Gibson, 1977). This clearly shows a strong interplay and dependency among the affordances available in an animal-environment system. This issue also becomes apparent in the case of the hierarchy of affordances (see Section 3) which, at the end of the day, is implicitly shaped by these dynamics by virtue of affordances being acted upon. We capture these dynamics by introducing a third and final relation

$$\text{Changes–}\phi(\text{Affords–}\phi, \text{ability}),$$

where again ϕ relates to a specific affordance. Further, we also relate this predicate to a specific ability indicating what skill is necessary to change the dynamic state of an object or the environment that is bound to ϕ. Observe also that this predicate is not applicable to Φ in terms of higher-level action goals. This is by virtue of the fact that the dynamics of affordances eventually are nested at the feature level of an object or the environment relative to a temporary property.

Summarizing the above discussion, the competitive and dynamic hierarchical structure of affordances can thence be described by the four relations

$$\begin{aligned}
&\text{Affords–}\phi(\text{feature, ability})\\
&\text{Affords–}\Phi(\text{Affords–}\phi_1, \text{Affords–}\phi_2, \ldots, \text{Affords–}\phi_n,\ \text{ability})\\
&\text{Competes–}\Upsilon(\text{Affords–}\Upsilon_1, \text{Affords–}\Upsilon_2, \ldots, \text{Affords–}\Upsilon_n)\\
&\text{Changes–}\phi(\text{Affords–}\phi, \text{ability})
\end{aligned}$$

where $\Upsilon \in \{\phi, \Phi\}$. Our novel formalization of the structure of affordances does not break with Gibson's argument that affordances are directly perceived. In Chemero's theory, direct perception is expressed by the relation

$$\text{Perceives}[\text{animal, Affords–}\phi(\text{feature, ability})].$$

Clearly, in this form an animal could only perceive affordances related to low-level action goals. To allow for direct perception of both affordances relating to low-level as well as higher-level action goals, we refine Chemero's definition to

$$\text{Perceives}[\text{animal, Affords–}\Upsilon]$$

where $\Upsilon \in \{\phi, \Phi\}$. This readily makes affordances relating to higher-level action goals directly perceivable by an animal.

[†]Observe that this does not contradict Chemero's notion of existence (Chemero, 2003).

As a final remark, let us briefly elaborate on Affords–Φ. One might argue that this relation, despite solidifying the hierarchical structure among affordances, does not account for their competitive and dynamic aspects as expressed by Competes–Υ and Changes–ϕ. However, we argue that Affords–Φ is implicitly dependent (i) on the availability (e.g., some temporal property) of an affordance to be grounded, and (ii) only perceivable affordances, i.e., Affords–Φ holds, can compete. Moreover, only affordances that appear more beneficial (e.g., won the competition) are considered in constructing a hierarchy that depicts a higher-level action goal.

To conclude, we want to clarify that we do not want to claim to redefine the notion of affordances. We rather see our contribution as a comprehensive summary of how the this notion has evolved over the years by extending Chemero's original formalism. Interestingly, these novel facets of affordances do not contradict Gibson's original formulation. In fact, they are just refined explanations of his original thoughts on which he did not elaborate more thoroughly during his life.

4 Classification Criteria for Computational Models of Affordances

It is obvious that the underlying notion of an affordance is paramount in a computational affordance model for it to succeed. We thus represent a computational affordance model as the union of both an abstract *affordance model* and an underlying *computational model*. In this spirit the former model then deals with perceptive, structural and developmental aspects, that is, the embodiment of affordances. Complementary to this, the computational model addresses relevant aspects related to the concrete implementation of the mechanics of affordance perception.

Figure 1 gives an overview of our taxonomy and its classification criteria.

For each of the criteria discussed in Sections 4.1 and 4.2, if a clear choice cannot be made due to missing information in a paper, the criterion is categorized as *not specified*.

Affordance Model Criteria

Affordance model criteria provide means to study the structure of affordances for a specific computational model (see Computational Model Criteria). Thus, these criteria describe perceptive, representational and developmental aspects of the model under evaluation. Naturally, both the perceptive and developmental aspects could be regarded as being part of the implementation of a computational affordance model rather than of the affordance model itself. We however justify their inclusion as part of the affordance model by arguing that, in both cases, they ultimately ground the means by which an understanding of affordances can be developed in the first place given the close association between the embodied aspects of a computational affordance model, *what* affordances are modeled, and *how* they end up actually being modeled.

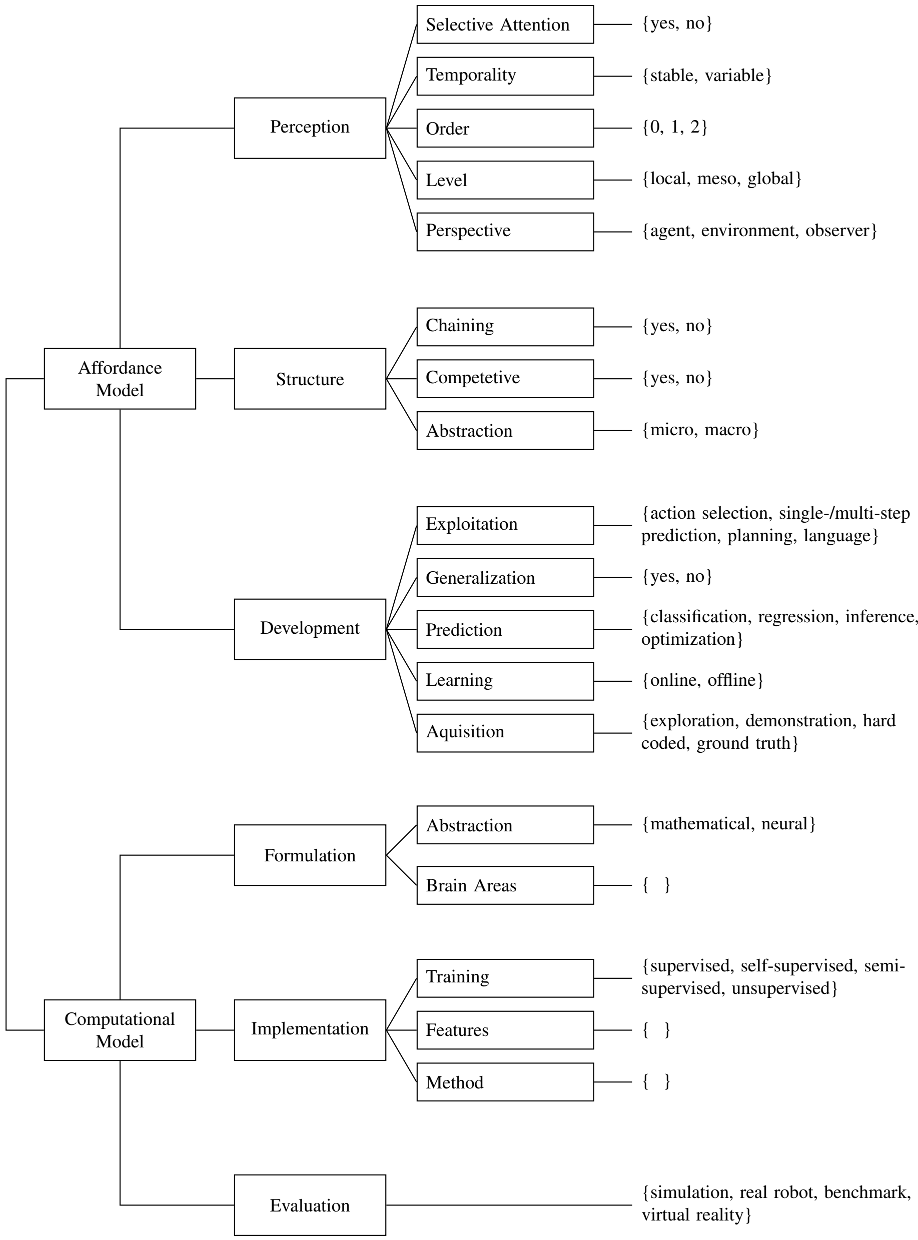


Figure 1. Overview of our taxonomy for classifying computational models of affordances in robotics. For the sake of clarity, the choice *not specified* is excluded.

Perception Perceptive aspects of our affordance model study the means by which the agent[‡] that embodies the model under study perceives and understands affordances.

[‡]Observe that an agent may also comprise an animal. However, since we are studying affordances in a robotics context we deliberately replace the animal in the animal-environment system by an agent to put more emphasis on the artificiality of the primarily interacting entity.

These are essential drivers for the development of a meaningful interpretation of affordances, as they identify the main input modalities in terms of affordance perception.

Perspective The perspective identifies the view from which affordances are perceived. Gibson originally did not consider such an aspect, which initially led to much confusion. Obviously, the prevalent perspective in robotics is that of an agent's point of view. Yet, to allow for a precise distinction, according to Sahin (Şahin et al., 2007) three perspectives are necessary:

- The *agent*'s perspective, as just stated, is the prevalent one in robotics, where an agent perceives an affordance from their own perspective depending on the embodiment of its skills and capabilities (see Order) (see Table 2, rows 1-15).
- From the perspective of an *environment*, affordances are perceived as extended properties of the environment itself. This yields an agent that *attends* to the environment until it is offered an affordance. This is, in contrast to the agent's perspective, a sense of passive affordance perception, as the agent does not actively *seek* them.
- The *observer* perspective essentially classifies a computational affordance model as a multi-agent model (see Table 2, rows 16, 21, and 48). In this context, an agent does not perceive affordances at all but instead relies on an additional agent — the *observer* — that provides relevant information by embodying relevant affordance perception modalities.

Level In the context of a concrete scene of the environment, affordances can be perceived at different levels of the scene. That is, an agent may perceive affordance at three different levels:

- At the *local* level, an agent perceives affordances at small object regions, e.g., specific object patches (see Table 2, rows 1, 8, and 22). This has the advantage that an agent can immediately predict how to interact with specific objects, yet at the expense of discarding any additional semantic or contextual information on the concrete object or its situatedness in the environment, respectively.
- At the *meso* level, an agent perceives affordances in terms of complete objects instead of only small patches (see Table 2, rows 15-21). This has the advantage that the agent can use information about object actual state, its physical properties, and relationships between *patches* to perceive more complex affordances such as task-based grasping, pounding, and pouring.
- At the *global* level, an agent perceives affordances in the context of its environment (see Table 2, rows 3-6). This has the advantage that an agent can situate itself to create a more complete view of the world and its current action possibilities. This then opens up the possibility of chaining affordances (see Chaining) to achieve higher-level cognitive tasks. However, alternating situatedness may introduce a substantial degree of uncertainty, such as that produced by occlusion and clutter, that may tamper with accurate perception.

Observe that an agent may perceive both stable and variable affordances (see Temporality) at all levels, as well as of all orders (see Order).

Order The *order* of affordances (Aldoma et al., 2012) is related to the set of skills embodied as well as the actual embodiment of an agent. In this sense, this criterion classifies which kinds of affordances the agent—given its skills and embodiment—can perceive and make use of:

- *0th-order* affordances are those which are found on objects *irrespective of their current state in the world* (see Table 2, rows 1, 24, and 61). From a cognitive point of view, this allows for the creation of a semantic model of an object in order to understand its functionality. That said, 0th-order affordances ultimately comprise a specific functionality (see Section 3).
- *1st-order* affordances can be seen as immediate action possibilities offered to the agent *with respect to its embodiment and skills* (see Table 2, rows 5-23). In this way, not only the affordances depend on the embodiment and skills of the agent but also on the current state of the object and the environment.
- *2nd-order* affordances are those that appear as a consequence of an agent acting upon a specific affordance (see Table 2, rows 4, 28, and 40). Perceiving such affordances is crucial for affordance chaining (see Chaining) as well as for achieving higher-level cognitive tasks.

Observe that this criterion has an intrinsic hierarchical structure as both 1st- and 2nd-order affordances are a subset of 0th-order affordances. Further, 2nd-order affordances can be seen as a subset of 1st-order affordances, as they also represent immediate action possibilities, yet are only applicable if a necessary action has been taken previously.

Temporality Temporality relates affordances to the statics and dynamics of the environment or object with respect to an agent's behavior (Thill et al., 2013; Borghi & Riggio, 2015). This aspect is important given the nature of the world we live in, which essentially is a highly dynamical system that may rapidly change over time (see Section 3, *The Dynamics of Affordances*). To account for such changes, it is important to both be able to recognize them and to be able to react to them. We reflect this capability within an affordance model in terms of an agent's ability to differentiate between both *stable* and *variable* affordances:

- *Stable* affordances are affordances that are related to static object properties, e.g., a mug's handle usually affords graspability, which normally does not change over time (except in the case where the handle breaks off or is occluded) (see Table 2, rows 1-3).
- *Variable* affordances, in contrast, are affordances that depend on dynamic object properties, that is, properties that change over time (see Table 2, rows 59, 67, and 139). Considering again the concrete case of a mug, a variable affordance is *fillability*, which clearly changes over time. A mug that is already full cannot be filled any further.

Observe that this criterion is open to *multiple selection*. Perceiving stable affordances does not exclude perception of variable ones. Equally, perceiving variable ones also does not exclude perception of stable ones.

Selective Attention Selective attention comprises an interesting aspect when it comes to perceiving affordances. If an agent embodies selective attention it essentially is capable of focusing on those affordances only needed to reach a specific goal. Such a behavior is desirable to efficiently learn and achieve higher-level actions and action goals, respectively, by focusing on the task at hand and inhibiting any disturbances and noise from the surrounding world. This criterion supports two choices, viz. *yes* or *no*.

Structure Structural aspects of our affordance model study how the perceived affordances are organized in the environment. That is, they describe the relationship and interplay between the different affordances offered to an agent. From a cognitive point of view, this defines the means that ultimately allow an agent to *reason* about its actions and the resulting effects in the world around it.

Abstraction Generally affordances are considered at an atomic level, that is, they describe a single action possibility offered by the environment. However, treating affordances only at such an atomic level hinders an agent in constructing a more thorough picture of the world and in developing a deeper understanding of its embodiment. Thus, our taxonomy considers both atomic affordances, and also combinations of these atomic affordances, that allow an agent to *think* in terms of higher-level action goals (see Section 3, *The Hierarchy of Affordances*).

- *Micro* affordances represent single, atomic affordances (see Table 2, rows 70-88). The term atomic shall indicate that such an affordance cannot be deconstructed beyond representing a single, atomic action. From a developmental point of view, micro affordances correspond to low-level action goals that are achieved using a single, primitive motor skill.
- Contrary to micro affordances, *macro* affordances represent, from a cognitive point of view, higher-level action goals that result from achieving a series of low-level action goals (see Table 2, rows 57, 60, and 62). From a developmental point of view this corresponds to learning that performing specific actions in the correct order can result in achieving higher-level action goals.

Competitive Competition among affordances introduces to an agent a certain degree of deliberate choice (see Section 3, *The Competitiveness of Affordances*), by enabling it to choose between identical (or infinitesimally similar) affordances that yield identical (or infinitesimally similar) outputs. Clearly, competition among semantically unrelated affordances is to be neglected as this is driven by the agent's homeostasis and happens at a behavioral level. Thus, by allowing identical (or infinitesimally similar) affordances to compete with each other, an agent can not only react properly based on the offered action possibilities, but further rely on a weak notion of free will in terms of deciding which affordance to choose. More formally, if and only if an agent decides deliberately between similar but spatially separated affordances (e.g., different chairs), then the perceived affordances in the environment are competing. This criterion supports two choices, viz. *yes* or *no*.

Chaining Chaining of affordances is essential for an agent to learn and reason about higher-level action goals (see Abstraction). Obviously, an agent needs to have the ability to chain affordances by relating actions to effects and in turn, effects to novel action possibilities, not only to be able to understand a complex task overall, but also how to achieve it given its embodiment. This resembles Gibson's idea of behavior affording behavior (Gibson, 1977). Observe however that being able to chain affordances does not yet imply that an agent understands higher-level action goals or how to reason about macro affordances (see Abstraction). Further, we want to clarify that chaining does not only refer to strict sequential application of affordances, but may also refer to overlapping affordances, e.g., in bimanual manipulation tasks. Chaining of affordances can also happen at a low level in terms of repetitive execution of some primitive motor skill. This criterion supports two choices, viz. *yes* or *no*.

Development Finally, developmental criteria of our affordance taxonomy analyze how an agent learns about new affordances. Although we previously discussed perceptual aspects of the taxonomy, these should not be confused with developmental aspects. In order to develop knowledge by learning about the world, an agent requires some means of perception. Thus, this dimension of our taxonomy of affordance model relates to developmental processes that allow agents to *learn* how to interact with the world.

Acquisition For an agent to be able to make meaningful use of its knowledge of the affordances in the environment, it must first be endowed with a systematic means of acquiring that knowledge, not only in terms of the data involved, but also in terms of the meaning behind the data. This acquisition can happen in multiple ways, from fully autonomous learning to bootstrapped learning in terms of hardcoded knowledge:

- *Exploration* is the most common strategy for affordance acquisition (see Table 2, rows 5-9). As inspired by the cognitive development of children, an agent learns about affordances by (i) recognizing objects, (ii) performing simple interactions with them, i.e., pushing, pulling, and the like, and (iii) observing the effect triggered by the performed action. By repeating this process with different objects an agent then gradually learns about the affordances offered in its environment.
- Programming by *Demonstration* (PbD) is a popular concept in robotics that is used to teach a robot a manipulative task (see Table 2, rows 117-120). In such a setting, an agent is usually guided by some teacher while remembering the specific movements and interactions it is taught. Compared with exploratory learning of affordances, this kind of learning forces an agent to adopt the role of an observer as it does not perform any movements autonomously.
- Providing an agent with *ground truth* essentially boils down to supervised learning where the agent learns affordances from examples in terms of features in its environment (see Table 2, rows 3, 13, and 14). These examples are generally provided by means of 2D video or image data, but also sometimes 3D scene data, and are usually annotated with *0^{th}-order* affordances (see Order).
- *Hardcoded* affordance information avoids the acquisition and learning process altogether (see Table 2, rows 53, 54, and 66). In such a setting, the agent is initially provided with a set of known affordances and rules with which to detect them in the environment. Using annotated action information it then interacts with the world using corresponding, also hardcoded, motor skills.

It should be noted that, apart from exploratory learning of affordances, in all cases the agent is essentially forced to "live" with what is provided and lacks any possibility to learn more about the world autonomously. We believe that the latter circumstance is not useful for affordance learning, as affordances, in general, depend on the embodiment of the agent itself. It thus should learn on its own what it is able to do instead of being told what it can do in order to foster an understanding of its own embodiment. However, for the sake of completeness, we also include these choices in our taxonomy.

Prediction After learning, in order to make use of its knowledge, an agent needs to be skilled in predicting the affordances in its environment. That is, given its perceptions, it needs some way of predicting which affordances will emerge in the environment. We are well aware that this criterion has a strong intersection with the computational model of our taxonomy (see Computational Model Criteria) as it entirely relies on its implementation (see Implementation). However, we argue that for the sake of completeness of the developmental dimension of our taxonomy, the means of predicting affordances are necessary, as this is also an important aspect of the development of affordances in terms of how the agents recognize them:

- *Classification* is used if an agent learns to relate specific input patterns to discrete categorical outputs (see Table 2, rows 17-19). In this spirit, an agent identifies classes of affordances which it relates to similar patterns in its environment. Observe that classification readily allows for generalization. Formally, this defines a mapping from continuous to discrete space.
- *Regression* relates to estimating the relationships among variables (see Table 2, rows 8-11). In this sense, an agent learns to predict an outcome given its sensory inputs by learning descriptive regression functions. Formally, this defines a mapping from continuous to continuous space.
- *Inference* is a natural scheme to logically infer new knowledge from existing knowledge (see Table 2, rows 4-7). In such a setting an agent usually employs hard-coded inference rules, e.g., logical formulas, connections within graphs, and the like, to derive new knowledge from existing knowledge using observable data. Formally, this defines a mapping from discrete to discrete space.
- *Optimization* is another popular tool to learn the best expected outcome given some input (see Table 2, rows 25-27). Framed from a mathematical point of view, any learning problem essentially describes an optimization problem, whether or not one wants to minimize a loss or maximize accuracy or a reward. Formally, this defines a mapping from either discrete or continuous to continuous space.

Generalization Being able to generalize affordances enables an agent to use its acquired knowledge to adapt to new situations it has not yet faced. However, simply being able to recognize a known object in a different pose, or a similar object, does not account for generalization on its own. We claim that in the context of affordances, generalization is required to happen at a cognitive level. As a concrete example, consider an agent that has learned that a ball affords rollability. Now, if faced with a sideways lying cylinder, in order to be able to generalize, the agent should also be able to perceive the rollability offered by this very cylinder. Being able to generalize at such a cognitive level ultimately results in faster development, as an agent is not required to learn each and every affordance for each and every environment. Instead, the agent exploits various means of similarity to map existing knowledge to a new environment. This criterion supports two choices, viz. *yes* or *no*.

Exploitation The concept of affordances was originally formulated in a visual perception theory where action possibilities are thought to be 'directly' detected for actuation without further recognition or reasoning. In robotics, the idea of direct perception of affordances has already been studied and fits well to behavior-based or reactive architectures where the tight coupling between perception and action is the main emphasis. Robots, on the other hand, are expected to perform in ways that are much more intelligent and complicated than those of reactive systems. Roboticists have thus exploited the affordances concept for different tasks which require different levels of decision making: the range varies from detecting affordances for reactive action selection to reasoning over affordances for higher level cognition.

- *Action Selection:* Given the percept of the environment, use affordances to select an action to achieve a goal (see Table 2, rows 41-54).
- *Single-/Multi-step Prediction:* Given the percept of the environment, use affordances first to detect action possibilities, and then to predict the changes obtained through execution of the afforded actions (see Table 2, rows 21, 35, and 41). These studies typically chain the predicted changes to make multi-step predictions on the future states of the environment that are expected to be obtained by sequences of actions.
- *Planning:* Multi-step prediction can be achieved through simple search mechanisms such as forward chaining that forms a search tree starting from the current state. Established AI planning methods, on the other hand, run on rules defined over symbols. Some researchers have been able to form such symbols and rules by exploiting affordances and have achieved symbolic planning using affordance-based structures.
- *Language:* Finally, affordances are extended to natural language constructs such as verbs and nouns, enabling communication between different agents (see Table 2, rows 11, 143, and 144).

Learning In order for an agent to proceed in its development and acquire knowledge, it must implement a learning procedure. The learning procedure is crucially dependent on the availability of training data, that is, how and when the data is acquired relative to when the machinery of the learning procedure is triggered. How this process proceeds can heavily influence the construction of the model, as well as the generalizability and developmental aspects of the system. Broadly, this criterion breaks down into two categories: offline vs. online learning.

- *Offline learning* or *batch learning* refers to a type of learning procedure in which the entire training data set is available at a given time as a single batch and in which the learning procedure is applied to only that batch of training data, typically just once without being updated (see Table 2, rows 13-16). Thus, in an evaluation setting, the model might be trained using a batch of training data and subsequently tested using a batch of test data, both of which have been gathered in advance, or, in live robotics setting, the model might be trained in advance of deployment and subsequently used to evaluate individual test cases as they become available to the robot.

- *Online learning* or *continuous learning*, by comparison, refers to a type of learning procedure in which the model may be trained using incremental updates, training-sample-by-training-sample (see Table 2, rows 4-7). The model may be initiated either via offline learning using a batch of training data, where the model is subsequently updated incrementally, or from scratch, perhaps using some form of random initialization, and subsequently updated incrementally using individual training samples. In the former case, when the model is being updated, it should not have access to the original batch of training data. If the model were to require the entire original batch of training data as well as the novel training data in order to make an update, then, although this could be considered to be a type of incremental updating, we consider it to be batch learning under our taxonomy.

Computational Model Criteria

Computational model criteria capture the underlying implementation of a computational affordance model. In this spirit, these criteria deal with the mathematical and theoretical underpinnings of the computational aspects of an affordance model. Observe that the selections for these criteria usually depend on the formulation of the affordance model itself. This is by virtue of the perceptive and developmental aspects of the affordance model which ultimately define the necessary mechanisms for perception and learning.

Formulation The formulation of the computational model specifies its theoretical origins in terms of being neurally or mathematically motivated. These are the two main strains of computational models of affordances that are prevalent in current robotics research. Further, in the case of a neural model, this criterion also captures the neural counterparts of the model for specific brain areas.

Abstraction The *abstraction* of the computational model classifies a model as being either mathematically or neurally motivated. In this sense, the abstraction decides whether a model *approximates* or *emulates* cognitive processes of the human brain. Obviously, a mathematically motivated model can only aim at approximating cognitive processes by definition of its underlying mathematics, e.g., statistical inference. On the other hand, neurally motivated models aim at emulating cognitive processes by rebuilding specific neuronal structures. Though the latter one may also be described using mathematical formalisms, the underlying mechanisms are not built on mathematical concepts but rather on biological processes.

Brain Areas In the case of neural models that aim at imitating cognitive processes, there is a strong relationship to specific areas of the mammal brain that are emulated using software models. This criterion tries to capture this relatedness. It is an open choice criterion in the sense that we do not restrict the possible set of choices. This is motivated by the large number of brain areas that may by involved. In the event of a mathematically formulated model, this criterion may be left empty.

Implementation Implementational aspects of the computational affordance model finally specify various characteristics at the software level. This subsumes (i) the concrete mathematical tools that are employed for learning and prediction, (ii) the environmental features that are used by the model, and (iii) the kind of training that is applied to the model. Clearly, this is a purely technical dimension of our taxonomy, yet it is evidently relevant in order to discuss the capabilities of concrete models and compare them with other models.

Method The method describes the concrete mathematical tools that are used to implement the computational model corresponding to an affordance model. This, again, is an open choice criterion as the number of available tools and their combination is by no means limited.

Features In order to be able to work on meaningful inputs, a computational affordance model generally employs various features to understand the environment. This essentially subsumes any kind of feature that is employed, from plain points to complex 3D surface features. Similar to the case with the selected method, this criterion is also an open choice one. The reason is that the number of possible features and their potential combinations are vast.

Training The penultimate dimension of the implementational aspects of a computational affordance model addresses the kind of training that is used. This depends on the predictive aspects of the developmental dimension of the affordance model (see Development). Clearly, using classification for prediction requires the training of a classifier. This criterion captures the four prevalent types of training paradigms in robotic affordance learning research.

- *Unsupervised* learning generally refers to learning procedures that involve finding the underlying structure in data, e.g. clustering or density estimation, without the provision of ground-truth labels or the specification of a desired outcome beyond certain constraints on the structure itself (see Table 2, rows 2, 12, and 18). With respect to developmental robotics, this corresponds well conceptually to the idea of the autonomous discovery of patterns or concepts from sensory information.
- *Supervised* learning refers to learning procedures that employ training data consisting of two components– input samples and associated target values, e.g. ground-truth labels– in order to learn to predict target values for unknown input samples (see Table 2, rows 1, 3, and 10). It generally functions well as a model for robot instruction, e.g. a tutor might show a robot a series of objects while telling it what the objects afford and the robot must learn to correctly identify the affordances of those objects so that it can accurately predict the affordances of similar objects in future.
- *Self-supervised* learning is a form of learning where information from one data view, e.g. the feature space associated with a particular sensory modality, is used to direct the learning procedure in another data view (see Table 2, rows 5-9). Usually this entails unsupervised learning, e.g. clustering, in the first data view forming concepts that can be used as target values, e.g. category labels, for supervised learning in the second view. This is termed *self*-supervised learning

because the supervision comes from the internal representations of a learning agent as opposed to some external source.

- *Semi-supervised* is a hybrid form of learning that employs both labeled and unlabeled training data (see Table 2, rows 11, 32, and 35). It most naturally resembles human learning which similarly builds on initial training as supervised by a caregiver with subsequent unsupervised learning via autonomous exploration.

Evaluation The last dimension of the implementational aspects of a computational affordance model in our taxonomy captures the means of evaluation that have been used to study the model. Observe that this dimension still has a high relevancy for a computational affordance model, as it allows for the judgment of whether or not a model is useful in practice or is simply a theoretical idea that lacks any grounding in the real world. Further, it allows for a discussion of the maturity of an approach.

- *Simulation* implies that a computational affordance model has only been evaluated in a robotic simulation environment (see Table 2, rows 7-9). Clearly, this renders an approach rather immature as the crude approximation of a simulated world differs drastically from the real physical world and its uncertainty. Evaluating an approach only in simulation does not allow for more than stating that its potential feasibility.
- *Real robot* classifies an approach as mature, since the corresponding computational affordance model has been evaluated in the physical world using a real robot (see Table 2, rows 63-68). This immediately implies that the proposed model can deal with problems of the real world like noise and uncertainty.
- *Benchmark* applies to computational models of affordances for which benchmark-like evaluation schemes were devised by the authors (see Table 2, rows 1, 3, and 10). Generally, these can fall into two categories. On the one side, a baseline is computed from a training data set which is then compared to the outcomes of a test data set. On the other side, the baseline is established from results from similar studies investigating the same research question, and then compared against the authors' own model using the same data as the reference study.
- *Virtual Reality* applies to computational models of affordances for which are evaluated in a virtual environment by a human providing non-simulated interactions in an otherwise simulated environment with a simulated artificial agent.

Finally, as a last note, this criterion is also open for multiple selection. Clearly, there are models that have been evaluated both in simulation and using a real robot.

The taxonomy as outlined in this section ought not represent ruling guidelines on how to implement computational models of affordances in robotics. Rather, it comprises a pivotal attempt to outline, in an organized manner, the manifold aspects that may comprise a computational affordance model. Hence, we claim that it represents a metamodel for computational models of affordances that, while aiming to avoid impeding research by advocating for its adherence, shall motivate the study of what actually comprises a computational affordance model in robotics.

5 Selection and Classification of Publications on Computational Affordance Models

This section describes the systematic search and selection procedure that we followed when selecting relevant papers for classification. Further, we discuss relevant threats to the validity of our study. The resulting classification of computational models of affordances covered in the selected publications is then used in the next section to indicate the adequacy of the defined criteria and for further discussions.

Selection of Publications

The selection of relevant peer-reviewed primary publications comprises the definition of a search strategy and paper selection criteria as well as the selection procedure applied to the collected papers.

Search Strategy The initial search conducted to retrieve relevant papers was performed automatically on the 31st December 2016 by consulting the following digital libraries:

- IEEE Digital Library (`http://ieeexplore.ieee.org/`),
- ScienceDirect (`http://www.sciencedirect.com/`),
- SpringerLink (`http://link.springer.com/`), and
- SAGE (`http://journals.sagepub.com`).

These libraries were chosen as they cover most of the relevant research on robotics. The search string was kept simple, i.e.,

`affordance AND robot`

in order to keep the search very general and to avoid missing any publications featuring more precise terminology. Observe that the search was applied to all of the following search fields: (i) paper title, (ii) abstract, (iii) body, and (iv) keywords. The search produced a set of 1980 retrieved papers, thus a paper selection process was subsequently employed to further filter the results.

Paper Selection Figure 2 gives an overview of the paper selection process, which occurred in three phases. In the first phase, papers were rejected based on their title: if the title did not indicate any relevance to robotics and affordances, papers were discarded from the classification. This reduced the initial set of 1980 papers to 987 remaining papers. In the second phase, papers were rejected based on their abstract, reducing the number of relevant papers to 398. In the third and final phase, papers were rejected based on their content, reducing the set of relevant papers to 146. Thus, our final classification, as discussed in Section 6, consists of a total of 146 papers. Note that during the last iteration, a number of relevant papers were rejected on the basis that they either failed to introduce a novel model or failed to sufficiently reevaluate an existing model.

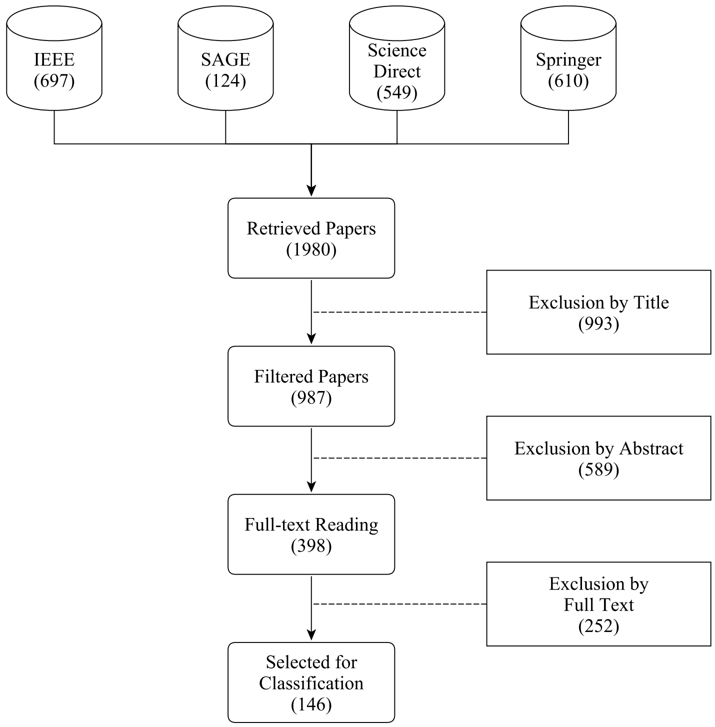


Figure 2. Selection of publications studied in this survey.

Paper Classification

The 146 remaining publications were classified according to the classification criteria as defined and discussed in Section 4 by five researchers. For this purpose, the remaining set of primary publications was randomly split into five sets of equal size for data extraction and classification. A classification spreadsheet was created for this purpose. Besides bibliographic information (title, authors, year, publisher) this sheet contains *classification fields* for each of the defined criteria. To avoid misclassification, the scale and characteristics of each classification criterion were additionally implemented as a selection list for each criterion. As previously mentioned, the list also contained the item 'not specified', to cater for situations where a specific criterion is not defined or could not be inferred from the contents of a paper. Problems encountered during the classification process were remarked upon in an additional comment field. The resulting classification of all publications was then reviewed independently by all five researchers. Finally, in multiple group sessions, all comments were discussed and resolved among all five researchers.

Threats to Validity

Obviously there exist different factors that can influence the results of our study, e.g., the defined search string as discussed previously. Threats to validity include publication bias as well as the identification and classification of publications.

Publication Bias This threat relates to the circumstance that only certain approaches, i.e., those producing promising results or promoted by influential organizations are published (Kitchenham, 2004). We regard this threat as moderate since the sources of publications were not restricted to a certain publisher, journal or conference. Therefore, we claim that our results sufficiently cover existing work in the field of affordances and robotics. However, to balance the trade-off between reviewing as much literature as possible and, at the same time, accumulating reliable and relevant information, gray literature (technical reports, work in progress, unpublished or not peer-reviewed publications) was excluded (Kitchenham, 2004). Further, the required number of pages was set to four to guarantee that papers contained enough information in order to classify them appropriately.

Threats to the Identification of Publications This threat is related to the circumstance that, during the search and selection of publications, relevant papers may have been missed. This is why we employed a rather general search string to avoid missing potentially relevant publications during the automated search. However, in order to additionally reduce the threat of missing important publications, we informally checked papers referenced by the selected papers. We did not become aware of any frequently cited papers that were missed.

Threats to the Classification of Publications Given the high volume of publications that needed to be classified according to a substantial number of defined criteria, the threat of misclassification needed to be addressed. Various measures were taken in order to mitigate this threat. First of all, all criteria were precisely defined, as presented and discussed in Section 4, prior to the commencement of the classification process. There was scope for the refinement of the concepts by the researchers during the process, but this was restricted to mainly descriptive adjustments. Secondly, for each of the criteria we added a list of possible selections in the classification sheet to avoid misclassification. Third, the classification was conducted in parallel by five researchers who are experts in the field and who repeatedly cross-checked the classification independently. Finally, bi-weekly meetings were held by the five researchers to discuss and resolve any comments that arose during independent classification.

Terminology We are aware that the way we use specific terminology, e.g., *learning*, *inference*, or *understanding* may not be perfectly in-line with their use in different areas of research. However, this survey has been written with a robotics research background, which is why we stick to the terminology as used in this field. However, readers from different fields should not face any problems in properly interpreting the content of this work, as the terminology as used in robotics research—to a high degree—has been coined by relevant concepts from psychology and neuroscience.

6 Results and Discussion

This section presents and discusses the classification results of the selected publications (see Section 5). The complete classification of all 146 papers in terms of the proposed taxonomy (see Section 4) is shown in the supplementary material of this article (see Table 2 and is also available online[§].

For each of the selected papers it was possible to classify the covered computational affordance model approach uniquely according to the defined criteria from Section 4. This indicates the adequacy of the defined criteria for the classification of computational models of affordances, thus providing a framework for understanding, categorizing, assessing, and comparing computational models of affordances in robotics. Besides validation of this paper's criteria, the classification enables, as it was performed in a systematic and comprehensive way, an aggregated view and analysis of the state of the art of the use of affordances in robotics.

Figure 4 gives a correlated view on our classification by providing statistics for instances in which each of the criteria coincide with each of the other criteria, thereby providing the foundation for the following discussions. Looking at the leaves of the taxonomy one can see the distribution of the papers in Figure 3.

Perception										Structure					
Perspective		Level		Order		Temporality		Selective Attention		Abstraction		Competetive		Chaining	
agent	138	global	27	0th	15	stable	128	yes	10	micro	135	yes	44	yes	19
observer	8	meso	104	1st	120	variable	4	no	132	micro+macro	7	no	102	no	127
environment	0	local	15	2nd	11	stable+variable	14	not specified	4	macro	4				

Development								Formulation and Implementation				Evaluation	
Acquisition		Prediction		Generalization		Exploitation		Abstraction		Training		real robot	82
ground truth	36	classification	59	yes	106	action selection	78	math	139	supervised	64	simulation	29
exploration	77	regression	22	no	40	single/m.s. pred.	23	neural	7	unsupervised	40	benchmark	25
demonstration	19	inference	26	Learning		planning	32			self-supervised	25	simulation+benchmark	0
hard coded	8	optimization	39	online	53	language	3			semi-supervised	10	simulation+real robot	9
dem.+expl.	6			offline	91	not specified	10			not specified	7	benchmark+real robot	1
				not specified	2							virtual reality	0

Figure 3. Summary table of all criteria for all classified papers. (numbers missing to 146: not specified.)

Maturity of the Model One of the prime criteria for a computational affordance model is its applicability in terms of the chosen evaluation scenario. Obviously, models evaluated on a real robot are much more expressive and powerful than models only evaluated, e.g., in simulation or using some benchmark. Figure 4 shows that the majority of developed models were evaluated on a *real robot* (63% of total papers) thus providing valuable examples in which Gibson's original formulation of affordances, whereby they are directly related to an agent's skills, is followed. Interestingly, of the works that were evaluated using a real robot, a higher proportion were likely to have acquired their affordance knowledge via pure exploration than via all other means

[§] https://iis.uibk.ac.at/public/ComputationalAffordanceModels/

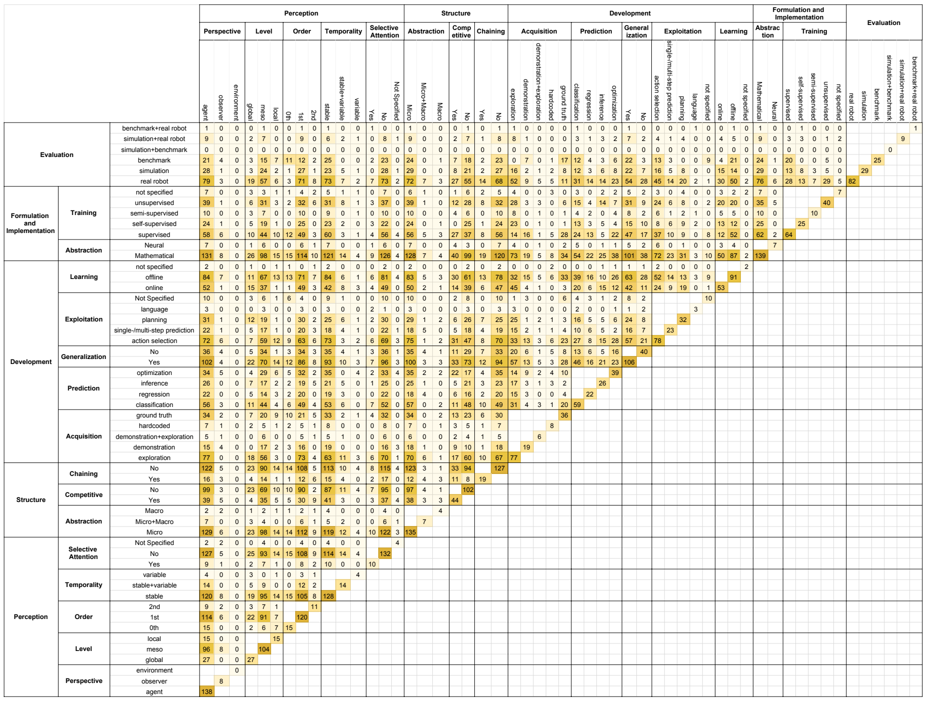


Figure 4. Correlation matrix of all criteria for all classified papers (best viewed on a computer display).

combined:

$$P\,(\textit{exploration}\,|\,\textit{real robot} \text{ or } \textit{simulation} + \textit{real robot} \text{ or } \textit{benchmark} + \textit{real robot})$$
$$\approx 0.66, \tag{1}$$
$$P\left(\textit{exploration}^{\complement}\,\middle|\,\textit{real robot} \text{ or } \textit{simulation} + \textit{real robot} \text{ or } \textit{benchmark} + \textit{real robot}\right)$$
$$\approx 0.34^{\P}. \tag{2}$$

This indicates the importance of following an acquisition approach that allows an agent to implicitly relate affordances to both itself and its skills when working with real robot systems. Though there also exist works with real robots that instead use *hard-coded* information (5 papers) or *ground truth* data (11 papers) as their basis for knowledge, we argue that successful models generally should be built with the embodiment and autonomy of the agent in mind, so that the agent has the potential to gain an enhanced understanding of an affordance. Unfortunately affordance learning from demonstration

¶The derivations for these conditional probability estimates, as well as for other probabilistic estimates herein, may be found in the supplemental material.

(17% of total papers) has not yet received much attention, as also implied by di Pellegrino *et al.* (1992), Rizzolatti et al. (1996) and Thill *et al.* (2013). Last but not least, this general trend in learning by exploration is also expressed in terms of the perspective used for the affordance model that is, the *agent* perspective is prevalent (138 papers) whereas the other two perspectives (*observer*–8 papers, *environment*–0 papers) have seldom been applied or have not been applied at all. We argue that this stems from the as yet unsolved *correspondence problem* in robotics research, which asks how the perception and actions of a demonstrator should be mapped to the corresponding state spaces of a learner.

More mature and expressive models provide better opportunities for exploitation. From Figure 4 we can see a strong trend towards *action selection* (78 total papers), *single-/multi-step prediction* (23 total papers) and *planning* (32 total papers). However, the exploitation of learned models in *language* understanding has received little to no attention at all (3 total papers). This is surprising as high-level cognitive representations that allow symbolic manipulation for planning can, in theory, also be used to support language skills. The deficit in the numbers between planning models and language models is probably due to the fact that, while the structures used for planning can be acquired solely via self-exploration, to achieve communication via language, robots should develop further from social interaction, a learning paradigm which requires much additional implementation effort. Unfortunately, there also exist a few models where the exploitation strategy has not been explicitly mentioned (10 total papers). These models are generally implemented using purely vision-based approaches that do not take an agent or its embodiment into account. It remains an open question for now how well these models can explain relevant affordances and their relationships to agents.

Another important aspect related to the maturity and exploitability of a model is its capacity for generalization. Figure 4 reveals that generally, models that take the *agent*'s perspective into account come with strong generalization capabilities (102 of 138 papers). This connection can be understood by once again thinking about how affordances are defined, that is, as available action opportunities that are related to, and constrained by, an agent's skills. Clearly, learning with respect to one's own perceptual capabilities unlocks stronger generalization capabilities than learning with respect to someone else's, as generally, subjective perception differs. Thus, an observing agent will be able to learn from demonstrations by a teacher if and only if they come equipped with the same capabilities and skills. An interesting discrepancy arises, however, when the generalization capabilities of a model and the level at which affordances are learned are compared. It is evident from Figure 4 that, of the models that learn affordances either at the *local* level or at the *global* level, there are proportionally more of them that exhibit generalization capabilities than there are with models that learn affordances at the *meso* level (93% and 81% vs. 67% respectively). We argue that this is because affordances at the meso level are generally learned in a category specific manner, thereby hampering generalization capabilities by restricting them to those categories. On the other hand, learning affordances at the local level enables agents to form direct relationships between parts and the afforded actions and to generalize the relationships to novel objects with known/similar parts. In the case of learning affordances at the global level, agents must again learn the affordance relationships between items in the environment. Generalization to novel environments is thus possible as long as the novel environment includes similar items.

Learning a Model The maturity of a computational affordance model is directly related to the means by which relevant information is acquired to train the model. An interesting aspect worthy of consideration here is the question of what kinds of affordances are learned. Most existing models either focus solely on *grasping* or take into account a broader range of *manipulation* affordances, e.g., pulling, dragging, pushing, and the like. Also, *traversability* and *locomotion* have received strong attention (*grasping*–29, *manipulation*–39, *traversability* and *locomotion*–18). Being able to traverse between different locations is as important as being able to manipulate objects for robots that are to be smoothly integrated into our daily lives, and so, tackling a wide variety of affordance problems is of great necessity and ought to be encouraged in the research community.

Taking a closer look at the perceptual features used for affordance learning and detection reveals the use of visual features without much reference to physical or material properties of the objects (42 papers that use properties like size or color). A complete understanding of affordances entails forming representations over multiple sensory modalities such that more diverse and robust representations can be formed. The potential importance of such multi-modal sensory feedback, whereby affordances might come to be represented in terms of the physical or material properties of objects, is emphasized by the work of Johansson and Flanagan (2009), who have shown experimentally that people with impaired tactile sensibility, e.g. due to fingertip anesthesia, can experience serious difficulties with simple manipulation tasks because their brains lack the necessary information about mechanical contact states that is needed to plan and control object manipulations.

Despite the fact that many models do *exploration*-based acquisition of affordances, there are a substantial number of models within that category that apply *offline* learning (32 of 77 papers) of affordances instead of *online* learning (45 of 77 papers). Online learning often requires more effort in terms of theory, implementation and experiment compared to offline learning and the extra effort does not necessarily pay off in terms of the publication of results. It is worth noting however that, while the data acquisition aspect of the experimental component is often substantially simplified in offline settings, this does not preclude the possibility of using offline datasets for training and analyzing online models that may subsequently prove useful in developmental robotics scenarios. In addition, while the basic idea of most online learning methods, that of incrementally updating a model sample-by-sample, is relatively well understood, there are still many open questions surrounding related ideas that are broader in scope, such as *life-long learning* or *structural bootstrapping* (Wörgötter et al., 2015), ideas which, when considered in full detail, are far beyond the capabilities of current robotic applications. Researchers would therefore be well advised to aim to either explicitly show the benefits of particular online learning methods, for example active selection of future learning targets, or to try to use the available methods to implement achievable aspects of more long-term projects, such as life-long learning in robots, in practical settings.

A final point worth discussing with regard to model learning is the acquisition of affordances. As can be seen in Figure 4, the most common approach is to use *exploration* (77 total papers). As already argued earlier, this is only natural as it resembles the primary way in which humans acquire their affordance knowledge. Although learning from exploration is more intuitive in this sense, it is unfortunately more time-consuming to set up and difficult implement than other simpler approaches. It does, however, show its strength for affordances whose effected behavior on the manipulated object is hard to predict, e.g., pushing, tapping, and the like, whereas for affordances like (task-independent) grasping or lifting, exploration is not a necessity, as these could instead be predicted more easily from static object features. Looking at other means of acquisition, it can be seen that—despite not yet having received the relevant attention— learning affordances from *demonstration* or *demonstration and exploration* collectively still account for 17% of all approaches. This shows that there is ongoing research that is attempting to tackle what is, given its necessary consideration of multiple agent perspectives, effectively one of the more difficult branches of affordance learning research. In our estimation, it is clear that research in this direction, directly targeting the construction of robots that are capable of learning from other robots and humans, must be intensified in order to boost the capacity of robots to developmentally acquire affordance knowledge more generally. Outside of exploration and demonstration-based models, there also exist a few approaches that rely on labeled *ground truth* or *hard-coded* data. We argue that these are not ideal learning paradigms as they do not take into account the agent's skills, thus forgoing the acquisition of grounded agent-environment relationships, and ultimately making the prospect of agents autonomously expanding their own affordance conceptualization difficult, if not impossible.

Perceiving and Structuring Affordances Clearly perception plays a major role in detecting and learning affordances in the environment. Regarding this, it is interesting that so few existing models take *selective attention* into account (10 total papers), considering that it would provide the means to selectively focus and subsequently analyze a distinct object in the scene by removing noise and irrelevant information. While there has been a history of research into the development of such models in the computer vision community, e.g. (Tsotsos et al., 1995), the ideas and models developed there have perhaps been slow to transfer into the domain of robotics research.

Another important aspect of affordance perception is the *order* of perceived affordances. Most existing models perceive affordances of the *1st*-order (120 total papers), that is, affordances that are immediately related to an agent's skills. Once again, this focus seems motivated by Gibson's characterization of affordances, in which the relationships between an agent's skills and associated action opportunities in the environment are a core underpinning. Additionally, learning 1st-order affordances requires forming relations from an interaction instance, while learning higher-order affordances requires propagating the effects of interactions. Interestingly there are relatively few works which perceive *2nd*-order affordances (11 total papers). This low number directly relates to the fact that, currently, most models also do not support *chaining* of affordances which is, in many models, predicated on the perception of

*2^nd^*-order affordances:

$$P(chaining \mid \mathit{0^{th}}\text{-}order) \approx 0.01, \tag{3}$$
$$P(chaining \mid \mathit{1^{st}}\text{-}order) \approx 0.1, \tag{4}$$
$$P(chaining \mid \mathit{2^{nd}}\text{-}order) \approx 0.55. \tag{5}$$

Focusing on the perception of *0*-order affordances is rare, as noted previously, since they only provide very coarse information on the available affordances in the environment without specifying whether an agent can actually utilize them or not. A further interesting insight can be gained by analyzing the statistics in Figure 4 and relating the *order* of affordances to the way in which they are acquired. The use of 1^{st}-order affordances is more likely when exploration-based acquisition strategies are employed than when other strategies are used:

$$P\left(1^{st}\text{-order} \mid (\textit{exploration} \text{ or } \textit{demonstration + exploration})\right) \approx 0.94, \tag{6}$$
$$P\left(1^{st}\text{-order} \,\middle|\, (\textit{exploration} \text{ or } \textit{demonstration + exploration})^{\complement}\right) \approx 0.67. \tag{7}$$

A final interesting aspect of perceiving affordances to be discussed is how well existing techniques relate to Gibson's idea of *behavior affording behavior*. This, however, immediately requires that a model fulfill various requirements, viz., being able to perceive both 2^{nd}-order and macro affordances, but also being able to perceive not only *stable* but also *variable* affordances that encode possible effects and changes in the environment. Lacking these immediately omits the possibility of reasoning in terms of behavior that affords novel behavior. Again, the numbers from Figure 4 are revealing here. Generally most models only perceive *stable* affordances (128 total papers). If we look at models that perceive *macro* affordances, the number is very low (4 total papers). The figures further show that there are few to no papers that jointly deal with all three of these issues, thus showing a deficit in the field of models that have the potential to enable behavior affording behavior:

$$P\left(macro, variable\right) \approx 0, \tag{8}$$
$$P\left(macro, 2^{nd}\text{-order}\right) \approx 0.007, \tag{9}$$
$$P\left(variable, 2^{nd}\text{-order}\right) \approx 0.007, \tag{10}$$
$$\implies P\left(macro, variable, 2^{nd}\text{-order}\right) \approx 0. \tag{11}$$

Formalizing a Model Regarding the question of how computational models of affordances are generally formalized, there is a clear trend towards *mathematical* formalization (139 total papers) as opposed to *neural* formalization. The crucial difference between these kinds of computational models is that *neurally* inspired models attempt to model distinct brain areas whereas *mathematically* models just apply pure machine learning without any mapping to brain areas. We assume that this again stems from the fact that our understanding of the mammalian brain at this time of writing is still very limited and we do not yet fully understand how concepts are formed and causalities are employed for decision making. And of course inspiration from the brain is important, but generally this is limited to the inspirational level. A further argument for using mathematical formalisms is simply the fact that math allows for aspects of the model to be explained more easily than when modeling biological processes using tools from synthetic biology. However, using neural formalisms would not only allow us to define more accurate models, but also bolster our understanding of the concept of affordances resulting in more powerful models for affordances in robotics. What is also interesting about neural models is that in the early days of affordance research in robotics they received quite some attention. Yet using such models is very difficult as exploratory learning is generally not possible and thus has to be done offline, i.e., after collecting all data and preprocessing it.

Finally, as a last point to discuss here we want to take a closer look at the ways in which models are trained. The evidence shows that most models are trained in either a *supervised* or *unsupervised* manner (64 and 40 total papers, respectively), and that far fewer models have pursued *semi-supervised* or *self-supervised* training strategies (10 and 25 total papers, respectively) by comparison. An interesting relationship emerges upon analyzing the statistics regarding supervised learning. When supervised learning methods are employed, they are far more likely to be used in offline learning settings than in online ones:

$$P\left(\textit{offline} \,|\, \textit{supervised}\right) \approx 0.81, \tag{12}$$

$$P\left(\textit{online} \,|\, \textit{supervised}\right) \approx 0.19. \tag{13}$$

The use of unsupervised learning is also rather easy to explain by looking at the employed mathematical models. Now, looking at the way in which training is conducted in light of the way in which affordance data is acquired, it becomes clear that unsupervised learning has an important role in the context of online learning, being almost twice as likely to be used in such a setting than supervised learning:

$$P\left(\textit{unsupervised} \,|\, \textit{online}\right) \approx 0.38, \tag{14}$$

$$P\left(\textit{supervised} \,|\, \textit{online}\right) \approx 0.23. \tag{15}$$

Online learning of affordances by concurrently extending existing knowledge is often implemented by using unsupervised clustering with *self-organizing maps (SOMs)* as these can be seen as parameter-free models that aim to learn the best representation just from data. From an ecological psychologist's point of view it would be rather obvious to use self-supervised or semi-supervised learning for this purpose, which seem to more plausibly resemble animal learning in the sense of building one's own internal representations or learning from a teacher (observe that this may also explain the lack of demonstration-based approaches for learning affordance models), but these models are more difficult to implement in practice.

7 Open Research Challenges

Our discussion in the previous section showed that there is a strong interest in the robotics research community to apply the concept of affordances. However, it also revealed that current affordance-related research in robotics is still at an early stage, admitting various open research challenges. We believe that addressing these is paramount to advancing affordance related research in robotics.

- *Solving the Correspondence Problem* We claim that it is of utter importance to solve the correspondence problem in robotics, i.e., mapping of observed motions. This would address current drawbacks in both learning from demonstration and in the perception of affordances from an observer's point of view.
- *Improving Language Understanding* Humans use language as one of the main forms of communication to exchange information. Improved language understanding and synthesis capabilities will drastically boost affordance research in robotics for learning affordances from natural action descriptions as well for exploiting learned affordance knowledge more naturally for both robotic teaching as well as for robotic assistance in our daily lives.
- *Behavior Affording Behavior* Most existing work currently only addresses the detection of micro affordances. However, to make robots more autonomous in terms of their planning and reasoning capabilities, we conjecture that future research in affordances and robotics ought to revisit Gibson's idea of behavior affording behavior. This requires taking a closer look at the effects of performing a specific action in order to reason about potential future world states and novel resultant emerging action possibilities.
- *Sensing Physical Properties* Looking only at the shape of an object already provides a lot of information on the probable purpose and use of an object. However, we argue that in order to fully understand an object's affordances it is necessary to also take into account the physical properties of the object (e.g., material, weight, center of gravity) to properly perceive their applicability as tools.
- *The Need for Neurally Inspired Models* Given that animals are biological creatures, understanding the ways in which they perceive and act upon affordances requires more detailed neural models of the workings of their visuomotor systems. From primitive organisms all the way up to higher-order mammalian species, they provide the quintessential examples of working affordance perception systems in the natural world and the affordance research community would do well to understand and selectively imitate their core characteristics. We thus argue that there is an urgent need for novel, well-informed neural models for affordance perception in order to make our robots more powerful and autonomous.
- *Intensifying Semi-supervised and Self-Supervised Learning of Affordances* A learning agent will be able to build a strong cognitive understanding about the meaning of affordances and how to apply them only in so far as it is enabled to learn relevant concepts with minimal external guidance and by learning to represent them with respect to itself with a progressive degree of autonomy. Semi-supervised learning allows for a relatively small amount of supervised labeling of the data, and self-supervised learning endows the agent with the ability to use its autonomously formed concepts to drive its own supervision. We note that, given the relative lack of publications with the relevant focus, there is an intensified need for developing novel computational affordance models that specifically tackles these areas.
- *Selective Attention* Selective attention is an important aspect for focused perception. It is imperative for blocking out clutter and noise. Moreover, it would

not only help solve the problem of learning affordances from demonstration, but would provide agents with the basic building blocks necessary for affordance chaining as well as for planning in cluttered environments with more than one object.

8 Conclusions

In this article we set forth three major contributions. After an introductory part (see Sections 1-3) where we discussed existing interpretations and formalizations of Gibson's theory of affordances, we, as a first major contribution, presented a taxonomy for computational models of affordances in robotics by discussing a battery of evaluation criteria relevant to their study (see Section 4). Following that, our second main contribution was to perform a systematic literature review on affordance research in robotics to subsequently classify selected publications using the taxonomy. Given this classification, our third and final contribution was to give a detailed overview of existing research on affordances in robotics based on analysis of the resultant data, followed by an elicitation of open research questions. We claim that addressing these questions (see Section 7) is paramount to advancing affordance research in robotics.

Supplemental material

Derivation of Probability Estimates

In the following, we detail our empirical data analysis for calculating the equations mentioned in Section 6. Our analysis is based on conditional probability estimates using proportional statistics from the data in Figure 4.

Derivation of Equations 1 and 2:

Letting

$X =$ the event of a paper being classified with the label *exploration*,

$A =$ the event of a paper being classified with the label *real robot*,

$B =$ the event of a paper being classified with the label *simulation* + *real robot*,

$C =$ the event of a paper being classified with the label *benchmark* + *real robot*,

using the laws of probability as well as relative frequency probability estimates, we derive the following:

$$\begin{aligned} P(X|A \cup B \cup C) &= \frac{P(X \cap (A \cup B \cup C))}{P(A \cup B \cup C)} && \text{(conditional probability)} \\ &= \frac{P(A \cup B \cup C \,|\, X)P(X)}{P(A \cup B \cup C)} && \text{(numerator follows from conditional probability)} \\ &= \frac{(P(A \,|\, X) + P(B \,|\, X) + P(C \,|\, X))P(X)}{P(A) + P(B) + P(C)} && \text{(numerator and denominator follow from independence)} \\ &\approx \frac{\left(\frac{52}{77} + \frac{8}{77} + \frac{1}{77}\right)\frac{77}{146}}{\frac{82}{146} + \frac{9}{146} + \frac{1}{146}} \\ &\approx 0.6630 \end{aligned}$$

and

$$\begin{aligned} P\left(X^{\complement} \middle| A \cup B \cup C\right) &= 1 - P\left(X | A \cup B \cup C\right) \\ &\approx 1 - 0.6630 \\ &\approx 0.3370, \end{aligned}$$

where $X^{\complement}$ denotes the complement of X.

Derivation of Equations 3, 4 and 5:

Using the relevant classification labels to denote the events of papers being classified as such, and again using relative frequency probability estimates, we have

$$\begin{aligned} P(\textit{chaining} \,|\, \textit{0}^{\textit{th}}\textit{-order}) &\approx \frac{1}{15} \approx 0.0067 \\ P(\textit{chaining} \,|\, \textit{1}^{\textit{st}}\textit{-order}) &\approx \frac{12}{120} \approx 0.1000 \\ P(\textit{chaining} \,|\, \textit{2}^{\textit{nd}}\textit{-order}) &\approx \frac{6}{11} \approx 0.5455. \end{aligned}$$

Derivation of Equations 6 and 7:

Letting

$$
\begin{aligned}
Z &= \text{the event of a paper being classified with the label } \textit{1st-order}, \\
A &= \text{the event of a paper being classified with the label } \textit{exploration}, \\
B &= \text{the event of a paper being classified with the label } \textit{demonstration}, \\
C &= \text{the event of a paper being classified with the label } \textit{demonstration + exploration} \\
D &= \text{the event of a paper being classified with the label } \textit{hardcoded}, \\
E &= \text{the event of a paper being classified with the label } \textit{ground truth},
\end{aligned}
$$

again, using conditional probability and relative frequency probability estimates, we have

$$
\begin{aligned}
P(Z \mid A \cup C) &= \frac{P(Z \cap (A \cup C))}{P(A \cup C)} \\
&\approx \frac{((73+5)/146)}{((77+6)/146)} \\
&\approx 0.940,
\end{aligned}
$$

and

$$
\begin{aligned}
P\left(Z \,\middle|\, (A \cup C)^{\complement}\right) &= \frac{P(Z \cap (B \cup D \cup E))}{P(B \cup D \cup E)} \\
&\approx \frac{((16+5+21)/146)}{((19+8+36)/146)} \\
&\approx 0.6667.
\end{aligned}
$$

Derivation of Equations 8, 9, 10 and 11:

Again, using the relevant classification labels to denote the events of papers being classified as such, and again using relative frequency probability estimates, we have

$$
\begin{aligned}
P\left(\textit{macro}, \textit{variable}\right) &\approx \frac{0}{146} = 0, \\
P\left(\textit{macro}, 2^{nd}\text{-order}\right) &\approx \frac{1}{146} \approx 0.007, \\
P\left(\textit{variable}, 2^{nd}\text{-order}\right) &\approx \frac{1}{146} \approx 0.007,
\end{aligned}
$$

and

$$
\begin{aligned}
& P\left(\textit{macro}, \textit{variable}, 2^{nd}\text{-order}\right) \leq P\left(\textit{macro}, \textit{variable}\right) \\
\implies & P\left(\textit{macro}, \textit{variable}, 2^{nd}\text{-order}\right) \approx 0.
\end{aligned}
$$

Derivation of Equations 12 and 13:

Using the classification labels and relative frequency probability estimates as before, we have

$$P\left(\textit{offline}\,|\,\textit{supervised}\right) \approx \frac{52}{64} \approx 0.8125,$$
$$P\left(\textit{online}\,|\,\textit{supervised}\right) \approx \frac{12}{64} \approx 0.1875.$$

Derivation of Equations 14 and 15:

Using the classification labels and relative frequency probability estimates as before, we have

$$P\left(\textit{unsupervised}\,|\,\textit{online}\right) \approx \frac{20}{53} \approx 0.3774,$$
$$P\left(\textit{supervised}\,|\,\textit{online}\right) \approx \frac{12}{53} \approx 0.2264.$$

Abbreviations for Classification Criteria

Figures 6 and 5 define the abbreviations for affordance and computational model criteria used for the classification of computational models of affordances shown in Table 2.

Formulation and Implementation	Abstraction	Ab	mathematical	M
			neural	N
	Training	Train	supervised	S
			self-supervised	SELF
			semi-supervised	SEMI
			unsupervised	U
			not specified	NS
Evaluation		Eval	real robot	RR
			simulation	S
			benchmark	B
			simulation+ benchmark	SB
			simulation+ real robot	SRR
			benchmark+ real robot	BRR
			virtual reality	VR

Figure 5. Abbreviations for computational model criteria.

Perception	Perspective	Per	agent	A
			observer	O
			environment	E
	Level	Lvl	global	G
			meso	M
			local	L
	Order	O	0th	0
			1st	1
			2nd	2
	Temporality	Tmp	stable	S
			stable+variable	SV
			variable	V
	Selective Attention	SA	yes	Y
			no	N
			not specified	NS
Structure	Abstraction	Abstr	micro	MI
			micro+macro	MIMA
			macro	MA
	Competitive	Co	yes	Y
			no	N
	Chaining	Ch	yes	Y
			no	N
Development	Acquisition	Acq	exploration	E
			demonstration	D
			ground truth	GT
			hardcoded	H
			demonstration+ exploration	DE
	Prediction	Pr	classification	C
			regression	R
			inference	I
			optimization	O
	Generalization	Ge	yes	Y
			no	N
	Exploitation	Exp	action selection	AS
			single-/multi-step prediction	SP
			planning	P
			language	L
			not specified	NS
	Learning	Lrn	online	ON
			offline	OFF
			not specified	NS

Figure 6. Abbreviations for affordance model criteria.

Classification of Selected Publications on Computational Affordance Models

Table 2 shows the classification of the selected publications sorted by author names and applying the abbreviations from Figures 6 and 5. The list of publications is also available online[‖] as an interactive table supporting filtering and sorting on each column.

Table 2. Classification of selected publications.

No	Paper	Per	Lvl	O	Tmp	SA	Abstr	Co	Ch	Acq	Pr	Ge	Exp	Lrn
1	(Abelha et al., 2016)	A	L	0	S	N	MI	Y	N	GT	O	Y	AS	OFF
2	(Akgun et al., 2009)	A	M	1	S	N	MI	N	N	E	C	Y	AS	OFF
3	(Aldoma et al., 2012)	A	G	1	S	N	MI	N	N	GT	C	Y	NS	OFF
4	(Antunes et al., 2016)	A	G	2	SV	N	MIMA	Y	Y	E	I	Y	P	ON
5	(Baleia et al., 2015)	A	G	1	S	N	MI	N	N	E	I	Y	P	ON
6	(Baleia et al., 2014)	A	G	1	S	N	MI	N	N	E	I	Y	P	ON
7	(Barck-Holst et al., 2009)	A	M	1	S	N	MI	N	N	E	I	Y	P	ON
8	(Bierbaum et al., 2009)	A	L	1	S	N	MI	N	N	E	R	Y	P	OFF
9	(Carvalho & Nolfi, 2016)	A	G	1	S	N	MI	N	N	E	R	Y	AS	ON
10	(Castellini et al., 2011)	A	M	1	S	N	MI	N	N	D	R	Y	AS	OFF
No	Paper	Per	Lvl	O	Tmp	SA	Abstr	Co	Ch	Acq	Pr	Ge	Exp	Lrn

[‖] https://iis.uibk.ac.at/public/ComputationalAffordanceModels/

Table 2. Classification of selected publications.

No	Paper	Per	Lvl	O	Tmp	SA	Abstr	Co	Ch	Acq	Pr	Ge	Exp	Lrn
11	(Çelikkanat et al., 2015)	A	M	1	S	N	MI	N	N	E	O	N	L	OFF
12	(Chan et al., 2014)	A	M	1	S	N	MI	Y	N	D	O	Y	AS	ON
13	(Chang, 2015)	A	M	1	S	N	MI	N	N	GT	C	Y	AS	OFF
14	(Chen et al., 2015)	A	G	1	SV	N	MI	N	N	GT	R	Y	P	OFF
15	(Chu et al., 2016a)	A	M	1	S	N	MI	N	N	DE	C	Y	AS	OFF
16	(Chu et al., 2016b)	O	M	1	S	N	MI	N	N	DE	O	N	AS	OFF
17	(Cos et al., 2010)	A	M	1	S	N	MI	N	N	E	C	Y	P	ON
18	(Cruz et al., 2016)	A	M	1	SV	N	MI	N	N	GT	C	N	AS	OFF
19	(Cruz et al., 2015)	A	M	1	S	N	MI	N	N	GT	C	N	AS	OFF
20	(Cutsuridis & Taylor, 2013)	A	M	1	S	Y	MI	Y	N	E	I	Y	AS	ON
21	(Dag et al., 2010)	O	M	1	S	N	MI	N	N	D	C	N	SP	OFF
22	(Dang & Allen, 2014)	A	L	1	S	N	MI	Y	N	E	O	Y	AS	OFF
23	(Dehban et al., 2016)	A	M	1	S	N	MI	N	N	E	I	Y	AS	ON
24	(Desai & Ramanan, 2013)	A	G	0	S	N	MI	N	N	GT	O	Y	NS	OFF
25	(Detry et al., 2009)	A	M	1	S	N	MI	N	N	E	O	N	P	ON
26	(Detry et al., 2011)	A	M	1	S	N	MI	N	N	E	O	N	P	ON
27	(Detry et al., 2010)	A	M	1	S	N	MI	N	N	E	O	N	AS	ON
28	(Dogar et al., 2007)	A	M	2	S	Y	MI	Y	Y	E	C	Y	AS	OFF
29	(Duchon et al., 1998)	A	G	1	V	N	MI	N	N	E	O	Y	P	ON
30	(Dutta & Zielinska, 2016)	A	M	1	S	N	MI	Y	N	GT	O	Y	AS	OFF
31	(Eizicovits et al., 2012)	A	M	1	S	N	MI	Y	N	D	O	N	P	OFF
32	(Erdemir et al., 2012)	A	G	1	S	N	MI	N	N	E	C	Y	P	OFF
33	(Erdemir et al., 2008)	A	M	1	S	N	MI	N	N	E	R	N	P	OFF
34	(Erkan et al., 2010)	A	M	1	S	N	MI	Y	N	GT	O	Y	AS	ON
35	(Fichtl et al., 2016)	A	M	1	S	N	MI	Y	N	E	C	Y	SP	ON
36	(Fitzpatrick et al., 2003)	A	M	1	S	N	MI	N	N	E	O	N	AS	OFF
37	(Fleischer et al., 2008)	A	M	1	S	N	MI	Y	N	GT	C	Y	AS	OFF
38	(Fritz et al., 2006a)	A	M	1	S	N	MI	N	N	GT	O	Y	AS	OFF
39	(Fritz et al., 2006b)	A	M	1	S	N	MI	N	N	GT	C	N	AS	OFF
40	(Geng et al., 2013)	A	M	2	S	N	MI	Y	N	DE	C	Y	P	OFF
41	(Gijsberts et al., 2010)	A	M	1	S	N	MI	N	N	GT	R	Y	SP	OFF
42	(Glover & Wyeth, 2016)	A	M	1	S	N	MI	N	N	E	I	Y	AS	ON
43	(Gonçalves et al., 2014b)	A	M	1	S	N	MI	N	N	E	I	Y	AS	OFF
44	(Gonçalves et al., 2014a)	A	M	1	S	N	MI	N	N	E	I	Y	AS	OFF
45	(Griffith et al., 2012)	A	M	1	S	N	MI	N	N	E	C	Y	AS	OFF
46	(Hakura et al., 1996)	A	M	1	S	N	MI	N	N	E	C	Y	AS	ON
47	(Hart & Grupen, 2013)	A	M	1	S	N	MI	Y	N	E	O	Y	AS	ON
48	(Hassan & Dharmaratne, 2016)	O	M	1	S	N	MI	N	N	D	C	Y	AS	OFF
49	(Hendrich & Bernardino, 2014)	A	M	1	S	N	MI	N	N	D	R	Y	AS	OFF
50	(Hermans et al., 2013b)	A	M	1	S	N	MI	N	N	E	R	Y	AS	ON
51	(Hermans et al., 2013a)	A	M	1	S	N	MI	Y	N	E	R	Y	AS	OFF
52	(Jiang et al., 2013)	A	M	1	S	N	MI	N	N	D	O	Y	NS	OFF
53	(Kaiser et al., 2014)	A	G	1	S	N	MI	N	N	H	I	Y	AS	OFF
54	(Kaiser et al., 2015)	A	G	2	S	N	MI	N	N	H	I	N	AS	NS
55	(Kamejima, 2002)	A	G	1	S	N	MI	N	N	E	R	Y	P	ON
56	(Kamejima, 2008)	A	G	1	SV	N	MI	N	N	E	C	Y	P	ON
57	(Katz et al., 2014)	A	L	0	S	N	MA	Y	Y	GT	C	Y	AS	OFF
58	(Kim & Sukhatme, 2014)	A	L	0	S	N	MI	N	N	GT	R	Y	NS	OFF
59	(Kim & Sukhatme, 2015)	A	L	2	V	N	MI	N	N	GT	O	Y	SP	OFF
60	(Kim et al., 2006)	A	G	1	S	N	MA	N	N	E	C	Y	P	ON
61	(Kjellström et al., 2011)	A	M	0	S	N	MI	N	N	D	C	N	NS	OFF
62	(Koppula & Saxena, 2014)	O	M	2	S	N	MA	Y	Y	GT	O	Y	AS	OFF
63	(Koppula et al., 2013)	O	M	2	S	Y	MI	Y	Y	GT	C	Y	AS	ON
64	(Kostavelis et al., 2012)	A	G	1	S	N	MI	N	N	GT	C	Y	SP	OFF
65	(Kroemer et al., 2012)	A	L	1	S	N	MI	Y	N	D	R	Y	AS	OFF
66	(Kroemer & Peters, 2011)	O	M	1	S	N	MA	Y	Y	H	O	N	P	OFF
67	(Kubota et al., 2003)	A	G	1	V	N	MI	N	N	E	O	Y	AS	ON
68	(Lee & Suh, 2010)	A	M	1	S	N	MI	Y	Y	GT	C	Y	AS	OFF
69	(Lee & Suh, 2013)	A	M	1	S	N	MIMA	Y	Y	D	O	N	AS	OFF
70	(Lewis et al., 2005)	A	G	1	S	N	MI	Y	N	E	C	Y	P	ON
71	(Lopes et al., 2007)	A	M	1	S	N	MI	Y	N	DE	O	Y	AS	OFF
72	(MacDorman, 2000)	A	M	1	S	N	MI	N	Y	E	C	N	P	ON
73	(Mar et al., 2015a)	A	M	1	SV	N	MI	Y	N	E	R	Y	NS	OFF
74	(Mar et al., 2015b)	A	M	1	SV	N	MI	N	N	E	C	Y	SP	OFF
75	(Maye & Engel, 2013)	A	G	1	SV	N	MI	N	N	E	I	Y	SP	ON
76	(Metta & Fitzpatrick, 2003)	A	M	1	S	N	MI	N	N	E	C	N	AS	ON
77	(Min et al., 2015)	A	M	1	SV	N	MI	N	N	E	I	N	AS	ON
78	(Modayil & Kuipers, 2008)	A	M	1	S	Y	MI	N	N	E	O	Y	P	ON
79	(Mohan et al., 2014)	A	M	1	S	N	MI	N	N	E	C	Y	P	ON
80	(Moldovan et al., 2012)	A	G	2	S	N	MI	Y	Y	GT	I	N	SP	OFF
81	(Moldovan & Raedt, 2014)	A	M	0	S	N	MI	Y	N	H	O	N	AS	OFF
82	(Montesano & Lopes, 2009)	A	L	1	S	N	MI	N	N	E	I	Y	AS	OFF
83	(Montesano et al., 2007b)	A	M	1	S	N	MI	N	N	E	I	N	AS	OFF
84	(Montesano et al., 2008)	A	M	2	S	N	MI	Y	N	E	I	Y	SP	OFF
85	(Montesano et al., 2007a)	A	M	1	S	N	MI	Y	N	D	O	N	AS	OFF
86	(Murphy, 1999)	A	M	1	S	N	MI	N	N	H	C	N	AS	OFF
87	(Mustafa et al., 2016)	A	M	2	S	N	MI	Y	N	GT	C	Y	AS	OFF
88	(Myers et al., 2015)	A	L	0	S	N	MI	N	N	GT	I	Y	AS	OFF
89	(Nishide et al., 2008a)	A	M	1	S	N	MIMA	N	N	E	R	N	SP	OFF
No	Paper	Per	Lvl	O	Tmp	SA	Abstr	Co	Ch	Acq	Pr	Ge	Exp	Lrn

Table 2. Classification of selected publications.

No	Paper	Per	Lvl	O	Tmp	SA	Abstr	Co	Ch	Acq	Pr	Ge	Exp	Lrn
90	(Nishide et al., 2008b)	A	M	1	S	N	MIMA	N	N	E	R	N	SP	OFF
91	(Nishide et al., 2009)	A	G	1	S	N	MIMA	N	Y	E	R	Y	SP	OFF
92	(Nishide et al., 2012)	A	M	1	S	N	MI	N	N	E	R	Y	AS	OFF
93	(Ogata et al., 1997)	A	G	1	S	N	MI	N	N	E	C	Y	P	ON
94	(Oladell & Huber, 2012)	A	M	1	S	N	MI	Y	N	H	O	N	AS	OFF
95	(Omrčen et al., 2009)	A	M	1	S	NS	MIMA	Y	N	E	O	Y	SP	OFF
96	(Paletta & Fritz, 2008)	A	M	1	SV	N	MI	N	N	E	C	Y	P	ON
97	(Paletta et al., 2007)	A	L	1	S	Y	MI	N	N	GT	O	N	AS	OFF
98	(Price et al., 2016)	A	M	1	S	N	MI	N	Y	E	C	Y	P	OFF
99	(Ramirez & Ridel, 2006)	A	G	1	S	N	MI	Y	N	E	C	Y	AS	ON
100	(Richert et al., 2008)	A	M	1	S	N	MI	Y	N	E	C	Y	AS	ON
101	(Ridge & Ude, 2013)	A	M	1	S	N	MI	N	N	E	C	Y	SP	ON
102	(Ridge et al., 2010)	A	M	1	S	N	MI	N	N	E	C	Y	SP	ON
103	(Ridge et al., 2015)	A	M	1	S	N	MI	N	N	E	C	Y	SP	ON
104	(Rome et al., 2008)	A	G	1	S	Y	MI	N	N	GT	C	N	P	OFF
105	(Roy & Todorovic, 2016)	A	G	0	S	N	MI	N	N	GT	C	Y	NS	OFF
106	(Rudolph et al., 2010)	A	M	1	S	N	MI	N	N	D	I	Y	SP	ON
107	(Şahin et al., 2007)	A	M	1	S	N	MI	N	N	E	C	Y	P	ON
108	(Sánchez-Fibla et al., 2011)	A	M	1	S	N	MI	N	N	E	R	Y	AS	ON
109	(Sarathy & Scheutz, 2016)	A	M	1	SV	N	MI	N	N	DE	I	Y	P	ON
110	(Schoeler & Wörgötter, 2016)	A	L	1	S	N	MI	N	N	GT	C	Y	AS	OFF
111	(Shinchi et al., 2007)	A	M	0	S	N	MI	N	N	D	I	Y	NS	ON
112	(Sinapov & Stoytchev, 2007)	A	M	1	SV	N	MI	N	N	E	C	Y	SP	OFF
113	(Sinapov & Stoytchev, 2008)	A	M	1	S	N	MI	N	N	E	C	N	SP	ON
114	(Song et al., 2016)	A	M	1	S	N	MI	N	N	GT	C	Y	AS	OFF
115	(Song et al., 2010)	A	M	1	S	N	MI	Y	N	GT	O	Y	AS	OFF
116	(Song et al., 2011b)	A	M	0	S	N	MI	Y	N	GT	R	N	NS	OFF
117	(Song et al., 2013)	O	M	1	S	NS	MI	Y	N	D	O	N	AS	OFF
118	(Song et al., 2011a)	O	M	1	S	NS	MI	Y	N	D	O	Y	AS	OFF
119	(Song et al., 2015)	A	M	1	S	NS	MI	Y	N	D	O	Y	AS	OFF
120	(Stark et al., 2008)	A	M	0	S	N	MI	N	N	D	C	Y	AS	OFF
121	(Stoytchev, 2008)	A	M	2	SV	N	MI	Y	Y	E	I	N	AS	ON
122	(Stoytchev, 2005)	A	M	1	S	N	MI	N	N	E	C	N	AS	OFF
123	(Stoytchev, 2005)	A	G	1	SV	N	MIMA	N	Y	E	R	N	SP	ON
124	(Stramandinoli et al., 2015)	A	M	1	S	N	MI	N	N	E	O	N	AS	OFF
125	(Sun et al., 2010)	A	M	1	S	N	MI	N	Y	GT	C	N	P	OFF
126	(Sweeney & Grupen, 2007)	A	M	1	S	N	MI	N	N	D	I	Y	AS	OFF
127	(Szedmak et al., 2014)	A	M	1	S	N	MI	N	N	GT	C	Y	NS	OFF
128	(Tagawa et al., 2002)	A	M	1	S	N	MI	N	N	E	I	Y	AS	ON
129	(Pas & Platt, 2016)	A	L	0	S	N	MI	N	N	H	O	Y	AS	NS
130	(Tikhanoff et al., 2013)	A	M	1	S	N	MI	Y	N	E	R	N	SP	ON
131	(Ugur & Şahin, 2010)	A	G	1	S	Y	MI	N	N	E	C	Y	AS	OFF
132	(Ugur et al., 2011)	A	M	1	S	N	MI	N	Y	E	C	Y	P	OFF
133	(Ugur et al., 2015)	A	M	1	S	N	MI	N	Y	DE	C	Y	SP	OFF
134	(Ugur & Piater, 2016)	A	M	1	S	Y	MI	N	N	GT	C	Y	AS	ON
135	(Ugur & Piater, 2015)	A	M	1	SV	N	MI	N	Y	E	C	Y	P	OFF
136	(Varadarajan & Vincze, 2013)	A	L	0	S	N	MI	N	N	GT	C	Y	AS	OFF
137	(Varadarajan & Vincze, 2012)	A	L	0	S	N	MI	N	N	GT	C	Y	AS	OFF
138	(Viña et al., 2013)	A	M	1	S	N	MI	Y	N	E	R	Y	P	OFF
139	(Wang et al., 2013)	A	G	1	V	N	MI	N	N	E	O	N	AS	ON
140	(Windridge et al., 2008)	A	M	1	S	N	MI	Y	Y	E	O	Y	AS	ON
141	(Yi et al., 2012)	A	M	1	S	N	MI	N	N	E	I	Y	AS	ON
142	(Yu et al., 2015)	A	M	0	S	N	MI	Y	N	GT	O	Y	AS	OFF
143	(Yürüten et al., 2013)	A	M	1	S	Y	MI	N	N	E	C	Y	L	OFF
144	(Yürüten et al., 2012)	A	M	1	S	Y	MI	N	N	E	C	N	L	OFF
145	(Zhu et al., 2014)	A	M	1	S	N	MI	N	N	H	I	Y	SP	OFF
146	(Zhu et al., 2015)	A	L	1	S	N	MI	Y	N	D	O	Y	AS	ON
No	Paper	Per	Lvl	O	Tmp	SA	Abstr	Co	Ch	Acq	Pr	Ge	Exp	Lrn

No	Paper	Kind of Affordance	Features	Ab	Train	Eval
1	(Abelha et al., 2016)	tool-use	point clouds, superquadrics	M	S	B
2	(Akgun et al., 2009)	rollability	shape, size	M	U	RR
3	(Aldoma et al., 2012)	general	SEE, SHOT, NDS, SI, PFH	M	S	B
4	(Antunes et al., 2016)	pulling, dragging, grasping	2D geom feat., 2D tracked object displacement	M	NS	RR
5	(Baleia et al., 2015)	traversability	depth, haptic	M	SELF	RR
6	(Baleia et al., 2014)	traversability	depth, haptic	M	SELF	RR
7	(Barck-Holst et al., 2009)	grasping	shape, size, grasp region, force	M	SELF	S
8	(Bierbaum et al., 2009)	grasping	planar faces of object	M	SELF	S
9	(Carvalho & Nolfi, 2016)	traversability	depth, haptic	M	SELF	S
10	(Castellini et al., 2011)	grasping	SIFT BoW, contact joints	M	S	B
11	(Çelikkanat et al., 2015)	pushing, grasping, throwing, shaking	depth, haptic, proprioceptive and audio	M	SEMI	RR
12	(Chan et al., 2014)	grasping	pose, action-object relation	M	U	RR
13	(Chang, 2015)	cutting, painting	edges, TSSC	N	S	RR
14	(Chen et al., 2015)	traversability	RGB images, motor controls	M	S	S
15	(Chu et al., 2016a)	openable, scoopable	forces and torques	M	S	RR
16	(Chu et al., 2016b)	pushing, openning, turning	color, size, pose, force torque, robot arm pose	M	SELF	RR
17	(Cos et al., 2010)	general	illumination	M	S	S
No	Paper	Kind of Affordance	Features	Ab	Train	Eval

No	Paper	Kind of Affordance	Features	Ab	Train	Eval
18	(Cruz et al., 2016)	manipulation, locomotion	agent state, action, object	M	U	S
19	(Cruz et al., 2015)	graspable, dropable, moveable, cleanable	robot state, intended action, object information	M	S	S
20	(Cutsuridis & Taylor, 2013)	grasping	shape	N	U	RR
21	(Dag et al., 2010)	manipulation	3D position, orientation, shape, size	M	U	B
22	(Dang & Allen, 2014)	grasping	grasp, shape context	M	S	RR
23	(Dehban et al., 2016)	pulling, dragging	2D shape, object displacement	M	U	SRR
24	(Desai & Ramanan, 2013)	grasping, support	HOG	M	S	B
25	(Detry et al., 2009)	grasping	ECV	M	SELF	RR
26	(Detry et al., 2011)	grasping	ECV	M	SELF	RR
27	(Detry et al., 2010)	grasping	ECV	M	SELF	RR
28	(Dogar et al., 2007)	traversability	shape, distance	M	U	RR
29	(Duchon et al., 1998)	locomotion, survival	optical flow	M	NS	SRR
30	(Dutta & Zielinska, 2016)	reachable, pourable, movable, drinkable	angular, location + dist. to object, sematic labels	M	S	S
31	(Eizicovits et al., 2012)	grasping	wrist location, roll angle	M	S	RR
32	(Erdemir et al., 2012)	crawling	fixation point, motor values	M	SEMI	RR
33	(Erdemir et al., 2008)	traversabilty	object edges	M	SELF	SRR
34	(Erkan et al., 2010)	grasping	ECV	M	SEMI	RR
35	(Fichtl et al., 2016)	rake, pull/sh, move, lift, take, pour, slide	pose, size; relational hist.feat./PCA on PCL	M	SEMI	S
36	(Fitzpatrick et al., 2003)	general	shape, identity	N	U	RR
37	(Fleischer et al., 2008)	grasping	orientation, object + hand shape, saliency of feat.	N	U	B
38	(Fritz et al., 2006a)	lifting	SIFT, color, mass-center, shape descr., actuator	M	S	S
39	(Fritz et al., 2006b)	lifting	SIFT	M	SELF	S
40	(Geng et al., 2013)	grasping	not specified	N	U	RR
41	(Gijsberts et al., 2010)	grasping	SIFT	M	S	B
42	(Glover & Wyeth, 2016)	grasping	object pose, tactile readings	M	U	RR
43	(Gonçalves et al., 2014b)	pulling, dragging	2D geom. feat., 2D tracked object displacement	M	SELF	S
44	(Gonçalves et al., 2014a)	pulling, dragging	2D geom. feat., 2D tracked object displacement	M	SELF	SRR
45	(Griffith et al., 2012)	drop, move, grasping, /shake	auditory and visual feature trajectories, depth	M	U	RR
46	(Hakura et al., 1996)	traversability	pulse sensor readings	M	SEMI	S
47	(Hart & Grupen, 2013)	grasping	hue, shape, pose of object	M	U	RR
48	(Hassan & Dharmaratne, 2016)	general	SIFT, HOG, textons, color hist., object attributes	M	S	B
49	(Hendrich & Bernardino, 2014)	grasping	shape, size	M	S	RR
50	(Hermans et al., 2013b)	pushing, pulling	pose, depth	M	S	RR
51	(Hermans et al., 2013a)	pushing	Histogram of points (in pose space)	M	SEMI	RR
52	(Jiang et al., 2013)	general	human pose, object pose	M	S	B
53	(Kaiser et al., 2014)	support, lean, grasping, hold	surface characteristics	M	S	RR
54	(Kaiser et al., 2015)	pushing, lifting	surface normals, area	M	NS	RR
55	(Kamejima, 2002)	maneuverability	scene image	M	U	RR
56	(Kamejima, 2008)	maneuverability	scene image	M	U	RR
57	(Katz et al., 2014)	pushing, pulling, grasping	PCA axes, size, center of gravity	M	S	RR
58	(Kim & Sukhatme, 2014)	pushing, lifting, grasping	geometric features	M	S	B
59	(Kim & Sukhatme, 2015)	pushing	geometric features	M	S	S
60	(Kim et al., 2006)	traversability	3D pixel information, texture	M	SELF	RR
61	(Kjellström et al., 2011)	open, pour, hammer	spatial pyramids of HoG	M	S	B
62	(Koppula & Saxena, 2014)	general	human pose, feat. w.r.t. skeleton joints / objects	M	S	B
63	(Koppula et al., 2013)	general	BB, centroid, SIFT	M	S	B
64	(Kostavelis et al., 2012)	traversability	dispartiy maps, hist. of pixel distribution	M	S	RR
65	(Kroemer et al., 2012)	pouring, grasping	pointclouds	M	S	RR
66	(Kroemer & Peters, 2011)	grasping, pushing, striking	pose	M	S	RR
67	(Kubota et al., 2003)	traversability	optical flow	M	U	RR
68	(Lee & Suh, 2010)	general	not specified	M	S	RR
69	(Lee & Suh, 2013)	general	trajectories (joints and end-effectors)	M	S	RR
70	(Lewis et al., 2005)	locomotion	color, texture	M	S	SRR
71	(Lopes et al., 2007)	grasping, tapping, touching	shape, color, scale	M	SEMI	RR
72	(MacDorman, 2000)	navigation	color	M	SELF	S
73	(Mar et al., 2015a)	pulling/dragging	OMS-EGI (3D)	M	U	RR
74	(Mar et al., 2015b)	pulling, dragging	2D geometrical features	M	SELF	SRR
75	(Maye & Engel, 2013)	traversability	not specified	M	U	RR
76	(Metta & Fitzpatrick, 2003)	rollability	color	N	U	RR
77	(Min et al., 2015)	locomotion	shape	M	S	SRR
78	(Modayil & Kuipers, 2008)	manipulability	shape	M	U	RR
79	(Mohan et al., 2014)	reach, grasp, push, search	size, color, shape, world map	M	U	RR
80	(Moldovan et al., 2012)	general	not specified	M	U	RR
81	(Moldovan & Raedt, 2014)	general	geometric properties	M	S	S
82	(Montesano & Lopes, 2009)	grasping	Gaussian, Sobel, Laplacian Filters	M	U	RR
83	(Montesano et al., 2007b)	general	color, shape, size, position; robot gripper pose	M	U	RR
84	(Montesano et al., 2008)	general	convexity, compactness, circleness, squareness	M	U	RR
85	(Montesano et al., 2007a)	grasping, taping	color, shape, size	M	S	RR
86	(Murphy, 1999)	docking, path following, picking	HC perceptual affordance detectors	M	NS	RR
87	(Mustafa et al., 2016)	general	3D texlets	M	U	RR
88	(Myers et al., 2015)	general	Depth, SNorm, PCurv, SI+CV	M	U	B
89	(Nishide et al., 2008a)	pushing	shape, motion	M	S	RR
90	(Nishide et al., 2008b)	pushing	shape, motion	M	S	RR
91	(Nishide et al., 2009)	pulling, dragging	SOM object feature from image	M	S	RR
92	(Nishide et al., 2012)	manipulability	SOM output	M	SEMI	RR
93	(Ogata et al., 1997)	traversability	SOM output	M	SEMI	S
94	(Oladell & Huber, 2012)	lifting, dropping, stacking	location, shape, color	M	S	S
95	(Omrčen et al., 2009)	grasping, pushing	objecet image	M	S	RR
96	(Paletta & Fritz, 2008)	liftability	SIFT	M	S	S
97	(Paletta et al., 2007)	lifting	SIFT, color, shape	M	S	RR
No	**Paper**	**Kind of Affordance**	**Features**	**Ab**	**Train**	**Eval**

No	Paper	Kind of Affordance	Features	Ab	Train	Eval
98	(Price et al., 2016)	grasping	wrenches	M	NS	SRR
99	(Ramirez & Ridel, 2006)	traversability	color	N	S	RR
100	(Richert et al., 2008)	traversability	color, distance, angle	M	S	S
101	(Ridge & Ude, 2013)	pushing	action-grounded 3D shape	M	SELF	RR
102	(Ridge et al., 2010)	pushing	2D,3D shape, 2D motion	M	SELF	RR
103	(Ridge et al., 2015)	pushing	2D,3D shape and motion	M	SELF	RR
104	(Rome et al., 2008)	lifting, trabersability	SIFT	M	S	RR
105	(Roy & Todorovic, 2016)	walkable, sittable, lyable, and reachable	RGB+D, surface normals, semantic labels	M	S	B
106	(Rudolph et al., 2010)	general	object, world, meta (object-object) features	M	S	S
107	(Şahin et al., 2007)	general	not specified	M	NS	RR
108	(Sánchez-Fibla et al., 2011)	pushing	shape, position, orientation	M	U	RR
109	(Sarathy & Scheutz, 2016)	cognitive affordance	visual information	M	U	S
110	(Schoeler & Wörgötter, 2016)	general	SHOT, ESF	M	U	B
111	(Shinchi et al., 2007)	general	color, contour, barycentric pos., num. of objects	M	U	B
112	(Sinapov & Stoytchev, 2007)	pulling, dragging	changes in raw pixels	M	SELF	S
113	(Sinapov & Stoytchev, 2008)	pulling, dragging	raw pixels, trajectories	M	SELF	S
114	(Song et al., 2016)	grasping	BB, category, texture	M	S	RR
115	(Song et al., 2010)	grasping	size, convexity, grasp pose	M	S	S
116	(Song et al., 2011b)	grasping	local features, HOG	M	S	B
117	(Song et al., 2013)	grasping	grasp parameters, dimension	M	S	RR
118	(Song et al., 2011a)	grasping	grasp parameters, dimension	M	S	S
119	(Song et al., 2015)	general	shape, grasp parameters	M	S	SRR
120	(Stark et al., 2008)	grasping	k-adjacent segments, ISM	M	S	RR
121	(Stoytchev, 2008)	grasping	position, color	M	U	RR
122	(Stoytchev, 2005)	extend, slide, contract	position, color	M	SELF	RR
123	(Stoytchev, 2005)	pulling, dragging, pushing, grasping	object postion, tool color	M	U	RR
124	(Stramandinoli et al., 2015)	general	not specified	M	U	RR
125	(Sun et al., 2010)	locomotion	color, edge	M	S	RR
126	(Sweeney & Grupen, 2007)	grasping	moment feature	M	S	RR
127	(Szedmak et al., 2014)	object	shape, size	M	S	B
128	(Tagawa et al., 2002)	general (positive and negative)	object postion	M	U	S
129	(Pas & Platt, 2016)	grasping	curvature, circle fitting	M	NS	RR
130	(Tikhanoff et al., 2013)	pulling, dragging	SIFT, pull angle, tracked dist.	M	S	RR
131	(Ugur & Şahin, 2010)	traversability	shape, size	M	SELF	RR
132	(Ugur et al., 2011)	object	shape, size	M	U	RR
133	(Ugur et al., 2015)	object	shape, size	M	U	RR
134	(Ugur & Piater, 2016)	object	shape, size	M	S	B
135	(Ugur & Piater, 2015)	object	shape, size	M	U	RR
136	(Varadarajan & Vincze, 2013)	general	superquadrics	M	S	B
137	(Varadarajan & Vincze, 2012)	general	gradient image, superquadrics	M	S	B
138	(Viña et al., 2013)	grasping	hand-object relative pose	M	S	RR
139	(Wang et al., 2013)	moveability	color, size	M	SEMI	RR
140	(Windridge et al., 2008)	sorting	image point entropy	M	U	S
141	(Yi et al., 2012)	carryable, stackable, liftable, moveable	color, size	M	U	S
142	(Yu et al., 2015)	containability	voxels	M	S	B
143	(Yürüten et al., 2013)	manipulation	3D shape, size	M	SELF	BRR
144	(Yürüten et al., 2012)	manipulation	3D shape, size	M	SELF	RR
145	(Zhu et al., 2014)	general	pose, human-object pose info	M	S	B
146	(Zhu et al., 2015)	tool-use	material, volume, mass	M	S	B
No	Paper	Kind of Affordance	Features	Ab	Train	Eval

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007–2013 (Specific Programme Cooperation, Theme 3, Information and Communication Technologies) under grant agreement no.;610532, SQUIRREL; from the European Union's Horizon 2020 research and innovation programme under grant agreement no.;731761, IMAGINE; from the Slovenian Research Agency grant J2-7360; and was partially supported by TUBITAK BIDEB under agreement No 117CO16 and by Bogazici Research Fund (BAP) Startup project no 12022.

References

Abelha, P., Guerin, F. & Schoeler, M. (2016). A model-based approach to finding substitute tools in 3d vision data. In *2016 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2471–2478).

Akgun, B., Dag, N., Bilal, T., Atil, I. & Sahin, E. (2009). Unsupervised learning of affordance relations on a humanoid robot. In *2009 24th International Symposium on Computer and Information Sciences* (pp. 254–259).

Aldoma, A., Tombari, F. & Vincze, M. (2012). Supervised learning of hidden and non-hidden 0-order affordances and detection in real scenes. In *2012 IEEE International Conference on Robotics and Automation* (pp. 1732–1739).

Antunes, A., Jamone, L., Saponaro, G., Bernardino, A. & Ventura, R. (2016). From human instructions to robot actions: Formulation of goals, affordances and probabilistic planning. In *2016 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5449–5454).

Baleia, J., Santana, P. & Barata, J. (2014). Self-supervised learning of depth-based navigation affordances from haptic cues. In *2014 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)* (pp. 146–151).

Baleia, J., Santana, P. & Barata, J. (2015). On exploiting haptic cues for self-supervised learning of depth-based robot navigation affordances. *Journal of Intelligent & Robotic Systems*, *80(3)*, 455–474.

Barck-Holst, C., Ralph, M., Holmar, F. & Kragic, D. (2009). Learning grasping affordance using probabilistic and ontological approaches. In *2009 International Conference on Advanced Robotics* (pp. 1–6).

Barwise, J. (1989). *The Situation in Logic*. Center for the Study of Language and Information.

Bierbaum, A., Rambow, M., Asfour, T. & Dillmann, R. (2009). Grasp affordances from multi-fingered tactile exploration using dynamic potential fields. In *2009 9th IEEE-RAS International Conference on Humanoid Robots* (pp. 168–174).

Borghi, A. M. & Riggio, L. (2015). Stable and Variable Affordances Are Both Automatic and Flexible. *Frontiers in Human Neuroscience*, *9*, 351.

Cangelosi, A. & Schlesinger, M. (2015). *Developmental Robotics: From Babies to Robots*. MIT Press.

Carvalho, J. T. & Nolfi, S. (2016). Behavioural plasticity in evolving robots. *Theory in Biosciences*, *135(4)*, 201–216.

Castellini, C., Tommasi, T., Noceti, N., Odone, F. & Caputo, B. (2011). Using object affordances to improve object recognition. *IEEE Transactions on Autonomous Mental Development*, *3(3)*, 207–215.

Chan, W. P., Kakiuchi, Y., Okada, K. & Inaba, M. (2014). Determining proper grasp configurations for handovers through observation of object movement patterns and inter-object interactions during usage. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1355–1360).

Chang, O. (2015). *A Bio-Inspired Robot with Visual Perception of Affordances* (pp. 420–426). Cham: Springer International Publishing.

Chemero, A. (2003). An Outline of a Theory of Affordances. *Ecological Psychology*, *15(2)*, 181–195.

Chemero, A. & Turvey, M. T. (2007). Gibsonian Affordances for Roboticists. *Adaptive Behavior*, *15(4)*, 473–480.

Chen, C., Seff, A., Kornhauser, A. & Xiao, J. (2015). Deepdriving: Learning affordance for direct perception in autonomous driving. In *2015 IEEE International Conference on Computer Vision (ICCV)* (pp. 2722–2730).

Chu, V., Akgun, B. & Thomaz, A. L. (2016a). Learning haptic affordances from demonstration and human-guided exploration. In *2016 IEEE Haptics Symposium (HAPTICS)* (pp. 119–125).

Chu, V., Fitzgerald, T. & Thomaz, A. L. (2016b). Learning object affordances by leveraging the combination of human-guidance and self-exploration. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 221–228).

Cisek, P. (2007). Cortical Mechanisms of Action Selection: the Affordance Competition Hypothesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362(1485)*, 1585–1599.

Cos, I., Cañamero, L. & Hayes, G. M. (2010). Learning affordances of consummatory behaviors: Motivation-driven adaptive perception. *Adaptive Behavior*, *18(3-4)*, 285–314.

Cruz, F., Magg, S., Weber, C. & Wermter, S. (2016). Training agents with interactive reinforcement learning and contextual affordances. *IEEE Transactions on Cognitive and Developmental Systems*, *8(4)*, 271–284.

Cruz, F., Twiefel, J., Magg, S., Weber, C. & Wermter, S. (2015). Interactive reinforcement learning through speech guidance in a domestic scenario. In *2015 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8).

Cutsuridis, V. & Taylor, J. G. (2013). A cognitive control architecture for the perception–action cycle in robots and agents. *Cognitive Computation*, *5(3)*, 383–395.

Dag, N., Atil, I., Kalkan, S. & Sahin, E. (2010). Learning affordances for categorizing objects and their properties. In *2010 20th International Conference on Pattern Recognition* (pp. 3089–3092).

Dang, H. & Allen, P. K. (2014). Semantic grasping: planning task-specific stable robotic grasps. *Autonomous Robots*, *37(3)*, 301–316.

Dehban, A., Jamone, L., Kampff, A. R. & Santos-Victor, J. (2016). Denoising auto-encoders for learning of objects and tools affordances in continuous space. In *2016*

IEEE International Conference on Robotics and Automation (ICRA) (pp. 4866–4871).

Desai, C. & Ramanan, D. (2013). Predicting functional regions on objects. In *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 968–975).

Detry, R., Baseski, E., Popovic, M., Touati, Y., Kruger, N., Kroemer, O., Peters, J. & Piater, J. (2009). Learning object-specific grasp affordance densities. In *2009 IEEE 8th International Conference on Development and Learning* (pp. 1–7).

Detry, R., Kraft, D., Buch, A. G., Krüger, N. & Piater, J. (2010). Refining grasp affordance models by experience. In *2010 IEEE International Conference on Robotics and Automation* (pp. 2287–2293).

Detry, R., Kraft, D., Kroemer, O., Bodenhagen, L., Peters, J., Krüger, N. & Piater, J. (2011). Learning grasp affordance densities. *Paladyn*, *2(1)*, 1.

Dogar, M. R., Cakmak, M., Ugur, E. & Sahin, E. (2007). From primitive behaviors to goal-directed behavior using affordances. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 729–734).

Dotov, D. G., Nie, L. & de Wit, M. M. (2012). Understanding Affordances: History and Contemporary Development of Gibson's Central Concept. *Avant: Trends in Interdisciplinary Studies*, *2(3)*, 28–39.

Duchon, A. P., Kaelbling, L. P. & Warren, W. H. (1998). Ecological Robotics. *Adaptive Behavior*, *6(3-4)*, 473–507.

Dutta, V. & Zielinska, T. (2016). *Predicting the Intention of Human Activities for Real-Time Human-Robot Interaction (HRI)* (pp. 723–734). Cham: Springer International Publishing.

Eizicovits, D., Yaacobovich, M. & Berman, S. (2012). Discrete fuzzy grasp affordance for robotic manipulators. *IFAC Proceedings Volumes*, *45(22)*, 253–258.

Ellis, R. & Tucker, M. (2000). Micro-affordance: The Potentiation of Components of Action by Seen Objects. *British Journal of Psychology*, *91(4)*, 451–471.

Erdemir, E., Frankel, C. B., Thornton, S., Ulutas, B. & Kawamura, K. (2008). A robot rehearses internally and learns an affordance relation. In *2008 7th IEEE International Conference on Development and Learning* (pp. 298–303).

Erdemir, E., Wilkes, D. M., Kawamura, K. & Erdemir, A. (2012). Learning structural affordances through self-exploration. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication* (pp. 865–870).

Erkan, A. N., Kroemer, O., Detry, R., Altun, Y., Piater, J. & Peters, J. (2010). Learning probabilistic discriminative models of grasp affordances under limited supervision. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1586–1591).

Fichtl, S., Kraft, D., Kruger, N. & Guerin, F. (2016). Bootstrapping relational affordances of object pairs using transfer. *IEEE Transactions on Cognitive and Developmental Systems*, *PP(99)*, 1–1.

Fitzpatrick, P., Metta, G., Natale, L., Rao, S. & Sandini, G. (2003). Learning about objects through action - initial steps towards artificial cognition. In *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, Volume 3 (pp. 3140–3145 vol.3).

Fleischer, F., Casile, A. & Giese, M. A. (2008). *Neural Model for the Visual Recognition of Goal-Directed Movements* (pp. 939–948). Berlin, Heidelberg: Springer Berlin Heidelberg.

Fritz, G., Paletta, L., Breithaupt, R., Rome, E. & Dorffner, G. (2006a). Learning predictive features in affordance based robotic perception systems. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3642–3647).

Fritz, G., Paletta, L., Kumar, M., Dorffner, G., Breithaupt, R. & Rome, E. (2006b). *Visual Learning of Affordance Based Cues* (pp. 52–64). Berlin, Heidelberg: Springer Berlin Heidelberg.

Geng, T., Wilson, J., Sheldon, M., Lee, M. & Hülse, M. (2013). Synergy-based affordance learning for robotic grasping. *Robotics and Autonomous Systems*, *61(12)*, 1626–1640.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Houghton Mifflin.

Gibson, J. J. (1977). The Theory of Affordances. *Perceiving, Acting, and Knowing: Toward an Ecological Psychology* (pp. 67–82).

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Psychology Press.

Gijsberts, A., Tommasi, T., Metta, G. & Caputo, B. (2010). Object recognition using visuo-affordance maps. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1572–1578).

Glover, A. & Wyeth, G. (2016). Towards lifelong affordance learning using a distributed markov model. *IEEE Transactions on Cognitive and Developmental Systems*, *PP(99)*, 1–1.

Gonçalves, A., Abrantes, J., Saponaro, G., Jamone, L. & Bernardino, A. (2014a). Learning intermediate object affordances: Towards the development of a tool concept. In *4th International Conference on Development and Learning and on Epigenetic Robotics* (pp. 482–488).

Gonçalves, A., Saponaro, G., Jamone, L. & Bernardino, A. (2014b). Learning visual affordances of objects and tools through autonomous robot exploration. In *2014 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)* (pp. 128–133).

Griffith, S., Sinapov, J., Sukhoy, V. & Stoytchev, A. (2012). A behavior-grounded approach to forming object categories: Separating containers from noncontainers. *IEEE Transactions on Autonomous Mental Development*, *4(1)*, 54–69.

Hakura, J., Yokoi, H. & Kakazu, Y. (1996). *Affordance in Autonomous Robot* (pp. 156–167). Tokyo: Springer Japan.

Hart, S. & Grupen, R. (2013). *Intrinsically Motivated Affordance Discovery and Modeling* (pp. 279–300). Berlin, Heidelberg: Springer Berlin Heidelberg.

Hassan, M. & Dharmaratne, A. (2016). *Attribute Based Affordance Detection from Human-Object Interaction Images* (pp. 220–232). Cham: Springer International Publishing.

Heft, H. (2003). Affordances, Dynamic Experience, and the Challenge of Reification. *Ecological Psychology*, *15(2)*, 149–180.

Hendrich, N. & Bernardino, A. (2014). *Affordance-Based Grasp Planning for Anthropomorphic Hands from Human Demonstration* (pp. 687–701). Cham: Springer International Publishing.

Hermans, T., Li, F., Rehg, J. M. & Bobick, A. F. (2013a). Learning contact locations for pushing and orienting unknown objects. In *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (pp. 435–442).

Hermans, T., Rehg, J. M. & Bobick, A. F. (2013b). Decoupling behavior, perception, and control for autonomous learning of affordances. In *2013 IEEE International Conference on Robotics and Automation* (pp. 4989–4996).

Horton, T. E., Chakraborty, A. & Amant, R. S. (2012). Affordances for Robots: A Brief Survey. *Avant: Journal of Philosophical-Interdisciplinary Vanguard*, *3(2)*, 70–84.

Jamone, L., Ugur, E., Cangelosi, A., Fadiga, L., Bernardino, A., Piater, J. & Santos-Victor, J. (2016). Affordances in Psychology, Neuroscience and Robotics: A Survey. *IEEE Transactions on Cognitive and Developmental Systems*, *PP(99)*, 1–1.

Jiang, Y., Koppula, H. & Saxena, A. (2013). Hallucinated humans as the hidden context for labeling 3d scenes. In *2013 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2993–3000).

Johansson, R. S. & Flanagan, J. R. (2009). Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature Reviews Neuroscience*, *10(5)*, 345–359.

Jones, K. S. (2003). What Is an Affordance? *Ecological Psychology*, *15(2)*, 107–114.

Kaiser, P., Gonzalez-Aguirre, D., Schültje, F., Borràs, J., Vahrenkamp, N. & Asfour, T. (2014). Extracting whole-body affordances from multimodal exploration. In *2014 IEEE-RAS International Conference on Humanoid Robots* (pp. 1036–1043).

Kaiser, P., Grotz, M., Aksoy, E. E., Do, M., Vahrenkamp, N. & Asfour, T. (2015). Validation of whole-body loco-manipulation affordances for pushability and liftability. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)* (pp. 920–927).

Kamejima, K. (2002). Representation and extraction of image feature associated with maneuvering affordance. In *Proceedings of the 41st SICE Annual Conference. SICE 2002.*, Volume 1 (pp. 106–111 vol.1).

Kamejima, K. (2008). Anticipative generation and in-situ adaptation of maneuvering affordance in naturally complex scene. In *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 27–32).

Katz, D., Venkatraman, A., Kazemi, M., Bagnell, J. A. & Stentz, A. (2014). Perceiving, learning, and exploiting object affordances for autonomous pile manipulation. *Autonomous Robots*, *37(4)*, 369–382.

Kim, D., Sun, J., Oh, S. M., Rehg, J. M. & Bobick, A. F. (2006). Traversability classification using unsupervised on-line visual learning for outdoor robot navigation. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* (pp. 518–525).

Kim, D. I. & Sukhatme, G. S. (2014). Semantic labeling of 3d point clouds with object affordance for robot manipulation. In *2014 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5578–5584).

Kim, D. I. & Sukhatme, G. S. (2015). Interactive affordance map building for a robotic task. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 4581–4586).

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, *33(2004)*, 1–26.

Kjellström, H., Romero, J. & Kragić, D. (2011). Visual object-action recognition: Inferring object affordances from human demonstration. *Computer Vision and Image Understanding*, *115(1)*, 81–90.

Koppula, H. S., Gupta, R. & Saxena, A. (2013). Learning human activities and object affordances from RGB-d videos. *The International Journal of Robotics Research*, *32(8)*, 951–970.

Koppula, H. S. & Saxena, A. (2014). *Physically Grounded Spatio-temporal Object Affordances* (pp. 831–847). Cham: Springer International Publishing.

Kostavelis, I., Nalpantidis, L. & Gasteratos, A. (2012). Collision risk assessment for autonomous robots by offline traversability learning. *Robotics and Autonomous Systems*, *60(11)*, 1367–1376.

Kroemer, O. & Peters, J. (2011). A flexible hybrid framework for modeling complex manipulation tasks. In *2011 IEEE International Conference on Robotics and Automation* (pp. 1856–1861).

Kroemer, O., Ugur, E., Oztop, E. & Peters, J. (2012). A kernel-based approach to direct action perception. In *2012 IEEE International Conference on Robotics and Automation* (pp. 2605–2610).

Kubota, N., Niwa, M., Azuma, H. & Ueda, A. (2003). A perceptual system for vision-based evolutionary robotics. In *International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003*, Volume 1 (pp. 841–844 Vol.1).

Lee, S. H. & Suh, I. H. (2010). Goal-oriented dependable action selection using probabilistic affordance. In *2010 IEEE International Conference on Systems, Man and Cybernetics* (pp. 2394–2401).

Lee, S. H. & Suh, I. H. (2013). Skill learning and inference framework for skilligent robot. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 108–115).

Lewis, M. A., Lee, H.-K. & Patla, A. (2005). Foot placement selection using non-geometric visual properties. *The International Journal of Robotics Research*, *24(7)*, 553–561.

Lopes, M., Melo, F. S. & Montesano, L. (2007). Affordance-based imitation learning in robots. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1015–1021).

MacDorman, K. F. (2000). Responding to affordances: learning and projecting a sensorimotor mapping. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, Volume 4 (pp. 3253–3259 vol.4).

Mar, T., Tikhanoff, V., Metta, G. & Natale, L. (2015a). Multi-model approach based on 3d functional features for tool affordance learning in robotics. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)* (pp. 482–489).

Mar, T., Tikhanoff, V., Metta, G. & Natale, L. (2015b). Self-supervised learning of grasp dependent tool affordances on the icub humanoid robot. In *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3200–3206).

Maye, A. & Engel, A. K. (2013). Extending sensorimotor contingency theory: prediction, planning, and action generation. *Adaptive Behavior*, *21(6)*, 423–436.

Metta, G. & Fitzpatrick, P. (2003). Better vision through manipulation. *Adaptive Behavior*, *11(2)*, 109–128.

Michaels, C. F. (2003). Affordances: Four Points of Debate. *Ecological Psychology*, *15(2)*, 135–148.

Min, H., Yi, C., Luo, R., Bi, S., Shen, X. & Yan, Y. (2015). Affordance learning based on subtask's optimal strategy. *International Journal of Advanced Robotic Systems*, *12(8)*, 111.

Min, H., Yi, C., Luo, R., Zhu, J. & Bi, S. (2016). Affordance Research in Developmental Robotics: A Survey. *IEEE Transactions on Cognitive and Developmental Systems*, *8(4)*, 237–255.

Modayil, J. & Kuipers, B. (2008). The initial development of object knowledge by a learning robot. *Robotics and Autonomous Systems*, *56(11)*, 879–890.

Mohan, V., Bhat, A., Sandini, G. & Morasso, P. (2014). From object-action to property-action: Learning causally dominant properties through cumulative explorative interactions. *Biologically Inspired Cognitive Architectures*, *10*, 42–50.

Moldovan, B., Moreno, P., van Otterlo, M., Santos-Victor, J. & Raedt, L. D. (2012). Learning relational affordance models for robots in multi-object manipulation tasks. In *2012 IEEE International Conference on Robotics and Automation* (pp. 4373–4378).

Moldovan, B. & Raedt, L. D. (2014). Occluded object search by relational affordances. In *2014 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 169–174).

Montesano, L. & Lopes, M. (2009). Learning grasping affordances from local visual descriptors. In *2009 IEEE 8th International Conference on Development and Learning* (pp. 1–6).

Montesano, L., Lopes, M., Bernardino, A. & Santos-Victor, J. (2007a). Affordances, development and imitation. In *2007 IEEE 6th International Conference on Development and Learning* (pp. 270–275).

Montesano, L., Lopes, M., Bernardino, A. & Santos-Victor, J. (2007b). Modeling affordances using bayesian networks. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 4102–4107).

Montesano, L., Lopes, M., Bernardino, A. & Santos-Victor, J. (2008). Learning object affordances: From sensory–motor coordination to imitation. *IEEE Transactions on Robotics*, *24(1)*, 15–26.

Murphy, R. R. (1999). Case studies of applying gibson's ecological approach to mobile robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *29(1)*, 105–111.

Mustafa, W., Waechter, M., Szedmak, S. & Agostini, A. (2016). Affordance estimation for vision-based object replacement on a humanoid robot. In *Proceedings of ISR 2016: 47st International Symposium on Robotics* (pp. 1–9).

Myers, A., Teo, C. L., Fermüller, C. & Aloimonos, Y. (2015). Affordance detection of tool parts from geometric features. In *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1374–1381).

Nishide, S., Nakagawa, T., Ogata, T., Tani, J., Takahashi, T. & Okuno, H. G. (2009). Modeling tool-body assimilation using second-order recurrent neural network. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp.

5376–5381).

Nishide, S., Ogata, T., Yokoya, R., Komatani, K., Okuno, H. G. & Tani, J. (2008a). Structural feature extraction based on active sensing experiences. In *International Conference on Informatics Education and Research for Knowledge-Circulating Society (icks 2008)* (pp. 169–172).

Nishide, S., Ogata, T., Yokoya, R., Tani, J., Komatani, K. & Okuno, H. G. (2008b). Active sensing based dynamical object feature extraction. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1–7).

Nishide, S., Tani, J., Takahashi, T., Okuno, H. G. & Ogata, T. (2012). Tool - ;body assimilation of humanoid robot using a neurodynamical system. *IEEE Transactions on Autonomous Mental Development*, *4(2)*, 139–149.

Ogata, T., Hayashi, K., Kitagishi, I. & Sugano, S. (1997). Generation of behavior automaton on neural network. In *Intelligent Robots and Systems, 1997. IROS '97., Proceedings of the 1997 IEEE/RSJ International Conference on*, Volume 2 (pp. 608–613 vol.2).

Oladell, M. & Huber, M. (2012). Symbol generation and feature selection for reinforcement learning agents using affordances and u-trees. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 657–662).

Omrčen, D., Böge, C., Asfour, T., Ude, A. & Dillmann, R. (2009). Autonomous acquisition of pushing actions to support object grasping with a humanoid robot. In *2009 9th IEEE-RAS International Conference on Humanoid Robots* (pp. 277–283).

Paletta, L. & Fritz, G. (2008). *Reinforcement Learning of Predictive Features in Affordance Perception* (pp. 77–90). Berlin, Heidelberg: Springer Berlin Heidelberg.

Paletta, L., Fritz, G., Kintzler, F., Irran, J. & Dorffner, G. (2007). *Perception and Developmental Learning of Affordances in Autonomous Robots* (pp. 235–250). Berlin, Heidelberg: Springer Berlin Heidelberg.

Pas, A. t. & Platt, R. (2016). *Localizing Handle-Like Grasp Affordances in 3D Point Clouds* (pp. 623–638). Cham: Springer International Publishing.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91(1)*, 176–180.

Price, A., Balakirsky, S., Bobick, A. & Christensen, H. (2016). Affordance-feasible planning with manipulator wrench spaces. In *2016 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3979–3986).

Ramirez, A. B. & Ridel, A. W. (2006). Bio-inspired model of robot adaptive learning and mapping. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 4750–4755).

Reed, E. S. (1996). *Encountering the World*. Oxford University Press.

Richert, W., Lüke, O., Nordmeyer, B. & Kleinjohann, B. (2008). Increasing the autonomy of mobile robots by on-line learning simultaneously at different levels of abstraction. In *Fourth International Conference on Autonomic and Autonomous Systems (ICAS'08)* (pp. 154–159).

Ridge, B., Leonardis, A., Ude, A., Deniša, M. & Skočaj, D. (2015). Self-supervised online learning of basic object push affordances. *International Journal of Advanced Robotic Systems*, *12(3)*, 24.

Ridge, B., Skočaj, D. & Leonardis, A. (2010). Self-supervised cross-modal online learning of basic object affordances for developmental robotic systems. In *2010 IEEE International Conference on Robotics and Automation* (pp. 5047–5054).

Ridge, B. & Ude, A. (2013). Action-grounded push affordance bootstrapping of unknown objects. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2791–2798).

Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. (1996). Premotor Cortex and the Recognition of Motor Actions. *Cognitive Brain Research*, *3(2)*, 131–141.

Rome, E., Paletta, L., Şahin, E., Dorffner, G., Hertzberg, J., Breithaupt, R., Fritz, G., Irran, J., Kintzler, F., Lörken, C., May, S. & Uğur, E. (2008). *The MACS Project: An Approach to Affordance-Inspired Robot Control* (pp. 173–210). Berlin, Heidelberg: Springer Berlin Heidelberg.

Roy, A. & Todorovic, S. (2016). *A Multi-scale CNN for Affordance Segmentation in RGB Images* (pp. 186–201). Cham: Springer International Publishing.

Rudolph, M., Muhlig, M., Gienger, M. & Bohme, H. J. (2010). Learning the consequences of actions: Representing effects as feature changes. In *2010 International Conference on Emerging Security Technologies* (pp. 124–129).

Şahin, E., Çakmak, M., Doğar, M. R., Uğur, E. & Üçoluk, G. (2007). To afford or not to afford: A new formalization of affordances toward affordance-based robot control. *Adaptive Behavior*, *15(4)*, 447–472.

Sarathy, V. & Scheutz, M. (2016). A logic-based computational framework for inferring cognitive affordances. *IEEE Transactions on Cognitive and Developmental Systems*, *PP(99)*, 1–1.

Schoeler, M. & Wörgötter, F. (2016). Bootstrapping the semantics of tools: Affordance analysis of real world objects on a per-part basis. *IEEE Transactions on Cognitive and Developmental Systems*, *8(2)*, 84–98.

Shaw, R. (1982). Ecological Psychology: The Consequence of a Commitment to Realism. *Cognition and the Symbolic Processes*.

Shaw, R. (2001). Processes, Acts, and Experiences: Three Stances on the Problem of Intentionality. *Ecological Psychology*, *13(4)*, 275–314.

Shinchi, Y., Sato, Y. & Nagai, T. (2007). Bayesian network model for object concept. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, Volume 2 (pp. II–473–II–476).

Sinapov, J. & Stoytchev, A. (2007). Learning and generalization of behavior-grounded tool affordances. In *2007 IEEE 6th International Conference on Development and Learning* (pp. 19–24).

Sinapov, J. & Stoytchev, A. (2008). Detecting the functional similarities between tools using a hierarchical representation of outcomes. In *2008 7th IEEE International Conference on Development and Learning* (pp. 91–96).

Song, D., Ek, C. H., Huebner, K. & Kragic, D. (2011a). Embodiment-specific representation of robot grasping using graphical models and latent-space discretization. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 980–986).

Song, D., Ek, C. H., Huebner, K. & Kragic, D. (2015). Task-based robot grasp planning using probabilistic inference. *IEEE Transactions on Robotics*, *31(3)*, 546–561.

Song, D., Huebner, K., Kyrki, V. & Kragic, D. (2010). Learning task constraints for robot grasping using graphical models. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1579–1585).

Song, D., Kyriazis, N., Oikonomidis, I., Papazov, C., Argyros, A., Burschka, D. & Kragic, D. (2013). Predicting human intention in visual observations of hand/object interactions. In *2013 IEEE International Conference on Robotics and Automation* (pp. 1608–1615).

Song, H. O., Fritz, M., Goehring, D. & Darrell, T. (2016). Learning to detect visual grasp affordance. *IEEE Transactions on Automation Science and Engineering*, *13(2)*, 798–809.

Song, H. O., Fritz, M., Gu, C. & Darrell, T. (2011b). Visual grasp affordances from appearance-based cues. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)* (pp. 998–1005).

Stark, M., Lies, P., Zillich, M., Wyatt, J. & Schiele, B. (2008). *Functional Object Class Detection Based on Learned Affordance Cues* (pp. 435–444). Berlin, Heidelberg: Springer Berlin Heidelberg.

Steedman, M. (2002). Plans, Affordances, And Combinatory Grammar. *Linguistics and Philosophy*, *25(5)*, 723–753.

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

Stoytchev, A. (2005). Behavior-grounded representation of tool affordances. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation* (pp. 3060–3065).

Stoytchev, A. (2008). *Learning the Affordances of Tools Using a Behavior-Grounded Approach* (pp. 140–158). Berlin, Heidelberg: Springer Berlin Heidelberg.

Stramandinoli, F., Tikhanoff, V., Pattacini, U. & Nori, F. (2015). A bayesian approach towards affordance learning in artificial agents. In *2015 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 298–299).

Sun, J., Moore, J. L., Bobick, A. & Rehg, J. M. (2010). Learning visual object categories for robot affordance prediction. *The International Journal of Robotics Research*, *29(2-3)*, 174–197.

Sweeney, J. D. & Grupen, R. (2007). A model of shared grasp affordances from demonstration. In *2007 7th IEEE-RAS International Conference on Humanoid Robots* (pp. 27–35).

Szedmak, S., Ugur, E. & Piater, J. (2014). Knowledge propagation and relation learning for predicting action effects. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 623–629).

Sánchez-Fibla, M., Duff, A. & Verschure, P. F. M. J. (2011). The acquisition of intentionally indexed and object centered affordance gradients: A biomimetic controller and mobile robotics benchmark. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1115–1121).

Tagawa, K., Konishi, K., Ito, D. & Haneda, H. (2002). Perception driven robotic assembly based on ecological approach. In *Proceedings of 2002 International Symposium on Micromechatronics and Human Science* (pp. 259–265).

Thill, S., Caligiore, D., Borghi, A. M., Ziemke, T. & Baldassarre, G. (2013). Theories and Computational Models of Affordance and Mirror Systems: An Integrative Review. *Neuroscience & Biobehavioral Reviews*, *37(3)*, 491–521.

Tikhanoff, V., Pattacini, U., Natale, L. & Metta, G. (2013). Exploring affordances and tool use on the icub. In *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (pp. 130–137).

Tsotsos, J. K., Culhane, S. M., Kei Wai, W. Y., Lai, Y., Davis, N. & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, *78(1)*, 507–545.

Turvey, M. (1992). Affordances and Prospective Control: An Outline of the Ontology. *Ecological Psychology*, *4(3)*, 173–187.

Ugur, E., Nagai, Y., Sahin, E. & Oztop, E. (2015). Staged development of robot skills: Behavior formation, affordance learning and imitation with motionese. *IEEE Transactions on Autonomous Mental Development*, *7(2)*, 119–139.

Ugur, E., Oztop, E. & Sahin, E. (2011). Goal emulation and planning in perceptual space using learned affordances. *Robotics and Autonomous Systems*, *59(7-8)*, 580–595.

Ugur, E. & Piater, J. (2015). Bottom-up learning of object categories, action effects and logical rules: From continuous manipulative exploration to symbolic planning. In *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2627–2633).

Ugur, E. & Piater, J. (2016). Emergent structuring of interdependent affordance learning tasks using intrinsic motivation and empirical feature selection. *IEEE Transactions on Cognitive and Developmental Systems*, *PP(99)*, 1–1.

Ugur, E. & Şahin, E. (2010). Traversability: A case study for learning and perceiving affordances in robots. *Adaptive Behavior*, *18(3-4)*, 258–284.

Varadarajan, K. M. & Vincze, M. (2012). Afrob: The affordance network ontology for robots. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1343–1350).

Varadarajan, K. M. & Vincze, M. (2013). *AfNet: The Affordance Network* (pp. 512–523). Berlin, Heidelberg: Springer Berlin Heidelberg.

Vera, A. H. & Simon, H. A. (1993). Situated Action: A Symbolic Interpretation. *Cognitive science*, *17(1)*, 7–48.

Viña, F. E., Bekiroglu, Y., Smith, C., Karayiannidis, Y. & Kragic, D. (2013). Predicting slippage and learning manipulation affordances through gaussian process regression. In *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (pp. 462–468).

Wang, C., Hindriks, K. V. & Babuska, R. (2013). Robot learning and use of affordances in goal-directed tasks. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2288–2294).

Warren, W. H. (1988). Action Modes and Laws of Control for the Visual Guidance Of Action. In O. G. Meijer & K. Roth (Eds.), *Complex Movement Behaviour: 'The' Motor-Action Controversy*, Advances in Psychology chapter 14, (pp. 339–379). Amsterdam: Elsevier Science/North-Holland.

Wells, A. J. (2002). Gibson's Affordances and Turing's Theory of Computation. *Ecological Psychology*, *14(3)*, 140–180.

Windridge, D., Shevchenko, M. & Kittler, J. (2008). *An Entropy-Based Approach to the Hierarchical Acquisition of Perception-Action Capabilities* (pp. 79–92). Berlin, Heidelberg: Springer Berlin Heidelberg.

Wörgötter, F., Geib, C., Tamosiunaite, M., Aksoy, E. E., Piater, J., Xiong, H., Ude, A., Nemec, B., Kraft, D., Krüger, N., Wächter, M. & Asfour, T. (2015). Structural Bootstrapping – A Novel, Generative Mechanism for Faster and More Efficient Acquisition of Action-Knowledge. *IEEE Transactions on Autonomous Mental Development*, *7(2)*, 140–154.

Yi, C., Min, H., Luo, R., Zhong, Z. & Shen, X. (2012). A novel formalization for robot cognition based on affordance model. In *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (pp. 677–682).

Yu, L. F., Duncan, N. & Yeung, S. K. (2015). Fill and transfer: A simple physics-based approach for containability reasoning. In *2015 IEEE International Conference on Computer Vision (ICCV)* (pp. 711–719).

Yürüten, O., Sahin, E. & Kalkan, S. (2013). The learning of adjectives and nouns from affordance and appearance features. *Adaptive Behavior*, *21(6)*, 437–451.

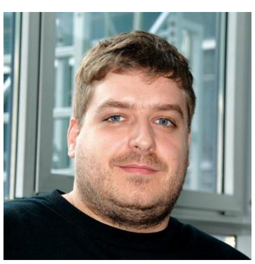
Yürüten, O., Uyanık, K. F., Çalışkan, Y., Bozcuoğlu, A. K., Şahin, E. & Kalkan, S. (2012). *Learning Adjectives and Nouns from Affordances on the iCub Humanoid Robot* (pp. 330–340). Berlin, Heidelberg: Springer Berlin Heidelberg.

Zhu, Y., Fathi, A. & Fei-Fei, L. (2014). *Reasoning about Object Affordances in a Knowledge Base Representation* (pp. 408–424). Cham: Springer International Publishing.

Zhu, Y., Zhao, Y. & Zhu, S. C. (2015). Understanding tools: Task-oriented object modeling, learning and recognition. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2855–2864).

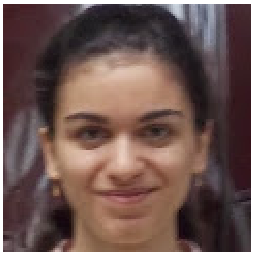
Çelikkanat, H., Orhan, G. & Kalkan, S. (2015). A probabilistic concept web on a humanoid robot. *IEEE Transactions on Autonomous Mental Development*, *7(2)*, 92–106.

Author Biographies



Philipp Zech is a postdoctoral researcher at the University of Innsbruck. He received his Ph.D. degree in Computer Science from the University of Innsbruck in 2014. He is interested in developmental and cognitive robotics, affordance learning and intelligent and adaptive manipulation.

Simon Haller is with the Department of Computer Science, University of Innsbruck, from where he received his B.Sc. in 2006. In 2012 he was granted the professional title Ing. by the Federal Ministry of Economics, Austria. He is a researcher and an engineering technician. His work focuses on problem solving in the field of robotic hard- and software.



Safoura Rezapour Lakani obtained her MSc in 2013 in software systems engineering from the RWTH Aachen University, Germany. She is currently a PhD student in the intelligent and interactive systems group, University of Innsbruck. She is working on visual object representation that can be used in robotic manipulation tasks.

Barry Ridge is a postdoctoral researcher at the Department of Automatics, Biocybernetics and Robotics, Jožef Stefan Institute, Ljubljana, Slovenia. He received a B.Sc. in Computer Applications from the School of Computing, Dublin City University, Ireland in 2002. He went on to study fractal geometry and measure theory at the School of Mathematics and Statistics, University of St. Andrews, Scotland, where he received an M.Phil. in Pure Mathematics in 2006. After being awarded a Marie Curie Fellowship to pursue doctoral studies at the Faculty of Computer and Information Science, University of Ljubljana, Slovenia, with a focus on robotic learning of object affordances, he defended his Ph.D. in 2014. His research interests include cognitive robotics, computer vision and machine learning.

Emre Ugur is an Assistant Professor in Department of Computer Engineering Department, Bogazici University, Turkey. After receiving his PhD in Computer Engineering from Middle East, he worked at ATR Japan as a researcher (2009-2013), at University of Innsbruck as a senior researcher (2013-2016), and at Osaka Univ. as specially appointed Assistant Professor (2015, 2016). He is interested in developmental and cognitive robotics, affordances and intelligent and adaptive manipulation.

Justus Piater is a professor of computer science at the University of Innsbruck, Austria, where he leads the Intelligent and Interactive Systems group. He holds a M.Sc. degree from the University of Magdeburg, Germany, and M.Sc. and Ph.D. degrees from the University of Massachusetts Amherst, USA, all in computer science. Before joining the University of Innsbruck in 2010, he was a visiting researcher at the Max Planck Institute for Biological Cybernetics in Tübingen, Germany, a professor of computer science at the University of Liège, Belgium, and a Marie-Curie research fellow at GRAVIR-IMAG, INRIA Rhône-Alpes, France. His research interests focus on visual perception, learning and inference in sensorimotor systems. He has published more than 160 papers in international journals and conferences, several of which have received best-paper awards, and currently serves as Associate Editor of the IEEE Transactions on Robotics.