A Measure of Good Motor Actions for Active Visual Perception

Jonas Ruesch
Institute for Systems and Robotics
Instituto Superior Técnico
Portugal, Lisbon 1049-001
Email: jruesch@isr.ist.utl.pt

Ricardo Ferreira
Institute for Systems and Robotics
Instituto Superior Técnico
Portugal, Lisbon 1049-001
Email: ricardo@isr.ist.utl.pt

Alexandre Bernardino
Institute for Systems and Robotics
Instituto Superior Técnico
Portugal, Lisbon 1049-001
Email: alex@isr.ist.utl.pt

***Abstract*—The morphology of an embodied agent heavily influences the characteristics of sensory signals induced during sensorimotor activity. In order to facilitate later information processing, it is desirable to couple sensors and actuators in a way such that the agent's actions induce well structured sensory feedback. In this work, we empirically investigate the meaning of "well structured" with respect to the predictability of the induced stimulus and the complexity of the model required to predict the change. For this purpose, we consider a visual sensor consisting of an unknown topology composed of light sensitive receptors and we investigate a number of different actions. For each action, we learn a stimulus prediction function and we propose to qualify the action depending on the complexity of the prediction model. By visualizing this measure for two different sensors and over different action spaces, we observe that a given sensor topology implicitly imposes actions for which the provoked change in stimulus can be described with less effort, which we feel has direct implications for the behavior of an agent using that sensor.**

I. Introduction

Every motor command triggered by our brain provokes a sequence of sensorimotor reactions which induce highly context sensitive stimuli. Usually, we merely attribute the observer's environment to be the context of these signals. This seems natural, because of course, it is something in the outside world which is supposed to cause the perceived sensor activation. However, having a closer look at the *process* which generates the sensor signals, we realize that the final signal is not only defined by the state of the environment with respect to the observer, but that the generation of the stimulus is heavily influenced by the structure of the sensor and the actions provoking it. In fact, all of these ingredients contribute conjointly when input signals for the visual system are formed. Drawing inspiration from this perspective we feel that, in the same way in which early visual stimuli are acquired and shaped by the recording sensor and its interaction patterns, any higher level visual percept is likewise strongly dependent on sensorimotor interaction. This point of view has a long-lasting history in developmental psychology. Classical work on sensorimotor coordination in perception include for example [1], [2] and [3]. There the idea is pursued that percepts are built during sensorimotor interaction which, inversely, leads to the understanding that the recognition of experienced phenomena is tied to applicable actions [3]. A more recent contribution especially with respect to visual perception can be found in [4], where it is argued that visual percepts are acquired through the training and execution of sensorimotor "skills".

In the work described here, we address the influence of sensor morphology and sensor movements on the structure of an induced stimulus signal. Our motivation is to understand how actions and sensor structure interplay in shaping primary stimuli recorded by a visual system. Our final intention is to qualify a sensor repositioning action depending on characteristics present in the stimulus it provokes. With this objective, we formulate a concept for *good motor actions* driven by two lines of thought: first of all, we prefer actions which lead to a change in sensory stimulus which is predictable; secondly, we prefer actions which lead to a change in stimulus which can be described using a simpler model. Following these criteria we develop an environment-independent measure specifying "how simple" it is to predict the outcome of a given action under a given sensor topology. We motivate these two central ideas related to prediction more in depth in Section II. In the results section of this paper, we evaluate two commonly considered sensor layouts using three different types of actions. By visual inspection of obtained measurements, we are able to detect what we consider to be good actions specific to a given sensor topology. Based on the assumption that actions leading to predictable sensor feedback and simpler prediction models are preferred, we are able to deduce how an agent's actions and behaviors are coupled with its sensor topology. We find that the resulting conclusions match well with observations made for biological systems.

Previous work related to the analysis of sensory feedback generated during sensorimotor interaction has been published in [5]. The authors describe results obtained when applying information theoretic measures between different actuator and sensor variables in real and simulated robots. While the results presented in [5] also address visual sensors and sensor repositioning actions, correlation and information transfer measures are computed between specific image locations and particular actuator variables during a specific task. The obtained results report measures depending on morphology, action *and* experienced situation. On the contrary, in our work we propose a measure based on a learnt prediction model which is valid for any situation the agent encounters in the sampled environment.

978-1-61284-990-4/11/$26.00

This approach explicitly excludes the influence of a particular situation and creates an environment-independent measure which qualifies an agent's action with respect to the agent's morphology.

Work described in [6], in some way inversely, shows how the topological structure of a visual sensor can be discovered using an entropy maximization algorithm, a process which is related to our work in the sense that eventually an agent-specific but environment-independent result is obtained.

In a more general form, the paradigm of sensorimotor coordination has also been promoted in [7], where sensorimotor interaction is advocated as a fundamental design principle for perception in embodied artificial agents. For example, based on work done in [8], it is sketched out how a robot can use sensory signal patterns generated during sensorimotor activity to facilitate recognition of objects.

The remaining part of this paper is organized as follows. In the next section we will motivate more in depth why we choose prediction and prediction model complexity to compose the proposed measure. We then introduce our definition of good motor actions in Section III, after describing the agent's observation and action model. In Section IV, we present and interpret results obtained from sampling two different sensor topologies.

II. What are Good Motor Actions?

The first, and more intuitive criterion we propose for good motor actions, is whether the action leads to an accurately predictable change in sensory stimulus. We choose prediction as an important trait of a good motor action since prediction is a fundamental mean to decide whether a change in stimulus complies or deviates from an expected feedback. This capability can be found in many animals, including organisms featuring a nervous system of a few hundred neurons [9]. An extensive overview of how predicted sensory signals are used in biological systems is given in [10]. As outlined there, sensory stimulus prediction is ubiquitous in nature being used for reflex inhibition, sensory filtration, and stabilization of percepts.

The second, less obvious but not less important criterion for good motor actions in the proposed sense, is the complexity of the prediction model. In what follows, we motivate this idea by reviewing three independent lines of research. Although, we do not formally prove the necessity for simple models, and at first sight there seems to be limited overlap between the three areas outlined, we hope we can convince the reader of the common underlying concept and of the appeal to strive for simple prediction models which can be achieved by smartly shaping the morphology of a physical system.

a) Eye Morphology in Nature: The receptor distribution in the human eye is a particularly prominent example of a specifically shaped sensor topology. In [11], it is pointed out that the log-polar-like receptor distribution corresponds to a mapping function which transforms image rotations and dilations (zoom) into simple coordinate shifts in the log-polar coordinate system. Hence, if an eye featuring such a receptor distribution is focusing on an object and that object is rotated or scaled, the projected image is merely shifted along the log-polar coordinate axes. It was argued that this property results in an advantage for the human visual cortex, as it could achieve image invariance for these transformations at a low computational cost by simply shifting the image.

Other mammalians, for example sheep, pigs, horses or the red kangaroos (*Macropus rufus*) feature a horizontally elongated ganglion cell distribution in their fovea, also called *visual streak* [12], [13]. Such distributions could account for the fact that horizontal image translations are more frequently experienced by these animals, because as opposed to predators, they have a more limited binocular vision and their behavior is less "object oriented". On the other hand, it is important for these species to observe the horizon, a behavior which induces horizontal image shifts. Hence, the visual streak improves the "coordinate shift" property previously described.

b) Self-similar Point Distributions: In an inventive work described in [14], the following was proved: a set of points randomly distributed on a disk converges to a stable configuration given: i) points are conjointly transformed by rotations, dilations and translations which are applied according to a given probability distribution; and ii) after a transformation action is applied, each point is moved towards transformed points which are lying closest to the point under consideration. The paper shows, the final point distribution is the configuration where each point has on average the smallest possible distance to the next closest transformed point under the given action probability distribution. In [14], this quality is called the *self-similarity* of that point-set and action probability distribution. Configurations obtained with unrestricted uniformly distributed rotation and dilation actions, but restricted uniformly distributed translation actions show highly regular fovea-like point distributions. Furthermore, using different action probability distributions for horizontal and vertical translations, elliptic (visual streak-like) point distributions are obtained (for an illustration see Figure 10 in [14]).

c) Morphological Computation for Distance Estimation: In a setup described in [15], a robot equipped with a number of light sensitive receptors moves along a textured wall. The receptors form a horizontal, one-dimensional visual sensor which can be reconfigured online. The robot executes a optimization algorithm which searches for a sensor configuration which ensures that the time required for a dark / bright contrast to move from one receptor to the next is the same for all pairs of receptors. It is then shown, that the optimal configuration found linearizes the computation of the robot's distance to the wall, allowing a very simple computation to execute this task.

In the next section, we will formalize a definition which embraces the motivation given here. Prior to that, we would like to emphasize the following. If one is familiar with the concept of uncertainty sampling, e.g. for active learning where an action is considered a good action when it leads to still unknown feedback, then the paradigm introduced in this work might sound counterintuitive, as we precisely reward actions

which do not discover novelty [16]. However, our understanding of good motor actions is different and this distinction is deliberate. We consider actions which lead to well predictable changes advantageous for an embodied agent, because they support and improve for example the capability to distinguish between sensory stimuli which originate in the environment (e.g. moving predators or prey) and self-induced stimuli (provoked by movements of the agent's own body). Relying on better predictable actions improves forward modelling of self-generated stimuli which is a common concept implemented in nervous systems of many but the simplest organisms [10].

III. A Definition of Good Motor Actions

To test the so far introduced ideas, we consider an agent in state s placed in a 2-dimensional world described as a luminance function $i_s : \mathbb{R}^2 \to \mathbb{R}$, i.e. a grayscale image. The agent observes the world through a sensor composed of N receptors. Each of these receptors is described as a receptive field f_n which is used to integrate luminance like

$$o_n(i_s) = \int f_n(x) i_s(x)\, dx, \tag{1}$$

where $o_n(i_s)$ is the value observed by the n-th receptor and the receptive field function f_n is modeled as a multivariate Gaussian on $i(x)$. We choose Gaussians as the possible set of functions for f_n because Gaussians are particularly well suited to describe receptive field functions, both for biological plausibility as well as for their amenable mathematical properties [17].

After taking the observation at state s, the agent can change i_s using its actuators. Therefore, after performing action a, the agent observes

$$o_n(i_s \circ a) = \int f_n(x) i_s\left(a(x)\right)\, dx, \tag{2}$$

where we define $i_{s+1} = i_s \circ a$. For simplicity, these actions are considered well behaved in the sense that they are smooth bijections. The notation $\mathbf{D}_s$ and $\mathbf{D}_{s+1}$ is used to denote the sensor activation values before and after applying action a:

$$\mathbf{D}_s = \begin{bmatrix} o_1(i_s) \\ o_2(i_s) \\ \vdots \\ o_N(i_s) \end{bmatrix}, \qquad \mathbf{D}_{s+1} = \begin{bmatrix} o_1(i_{s+1}) \\ o_2(i_{s+1}) \\ \vdots \\ o_N(i_{s+1}) \end{bmatrix}. \tag{3}$$

Incorporating the ideas presented in Section II, we now formulate what we mean by prediction using a simple model. The general form of stimulus prediction with the given observation model refers to a function $p^a : \mathbb{R}^n \to \mathbb{R}^n$ which, when applied to the initially observed sensor values $\mathbf{D}_s$, is able to approximate the sensor values obtained after applying an action a:

$$\mathbf{D}_{s+1} \approx p^a\left(\mathbf{D}_s\right). \tag{4}$$

Considering a static world and a spatially rigid sensor layout, we observe that the class of functions from which p^a should be chosen can be restricted. In the appendix, we provide an argument that motivates a reduction of this set of functions to the linear function set. Having selected a linear p^a we are now left with the concern how to decide on the complexity of the prediction model. A common and natural approach to select simpler models is to evaluate the number of parameters required by the model. For a linear predictor, we can translate this requirement by forcing the predictor to be sparse. In this sense, we revise equation (4) as

$$\mathbf{D}_{s+1} \approx \mathbf{P}^a\, \mathbf{D}_s, \quad \mathbf{P}^a \text{ sparse}, \tag{5}$$

where $\mathbf{P}^a$ is the matrix representation of the linear prediction function p^a. This equation is still ill defined since the notion of sparsity is vague and nothing has been said about the prediction error. It is our belief that these cannot be canonically defined, so several alternatives can be proposed as a means of mixing the importance of sparsity and allowed error:

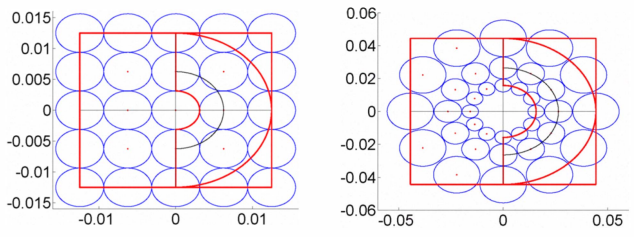
- Fix sparsity and minimize the prediction error. For example one can say that each row of matrix $\mathbf{P}^a$ must have a single non-zero entry (sparsity) and under this set minimize the norm error. Other sparsity sets can be chosen such as $\mathbf{P}^a$ having at most k non-zero entries or that $\mathbf{P}^a$ must be permutation matrices.
- Minimize the prediction error and infer sparsity. A strategy which first obtains the minimum norm error solution for $\mathbf{P}^a$ with a rule which grounds entries of this solution to zero and subsequently deduces sparsity.
- Simultaneously minimize both prediction error and sparsity. For example the well known LASSO algorithm allows for a single parameter to weight the importance of sparsity versus norm error [18].

Any of these methods will obtain a prediction matrix $\mathbf{P}^a$ as well as the associated prediction error E^a from several samples of the sensor values before ($\mathbf{D}_s$) and after ($\mathbf{D}_{s+1}$) applying action a from randomly chosen states s. An action performance score for the given sensor topology is therefore a function of the prediction matrix and the prediction error like $Q^a(\mathbf{P}^a, E^a)$. We call a good action one which induces a high performance score. The choice of this performance score is usually tied with the choice of the sparsity / prediction optimization algorithm. For example, if sparsity is fixed, then the performance should only depend on the final prediction error.

For the experiments described next, we chose the approach which first minimizes the prediction error while constraining entries of $\mathbf{P}^a$ to be greater or equal to zero. To provide the performance score we use the Gini index which is a well known sparsity measure complying with a number of desirable properties [19]. In short:

$$\begin{aligned} \mathbf{P}^a = \; & \operatorname{argmin} \quad \textstyle\sum_i \left\| \mathbf{D}^i_{s+1} - \mathbf{P}^a\, \mathbf{D}^i_s \right\|_2^2 \\ & \text{s.t.} \quad \mathbf{P}^a \geq \mathbf{0} \\ Q^a(\mathbf{P}^a) = \; & \text{Gini}\left(\mathbf{P}^a\right), \end{aligned}$$

where $\mathbf{P}^a$ is defined as an average over a number of actions i. We found this combination to provide good invariance to sampling noise and overall results consistent with what was



(a) Grid Layout (b) Foveal Layout

Fig. 1. Investigated sensor topologies. Circles visualize the standard deviation of the Gaussian receptive field of each receptor. In red, the covered action ranges; rectangular: the range of translational actions; c-shaped: the range of rotation and scale actions.

to be expected. Note that as privileged observers we, the reader and the authors, have access to the underlying sensor topology hence can judge what is to be expected, but this information is not given to the algorithm. Also, the constraint $\mathbf{P}^a \geq \mathbf{0}$ feels natural in a biological context and the chosen approach does not require any further parameter as it is the case when fixing sparsity (number of non-zero entries in $\mathbf{P}^a$) or when implementing a regularized version of the least squares solution such as the LASSO method (regularization parameter).

IV. Results

In this section we investigate two given sensor layouts, a regular grid-like configuration, and a non-uniform, fovea-inspired layout with a logarithmic parametrization, also commonly used to describe growth spirals found in nature [20]. We chose the grid configuration because of its relevance with respect to basically all artificial image sensors, e.g. CCD sensors available off-the-shelf. The logarithmic distribution was chosen to analyse stimulus change patterns for foveating visual systems as described in Section II. The two layouts are shown in Figure 1. We investigate three types of actions and their combinations: translations, rotations and dilations. For each sensor configuration and two action sub-spaces (shift-x / shift-y and rotation / scale), we applied the proposed measuring algorithm to 10 000 randomly chosen actions. Figure 1 shows the action ranges in red with respect to the sensor topologies. Each score was obtained by sampling a particular action 1.4-times the number of receptors from randomly chosen positions in the image. Figures 2 to 5 show the obtained scores. In what follows, we discuss the results for each sensor topology individually.

A. Uniform Grid Sensor Layout

Measurements for horizontal and vertical translation actions are presented in Figure 2. The plot shows clear peaks when a translation equals a combination of horizontal and vertical receptor distances. The peaks corresponding to larger action steps are slightly smaller because larger displacement actions lead to a bigger number of peripheral and unpredictable receptors, provoking noise in the prediction operator which is responsible for less sparse solutions. This is a desirable

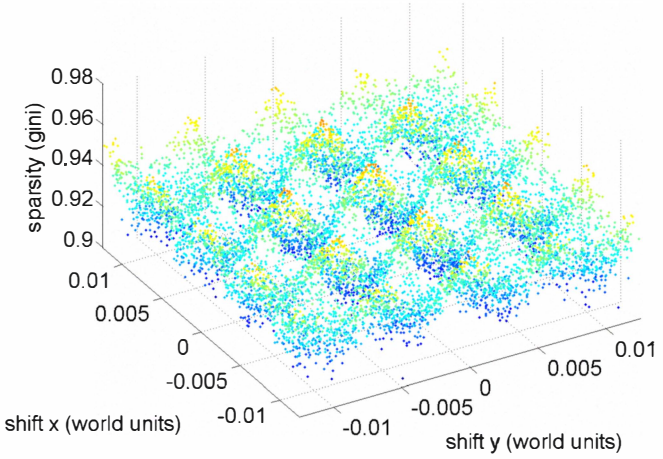


Fig. 2. Actions scores for the grid layout under horizontal and vertical translations. Translation units refer to the sampled world which extends over $[-1, 1]$ in both directions.

effect, as peripheral receptors do not qualify well in terms of our measure. The results for rotation and dilation are shown in Figure 3. Excluding the peak for the zero-action, only two significant peaks are visible located at zoom level $z = 1$ and $90°$ and $-90°$ rotation. This makes sense as these are the actions which achieve a perfect permutation of receptors. In between, the scores are significantly lower with certain angles achieving a slightly better score than others.

In summary, while this sensor achieves good measures for shifting actions with step-sizes related to receptor distances, the topology does not qualify well for zoom and rotation actions. Hence, while translation can be computed using less parameters and operations, compensation of actions inducing rotation and dilation is more complex for this layout.

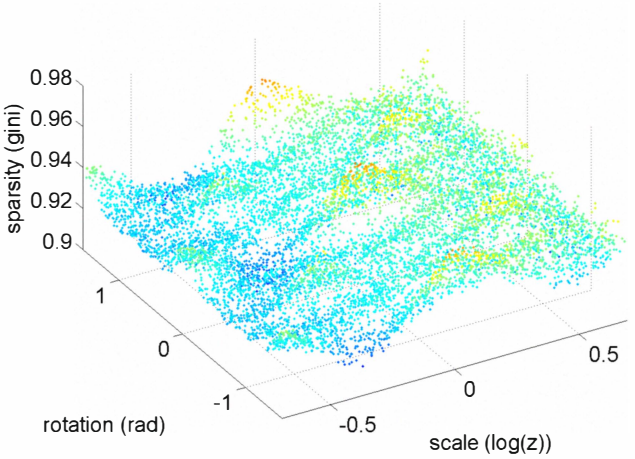


Fig. 3. Actions scores for the grid layout under rotation and dilation. The zoom factor z denotes a scaling of the sensor topology.

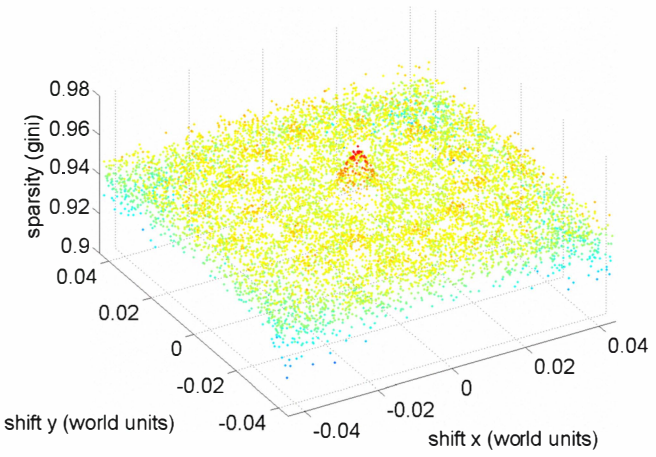


Fig. 4. Actions scores for the foveal layout under horizontal and vertical translations.Translation units refer to the sampled world which extends over $[-1, 1]$ in both directions.

B. Foveal Sensor Layout

Quite contrary to the grid-like layout, translational actions using the foveal sensor lead to the results shown in Figure 4. Disregarding the peak for the zero-action, the plot does not show clear good actions except for a ring of local optima corresponding to shifts of the center location to receptor positions on the second ring, although, these shifts are assigned a significantly lower score than the shift-peaks in the previously discussed layout (compare also Figure 1). Much different appear the measurements for the rotation and scale action space shown in Figure 5. Here we observe seven strongly expressed peaks at zoom level $z = 1$ distributed according to the number of receptors on the semicircle between $-90°$ and $90°$. We also note that at the same angular positions, the same number of peaks appears for zoom levels $z = 0.60$ and $z = 1.68$. These peaks correspond to scaling and rotation actions which map receptors from one ring to a neighbor ring of receptors.

We conclude, a foveal sensor favours simplified linear prediction operators for rotational transformations but sacrifices on the other hand sparsity under translational displacements.

C. Discussion

We observe a clear shift in the characteristics of the action scores when moving from the grid sensor layout to the foveal layout. While the grid layout achieves higher scores under translation, the foveal sensor requires a more descriptive prediction function for the same actions. This coincides with the fact that animals with a foveating visual system usually compensate image translations by appropriate sensor movements, e.g. saccadic gaze shifts or smooth pursuit of foveated objects. Hence, the handling of image translations is "outsourced" to active sensor repositioning allowing for an abatement of image transformations for which the sensor itself is badly optimized. Contrariwise, the grid-like sensor does not qualify well for rotation and scaling while the foveal sensor significantly facilitates the description of stimulus changes for these actions. Again, this matches well with observations in living organisms. Foveating visual systems can be assumed to frequently experience image rotation and scaling because foveal sensors are typically used to engage in object-oriented behavior ranging from prey-catching to in-hand manipulation; activities which typically involve self-induced actions resulting in approximation, adjustment or repositioning of an object. As the vision system compensates for translations of the target object, these behaviors mostly induce rotation and scaling of the object's projection on the observing sensor, which as we have seen are well supported transformations for the foveal sensor layout in the sense that they can be compensated by a computationally inexpensive post-processing step.

Giving consideration to these observations, it seems reasonable to assume that the morphology of a sensor has strong ties to the agent's behavior (and vice versa). This, however, also means that body and movements of a system striving for efficient functioning must be designed with these ties in mind. We believe, the measure introduced in this work can be of help in this respect as it reflects relationships between a given sensor morphology and given actions and it indicates ways of coupling the sensor with appropriate motor actions.

V. Future Work

Having confirmed that the introduced score aligns well with the reasoning introduced in Section II, we will tackle in future work a few shortcomings of the present implementation. Namely, improving the Gini approach such that it has better discriminatory behavior. In extended work, we will investigate a solution for equipping a system with the capability to autonomously select favourable actions for a given sensor by implicitly discovering favourable regions in the given action space without the need for repeated sampling of a single action. To enable such learning in the sense of an ontogenetic development approach, we will put the focus on on-line learning

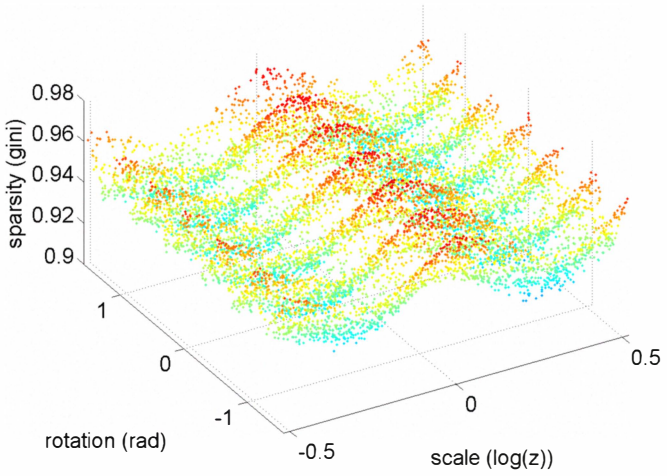


Fig. 5. Actions scores for the foveal layout under rotation and dilation. The zoom factor z denotes a scaling of the sensor topology.

techniques where sensorimotor experiences are sequentially acquired. Eventually, we will investigate in a more formal comparison the here introduced measure with related measures like the Bayesian (BIC) or Akaike (AIC) information criteria. The empirical data collected and presented in this work does not allow for a comparison, as these measures typically require a much bigger number of samples. Furthermore, the BIC / AIC requires the exact number of involved parameters, a quantity which is not directly available from $\mathbf{P}^a$ without setting a threshold.

APPENDIX

Linear Prediction Functions: The argument for linearity of prediction functions is a direct consequence of the world, action and sensor model. First note that the observation function (1) is linear in its argument i_s. If the observations are perfectly predictable, then equation (4) is perfectly satisfied, meaning that each receptor value satisfies

$$\mathbf{D}_{s+1} = p^a\left(\mathbf{D}_s\right) \tag{6}$$

$$\iff \begin{bmatrix} o_1\left(i_{s+1}\right) \\ o_2\left(i_{s+1}\right) \\ \vdots \\ o_N\left(i_{s+1}\right) \end{bmatrix} = p^a\left(\begin{bmatrix} o_1\left(i_s\right) \\ o_2\left(i_s\right) \\ \vdots \\ o_N\left(i_s\right) \end{bmatrix}\right). \tag{7}$$

Since o_i is linear, given any two images i_s^g and i_s^h and any two scale factors α and β the previous satisfies

$$\alpha \begin{bmatrix} o_1(i^g_{s+1}) \\ o_2(i^g_{s+1}) \\ \vdots \\ o_N(i^g_{s+1}) \end{bmatrix} + \beta \begin{bmatrix} o_1(i^h_{s+1}) \\ o_2(i^h_{s+1}) \\ \vdots \\ o_N(i^h_{s+1}) \end{bmatrix} = p^a\left(\alpha \underbrace{\begin{bmatrix} o_1(i^g_s) \\ o_2(i^g_s) \\ \vdots \\ o_N(i^g_s) \end{bmatrix}}_{\mathbf{D}^g_s} + \beta \underbrace{\begin{bmatrix} o_1(i^h_s) \\ o_2(i^h_s) \\ \vdots \\ o_N(i^h_s) \end{bmatrix}}_{\mathbf{D}^h_s}\right)$$

which, when equation (7) is replaced on the left hand side

$$\alpha p^a\left(\mathbf{D}^g_s\right) + \beta p^a\left(\mathbf{D}^h_s\right) = p^a\left(\alpha \mathbf{D}^g_s + \beta \mathbf{D}^h_s\right),$$

proves linearity of p^a whenever the action is perfectly predictable.

ACKNOWLEDGMENT

This work was supported by the European Commission, Project FP7-ICT-248366 RoboSoM, and by the Portuguese Government – Fundação para a Ciência e Tecnologia (ISR/IST plurianual funding) through the POS_Conhecimento Program that includes FEDER funds.

REFERENCES

[1] J. Dewey, "The reflex arc concept in psychology," *The Psychological Review*, vol. 3, pp. 357–370, 1896.

[2] J. Piaget, "The origins of intelligence in children," vol. 1953.

[3] J. J. Gibson, "The theory of affordances," *In Shaw et al., Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, pp. 67–82, 1977.

[4] J. K. O'Regan and A. Noë, "A sensorimotor account of vision and visual consciousness," *Behavioral and Brain Sciences*, vol. 24, no. 5, pp. 939–1031, 2001.

[5] M. Lungarella and O. Sporns, "Mapping information flow in sensorimotor networks," *PLoS Computational Biology*, vol. 2, pp. 1301–1312, 2006.

[6] L. Olsson, C. L. Nehaniv, and D. Polani, "From unknown sensors and actuators to actions grounded in sensorimotor perceptions," *Connection Science*, vol. 18, 2006.

[7] R. Pfeifer and J. Bongard, *How the Body Shapes the Way We Think*. MIT Press, 2006.

[8] G. Metta and P. Fitzpatrick, "Early integration of vision and manipulation," *Adaptive Behavior*, vol. 11, pp. 109—128, 2003.

[9] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner, "The structure of the nervous system of the nematode caenorhabditis elegans," *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, vol. 314, pp. 1–340, 1986.

[10] T. B. Crapse and M. A. Sommer, "Corollary discharge across the animal kingdom," *Nat. Rev. Neuroscience*, vol. 9, pp. 587–600, 2008.

[11] E. L. Schwartz, "Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding," *Vision Research*, vol. 20, no. 1, pp. 645–669, 1980.

[12] A. Hughes, "A comparison of retinal ganglion cell topography in the plains and tree kangaroo," *Physiology*, vol. 244, pp. 61–63, 1975.

[13] R. Hebel, "Distribution of retinal ganglion cells in five mammalian species (pig, sheep, ox, horse, dog)," *Anat. Embryol.*, vol. 150, pp. 45–51, 1976.

[14] S. M. Clippingdale and R. Wilson, "Self-similar neural networks based on a kohonen learning rule," *Neural Networks*, vol. 9, no. 5, pp. 747–763, 1996.

[15] L. Lichtensteiger and P. Eggenberger, "Evolving the morphology of a compound eye on a robot," in *Proc. 3rd Europ. Workshop on Advanced Mobile Robots*, 1999, pp. 127–134.

[16] D. Lewis and J. Catlett, "Heterogeneous uncertainty sampling for supervised learning," in *In Proc. Int. Conf. on Machine Learning (ICML)*, 1994, pp. 148–156.

[17] A. Pouget and L. Snyder, "Computational approaches to sensorimotor transformations," *Nature Neuroscience*, vol. 3, pp. 1192–1198, 2000.

[18] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Royal Statistical Society. Series B*, vol. 58, pp. 267–288, 1996.

[19] N. Hurley and S. Rickard, "Comparing measures of sparsity," *Transactions on Information Theory*, vol. 811, 2008.

[20] E. Clarkson, R. Levi-Setti, and G. Horváth, "The eyes of trilobites: The oldest preserved visual system," *Arthropod Structure & Development*, vol. 35, pp. 247–259, 2006.