How Direct is Perception of Affordances? A Computational Investigation of Grasping Affordances

Giovanni Tessitore (tessitore@na.infn.it), Marta Borriello (aiamano82@gmail.com), Roberto Prevete (prevete@na.infn.it), Guglielmo Tamburrini (tamburrini@na.infn.it)
University of Naples Federico II, Compl.Univ. M. S. Angelo, Naples, NA 80126 Italy

Abstract

The computational model presented here, Grasping Affordances (GA) model, provides a precise explication of the notion of affordance in the context of grasping actions carried out by monkeys. This explication is consistent with both direct perception theories and neuroscientific models of monkey brains, insofar as the identification of grasping affordances requires, according to this model, neither object recognition processes nor access to semantic memory. Nevertheless, this model posits a cascade of complicated computational processes, in the way of visuo-motor transformations, which suggest the advisability of qualifying and re-interpreting the claim that (grasping) affordances are directly available to an acting biological system. This re-interpretation undermines the alleged alternative between direct and indirect perception theories, to the extent that substantive visuo-motor transformations have to be posited in order to identify grasping affordances.

Keywords: affordances; visuo-motor transformations; direct perception; grasping

Background and Motivations

The notion of affordance was originally introduced by J. J. Gibson (Gibson, 1979) to single out perceived properties that enable one to interact with objects in the environment. Procedurally, the notion of affordance is framed in the context of *direct* perception theories, insofar as higher-level cognitive processes, such as access to semantic memory, logical inference, and object recognition processes are allegedly unnecessary to identify an affordance. Direct perception theories emerged in contrast with so-called indirect perception theories (Michaels & Carello, 1981). According to the latter, complex mental processing steps are needed to fill in the gap between impoverished descriptions of the world furnished by sensory inputs on the one hand, and the rich and accurate descriptions of the world delivered by perception on the other hand. Thus, in particular, perceiving a glass as a graspable object one can drink from is the final outcome of mental processes involving knowledge of what a glass is, what it can contain, and how one uses it.

A more precise understanding of the processes involved in identifying an affordance is crucial to isolate what is conceptually and empirically at stake in the controversy between direct and indirect perception theories. And an understanding of these processes is crucial for the modelling of specific sensory-motor control mechanisms in biological systems too. The existence of a particulary versatile sensory-motor control mechanism is witnessed by the wide range of sensory-motor associations that monkeys are able to perform. Notably, this behavioural ability persists upon presentation of many unknown/novel objects, thereby suggesting that a robust generalization process, based on perceived object properties, is at work there (Borghi, 2005).

In the context of grasping actions, neurophysiological data on the macaque's brain cortex are consistent with direct perception views of affordances. In particular, these data suggest that the anterior intraparietal area (AIP) is involved in the coding of object affordances (Rizzolatti & Sinigaglia, 2008), in the light of functional hypotheses concerning more extended brain circuits. The functional models of brain areas which have been found to deliver afferent signals to AIP include neither perceptual object recognition nor higher-level cognitive processes, such as planning and decision-making (Creem & Proffitt, 2000; Milner, 1998). Moreover, strong efferent pathways have been identified which connect AIP to pre-motor area F5 (Rizzolatti & Sinigaglia, 2008). Since F5 is prominently involved in the coding of object-oriented actions (such as grasping, holding, and manipulating), the AIP to F5 connections suggest the existence of some sort of *direct* functional link between perceptual feature detection and object-directed actions.

The computational model presented here, Grasping Affordances (GA) model, provides a precise explication of the notion of affordance in the context of grasping actions carried out by monkeys. This explication is consistent with both direct perception theories and neuroscientific models of the macaque's brain. It is consistent with direct perception theories, insofar as the identification of grasping affordances requires, according to the proposed computational model, neither object recognition processes nor access to semantic memory. It is consistent with neuroscientific models of the macaque's brain, insofar as (i) visual processes furnishing AIP inputs are modelled in accordance with the biological "Standard Model" proposed in (Riesenhuber & Poggio, 2000), and (ii) the overall system output does not conflict with neuroscientific data and modelling constraints insofar as inputs supplied by AIP to brain motor areas are concerned. Nevertheless, this model posits a cascade of complex computational processes, in the way of visuo-motor transformations, which suggest the advisability of qualifying and re-interpreting the claim that (grasping) affordances are directly available to an acting biological system. This re-interpretation undermines the alleged alternative between direct and indirect perception theories, to the extent that substantive visuo-motor transformations have to be posited in order to identify grasping affordances.

The paper is organized as follows. First, a selective overview of neurophysiological findings about sensory-motor circuits in the macaque's brain cortex is provided, and basic features of computational models accounting for some of these data are briefly recalled. Then, an explication of the notion of affordance in the context of grasping actions is advanced. This explication sets the basic functional requirements for a computational model of grasping affordances, whose architecture and basic functionalities are described in some detail, and whose performances are evaluated on the basis of some preliminary tests. The import of this model on direct perception theories and future developments are briefly outlined in the concluding remarks.

Relevant Neurophysiological Findings and Computational Models

Brain areas in the macaque parietal and motor cortex were shown to be involved in a series of sensory-motor transformations, such as the mapping into appropriate actions of visual information about objects and their location in the visual scene (Rizzolatti & Sinigaglia, 2008). In particular, the AIP-F5 parieto-frontal circuit appears to play a crucial role in the visual guidance of hand grasping and manipulation movements, where AIP (Rizzolatti & Sinigaglia, 2008) was identified as a prominent cortical area involved in the coding of grasping affordances. One should be careful to note, moreover, that along the cerebral pathway starting from primary visual cortex (V1), and reaching F5 via AIP, visual information is transformed into motor information apparently without the intervention of cortical areas involved in higher-level perceptual and cognitive functions, such as the recognition of objects and their uses (Creem & Proffitt, 2000; Milner, 1998)

Two main computational models have been proposed in order to account for these data, by modelling AIP functionalities in the context of more comprehensive brain circuits. These are the FARS model (Fagg & Arbib, 1998) and a computational model of AIP neurons introduced in (Oztop, Imamizu, Cheng, & Kawato, 2006).

FARS is a neural model of cortical processes involved in generating and executing grasping plans. This model focuses on the interaction between AIP and premotor area F5, without providing a computational account of how inputs to area AIP are produced. In fact, affordances are ”programmed“ into this model, by hard wiring connections from units representing neurons in areas PIP and IT and units which represent neurons of area AIP. The connectivity between these units is determined by behavioural compatibilities. For example, an AIP unit which is selective for a specific grasp type and hand aperture receives inputs from units which hold input parameters of objects at which this kind of grasp and aperture are usually directed. Moreover, the model does not specify how these input parameters are computed from visual input. Their availability is taken for granted, and therefore the processing that visual information undergoes along the path from V1 to AIP is presupposed too. This comprehensive presupposition is acceptable in the FARS model, which is chiefly concerned with the generation and execution of grasping plans. It is not equally acceptable in a computational model which aims at accounting for the processes enabling one to extract affordances from visual inputs. For this reason, we have outlined here a computational account of contextually significant visuo-motor transformations occurring on the path from V1 to AIP.

The model proposed in (Oztop et al., 2006) concerns the development of AIP neuron functionalities while an infant is learning to perform grasp actions. This model focuses on an account of how units with processing properties similar to those of AIP neurons emerge by visuo-motor learning. Interestingly, the model demonstrates that units with different kinds of object selectivity emerge. In particular, units were found which encode object dimensions independently of object shape. This model exhibits limited generalization capabilities with respect to novel objects which do not belong to the initial training set. In fact, this generalization capability is restricted to transformations with respect to the size of known objects.

The model of grasping affordance extraction presented below (GA model) provides - unlike the FARS model - a detailed account of significant steps in perceptual processing along the path from V1 to AIP. In addition to this, the GA model is endowed - unlike the model proposed in (Oztop et al., 2006) - with more extended generalization abilities in the way of novel/unknown objects.

GA Model Description

Affordances for Grasping

Affordances are not intrinsic properties of an object, but rather depend on the relationship between object and agent (Chemero, 2003). For example, differences in primate and feline effectors account to a large extent for the different affordances that objects convey to humans and cats, respectively. As one moves to consider more specifically grasping affordances for monkeys and humans, one should still be careful to note that graspable objects do not merely ’afford’ our grasping them. Indeed, multiple opportunities for grasping arise in connection with many graspable objects. For example, a mug can be grasped by handle, lateral side, and top. These grasps can be distinguished from each other in terms of hand shape and wrist rotation obtaining just before grasping the object (Tucker & Ellis, 2000). Accordingly, the grasping affordances associated to a graspable object will be identified in the GA model with a collection of (codes for) appropriate hand configurations assumed by a hand just prior to grasping the object (Oztop et al., 2006; Tsiotas, Borghi, & Parisi, 2005). Since a graspable object may be grasped in several ways, this means that multiple hand configurations can be associated to any given object in the GA model.

General GA Model Description

From the above discussion, three main requirements have emerged for a computational model of grasping affordances to be empirically adequate and to move beyond previous computational models which include affordance extraction func-

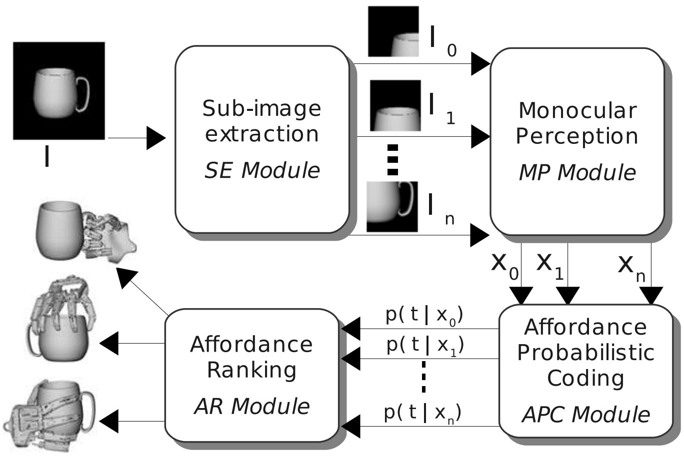


Figure 1: The GA model is formed by four modules: the SE Module, the MP Module, The APC Module, and the AR Module. This computational model receives an image depicting an object as input, and produces a list of affordances (appropriate grasps for the given object) as output.

tionalities: (a) the model must provide computational solutions for significant processing steps along the path from V1 to AIP; (b) the model must enable one to extract multiple hand-configurations from the same graspable object; (c) the model must possess generalization capabilities with respect to novel/unknown objects.

To accomplish (a), the visual pathway was modelled starting from primary visual cortex V1 and reaching, through areas V2 and V4, into the posterior infero-temporal area (PIT), which is identified as the cortical region supplying visual monocular information to AIP (Borra et al., 2007). A biologically plausible model of the ventral visual stream, named *Standard Model*, was proposed in (Riesenhuber & Poggio, 2000). A component of the Standard Model, the view-based Module, accounts for computations along the path from V1 to PIT which makes inputs available to AIP. Accordingly, the Monocular Perception (MP) Module (see Figure 1) which is an implementation of the view-based module was developed and included in the GA model.

To accomplish (b), that is, to provide a computational solution to the multiple affordance extraction problem, a probabilistic approach was pursued. In particular, this problem can be formalized as the problem of identifying and computing a multi-valued function which relates any visual input to a collection of hand-configurations. More precisely, let $X \subseteq \mathcal{R}^d$ be the d-dimensional space of visual inputs, and let $T \subseteq \mathcal{R}^c$ be the c-dimensional space of hand configurations. Then, one has to find a functional mapping f such that:

$$f : x \in X \longrightarrow \wp(T)$$

where $\wp(T)$ is the power set of T. A two-dimensional example of a multi-valued function is illustrated in Figure 2. This correspondence can be modelled by means of a probabilistic approach. More specifically, given $\bar{x}$, the output computed by the mapping f can be approximated by the unconditional probability density function $p(t)$. Thus, in general, the problem of modelling the functional mapping f can be viewed in terms of estimating the conditional distribution $p(t|x)$. A general framework for modelling conditional probability distributions makes use of mixture models whose parameters functionally depend on x (Bishop, 1995):

$$p(t|x) = \sum_{k=1}^{M} \alpha_k(x)\phi_k(t|x) \tag{1}$$

The $\phi_k(x)$ are kernel functions, which are usually Gaussian functions of the form

$$\phi_k(t|x) = \frac{1}{(2\pi)^{c/2}\sigma_k^c(x)} \exp\left\{ -\frac{\|t - \mu_k(x)\|^2}{2\sigma_k^2(x)} \right\} \tag{2}$$

The parameters $\alpha_k(x)$ can be regarded as prior probabilities of t generated from the k-th component of the mixture. The *Affordance Probabilistic Coding* (APC) Module was designed so as to provide a computational solution to (b), that is, to the multiple affordance extraction problem (see Figure 1).

To accomplish (c), that is, generalization capabilities enabling one to extract affordances from novel objects, a starting point was provided by the observation that the agent usually focuses its attention on the part of the object at which the grasping action is directed (Schiegg, Deubel, & Schneider, 2003). This behaviour suggests the possibility of associating parts of a graspable object to affordances, and to store this "mereological" information for use when novel graspable objects are presented. For example, one may learn to associate appropriate affordances to handles and cylinders, respectively, and to use this information when a cup (resulting from the "composition" of handle and cylinder) is presented. This process was actually implemented by sliding an "attention window" on the image of an object, and by extracting a collection of grasping affordances at each displacement step. This function is achieved by the Subimage Extraction (SE) Module (see Figure 1). Finally, a post-processing

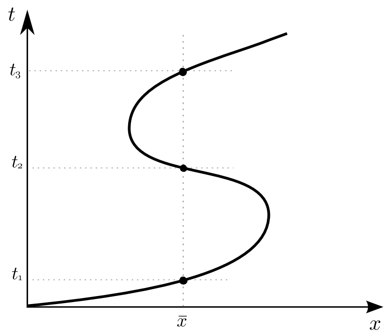


Figure 2: Two-dimensional example of a multi-valued function. Points on the x axis represent visual inputs, and points on the t axis represent hand-configurations. One may associate a x point with multiple t points.

step was required as well, in order to select the more plausible affordances. The post-processing step is accomplished by Affordance Ranking (AF) Module (see Figure 1). APC and AR modules account for the AIP affordance computation. The online learning of sensorimotor associations might be grounded onto a basic grasping ability such as described in (Oztop, Bradley, & Arbib, 2004). Learning of sensorimotor associations may occur by collecting pairs of visually presented ”object part” and related ”hand-configuration” every time a successful grasp is made. Since the focus of this work is not on the acquisition of sensorimotor associations, however, we suppose here that a series of such pairs is already available.

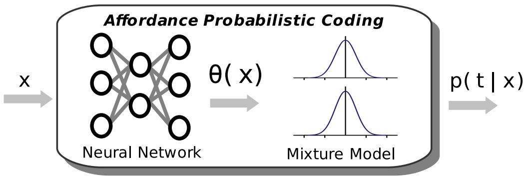


Figure 3: The APC Module is formed by a neural network and a Gaussian mixture model. Given an x vector, the neural network computes the required Gaussian parameters $\theta(x)$ to approximate $p(t|x)$ (see (Bishop, 1995) for more details).

GA Model specification and implementation

The GA model takes the image of an object as input and supplies the object's grasping affordances as output. It is composed by four modules, as shown in Figure 1. The input image I, represented in gray scale, is processed by the SE Module, which extracts n subimages $I_j,\ j = 1,...,n$. The number of subimages depends on the dimensions of the window W sliding on the image I, the image size, and the window displacement step DS.

Each subimage is then sent as input to the MP Module. The MP Module takes a sub-image I_j as input, and gives a 256 feature vector as output x_j. The latter is presented as input to the APC Module, which computes the corresponding $p(t|x_j)$.

To estimate $p(t|x_j)$, one uses a mixture model of the form expressed in eq. 1, whose parameters $\alpha_k(x)$, $\mu_k(x)$ and $\sigma_k(x)$ (for Gaussian kernel as in eq. 2) depend on the visual input x. The relationship between visual inputs x and corresponding mixture parameters is modelled by means of a two-layer, feed-forward neural network with H hidden nodes. Therefore, the ACP Module has a combined density model and neural network structure, as shown in Figure 3.

Since the APC Module receives n feature vectors x_j in input, its overall output is formed by n density functions $p(t|x_j)$. Note, however, that the desired output is a set $T = \{t_1, t_2, \ldots, t_L\}$ corresponding to the L distinct hand-configurations that enable one to grasp the viewed object. Therefore, a non-trivial output selection problem remains to be solved at this stage: one has to isolate hand-configurations which differ from each other as much as possible, and whose probability value is sufficiently high.

This requirement corresponds, for each single feature vector x and related $p(t|x)$, to choose as member of the set T the gaussians' centers $\mu_k(x)$ of the mixture associated to the higher values of $\alpha_k(x)$. In the case of n probability distributions $p(t|x_1), \ldots, p(t|x_n)$, in order to obtain a behaviour similar to the single distribution case, one may proceed as follow:

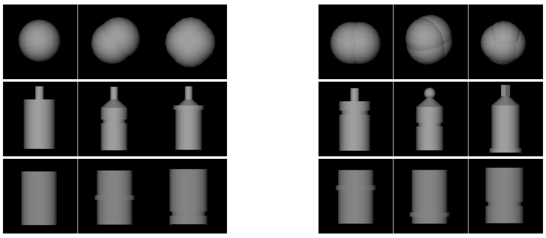
1. generate s sample points from each distribution, obtaining $n \times s$ points, each of which defines a hand configuration. Not every hand configuration thus obtained corresponds to grasps for the input object; only those gathering around the kernel's means do, while the other points are distributed in a sparse manner;
2. a clustering over the $n \times s$ points is performed;
3. the clusters are ranked according to the order of their variance values, and the first L clusters with lower variances are selected because a lower variance implies less uncertainty about the hand configurations;
4. finally, the set T will be formed by the centers of the selected clusters.

Test and Results

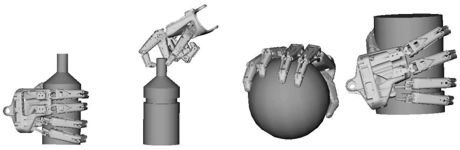
The GA model was designed so as to extract multiple hand-configurations, and to generalize its affordance-extraction capability with respect to novel objects. Two experiments were performed to test the extraction and generalization abilities, respectively. The results of these tests corroborate the possession of the extraction ability, in addition to the required generalization ability as far as novel objects obtained from the composition of known object parts are concerned. Let's see.

The first test, which is concerned with the extraction of multiple hand-configurations, makes use of three different prototypical object images: a sphere, a cylinder and a bottle. It is assumed that the first two objects can be grasped using a power grasp only, whereas the bottle can be grasped in two different ways, by precision and power grasps. For each of these prototypical object images, similar images were generated by means of small contour variations. For each prototype, the resulting training and test sets were composed by 20 and 10 images, respectively (Figure 4)

In order to generate target hand configurations, GraspIt! (Miller & Allen, 2004), a robotic grasping simulator, was used. In particular, the robotic hand called Robonaut, endowed with 14 degrees of freedom, was chosen. Consequently, in the GA model hand configurations are identified by a vector of 14 components, where each component represents just one hand joint's angle. Spherical and cylindrical objects are associated to a single hand configuration, generated manually by changing the Robonaut's degrees of freedom. Bottle objects are associated with two distinct hand configurations: a precision grasp, applied on the object's top part, and a power one applied on the lateral part (see fig. 4). Training set targets are generated adding some Gaussian noise to



(a) Some training objects (left) and test objects (right).



(b) Target hand-configurations.

Figure 4: Examples of spherical, cylindrical and bottle objects used to train and test the system, and target hand-configurations.

these hand configurations. In this test, the attention window encompasses the whole object. Thus, for each object there is a single feature vector x with an associated $p(t|x)$. Hand configurations are obtained by selecting $\mu_k(x)$ associated with the higher values of $\alpha_k(x)$. The model parameters are summarized in table 2. For the i-th degree of freedom, percentage error is defined as $\frac{|t^i - y^i|}{max_i - min_i} \times 100$, where y_i is the model output, and max_i and min_i are the max and the min value, respectively, for the i-th degree of freedom. *Average error* between model output hand configuration and target hand configuration is defined as the mean of percentage error over all degrees of freedom. For all test objects in each class, mean and standard deviation of average error is computed and showed in table 1.

Table 1: For each object class, the mean and standard deviation of the average error over all objects in the test set is reported here. Moreover, for each class mean hand-configuration over all objects in the class is exhibited.

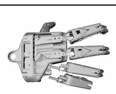
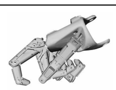
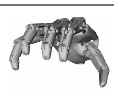
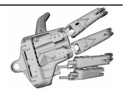
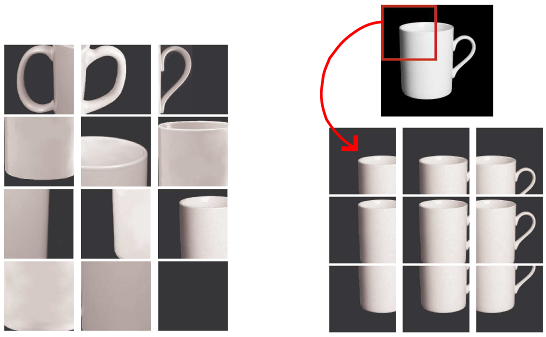
Bottle Grasp 1	Bottle Grasp 2	Spherical	Cylindrical
2% $\pm$ 0.4	1.9% $\pm$ 0.6	3.9% $\pm$ 1.4	1.3% $\pm$ 0.4
			

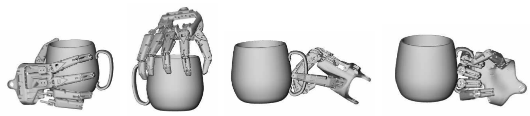
Table 2: Model parameters for each test. Image size, W and DS are expressed in pixels.

	H	M	Image size	W	DS	Cluster
Test 1	5	2	160×160	160×160	0	*None*
Test 2	5	5	500×500	160×160	30	5

The second experiment is meant to test generalization capabilities with respect to novel objects. To test this ability, the system was trained to associate *parts* of an object to hand-configurations. Subsequently, the system was given in input a novel object resulting from the ”composition“ of previously known parts. In this test, a cup is used, which is obtained from the composition of a cylinder and a handle. Examples of both training images and the cup used as test image are shown in figure 5. There are four kinds of training images: (a) cup handles; (b) upper and lower cup parts; (c) lateral cup parts; (d) non-graspable cup parts. Two target hand-configurations are associated with images (a); only one hand-configuration is associated to images (b) to (d). The training set targets are generated adding some Gaussian noise to hand configurations. Targets for non-graspable cup parts images are drawn from a Gaussian distribution with a large variance, so as to reflect the fact that in this case no plausible hand-configuration candidate exists. The K-Mean clustering algorithm is implemented by the AR Module, setting to 5 the number of clusters. In table 3, cluster centroids are shown together with cluster variance. The fifth cluster was discarded in view of its large variance. Note that the first four cluster centroids are very similar to target hand configurations (fig. 5) with respect to which mean percentage error was computed.



(a) Some training objects (left) and test object (right).



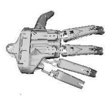
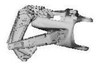
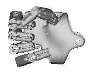
(b) Target hand-configurations.

Figure 5: (a) Examples of training and test images (see text). (b) Examples of target hand-configurations.

Concluding remarks

The architecture of the GA model is largely motivated by the goal of computationally investigating the allegedly direct link between perception and action established by the perception of affordances. One should be careful to note that the overall output of the GA model does not correspond to actions, but rather corresponds to hand configurations. Therefore, one may legitimately question the claim that the GA model computes a perception-action transformation. However, in

Table 3: The graph visualizes the obtained cluster centroids. Compare these images with target hand configurations of fig. 5. The fifth cluster was discarded in view of its large variance. The percentage error with respect to target was mediated over all degrees of freedom.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
				
$\sigma = 0.12$	$\sigma = 0.12$	$\sigma = 0.09$	$\sigma = 0.09$	$\sigma = 0.34$
Mean and standard deviation of percentage error				
$1.9\% \pm 2$	$2.5\% \pm 2$	$2\% \pm 1.2$	$1.8\% \pm 1.5$	*(discarded)*

the context of grasping actions, the model embodies the assumption that an appropriate hand configuration for grasping an object is a configuration assumed by a hand just prior to grasping that object. This configuration is closely related to the *goal* of the grasping action. Thus, the grasping action can be generated from the initial configuration, in terms of motor commands, by a forward model on the basis of such goal-related information. For this reason, one can meaningfully maintain that the computation of hand configurations from visual inputs performed by the GA model is the gist of a perception-action transformation.

As discussed in the first section, a more precise understanding of the processes involved in identifying an affordance is crucial to isolate conceptual and empirical differences between direct and indirect perception theories. The GA computational model is in agreement with the notion that the identification of affordances does not require higher cognitive processes, such as logical inference and object classification. However, the transformation performed in the GA model requires a cascade of fairly complicated processing stages, and the solution of non-trivial computational problems. Notably, in order to achieve significant generalization capabilities, the APC module was geared so as to produce in output a set of probability distributions each one of them expressed as a Gaussian mixture, coding hand configurations for just one part of the image. Here, the pertinent modelling question is: how one does choose the appropriate hand configurations for the object? In the case of just one probability distribution, a natural candidate are the centers of the Gaussians associated to the higher mixture coefficients. In the case of a set of probability distributions, various possibilities arise, only one of which was pursued in the GA model. This solution provides a significant proof-of-concept, together with a vivid illustration of the important qualifications that are needed when one makes use of the attribute direct in the expression direct perception of affordances. An alternative solution, which we are currently exploring, involves a unique probability distribution, which arises by taking as some sort of union over the set of distributions based on a similarity measure between gaussian mixture models (Hershey & Olsen, 2007).

References

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

Borghi, A. M. (2005). Grounding cognition: The role of perception and action in memory, language, and thinking. In (chap. Object concepts and action). Cambrige University Press.

Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., et al. (2007). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cerebral Cortex*.

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*.

Creem, S. H., & Proffitt, D. R. (2000). Defining the cortical visual systems: “what”, “where”, and “how”. *Acta Psychologica*.

Fagg, A., & Arbib, M. (1998). Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks*.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Hershey, J. R., & Olsen, P. A. (2007). Approximating the kullback leibler divergence between gaussian mixture models. In *Icassp 2007. ieee international conference on acoustics, speech, and signal processing.*

Michaels, C. F., & Carello, C. (1981). *Direct perception*. Englewood Cliffs, New Jersey 07632: Prentice-Hall, Inc.

Miller, A. T., & Allen, P. K. (2004). Graspit! a versatile simulator for robotic grasping. *Robotics & Automation Magazine, IEEE*.

Milner, A. D. (1998). Neuropsychological studies of perception and visuomotor control. *Philos Trans R Soc Lond B Biol Sci*.

Oztop, E., Bradley, N. S., & Arbib, M. A. (2004). Infant grasp learning: a computational model. *Experimental Brain Research*.

Oztop, E., Imamizu, H., Cheng, G., & Kawato, M. (2006). A computational model of anterior intraparietal (AIP) neurons. *Neurocomputing*.

Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, *3*, 1199–1204.

Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain. how our minds share actions and emotions.* Oxford University Press: Oxford.

Schiegg, A., Deubel, H., & Schneider, W. (2003). Attentional selection during preparation of prehension movements. *Visual Cognition*.

Tsiotas, G., Borghi, A., & Parisi, D. (2005). Objects and affordances: An artificial life simulation. In *Proceedings of the cognitive science society.*

Tucker, M., & Ellis, R. (2000). Micro-affordance: the potentiation of components of action by seen objects. *British Journal of Psychology*.