Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)
Survey Track

Building Affordance Relations for Robotic Agents - A Review

Paola Ardón, **Èric Pairet**, **Katrin S. Lohan**,
Subramanian Ramamoorthy and **Ronald P. A. Petrick**
Edinburgh Centre for Robotics
Edinburgh, Scotland, United Kingdom
paola.ardon@ed.ac.uk

Abstract

Affordances describe the possibilities for an agent to perform actions with an object. While the significance of the affordance concept has been previously studied from varied perspectives, such as psychology and cognitive science, these approaches are not always sufficient to enable direct transfer, in the sense of implementations, to artificial intelligence (AI)-based systems and robotics. However, many efforts have been made to pragmatically employ the concept of affordances, as it represents great potential for AI agents to effectively bridge perception to action. In this survey, we review and find common ground amongst different strategies that use the concept of affordances within robotic tasks, and build on these methods to provide guidance for including affordances as a mechanism to improve autonomy. To this end, we outline common design choices for building representations of affordance relations, and their implications on the generalisation capabilities of an agent when facing previously unseen scenarios. Finally, we identify and discuss a range of interesting research directions involving affordances that have the potential to improve the capabilities of an AI agent.

1 Introduction

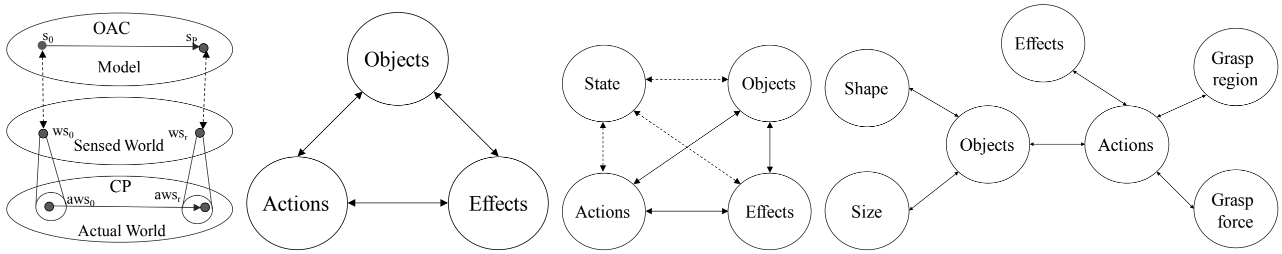
The psychologist James J. Gibson coined the term *affordance* as the ability of an agent to perform a certain action with an object in a given environment [Gibson and Carmichael, 1966]. In general, the concept contributes an encapsulated description of the different ways to comprehend the world [Hammond, 2010]. Nonetheless, [Gibson and Carmichael, 1966]'s understanding of affordances generated controversy among psychologists, resulting in a vast diversity of definitions [Norman, 1988; McGrenere and Ho, 2000].

In AI, affordances play a key role as intermediaries that organise the diversity of possible perceptions into tractable representations that can support reasoning processes to improve the generalisation of tasks. A number of examples of the use of affordances as a form of inductive bias for learning mechanisms can also be found in robotics. In this regard, robotics is a key frontier area for AI that allows for experimental and practical implementations of affordances, with applications to tasks such as action prediction, navigation and manipulation.

The idea of affordances has been studied from different perspectives. Early surveys [Chemero and Turvey, 2007; Şahin *et al.*, 2007; Horton *et al.*, 2012] summarise formalisms that attempt to bridge the controversial concept of affordances in psychology with mathematical representations. Other surveys discuss the connection of robotic affordances with other disciplines [Jamone *et al.*, 2016], and propose classification schemes to review and categorise the related literature [Min *et al.*, 2016; Zech *et al.*, 2017]. In contrast, we focus more on the implications of different design decisions regarding task abstraction and learning techniques that could scale up in physical domains, to address the need for generalisation in the AI sense. As a result, we attempt to capture and discuss the relationship between the requirements, implications and limitations of affordances in an intelligent artificial agent.

The goal of this paper is to provide guidance to researchers wanting to use the concept of affordances to improve generalisation in robotic tasks. As such, we focus on two key questions: *what aspects should be considered when including affordances as a bias for learning and policy synthesis in AI agents?* and *how does the combination of such aspects influence the generalisation capabilities of the system?* After a thorough literature analysis in Section 2, we discuss how, regardless of the underlying abstraction of the concept, using affordances usually refers to the problem of perceiving a *target object*, identifying what *action* is feasible with it and the *effect* of applying these actions on task performance. The relation of these three elements from now on will be referred to as the *affordance relation*. The diversity of techniques and their implications when building an affordance relation forms the backbone of this survey. As such, our contribution is twofold. First, given that we have identified common aspects that define the affordances concept from the point of view of an AI system, we outline the different types of data, processing, learning and evaluation techniques as found in the literature. Moreover, we identify the different levels of *a priori* knowledge on the affordance task. We find that this knowledge directly influences the generalisation capabilities of the system, i.e., the ability to broadly apply some knowledge by inferring from specific cases[1].

[1]The literature related to this survey is available at



(a) [Krüger *et al.*, 2011] (b) [Montesano *et al.*, 2007] (c) [Cruz *et al.*, 2016] (d) [Barck-Holst *et al.*, 2009]

Figure 1: Approaches that propose a mathematical formalism of the affordance relation. a) [Krüger *et al.*, 2011] propose object action complexes (OACs) as the relationship between the sensed and actual world (s_0 is the initial state and s_p is the predicted state). b) [Montesano *et al.*, 2007] represent affordances as a relation between objects, actions, and effects. Objects are entities which the agent is able to interact with, actions represent the behaviour that can be performed with the object, and effects are the results caused by applying an action. c) [Cruz *et al.*, 2016] represent the relations between state, objects, actions and effects, where the state is the current agent's condition and different effects could be produced for different situations. d) [Barck-Holst *et al.*, 2009] present an ontology affordance formalism for grasping, composed of three basic elements alongside object properties and grasping criteria.

2 Synopsis of Affordance Formalisms

This section summarises the evolution of formalisms that use the concept of affordances to improve an agent's performance. This evolution can be divided into two main stages. The first stage is characterised by mathematical conceptualisations as extensions from psychology theories. The second stage corresponds to formalisms that focus on the capabilities of the system rather than on recreations from psychology.

2.1 Psychology-centric Formalisms

In the early stages of the field, affordances were formalised from different perspectives. Namely, this work emphasised *where* the affordance resided following psychology theories. [Chemero and Turvey, 2007; Şahin *et al.*, 2007] extensively review and discuss existing approaches (up to 2007) that translate psychology perspectives into the robotics field. [Şahin *et al.*, 2007] classified the early affordance literature into three different parallel perspectives:

- Agent perspective: The affordance resides inside the agent's possibilities to interact with the environment.
- Environmental perspective: The affordance includes the perceived and hidden affordances in the environment.
- Observer perspective: The affordance relation is observed by a third party to learn these affordances.

The environmental perspective is the most abstract of the perspectives. However, more practically, given the nature of hidden affordances, current methods in applications such as robotics consider either the agent or the observer perspective to build tractable representations of affordance models.

2.2 Agent-centric Formalisms

[Şahin *et al.*, 2007], besides reviewing physchology-inspired work, proposed the first formalism focused on the agent perspective. They argued that an affordance relation between *effect* and an *(entity, behaviour)* tuple should be considered. As such, when the agent applies the behaviour on the *entity*, the *effect* is generated: (*effect*, (*entity, behaviour*)). From the agent's perspective, [Montesano *et al.*, 2007; Krüger *et al.*, 2011] define affordances as using symbolic representations obtained from sensory-motor experiences (see Fig 1(b) and Fig 1(a), respectively). Similarly, [Cruz *et al.*, 2016] consider that if an affordance exists and the agent has knowledge and awareness of it, the agent can choose to utilise it given the current state (see Fig. 1(c)). [Barck-Holst *et al.*, 2009] compare the reasoning engine used for learning in the ontological approach in contrast to the voting probabilistic function to examine the generalisation capabilities of the system (see Fig. 1(d)).

As observed in this section, the perspectives on how affordance should be included in a system vary significantly. Nonetheless, there are common aspects that can be identified across all formalisms that help us ground the basic requirements of affordances for a given task.

2.3 Formalism Influence on Review Criteria

In spite of the differences in approaching the problem, for both psychology and agent-centric formalisms, the purpose remains to achieve high levels of generalisation performance. Interestingly, work across both areas builds the affordance relation using the same three elements: a target object, an action to be applied to that target object and an effect that this action produces. Across formalisms, there are equivalences in both terminology (i.e., action $\leftrightarrow$ behaviour) and type of data (i.e., semantic labels, features). Especially from agent-centric formalisms, it is notable that depending on the nature of the data, the time when this data is processed and learned affects the agent's generalisation capabilities. We identify that, as in the formalisms, the literature on affordances for artificial agents follows the same three elements to build the affordance relation. Moreover, variations in the processing and learning of the data (as detailed in Section 3) constitute different levels of prior affordance relation knowledge, which closely influences the generalisation performance of the agent. Section 4 details this correlation.

Fig. 2 summarises a timeline of important stages in the field, as extracted from reviewing the literature. From the timeline, we can see that including affordances in robotic

https://paolaardon.github.io/affordance_in_robotic_tasks_survey/
It includes 152 papers from 2003 to 2020 in an interactive format.

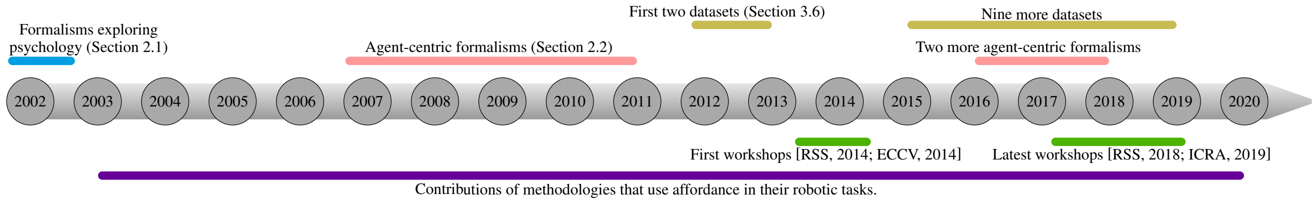


Figure 2: Timeline of significant events in the field of affordances for robotics tasks. Some of the highlights include the popularity of formalisms that focused on translating psychology theories to robotics, and those that proposed formalisms based solely on what the agent can perceive and do with the objects. Other important events are the release of datasets and events aimed at standardising the field. Contributions using affordances in robotics appear from 2003 onwards.

Figure 3: Design choices to build affordance relations as considered in this survey. Starting at the top-left branch and counter-clockwise, we outline the options found in the literature for collecting data in terms of sensory input (blue) and acquisition techniques (green). We also look at the type of actions used to build an affordance relation and perform a task (rose). Once the type of affordance elements have been chosen, we outline the different learning methods employed to relate those elements (orange). Finally, we review the different metrics for evaluation of the affordance task (purple).

tasks is a relatively new field. The agent-centric formalisms on which most of the methodologies base their approaches were not created until after 2007. Moreover, there are still recent proposals trying to ground the view of affordances in robotics. Regarding open source material and workshops, currently there are 11 online available datasets and there have been four publicly available workshops to discuss affordances in robotics [RSS, 2014; ECCV, 2014; RSS, 2018; ICRA, 2019]. Gatherings of researchers with similar interests open doors to discussion, and thus, create opportunities for grounding and advancing progress in the field. Expanded and detailed discussions on related issues are presented in Section 5 and 6.

3 Design Choices

We consider two aspects that influence the generalisation capabilities of an agent when using affordances: the acquisition and processing of the three elements composing an affordance relation model (i.e., target object, actions and effects), which is considered in this section; and the level of *a priori* knowledge of the affordance relation, with respect to the task start time, which is detailed in Section 4.

3.1 Sensory Input

We start by considering the data acquisition *sensory input*, which refers to the medium used to recognise all the physical and visual qualities that suggest a set of actions in the scene. For example, a ball contains the visual and physical features that suggest the affordance to `roll`. The features can be perceived with a different set of sensors. An extensive summary of perception interaction is presented in [Bohg *et al.*, 2017]. Common practices in robotic affordance tasks include using *visual* input [Zhu *et al.*, 2014; Saxena *et al.*, 2014], *tactile* sensors [Baleia *et al.*, 2015], *kinaesthetic* [Katz *et al.*, 2014] and *proprioceptive* sensory feedback [Bonaiuto and Arbib, 2015; Kaiser *et al.*, 2015]. The choice of sensory input is greatly influenced by the purpose of the target task, for example, using visual sensors for recognition tasks [Szedmak *et al.*, 2014; Myers *et al.*, 2015] or a combination of sensory inputs to understand the environment [Ardón *et al.*, 2020; Ardón *et al.*, 2021; Baleia *et al.*, 2015; Liu *et al.*, 2019; Veres *et al.*, 2020].

3.2 Data Collection

Together with the technique for collecting data, one usually also decides on the underlying data structure. For example, consider the task of making an object roll where there are two objects, a ball and an apple, both of which afford `roll`. The agent could use class labels that detect `ball` and `apple` or could identify features that make an object roll, in which case the agent could potentially generalise this behaviour to other objects. From the reviewed literature, the most common practice is to annotate pixel *labels*, through supervised or self-supervised methods, on RGB and RGB-D visual input that represent affordance features [Do *et al.*, 2018; Dehban *et al.*, 2016; Saxena *et al.*, 2014; Stoytchev, 2008]. Other work uses a combination of visual input with *demonstrations* from a tutor and *exploration* techniques [Gonçalves *et al.*, 2014; Antunes *et al.*, 2016; Wang *et al.*, 2013; Cruz *et al.*, 2016; Saxena *et al.*, 2014] to collect the affordance relation.

3.3 Deployed Actions

Tasks an agent might perform could range from pure affordance recognition on a target object to the deployment of an action to achieve a manipulation or navigation task. For example, an agent might recognise that a ball affords `roll` and then apply an action like `push` so that the ball rolls. A more complex action might be the compound task of `handing over` the ball, which would require the agent to `reach` and `grasp` the ball and then `approach` another agent. [Asada *et al.*, 2009] summarises work that imitates the cognitive development of an infant, dividing it into 12 stages according to the difficulty of the motions. Inspired by [Asada *et al.*, 2009] in this survey, we group actions in two sets: *primitive* actions, defined by simple motions, such as turning, moving forward, grasping or pushing [Abelha *et al.*, 2016; Baleia *et al.*, 2015; Seker *et al.*, 2019], and *compound* actions defined by combining multiple simple motions, such as pouring or handing over an object [Price *et al.*, 2016; Sun *et al.*, 2014; Zhu *et al.*, 2014].

3.4 Learning Affordance Relations

In addition to acquiring data to associate the affordance elements, it is important to consider the learning model that encapsulates the affordance relation. We identify four affordance relation learning strategies. In the rolling object example, one might have collected a sample showing that a ball affords `roll`. Given that the relation was learned with one specific rolling example, the agent knows one particular way of rolling. This *deterministic* approach results in a model without any randomness [Szedmak *et al.*, 2014]. If the collected data could instead contain overlapping examples, including 'randomness', then a *probabilistic* learning approach could be applied [Zhu *et al.*, 2014; Saxena *et al.*, 2014; Song *et al.*, 2015; Mar *et al.*, 2015; Nguyen *et al.*, 2017]. Another option is for the agent to have a prior set of logical rules that determines a round object in motion on a surface rolls, so it builds the affordance relation model based on *heuristics* [Baleia *et al.*, 2015]. A slightly more difficult scenario would be to make the ball roll and when it reaches a stationary state to make it roll back. Here the relation could be built as a task *planning* problem [Aksoy *et al.*, 2015; Cutsuridis and Taylor, 2013].

3.5 Metrics and Evaluation

Across the literature, different approaches might evaluate the same task differently. For instance, in the rolling ball example, an approach might say a task succeeds if the ball rolls, and fails if it does not move. In this case, the method determines if the task completed successfully based on *qualitative* metrics [Ardón *et al.*, 2019; Bonaiuto and Arbib, 2015; Cai *et al.*, 2019]. Another option might be to measure the success of rolling a ball by matching trajectory accuracy or measuring displacement on a surface. Work that uses a numeric-based metric *quantitatively* evaluates the task. Particularly for work using *quantitative* metrics, the evaluation is closely correlated to the application and the purpose of the task. To date, there exists little direct comparison across different applications (i.e., recognition, manipulation, navigation). However, popular metrics in the field include confusion matrices [Zhu *et al.*, 2014; Ye *et al.*, 2017; Song *et al.*, 2015; Aksoy *et al.*, 2015], mean square error (MSE) [Katz *et al.*, 2014], and accuracy of classification metrics that reflect intrinsic assessments [Myers *et al.*, 2015; Gonçalves *et al.*, 2014; Aldoma *et al.*, 2012; Cruz *et al.*, 2016]. Further discussion on the need for standard setups is presented in Sections 5.

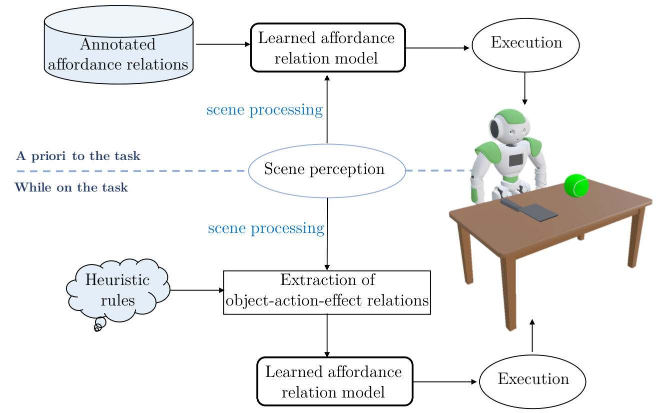


Figure 4: Flowcharts representing two contrasting approaches for building an affordance relation for the task of rolling an object in the scene. In both approaches, the first step is scene perception and processing (e.g., target object identification, feature extraction). In the first approach, the agent completely learns the affordance relation model prior to the task: after the data processing stage, the agent identifies a suitable affordance relation to roll the object. In the second approach, the agent has a set of rules allowing it to relate the environment perception with interactions and build an affordance relation while performing the task.

3.6 Datasets

Unlike other research fields which have many datasets available online, such as grasping that has over 30 online datasets as summarised in [Huang *et al.*, 2016], available datasets for affordance tasks are considerably fewer. The online interactive version of this survey in footnote 1 shows a summary of the available online datasets that collect data structures to build an affordance relation for robotic tasks. In particular, in this summary we identify the task that the dataset is intended for, summarise the dataset's content and its data type, as well as provide the online location of the dataset. Given the potential to improve the agent's understanding of the task, the affordance concept has been particularly popular for object recognition, manipulation and navigation in robotic applications [Ardón *et al.*, 2019; Shu *et al.*, 2016; Koppula *et al.*, 2016]. Thus, it is not surprising that the existing online datasets focus on one of these robotic tasks.

4 Deployment

In addition to the choice of data input, learning and evaluation technique reviewed in Section 3, the time when data is processed and learned, with respect to the beginning of the task, affects the generalisation capabilities of the robotic agent. For example, a robot can relate pushing a ball with making it roll because it knows the `push` $\rightarrow$ `roll` relation before the start of the task with some probability; or a series of heuristics indicating that the features of the ball that make it round might result in the ball rolling if pushed. In the latter case, the relation `push` $\rightarrow$ `roll` is built with some certainty until the agent starts the affordance task, allowing the robot to adapt to previously unseen scenarios. We consider that the combination of design choices (as detailed in Section 3) with the timing when the data is processed and learned determines the prior knowledge of the affordance relation. The different levels of this prior knowledge influence the system's performance in unknown environments and, as such, allow the agent to operate with different degrees of autonomy. We continue with the analogy of the robot rolling an object. In this section, we identify two general approaches to build an affordance relation in robotics and summarise them in Fig. 4.

4.1 Building Affordance Relations Prior to the Task

The first general approach that we identify corresponds to methods that have full *a priori task knowledge* about the possible affordance relations. Work that uses this approach usually requires less complexity in their design choices [Sun *et al.*, 2014; Moldovan and De Raedt, 2014; Byravan and Fox, 2017; Jiang *et al.*, 2013; Koppula *et al.*, 2016; Pieropan *et al.*, 2014]. Namely, the information encapsulated in the affordance relation models of these methods represents their generalisation capabilities. In the literature, we identify the following three main generalisation capabilities that consider full prior knowledge of the affordance relation.

Using Affordance Relations to Understand Surroundings

Consider the case when the robot learns a policy that indicates round objects roll. The agent is able to generalise the relation of every object in the scene that is round, by understanding object properties rather than mapping them to individual object classes (e.g., ball, apple) [Kim and Sukhatme, 2014; Aksoy *et al.*, 2015; Chu *et al.*, 2019b; Dutta and Zielinska, 2019; Fang *et al.*, 2018; Luddecke and Worgotter, 2017]. These approaches organise their features by their 'functionality'. For example, [Stark *et al.*, 2008] generalises that features representing handles offer the possibility of grasping, as do the ones representing the surface of a bottle. For this work, learning the affordance relation of the features with actions results in superior generalisation performance for object categorisation. [Cruz *et al.*, 2016] complete a cleaning task where the simulated robot uses reinforcement learning and a predefined set of contextual affordances with few starting actions. Having this prior information enables the system to reach higher rates of success, which is the case for [Cruz *et al.*, 2016; Wang *et al.*, 2013].

[Mar *et al.*, 2015; Abelha *et al.*, 2016] enable more precise affordance predictions for tool use scenarios. In [Mar *et al.*, 2015], instead of learning a single model that tries to relate all the possible variables in an affordance, the robot learns a separate affordance model for each set of tools and corresponding grasping poses sharing common functionality, thus categorising tool handles and poses. Along the same lines, [Castellini *et al.*, 2011] propose the use of grasping motor data (i.e., kinematic grasping data obtained from human demonstrations) to encode the affordances of an object, and then to use this representation on similar objects to improve object recognition.

Affordance Relation Models that Consider Perturbations

Work in this sub-category learns a single model for affordance relations and keeps it fixed for the rest of the robotics task. These approaches try to achieve the task with what they know about the affordance relation in the presence of perturbations or changes in the environment. For instance, say that a robot has a prior on the `push` $\rightarrow$ `roll` relationship, however, when attempting to perform the task it is not able to reach the target object. As a safety policy measure, the robot knows how to find a spatula that can help it reach the target object. In this case, there are no new relations learned [Pandey and Alami, 2013; Fallon *et al.*, 2015; Diana *et al.*, 2013; Marion *et al.*, 2017]. As a result, the actions become rules that are queried at execution time.

Some approaches in this category attempt to perform multi-step predictions based on known affordance relations, such as [Price *et al.*, 2016; Cutsuridis and Taylor, 2013; Veres *et al.*, 2020]. These methods often add a planning layer that allows them to achieve goal-oriented tasks by adapting to changes in the environment. Other work such as [Kostavelis *et al.*, 2012] and [Moldovan and De Raedt, 2014] perform navigation and grasping application tasks, respectively, in cluttered environments. They do so in a scenario with many objects where the purpose is to identify the most suitable object for a pre-defined task. In [Kostavelis *et al.*, 2012; Dogar *et al.*, 2007; Saputra *et al.*, 2019], the goal is to arrive at a destination while choosing to push or nudge objects on the way, while in [Moldovan and De Raedt, 2014], the goal is to find an object that might be occluded on a shelf to achieve a queried action. [Wang *et al.*, 2013; Pandey and Alami, 2013] use templates of interpretative triplets, containing the affordance relation components.

Affordance Relations for Multiple Objects and Agents

In an environment where there is a ball and a spatula, the robot associates the ball as the object that rolls, and the spatula as the object to grasp and use to push the ball so that it rolls. In this case, the agent has a prior affordance relation model that enables it to associate multiple object affordances to achieve one task [Koppula *et al.*, 2016; Pieropan *et al.*, 2014; Jiang *et al.*, 2013; Chan *et al.*, 2020; Shu *et al.*, 2017; Thermos *et al.*, 2017]. Work that considers *multi-object affordance relations* associates multiple objects in the scene [Ruiz and Mayol-Cuevas, 2018; Chu *et al.*, 2019a; Kaiser *et al.*, 2016; Sun *et al.*, 2014; Moldovan and De Raedt, 2014; Jiang *et al.*, 2013; Moldovan *et al.*, 2018]. They model the concept of object co-occurrence by calculating the probability of an object on a shelf being of a particular type and having a specific affordance, given that on the same shelf there are objects of a certain type. Others consider *multi-agent affordance relations* in the same environment [Price *et al.*, 2016; Song *et al.*, 2015]. For example, both approaches design a system where they consider the action capabilities of manipulating the objects among different agents and across places. [Song *et al.*, 2015] developed a framework by stages composed of [Song *et al.*, 2013; Song *et al.*, 2015] that makes sure a robotic end-effector properly hands over an object.

Building Affordance Relations While on the Task

The second general approach we found in the literature is to build the affordance relation *while on the task*. This approach requires less prior knowledge and design choices that adapt to partially new environments [Lopes *et al.*, 2007; Ugur *et al.*, 2015; Gonçalves *et al.*, 2014; Tikhanoff *et al.*, 2013; Dehban *et al.*, 2016; Baleia *et al.*, 2015]. Given that these approaches create affordance relations online, they are able to generalise to new environments. We identify two ways in which such methodologies generalise to novel scenarios.

Updating Models with New Affordance Relations

Following the example of the robot rolling an object on the scene, a possibility is that the robot knows the object rolls, from previous experience or from a tutor's demonstration. In this setup, the agent learns that poking the object from different directions affords pushing or pulling the object. As a result, the agent can extend the affordance relation model to include the new action possibilities [Gonçalves *et al.*, 2014; Tikhanoff *et al.*, 2013; Dehban *et al.*, 2016; Baleia *et al.*, 2015].

Methods in this sub-category learn and update a model using demonstrations from a tutor or trial and error techniques. Using *demonstrations*, especially for the robotics task of grasping, we find methods that exploit the benefits of learning by demonstration (LbD) to build the affordance relation model [Chan *et al.*, 2014; Ridge and Ude, 2013]. For example, [Ridge and Ude, 2013] proposes a self-supervised method that encapsulates features of the objects before and after being pushed. The features then serve as a base for the robot to know where to push new objects and create more affordance relations with other objects based on their ability to be pushed.

For methods using *trial and error*, they combine pre-learned affordance relation models with exploration to assess the effects of an action [Antunes *et al.*, 2016; Bozcuoğlu *et al.*, 2019; Seker *et al.*, 2019].

From Primitive Actions to Compound Behaviours

Another option available to the agent for making the object roll is to learn the combination of basic motions, such as reaching and pushing, that lead to the desired rolling outcome while performing the task, rather than having a prior affordance relation [Hermans *et al.*, 2013; Bonaiuto and Arbib, 2015]. The approaches in this section propose a framework that allows the robot to explore and learn an affordance relation model using primitive actions as the backbone. A set of heuristic rules are then put in place to guide the robot to compose actions and associate them with a target object and effects [Kaiser *et al.*, 2016; Ugur *et al.*, 2015]. These frameworks learn high-level behaviours, however, questions such as *how does a robot learn to pull an object towards itself?* or *how does the robot learn that spherical objects roll while a cube only slides when pushed?* concern learning of primitive actions at a control level. Some approaches learn the parameters to basic controller primitive actions to generalise to new robotic tasks by combining visual and tactile information and testing the heuristic model in a trial and error stage [Hermans *et al.*, 2013; Stoytchev, 2008; Bonaiuto and Arbib, 2015].

5 Limitations of Affordances in Robotics

As a summary, in Section 3 we outlined the different design choices for data input, learning methods and evaluation of the affordance relation. In Section 4 we described how the interaction of the design choices and the time when the affordance relation is built, with respect to the start of the task (i.e., prior knowledge of the affordance relation), influences the generalisation capabilities of the system. In particular, this correlation defines how well the methods perform on unseen environments. The diversity of design choices and prior knowledge of the affordance relation provides the field with great potential to adapt to many robotic applications. Nonetheless, this variety also makes it difficult to define standards in the field. The objective of this section is to identify and summarise the strengths (see Section 5.1) and weaknesses (see Section 5.2) as found in the reviewed literature.

5.1 Strengths

Generalisation to Novel Setups

The idea of identifying target objects by their functionality rather than by their categorical classes provides a system with the ability to perform the same task on similar objects. For example, an agent can learn that a bottle and a mug share features that afford pouring. Many approaches in the literature exploit the concept of generalisation to improve their object recognition and categorisation tasks [Aksoy *et al.*, 2015; Kroemer *et al.*, 2012; Mar *et al.*, 2015; Abelha *et al.*, 2016].

Potential for Autonomous Learning

As explained in Section 2, different formalisms define affordances as a co-defining relation between a target object, an action to be applied on the object and an effect that evaluates such actions. As such, the concept can be conceived as a closed loop that offers the potential for the system to learn by itself with little human intervention. In the reviewed literature, we find work that has some prior on the affordance relation and that employs it to learn new relations while on the task [Gonçalves *et al.*, 2014; Lopes *et al.*, 2007; Ugur *et al.*, 2015].

5.2 Weaknesses

Datasets

Using the concept of affordances to bias the learning of an artificial intelligent agent is relatively new. As detailed in Section 2, all the strategies account for target objects, actions and effects to create an affordance relation model. Nonetheless, it is difficult to find common ground for a dataset that contains these three elements and satisfies the needs of the different tasks and generalisation requirements. Unlike other self-contained research fields, such as grasping, manipulation and object recognition, affordances in robotics have few available online datasets. Most of the existing datasets are limited to (i) a single affordance associated with an object, and (ii) assume that this single affordance is true regardless of the context of the object. The usual approach to perform affordance tasks is to collect a motion that represents an action. A natural step towards fast-forwarding data collection would be the design and implementation of a data collection interface. This interface would facilitate the annotation of objects with the corresponding actions to perform a robotics task. Certainly, there needs to be a consensus in the field regarding the requirements of such centralised datasets for specific robotic applications. This type of agreement would help to facilitate benchmarking different types of affordance relation models.

Metrics

It is fundamental for the progress of the affordance concept as a learning bias to standardise metrics that reflect the performance of an affordance-aware agent in different tasks. A summary of the diversity of metrics used in the field is detailed in Section 5.2. Given that most of the contributions lay in collaborative tasks with other agents and improving generalisation performance, an interesting approach would be to measure the similarity of the actions taken by the system with those a human would execute. For example, it could be interesting to measure the differences in the trajectories executed to achieve a task. Such differences could be measured in terms of distances in the point distributions of the trajectories, or entropy of the trajectories as a whole. Options such as the Hausdorff distance and the Kullback-Leibler divergence are interesting to explore. The Hausdorff distance measures how similar or close two sets of points are, and the Kullback-Leibler divergence measures how one probability distribution is different from a second one. Including such evaluations could be a good assessment of performance on a robotic task in relation to ground truth data, regardless of learning.

6 Conclusions and Future Directions

In this survey, we explored the literature for approaches that included affordances in the execution of AI and robotic tasks, and identified common ground for building affordance relations. In contrast to previous reviews of affordances, we provide guidance on design decisions and how the concept can be used to guide policy learning to boost the agent's performance. First, we summarised affordance formalisms in Section 2, where we found that affordance relations are built using three common elements: target object, action and effect. Then, we outlined the design choices in Section 3 and the possible generalisation schemes in Section 4. In Section 5, we discussed several problems in the field. Given the relatively new usage of affordances to boost the agent's generalisation capabilities, there are interesting opportunities for future improvements. Next, we outline possible areas where research contributions can be made, based on our survey of the reviewed literature.

6.1 Design Choices and their influence

As previously mentioned, the choice of data input and learning time influence the performance of the methods, often in different ways. Nonetheless, some design choices have been explored more than others, thus leaving room for research into the generalisation capabilities that can be achieved. Fig. 5 shows a coverage map of the reviewed literature spread over the different design choices. The warmer the colour, the more the strategy is used across the literature. For example,

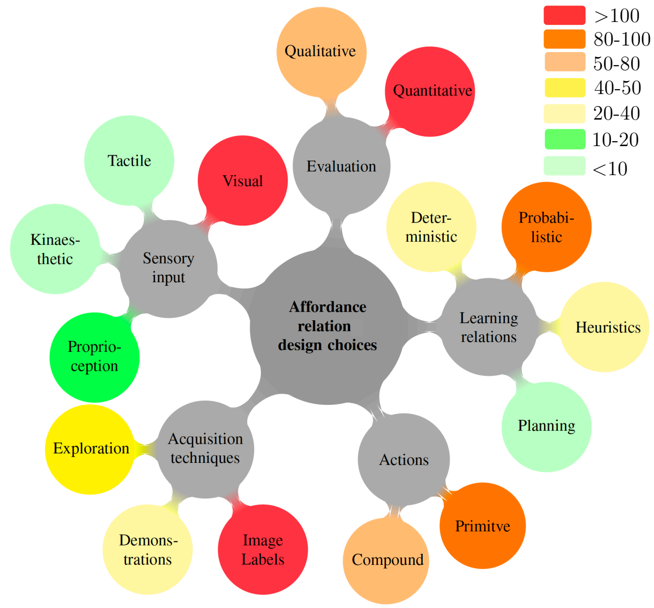


Figure 5: Population map of the design choices for work in the reviewed literature that includes affordances in robotic tasks. The warmer the colour (red) the more the work uses that element to design their affordance relation model, while the colder the colour (green), the less that element is used.

most of the work emphasises the learning of primitive actions as affordances (i.e., push, grasp, lift, among others), using visual perception and image labels to identify an affordance per target object, building the affordance relation probabilistically. On the opposite side, colder coloured elements indicate that there are very few approaches that exploit learning affordance trajectories in the form of motions (using kinaesthetic sensing), as well as those that exploit a multi-step prediction to achieve the tasks in a planning manner. Certainly, some of these components are highly dependent on hardware robustness more than others. Nonetheless, studying such aspects in greater depth would improve their inclusion in robotics tasks as well as provide valuable insights for collaboration activities and task replication across different agents.

6.2 Data Acquisition and Processing Correlation

Given that the field requires different types of data (i.e., target object, action and effects), there is an inherent need to cross-correlate data structures. These associations can be further enriched by diversifying data acquisition and processing. For example, knowing that a cup in a kitchen is likely to afford pouring liquid while in a bathroom might also serve as a toothbrush holder requires the agent to be able to relate not only object features but also features related to an agent's surroundings. Moreover, the data acquisition task should not be limited to visual object features alone but should also account for the object's material, texture and other physical properties to enhance the agent's interaction with the object. By using such cross-correlation, affordance relation models could provide the agents with the ability to generalise affordances as associations of objects' visual and physical characteristics, as well as with the surrounding context.

6.3 Autonomous Behaviour Learning

At present, the idea of including the concept of affordances in AI tasks has centred on performing one task at the time. For instance, many approaches detect one object affordance, such as a glass affords pouring, but rarely proceed with actually performing the pouring task. The field would benefit from a methodology that is able to unify associations of target objects with a library of actions and an online evaluation of the effects. This association would allow an agent to obtain feedback on the performance of the task and rectify it online. Such a method would open doors for autonomously concatenating goal-oriented tasks (e.g., preparing a recipe, cleaning dishes, etc.), and for exploring the feasibility of subsequent motion controllers and the interpretability of natural language instructions in large-scale tasks.

References

[Abelha *et al.*, 2016] Paulo Abelha, Frank Guerin, and Markus Schoeler. A model-based approach to finding substitute tools in 3D vision data. In *IEEE International Conference on Robotics and Automation*, pages 2471–2478. IEEE, 2016.

[Aksoy *et al.*, 2015] Eren Erdal Aksoy, Minija Tamosiunaite, and Florentin Wörgötter. Model-free incremental learning of the semantics of manipulation actions. *Robotics and Autonomous Systems*, 71:118–133, 2015.

[Aldoma *et al.*, 2012] Aitor Aldoma, Federico Tombari, and Markus Vincze. Supervised learning of hidden and non-hidden 0-order affordances and detection in real scenes. In *IEEE International Conference on Robotics and Automation*, pages 1732–1739. IEEE, 2012.

[Antunes *et al.*, 2016] Alexandre Antunes, Lorenzo Jamone, Giovanni Saponaro, Alexandre Bernardino, and Rodrigo Ventura. From human instructions to robot actions: Formulation of goals, affordances and probabilistic planning. In *IEEE International Conference on Robotics and Automation*, pages 5449–5454. IEEE, 2016.

[Ardón *et al.*, 2019] Paola Ardón, Èric Pairet, Ronald PA Petrick, Subramanian Ramamoorthy, and Katrin S Lohan. Learning grasp affordance reasoning through semantic relations. *IEEE Robotics and Automation Letters*, 4(4):4571–4578, 2019.

[Ardón *et al.*, 2020] Paola Ardón, Èric Pairet, Yvan Petillot, Ronald PA Petrick, Subramanian Ramamoorthy, and Katrin S Lohan. Self-assessment of grasp affordance transfer. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2020.

[Ardón *et al.*, 2021] Paola Ardón, Maria E Cabrera, Eric Pairet, Ronald PA Petrick, Subramanian Ramamoorthy, Katrin S Lohan, and Maya Cakmak. Affordance-aware handovers with human arm mobility constraints. *IEEE Robotics and Automation Letters*, 6(2):3136–3143, 2021.

[Asada *et al.*, 2009] Minoru Asada, Koh Hosoda, Yasuo Kuniyoshi, Hiroshi Ishiguro, Toshio Inui, Yuichiro Yoshikawa, Masaki Ogino, and Chisato Yoshida. Cognitive developmental robotics: A survey. *IEEE Transactions on Autonomous Mental Development*, 1(1):12–34, 2009.

[Baleia *et al.*, 2015] José Baleia, Pedro Santana, and José Barata. On exploiting haptic cues for self-supervised learning of depth-based robot navigation affordances. *Journal of Intelligent & Robotic Systems*, 80(3-4):455–474, 2015.

[Barck-Holst *et al.*, 2009] Carl Barck-Holst, Maria Ralph, Fredrik Holmar, and Danica Kragic. Learning grasping affordance using probabilistic and ontological approaches. In *International Conference on Advanced Robotics*, pages 1–6. IEEE, 2009.

[Bohg *et al.*, 2017] Jeannette Bohg, Karol Hausman, Bharath Sankaran, Oliver Brock, Danica Kragic, Stefan Schaal, and Gaurav S Sukhatme. Interactive perception: Leveraging action in perception and perception in action. *IEEE Transactions on Robotics*, 33(6):1273–1291, 2017.

[Bonaiuto and Arbib, 2015] James Bonaiuto and Michael A Arbib. Learning to grasp and extract affordances: the integrated learning of grasps and affordances (ILGA) model. *Biological cybernetics*, 109(6):639–669, 2015.

[Bozcuoğlu *et al.*, 2019] Asil Kaan Bozcuoğlu, Yuki Furuta, Kei Okada, Michael Beetz, and Masayuki Inaba. Continuous modeling of affordances in a symbolic knowledge base. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5452–5458. IEEE, 2019.

[Byravan and Fox, 2017] Arunkumar Byravan and Dieter Fox. Se3-nets: Learning rigid body motion using deep neural networks. In *IEEE International Conference on Robotics and Automation*, pages 173–180. IEEE, 2017.

[Cai *et al.*, 2019] Junhao Cai, Hui Cheng, Zhanpeng Zhang, and Jingcheng Su. Metagrasp: Data efficient grasping by affordance interpreter network. In *International Conference on Robotics and Automation*, pages 4960–4966. IEEE, 2019.

[Castellini *et al.*, 2011] Claudio Castellini, Tatiana Tommasi, Nicoletta Noceti, Francesca Odone, and Barbara Caputo. Using object affordances to improve object recognition. *IEEE Transactions on Autonomous Mental Development*, 3(3):207–215, 2011.

[Chan *et al.*, 2014] Wesley P Chan, Yohei Kakiuchi, Kei Okada, and Masayuki Inaba. Determining proper grasp configurations for handovers through observation of object movement patterns and inter object interactions during usage. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1355–1360. IEEE, 2014.

[Chan *et al.*, 2020] Wesley P Chan, Matthew KXJ Pan, Elizabeth A Croft, and Masayuki Inaba. An affordance and distance minimization based method for computing object orientations for robot human handovers. *International Journal of Social Robotics*, 12(1):143–162, 2020.

[Chemero and Turvey, 2007] Anthony Chemero and Michael T Turvey. Gibsonian affordances for roboticists. *Adaptive Behavior*, 15(4):473–480, 2007.

[Chu *et al.*, 2019a] Fu-Jen Chu, Ruinian Xu, and Patricio A Vela. Learning affordance segmentation for real-world robotic manipulation via synthetic images. *IEEE Robotics and Automation Letters*, 4(2):1140–1147, 2019.

[Chu *et al.*, 2019b] Vivian Chu, Reymundo A Gutierrez, Sonia Chernova, and Andrea L Thomaz. Real-time multisensory affordance-based control for adaptive object manipulation. In *International Conference on Robotics and Automation*, pages 7783–7790. IEEE, 2019.

[Cruz *et al.*, 2016] Francisco Cruz, Sven Magg, Cornelius Weber, and Stefan Wermter. Training agents with interactive reinforcement learning and contextual affordances. *IEEE Transactions on Cognitive and Developmental Systems*, 8(4):271–284, 2016.

[Cutsuridis and Taylor, 2013] Vassilis Cutsuridis and John G Taylor. A cognitive control architecture for the perception–action cycle in robots and agents. *Cognitive Computation*, 5(3):383–395, 2013.

[Dehban *et al.*, 2016] Atabak Dehban, Lorenzo Jamone, Adam R Kampff, and José Santos-Victor. Denoising auto-encoders for learning of objects and tools affordances in continuous space. In *IEEE International Conference on Robotics and Automation*, pages 4866–4871. IEEE, 2016.

[Diana *et al.*, 2013] Matteo Diana, Jean-Pierre de la Croix, and Magnus Egerstedt. Deformable-medium affordances for interacting with multi-robot systems. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5252–5257. IEEE, 2013.

[Do *et al.*, 2018] Thanh-Toan Do, Anh Nguyen, and Ian Reid. Affordancenet: An end-to-end deep learning approach for object affordance detection. In *IEEE international conference on robotics and automation*, pages 5882–5889. IEEE, 2018.

[Dogar *et al.*, 2007] Mehmet R Dogar, Maya Cakmak, Emre Ugur, and Erol Sahin. From primitive behaviors to goal-directed behavior using affordances. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 729–734. IEEE, 2007.

[Dutta and Zielinska, 2019] Vibekananda Dutta and Teresa Zielinska. Predicting human actions taking into account object affordances. *Journal of Intelligent & Robotic Systems*, 93(3-4):745–761, 2019.

[ECCV, 2014] ECCV. Second workshop on affordances: Visual perception of affordances and functional visual primitives for scene analysis, 2014.

[Fallon *et al.*, 2015] Maurice Fallon, Scott Kuindersma, Sisir Karumanchi, Matthew Antone, Toby Schneider, Hongkai Dai, Claudia Pérez D'Arpino, Robin Deits, Matt DiCicco, Dehann Fourie, et al. An architecture for online affordance-based perception and whole-body planning. *Journal of Field Robotics*, 32(2):229–254, 2015.

[Fang *et al.*, 2018] Kuan Fang, Te-Lin Wu, Daniel Yang, Silvio Savarese, and Joseph J Lim. Demo2Vec: Reasoning object affordances from online videos. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2139–2147, 2018.

[Gibson and Carmichael, 1966] James Jerome Gibson and Leonard Carmichael. *The senses considered as perceptual systems*, volume 2. Houghton Mifflin Boston, 1966.

[Gonçalves *et al.*, 2014] Afonso Gonçalves, João Abrantes, Giovanni Saponaro, Lorenzo Jamone, and Alexandre Bernardino. Learning intermediate object affordances: Towards the development of a tool concept. In *International Conference on Development and Learning and on Epigenetic Robotics*, pages 482–488. IEEE, 2014.

[Hammond, 2010] Michael Hammond. What is an affordance and can it help us understand the use of ICT in education? *Education and Information Technologies*, 15(3):205–217, 2010.

[Hermans *et al.*, 2013] Tucker Hermans, James M Rehg, and Aaron F Bobick. Decoupling behavior, perception, and control for autonomous learning of affordances. In *IEEE International Conference on Robotics and Automation*, pages 4989–4996. IEEE, 2013.

[Horton *et al.*, 2012] Thomas E Horton, Arpan Chakraborty, and Robert St Amant. Affordances for robots: a brief survey. *AVANT. Pismo Awangardy Filozoficzno-Naukowej*, 2:70–84, 2012.

[Huang *et al.*, 2016] Yongqiang Huang, Matteo Bianchi, Minas Liarokapis, and Yu Sun. Recent data sets on object manipulation: A survey. *Big data*, 4(4):197–216, 2016.

[ICRA, 2019] ICRA. 2nd edition of international workshop on computational models of affordance in robotic, 2019.

[Jamone *et al.*, 2016] Lorenzo Jamone, Emre Ugur, Angelo Cangelosi, Luciano Fadiga, Alexandre Bernardino, Justus Piater, and José Santos-Victor. Affordances in psychology, neuroscience, and robotics: A survey. *IEEE Transactions on Cognitive and Developmental Systems*, 10(1):4–25, 2016.

[Jiang *et al.*, 2013] Yun Jiang, Hema Koppula, and Ashutosh Saxena. Hallucinated humans as the hidden context for labeling 3d scenes. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2993–3000, 2013.

[Kaiser *et al.*, 2015] Peter Kaiser, Markus Grotz, Eren E Aksoy, Martin Do, Nikolaus Vahrenkamp, and Tamim Asfour. Validation of whole-body loco-manipulation affordances for pushability and liftability. In *IEEE-RAS International Conference on Humanoid Robots*, pages 920–927. IEEE, 2015.

[Kaiser *et al.*, 2016] Peter Kaiser, Eren E Aksoy, Markus Grotz, and Tamim Asfour. Towards a hierarchy of loco-manipulation affordances. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2839–2846. IEEE, 2016.

[Katz *et al.*, 2014] Dov Katz, Arun Venkatraman, Moslem Kazemi, J Andrew Bagnell, and Anthony Stentz. Perceiving, learning, and exploiting object affordances for autonomous pile manipulation. *Autonomous Robots*, 37(4):369–382, 2014.

[Kim and Sukhatme, 2014] David Inkyu Kim and Gaurav S Sukhatme. Semantic labeling of 3D point clouds with object affordance for robot manipulation. In *IEEE International Conference on Robotics and Automation*, pages 5578–5584, 2014.

[Koppula *et al.*, 2016] Hema S Koppula, Ashesh Jain, and Ashutosh Saxena. Anticipatory planning for human-robot teams. In *Experimental Robotics*, pages 453–470. Springer, 2016.

[Kostavelis *et al.*, 2012] Ioannis Kostavelis, Lazaros Nalpantidis, and Antonios Gasteratos. Collision risk assessment for autonomous robots by offline traversability learning. *Robotics and Autonomous Systems*, 60(11):1367–1376, 2012.

[Kroemer *et al.*, 2012] Oliver Kroemer, Emre Ugur, Erhan Oztop, and Jan Peters. A kernel-based approach to direct action perception. In *IEEE International Conference on Robotics and Automation*, pages 2605–2610. IEEE, 2012.

[Krüger *et al.*, 2011] Norbert Krüger, Christopher Geib, Justus Piater, Ronald Petrick, Mark Steedman, Florentin Wörgötter, Aleš Ude, Tamim Asfour, Dirk Kraft, Damir Omrčen, et al. Object–action complexes: Grounded abstractions of sensory–motor processes. *Robotics and Autonomous Systems*, 59(10):740–757, 2011.

[Liu *et al.*, 2019] Chunfang Liu, Bin Fang, Fuchun Sun, Xiaoli Li, and Wenbing Huang. Learning to grasp familiar objects based on experience and objects' shape affordance. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(12):2710–2723, 2019.

[Lopes *et al.*, 2007] Manuel Lopes, Francisco S Melo, and Luis Montesano. Affordance-based imitation learning in robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1015–1021. IEEE, 2007.

[Luddecke and Worgotter, 2017] Timo Luddecke and Florentin Worgotter. Learning to segment affordances. In *IEEE International Conference on Computer Vision Workshops*, pages 769–776, 2017.

[Mar *et al.*, 2015] Tanis Mar, Vadim Tikhanoff, Giorgio Metta, and Lorenzo Natale. Multi-model approach based on 3D functional features for tool affordance learning in robotics. In *IEEE-RAS International Conference on Humanoid Robots*, pages 482–489. IEEE, 2015.

[Marion *et al.*, 2017] Pat Marion, Maurice Fallon, Robin Deits, Andrés Valenzuela, Claudia Pérez D'Arpino, Greg Izatt, Lucas Manuelli, Matt Antone, Hongkai Dai, Twan Koolen, et al. Director: A user interface designed for robot operation with shared autonomy. *Journal of Field Robotics*, 34(2):262–280, 2017.

[McGrenere and Ho, 2000] Joanna McGrenere and Wayne Ho. Affordances: Clarifying and evolving a concept. In *Graphics interface*, volume 2000, pages 179–186, 2000.

[Min *et al.*, 2016] Huaqing Min, Chang'an Yi, Ronghua Luo, Jinhui Zhu, and Sheng Bi. Affordance research in developmental robotics: a survey. *IEEE Transactions on Cognitive and Developmental Systems*, 8(4):237–255, 2016.

[Moldovan and De Raedt, 2014] Bogdan Moldovan and Luc De Raedt. Occluded object search by relational affordances. In *IEEE International Conference on Robotics and Automation*, pages 169–174. IEEE, 2014.

[Moldovan *et al.*, 2018] Bogdan Moldovan, Plinio Moreno, Davide Nitti, José Santos-Victor, and Luc De Raedt. Relational affordances for multiple-object manipulation. *Autonomous Robots*, 42(1):19–44, 2018.

[Montesano *et al.*, 2007] Luis Montesano, Manuel Lopes, Alexandre Bernardino, and Jose Santos-Victor. Modeling affordances using bayesian networks. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4102–4107. IEEE, 2007.

[Myers *et al.*, 2015] Austin Myers, Ching L Teo, Cornelia Fermüller, and Yiannis Aloimonos. Affordance detection of tool parts from geometric features. In *IEEE International Conference on Robotics and Automation*, pages 1374–1381. IEEE, 2015.

[Nguyen *et al.*, 2017] Anh Nguyen, Dimitrios Kanoulas, Darwin G Caldwell, and Nikos G Tsagarakis. Object-based affordances detection with convolutional neural networks and dense conditional random fields. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5908–5915. IEEE, 2017.

[Norman, 1988] Donald A Norman. The psychology of everyday things, 1988.

[Pandey and Alami, 2013] Amit Kumar Pandey and Rachid Alami. Affordance graph: A framework to encode perspective taking and effort based affordances for day-to-day human-robot interaction. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2180–2187. IEEE, 2013.

[Pieropan *et al.*, 2014] Alessandro Pieropan, Carl Henrik Ek, and Hedvig Kjellström. Recognizing object affordances in terms of spatio-temporal object-object relationships. In *International Conference on Humanoid Robots*, pages 52–58. IEEE, 2014.

[Price *et al.*, 2016] Andrew Price, Stephen Balakirsky, Aaron Bobick, and Henrik Christensen. Affordance-feasible planning with manipulator wrench spaces. In *IEEE International Conference on Robotics and Automation*, pages 3979–3986. IEEE, 2016.

[Ridge and Ude, 2013] Barry Ridge and Ales Ude. Action-grounded push affordance bootstrapping of unknown objects. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2791–2798. IEEE, 2013.

[RSS, 2014] RSS. First workshop on affordances: Affordances in vision for cognitive robotics, 2014.

[RSS, 2018] RSS. International workshop on computational models ofaffordance in robotic, 2018.

[Ruiz and Mayol-Cuevas, 2018] Eduardo Ruiz and Walterio Mayol-Cuevas. Where can I do this? Geometric affordances from a single example with the interaction tensor. In *IEEE International Conference on Robotics and Automation*, pages 2192–2199. IEEE, 2018.

[Şahin *et al.*, 2007] Erol Şahin, Maya Çakmak, Mehmet R Doğar, Emre Uğur, and Göktürk Üçoluk. To afford or not to afford: A new formalization of affordances toward affordance-based robot control. *Adaptive Behavior*, 15(4):447–472, 2007.

[Saputra *et al.*, 2019] Azhar Aulia Saputra, Wei Hong Chin, Yuichiro Toda, Naoyuki Takesue, and Naoyuki Kubota. Dynamic density topological structure generation for real-time ladder affordance detection. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3439–3444. IEEE, 2019.

[Saxena *et al.*, 2014] Ashutosh Saxena, Ashesh Jain, Ozan Sener, Aditya Jami, Dipendra K Misra, and Hema S Koppula. Robobrain: Large-scale knowledge engine for robots. *arXiv preprint arXiv:1412.0691*, 2014.

[Seker *et al.*, 2019] M Yunus Seker, Ahmet E Tekden, and Emre Ugur. Deep effect trajectory prediction in robot manipulation. *Robotics and Autonomous Systems*, 119:173–184, 2019.

[Shu *et al.*, 2016] Tianmin Shu, M. S. Ryoo, and Song-Chun Zhu. Learning social affordance for human-robot interaction. In *International Joint Conference on Artificial Intelligence*, pages 3454–3461, 2016.

[Shu *et al.*, 2017] Tianmin Shu, Xiaofeng Gao, Michael S Ryoo, and Song-Chun Zhu. Learning social affordance grammar from videos: Transferring human interactions to human-robot interactions. In *IEEE International Conference on Robotics and Automation*, pages 1669–1676. IEEE, 2017.

[Song *et al.*, 2013] Dan Song, Nikolaos Kyriazis, Iason Oikonomidis, Chavdar Papazov, Antonis Argyros, Darius Burschka, and Danica Kragic. Predicting human intention in visual observations of hand/object interactions. In *IEEE International Conference on Robotics and Automation*, pages 1608–1615. IEEE, 2013.

[Song *et al.*, 2015] Dan Song, Carl Henrik Ek, Kai Huebner, and Danica Kragic. Task-based robot grasp planning using probabilistic inference. *IEEE Transactions on Robotics*, 31(3):546–561, 2015.

[Stark *et al.*, 2008] Michael Stark, Philipp Lies, Michael Zillich, Jeremy Wyatt, and Bernt Schiele. Functional object class detection based on learned affordance cues. In *International Conference on Computer Vision Systems*, pages 435–444. Springer, 2008.

[Stoytchev, 2008] Alexander Stoytchev. Learning the affordances of tools using a behavior-grounded approach. In *Towards Affordance-Based Robot Control*, pages 140–158. Springer, 2008.

[Sun *et al.*, 2014] Yu Sun, Shaogang Ren, and Yun Lin. Object-object interaction affordance learning. *Robotics and Autonomous Systems*, 62(4):487–496, 2014.

[Szedmak *et al.*, 2014] Sandor Szedmak, Emre Ugur, and Justus Piater. Knowledge propagation and relation learning for predicting action effects. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 623–629. IEEE, 2014.

[Thermos *et al.*, 2017] Spyridon Thermos, Georgios Th Papadopoulos, Petros Daras, and Gerasimos Potamianos. Deep affordance-grounded sensorimotor object recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 6167–6175, 2017.

[Tikhanoff *et al.*, 2013] V Tikhanoff, U Pattacini, L Natale, and G Metta. Exploring affordances and tool use on the iCub. In *IEEE-RAS International Conference on Humanoid Robots*, pages 130–137. IEEE, 2013.

[Ugur *et al.*, 2015] Emre Ugur, Yukie Nagai, Erol Sahin, and Erhan Oztop. Staged development of robot skills: Behavior formation, affordance learning and imitation with motionese. *IEEE Transactions on Autonomous Mental Development*, 7(2):119–139, 2015.

[Veres *et al.*, 2020] Matthew Veres, Ian Cabral, and Medhat Moussa. Incorporating object intrinsic features within deep grasp affordance prediction. *IEEE Robotics and Automation Letters*, 5(4):6009–6016, 2020.

[Wang *et al.*, 2013] Chang Wang, Koen V Hindriks, and Robert Babuska. Robot learning and use of affordances in goal-directed tasks. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2288–2294. IEEE, 2013.

[Ye *et al.*, 2017] Chengxi Ye, Yezhou Yang, Ren Mao, Cornelia Fermüller, and Yiannis Aloimonos. What can I do around here? Deep functional scene understanding for cognitive robots. In *IEEE International Conference on Robotics and Automation*, pages 4604–4611. IEEE, 2017.

[Zech *et al.*, 2017] Philipp Zech, Simon Haller, Safoura Rezapour Lakani, Barry Ridge, Emre Ugur, and Justus Piater. Computational models of affordance in robotics: a taxonomy and systematic classification. *Adaptive Behavior*, 25(5):235–271, 2017.

[Zhu *et al.*, 2014] Yuke Zhu, Alireza Fathi, and Li Fei-Fei. Reasoning about object affordances in a knowledge base representation. In *European Conference on Computer Vision*, pages 408–424. Springer, 2014.