Predicting visual stimuli from self-induced actions: an adaptive model of a corollary discharge circuit

Jonas Ruesch, Ricardo Ferreira, Alexandre Bernardino

***Abstract*—Neural circuits which route motor activity to sensory structures play a fundamental role in perception. Their purpose is to aid basic cognitive processes by integrating knowledge about an organism's actions and to predict the perceptual consequences of those actions. This work develops a biologically inspired model of a visual stimulus prediction circuit and proposes a mathematical formulation for a computational implementation. We consider an agent with a visual sensory area consisting of an unknown rigid configuration of light-sensitive receptive fields which move with respect to the environment according to a given number of degrees of freedom. From the agent's perspective, every movement induces a initially unknown change to the recorded stimulus. In line with evidence collected from studies on ontogenetic development and the plasticity of neural circuits, the proposed model adapts its structure with respect to experienced stimuli collected during the execution of a set of exploratory actions. We discuss the tendency of the proposed model to organize such that the prediction function is built using a particularly sparse feedforward network which requires a minimum amount of wiring and computational operations. We also observe a dualism between the organization of an intermediate layer of the network and the concept of self-similarity as introduced by [1].**

***Index Terms*—visual stimulus prediction, reafference, corollary discharge, plasticity, sparse neural networks, self-similarity.**

I. Introduction

The ability to learn and recognize causal relationships between motor actions and sensory feedback is fundamental to autonomous systems. In biological organisms, this mapping is done by neural circuits which are continuously trained and refined while the system interacts with its environment. Human infants for example learn to predict sensory feedback during playful interaction from the earliest period of their life on [2]. The skills acquired during this phase of sensorimotor learning allow them later to perceive the world in a spatially and temporally coherent manner where sensory experiences are interlaced over sequences of actions. As adults, we know from our own experience that predicted sensory feedback usually integrates with effectively experienced stimuli in such a seamless manner that we are hardly aware of its ubiquitous presence. For example, when we grasp an object, our senses report a vast amount of visual and tactile feedback which we – well accustomed to our body and environment – merely use to acknowledge anticipated sensations. Rather disturbing on the other hand, is the situation where expected feedback is inaccurately predicted: anybody climbing a stair in the dark expecting a nonexistent step is acquainted with that experience [3].

J. Ruesch, R. Ferreira and A. Bernardino are with the Institute for Systems and Robotics, Instituto Superior Técnico, 1049-001 Lisbon, Portugal. Mail: {jruesch,ricardo,alex}@isr.ist.utl.pt

In Neuroscience, studies carried out with a great number of different species showed that there is strong evidence to believe that neural pathways which route motor commands to sensory structures – also called corollary discharge circuits – are fundamental to nervous systems of all levels of complexity [4], [5]. In particular, the prediction of sensory stimuli plays an important role in early sensorimotor processing and is active much before we consciously execute a movement to grasp an object or climb a stair. In visual perception for example, a neural circuit discussed further in Sect. II, actively predicts changes in visual stimuli when we move our eyes.

In this work we develop an adaptive model of a corollary discharge circuit which learns to predict visual stimuli based on self-initiated displacement actions. The model consists of a layer of corollary discharge neurons (CDNs) integrating input from a visual motor area projecting to a region which processes visual sensory signals. A learning process is proposed, capable of minimizing the prediction error based on the adaptation of the spatial layout of corollary discharge neurons. Observing the resulting configurations of this model, we discuss that the introduced layer of corollary discharge neurons tends to cover the motor space according to the given (but unknown) topology of the recording visual sensor. We note that such an organization leads to a particularly simple prediction model able to predict a visual stimulus with less computational operations and physical connections.

A. Related Work

In robotics and artificial intelligence, the learning of sensorimotor relationships traditionally focuses on tuning a system's parameters such that sensory input can be translated into a motor action appropriate for a task at hand [6]. This typically involves learning a generally nonlinear coordinate transformation from a sensor related reference frame to motor space [7]. In this sense, work in visual sensorimotor learning often concentrates on the learning of oculomotor actions required to center a target stimulus on the visual sensor, or on how to translate visual signals into coordinated eye-hand movements for reaching [8], [9]. A recent review on sensorimotor learning focusing more in depth on mappings in the opposite direction – i.e. forward models like the one proposed in this work – can be found in [10].

In developmental robotics, prediction of sensory feedback has been previously addressed for example in work described in [11]. There, a general concept referred to as *expected perception* is advocated. Related, in [12] an implementation of predictive visuomotor coordination for visual tracking of a swinging pendulum is presented. By tuning the parameters of

a predefined motion model, a binocular visual system learns to follow and predict the sinusoidal movement of a pendulum.

Directly modeled after the same biological prototype we consider in this work, Quaia et al. propose in [13] a model describing a remapping process which can keep track of the location of saccade targets in eye-centered coordinates. The model is inspired by results obtained from recordings in the SC-MD-FEF pathway reviewed in Sect. II-B and aims at explaining observations made during a double saccade (a.k.a. double-step task in the literature).

In a broader context, [14] analyses the causal structure present in the information flow induced by sensorimotor activity using information theoretic measures. Their results essentially confirm that the characteristics of the generated signals have strong ties to spatiotemporal relationships defined by the physical embodiment and the movement strategies executed by the system under consideration. Based on this insight, it is concluded that a tight coupling between physical structure behavior and neural information processing is essential for embodied systems.

Work on explicitly deducing the topology of unknown visual sensor layouts has been described in [15]. There, an unknown sensoritopic map of a visual sensor is reconstructed using an entropy maximization method relying on information distance measures between sensor elements. Furthermore, in a second part of the work, a method to deduce sensorimotor laws is proposed. The approach relies on a second learning stage where the sensor topology learned in the first step is known and optical flow computed according to the Lucas-Kanade algorithm is used.

The authors of this article investigated in previous work the structure of linear stimulus prediction models for visual sensors composed of an unknown distribution of light sensitive receptors. We found that the pairing of a particular sensor topology and sensor actuation strategy has a profound impact on the complexity of the prediction operator [16]. This means, the prediction of visual stimuli can become computationally less complex when combining a sensor topology with a suitable action space or vice versa.

B. Contribution

We formulate an adaptive computational model of a visual stimulus prediction circuit primarily inspired by a well studied corollary discharge pathway found in primates. The model is developed from a representation which has a direct translation to a physical implementation of a signal transmitting circuit. The versatility of the model is demonstrated by training it to correctly predict future visual stimuli from visuomotor commands by minimizing the prediction error for a set of explorative motor actions and stimulus samples. We show that the model is able to learn a continuous prediction function covering the given motor space, and that the proposed optimization discovers configurations which are particularly easy to construct in a physical realization while at the same time minimizing the prediction error and computational operations.

Referring to the three key topics in sensorimotor learning as discussed in [10], this paper proposes a forward model (what is learned?), obtained by an error-based learning strategy (how is it learned?), stored as the spatial layout of a number of corollary discharge neurons and the strength of their feedforward connections (how are sensorimotor relationships represented?).

Despite similar inspirations, our model differs in important ways from the implementations presented e.g. in [11] and [12]. This is due to mainly two reasons. First, we believe it is in general desirable to keep knowledge required to train the agent or robot to a minimum. Secondly, we think it is absolutely crucial for a truly adaptive system that it does not require any externally designed model to solve a task at hand. This constraint is violated if for example a priori knowledge about the motion model of an observed object is assumed. Thus, different from the approach described in [12], the work presented in this paper attempts to implement the prediction of sensory stimuli with means *intrinsic* to the agent, i.e. by exploiting the structure of the agent's own sensorimotor system. Consequently, learning to predict sensory feedback involves the adaptation of the agent's internal structure and not merely the tuning of parameters of a model which is unrelated to the agent's embodiment. Compared to the model described in [13], our model aims at predicting the complete sensory feedback and does not focus on tracking a single target stimulus. Furthermore, we focus on learning the spatial organization of the prediction circuit which is not considered in [13]. On the other hand, we do not address the dynamics of an action, but consider a one-step ballistic movement which is elicited by the selection of a particular location in a given action space. This means, we consider actions to be signals denoting a change in motor position with respect to the current state. The biological analogy to this setup are motor spaces spanned by neural motor layers using population coding to select particular actions according to the sum of activated motor neurons. Such motor layers exist for example in the superior colliculus (SC) controlling visual saccades and body orientation as reviewed in Sect. II-B. We also note, we do not intend to present a model able to explain specific empirical data recorded in neuroscientific experiments; but we propose a general model – i.e. one based on a very limited set of assumptions - which is able to acquire the functionality provided by the SC-MD-FEF pathway as described in recent neuroscientific studies reviewed in Sect. II-B, and which is adaptive in the sense of neural plasticity as reviewed in Sect. II-C. In particular, this means that the proposed model adapts its stimulus prediction circuit such that a non-linear function is implemented capable of predicting visual stimuli for any action in the given action space. This function is optimized by observing visual stimuli experienced before and after executing actions defined through a set of randomly chosen action signals covering the given action space in the sense of an agent exploring its sensorimotor mapping by executing available motor actions and observing the resulting stimulus after taking the action.

Following ideas presented in [14], the adaptation process of the proposed model discovers and exploits spatiotemporal causalities induced by sensorimotor activities in order to shape the organization of the predictive circuit according to the given sensor layout and action space. In this way, the mapping of sensor and motor spaces is achieved in a topologically

coherent manner. As we will see, this means the given but completely unknown topology of the visual area influences the organization of the corollary discharge layer on the motor side. Such an adaptation procedure is partly related to [15], where the previously unknown topology of a visual sensor is deduced. Though, our methodology differs from [15] insofar as in our approach the learning of sensorimotor laws and the discovery of spatial relationships take place conjointly in one step, guided solely by the minimization of the prediction error.

The structure of this paper is organized as follows. To give an overview of forward models and their importance in living organisms, we first review in the next section a number of examples of stimulus prediction in nature. In particular, we discuss in Sect. II-B the corollary discharge circuit which we consider the closest biological prototype to the model proposed in this work. This circuit, found in primates, is believed to be responsible for the prediction of visual stimuli based on visuomotor commands. It leads from the motor layers of the superior colliculus (SC) to the frontal eye field (FEF, a visual sensory area in the frontal cortex). In Sect. II-C, a brief review of neural plasticity and ontogenetic adaptivity in the superior colliculus (SC) – the motor side of the previously discussed circuit – is presented. Plasticity in the SC is of major relevance to this work, as its motor layers represent the biological equivalent to the motor space from which corollary discharge neurons as proposed in this work integrate input. Based on the reviewed material, we deduce a graphical interpretation of the discussed circuit. Subsequently in Sect. III, we formulate a mathematical model of the introduced network architecture. Analogously to the biological prototype, this formulation links the action space of the considered neural motor layer (in the SC) to a visual sensory area (the FEF) and does not include any motor plant present between the neural motor layers and peripheral actuators, e.g. the pathway between SC and the oculomotor plant.

The proposed model features two important properties: 1) it is based on a reduced set of initial assumptions which allows us to consider given but unknown sensor topologies and action spaces, and 2) it is able to learn, thus, the circuit can adapt to different sensor topologies and sensor movement strategies. In Sect. V, we demonstrate the learning process for two different sensor topologies and action spaces. With a regular grid-like layout we address stimulus prediction for sensor configurations commonly found in artificial image sensors, e.g. CCD sensors available off-the-shelf. A non-uniform fovea-inspired layout serves to investigate the organization of the modeled circuit for sensors as found in animals relying on binocular vision. The obtained results are discussed in Sect. VI. We observe two interesting and interrelated properties: the circuit connecting motor and sensor areas converges to a particularly sparse configuration in terms of number of connections; and the geometry of the circuit's connection nodes shows strong self-similarity properties.

II. Stimulus Prediction in Living Organisms

Research targeted at corollary discharge circuits became especially popular during the last century when the *reafference principle* was proposed in [17].[1] The proposed concept provided an explanation for why sensory stimuli caused by self-initiated movements *(reafference)* can be distinguished from external signals *(exafference)*. Von Holst and Mittelstaedt suggested that a copy of a motor command, the *efference copy (EC)*, is used to distinguish the reafferent part of the signal from the exafferent part. At the same time, similar conclusions led Sperry coin the term *corollary discharge* (CD) [18]. Nowadays, despite the conceptual similarity between EC and CD, the commonly used terminology makes a slight difference between the two. CD is in general used to refer to signals which are transmitted along feedforward connections from the motor pathway to the sensory processing stream. These circuits can connect from any tier of the motor pathway to any other tier in the sensory processing stream. EC on the other hand is considered to be a motor signal affecting sensory channels close to the effector / sensor periphery. Fig. 1 illustrates the different levels of feedforward connections. For further reading on these circuits including this definition and an attempt to classify them, see [4], [5].

A. General Mechanisms of Stimulus Prediction

On a first level, prediction of sensory feedback is often used by an organism to distinguish between external signals (with origin in the environment) and signals induced by self-initiated actions. Male crickets for example filter their self-produced bursts of sound by generating a neural signal which anticipates the auditory stimulus [19], [20]. By doing so, the animal is able to suppress its own chirping while focusing on the response of female crickets. The same filtering strategy has been discovered in a number of other species which elicit escape reactions depending on whether a sensory signal is self-generated or results from an event in the environment [21], [22].

On a higher level, stimulus prediction is believed to be a basic mechanism used to achieve stable perception. For any organism, the acquisition of a coherent percept of the environment is not a passive one-step action but is the result of a continuous process of sensorimotor interactions which take place over a number of iterations. Rats for example explore objects by tactile whisking; bats "see" the world by listening to the echo of their self-produced ultrasonic waves, and many animals relying on vision continuously have to move their eyes and body to sweep their visual field over an observed scene [23], [24]. The motor actions involved in these exploration movements often induce drastic changes in sensory stimuli which we can understand when trying to focus for example on the image stream recorded by an abruptly moved camera. The question arises, how is the brain able to assemble a stable and coherent percept in light of such radical stimulus changes? Neuroscientists suggest that the nervous system takes advantage of knowing how motor actions affect sensory stimuli in order to relate sensory input recorded during action sequences. We will resume on this topic in particular for visual perception in the next section.

[1] In neuroscientific terms, *afference* comprises all the sensory signals coming from the periphery of the central nervous system. These signals are composed of reafferent signals which are self-induced stimuli and exafferent signals caused by changes in the environment.

Last but not least, stimulus prediction is important for fast action sequences from a dynamical point of view. If a signal providing sensory feedback for a previous action reaches the motor system only *after* the next action has to be initiated, then prediction is indispensable to plan accurate motor commands. This is for example the case for a sequence of fast eye movements (saccades). Studies with primates showed that the brain actually relies on a predictive control strategy: motor commands for subsequent saccades are issued before proprioceptive or visual sensory feedback from the previous movement is available to the motor system [25]–[29]. Furthermore, considering that primates execute up to three visual saccades per second during normal behavior, and knowing that neurons in the frontal and parietal visual areas of the primate brain record afferent signals with a latency of at least 60 ms, relying on purely passive afferent signals would mean that the cortical representation of the visual world is inaccurate during almost 20% of the time [13], [30], [31]. Such delay and accuracy is likely to be compensated by neural pathways like the one discussed in the next section.

B. The SC-MD-FEF Pathway

An important and well studied neural circuit for stimulus prediction in higher vertebrates is the SC-MD-FEF pathway in the primate cortex leading from the superior colliculus (SC) via the medial dorsal nucleus (MD) to the frontal eye field (FEF). This circuit is responsible for visual stimulus adaptation during eye saccades and serves as the biological prototype for our model.

At the beginning of the SC-MD-FEF pathway lies the superior colliculus, a phylogenetically ancient structure in the vertebrate midbrain known to be responsible for triggering eye and head movements [32]. The outer layers of the SC receive multi-modal somatosensory input, while the deeper layers of the SC are concerned with attention orienting motor movements. Optic input layers are retinotopically organized, meaning the topological map of the recording sensor (the retina) is still present in the SC. This topology is preserved through several layers and as far as down to the deeper motor areas. As a result, neural activity in the motor layers of the SC code eye saccades in a gaze-related retinotopic reference frame. This layout of movement fields in the SC was first revealed in studies done by [33]. Using microstimulation the spatial layout of the motor map was deduced in terms of relative gaze orientation angles. In [34] further evidence was collected which supports a population coding theory for the motor layers in the SC which remains valid until today. According to this theory, a blob-like activation of a number of neurons in the motor layer triggers a saccade to a target location which is encoded as the sum of all active motor neurons where each neuron acts like a spatial target vector weighted according to its activation.[2] The motor signals generated by this activation do not only command

[2]We note, to eventually move the eyes, retinotopic motor signals have to undergo a non-trivial transformation while travelling from the SC to the oculomotor nuclei [35]. However, in this work, we consider the remaining motor pathway as given.

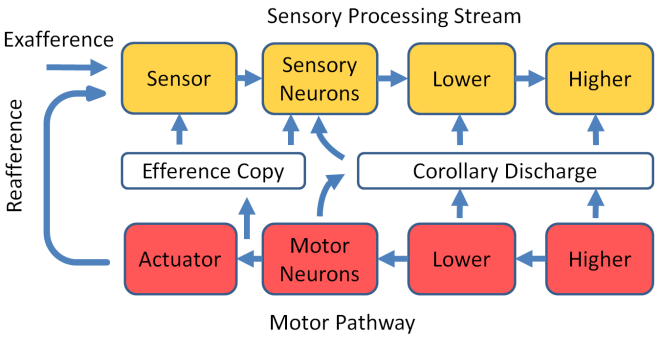


Fig. 1. Efference copy and corollary discharge circuits along the sensory processing stream and motor pathway. Adapted from [4].

eye movements via the motor pathway, they also travel in the opposite direction ascending the SC-MD-FEF pathway through which they eventually reach the frontal eye field, see Fig. 2.

In the frontal eye field, the corollary discharge signals from the SC are integrated with visual signals which reach the FEF through the main sensory processing stream. Typically, the receptive field (RF) of a stimulus processing neuron is spatially fixed with respect to the underlying input neurons. But, neuroscientific studies showed that in several areas in the visual system, there are neurons which feature so called *shifting* receptive fields. In the FEF these neurons modify their RFs influenced by the CD signal arriving from the SC-MD-FEF pathway. That is, when a saccade is executed, the RFs of these neurons are modulated to integrate visual signals from the target location of the saccade. A shifted RF is then called *future field (FF)*. Consequently, the presaccadic FF and the postsaccadic RF sample the same absolute location in visual space. Comparison of presaccadic and postsaccadic FEF neuron activation can therefore in principle be used for both, stabilization purposes, and to distinguish exafferent from reafferent stimuli (filtering). This view is held by a number of authors: [36]–[38] hypothesize that neurons with shifting RFs are able to perform comparative operations. The SC-MD-FEF pathway in general has been extensively studied by [25], [38]–[42]. Other areas in the visual system where neurons with shifting RFs have been found include the lateral intraparietal sulcus (LIP) [36], [43], [44] and extrastriate visual areas like V4 [45], [46].

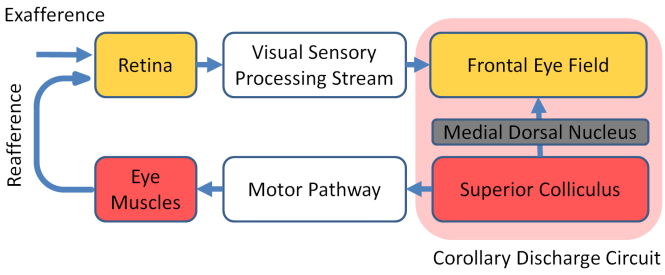


Fig. 2. The SC-MD-FEF corollary discharge circuit connecting the Superior Colliculus via the Medial Dorsal Nucleus with the Frontal Eye Field. Note the inverted (feedforward) direction with respect to the motor pathway and sensory processing stream.

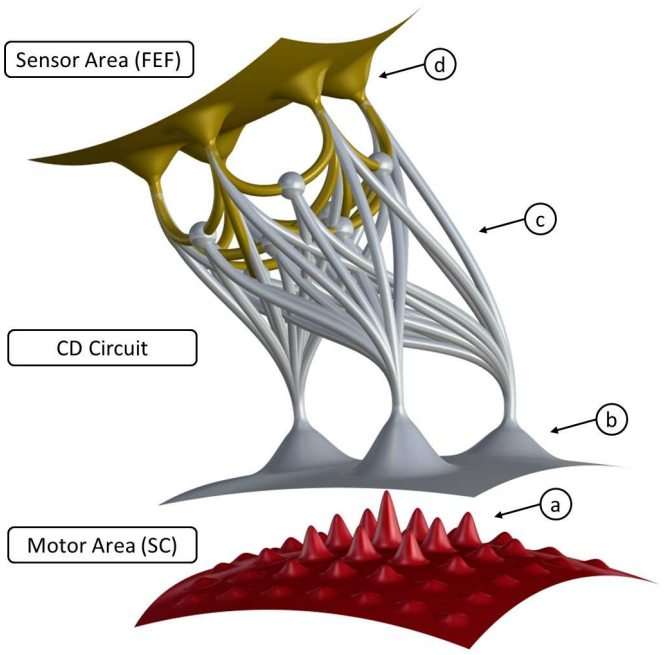


Fig. 3. Model of a visual corollary discharge circuit. A population of motor neurons in the superior colliculus (SC) codes visual saccades in a retinotopic reference frame (a). An intermediate layer of corollary discharge neurons (b) collects activation from the underlying motor layer and projects through feedforward connections (c) to the frontal eye field (FEF), a visual area (d). The corollary discharge signals modulate the activation of visual receptive fields and their connections such as to predict a future visual stimulus resulting from an activation in (a).

In Fig. 3, we introduce a graphical interpretation of the topological and functional relationships reviewed here. Signals travelling along the SC-MD-FEF pathway originate from a peak of activation in a layer of SC motor neurons coding eye movements in a retinotopic reference frame which is denoted (a) in Fig. 3. Along the corollary discharge pathway, this activation is integrated by corollary discharge neurons (CDNs, (b) in Fig. 3). The CDNs project in turn through feedforward connections (c), to visual neurons and their connections (d). We will resume on this interpretation when proposing the computational model of the SC-MD-FEF circuit in the next section.

C. Plasticity

In addition to highly dynamical prediction mechanisms, the visual cortex is long known for its plasticity during early life. The influential work described in [47], [48] showed that ontogenetic development of the visual areas V1–V5 involves a high degree of adaptivity. Based on single-unit recordings, the formation of orientation sensitive cells in striate and extrastriate cortices could be observed, clearly influenced by experienced stimuli [49]. Since then, a great body of work confirmed that plasticity is present along the entire visual sensory processing stream, see e.g. [50]. Moreover, plasticity also plays a major role in the superior colliculus. In [51] it is for example described how visual and auditory maps in the SC are topographically aligned during early life; and recent work reported in [52] studied alterations of the visual map in the SC due to lacking neuronal activity in the early retina. It was found that without previous visual stimuli, the projection layers in the SC are a coarse retinotopic map given by morphogenetic development. Subsequently, during growth, the spatially correlated firing of retinal ganglion cells refines the organization of the retinotopic layers in the SC. This organization is defective if natural neural activity due to lacking or disturbed visual input is not present after birth. Hence, like in the striate and extrastriate cortices, the fine tuning of the SC clearly depends on sensorimotor contingencies experienced while the animal interacts with its environment. This form of reafference exploration is a commonly found learning strategy in nature [53]. An equivalent methodology termed *motor babbling* can be applied in artificial embodied systems.

As described later, the CD circuit proposed in Fig. 3 incorporates plasticity by implementing a flexible network topology: corollary discharge neurons and their receptive fields (b) adapt in shape and position with respect to the underlying SC motor layer while feedforward connections (c) can alter their amount of discharge.

III. Computational Model

After compiling the graphical model shown in Fig. 3, we unfold our interpretation in this section into a mathematical model.

A. Observation and Action Model

We consider an agent with an arbitrary number of degrees of freedom and a given rigid visual sensor consisting of N spatially distributed visual receptive fields. The receptive fields are located on a surface which reflects a projection of the environment given as a function $i_s : \mathbb{R}^2 \to \mathbb{R}$ defining a luminance value for each point on the surface when the agent is in a certain state s. Given such a $i_s(x)$, we model visual receptive fields by a function f_n 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 receptive field and the receptive field function f_n is modeled as a multivariate Gaussian on $i_s(x)$. We choose Gaussians as a 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 [7].

After observing state s, the agent can perform an action which can change i_s. Here we assume that an action a is linear and changes i_s in a predictable way to $a(i_s)$. Appendix A describes the constraints posed on a and how it relates to a physical agent acting in a 3-dimensional world. Hence, after taking action a the agent observes

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

For convenience we define $i_{s+1} = a(i_s)$. In what follows, we use the notation $\mathbf{o}_s$ and $\mathbf{o}_{s+1}$ to denote the receptive field

activation values before and after applying action a:

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

In some occasions we write $\mathbf{a}$ instead of a when referring to a particular location in the action space which induces a transformation a as shown in Fig. 4. This notation directly relates to a peak of activation in the motor layer as shown in Fig. 3.

B. Prediction Model

The general form of stimulus prediction for the agent described refers to functions $p^a : \mathbb{R}^N \to \mathbb{R}^N$ which, when applied to the initially observed sensor values $\mathbf{o}_s$ are able to approximate the sensor values obtained after applying action a:

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

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 Appendix B, we provide an argument that motivates a reduction of these functions to the linear function set. The argument relies on the assumption that the actions executed by the agent lead to perfectly predictable changes of $i(x)$ on the sensor surface. Taking into account linear prediction functions, we can rewrite Eq. 4 as

$$\mathbf{o}_{s+1} \approx \mathbf{P}(\mathbf{a})\,\mathbf{o}_s. \tag{5}$$

In summary, we make the following assumptions:

- Predictability: The agent only executes actions inducing predictable transformations a on $i(x)$, i.e. of the form $i_{s+1} = a(i_s)$. See appendix A for a full description.
- Linearity: We model prediction of observations $\mathbf{o}_{s+1}$ with a linear prediction function p^a. This constrains the action transformation a to be linear. As the example at the end of appendix A shows, this is not as restrictive as it appears.

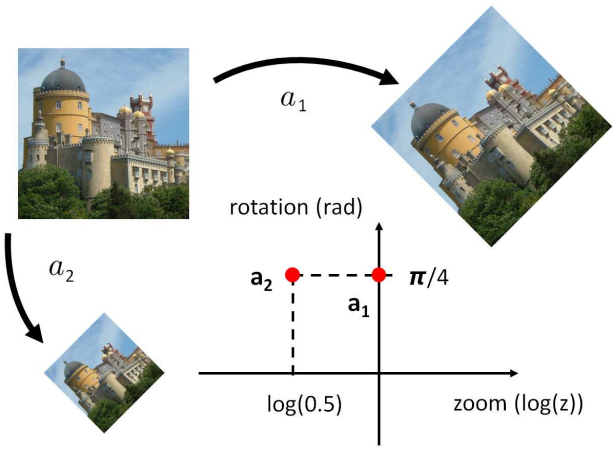


Fig. 4. Relation of locations in the action space $\mathbf{a}$ and transformation functions a: the selection of locations $\mathbf{a}_1$ or $\mathbf{a}_2$ induces the transformation functions a_1 or a_2. The action space corresponds to the motor area (a) in Fig. 3 where a location $\mathbf{a}$ is coded by a peak of activation.

- Gaussian Receptive Fields: We model receptive field functions with multivariate Gaussians.

Under these assumptions, we can now write the prediction model following directly Fig. 3. We interpret feedforward connections (gray) as manipulators which control the discharge rate of receptive field connections (yellow). Grouping the connections of a single corollary discharge neuron (CDN) together, we can write the feedforward connection weights of the j-th CDN as a matrix $\mathbf{P}_j$, where an entry (q, r) specifies how much the observation $o_q(i_s)$ of receptive field q contributes to the predicted observation $o_r(i_{s+1})$ of receptive field r. Combining contributions of different corollary discharge neurons according to the CDN layer (b) of Fig. 3, we compose the eventually predicted sensor stimulus as a linear combination

$$\mathbf{o}_{s+1} = \left[\sum_j \lambda_j(\mathbf{a})\mathbf{P}_j\right]\mathbf{o}_s, \tag{6}$$

where $\lambda_j(\mathbf{a})$ denotes the activation of a particular CDN depending on the action $\mathbf{a}$ coded by the underlying motor neurons. Hence, matrices $\mathbf{P}_j$ and their activation functions λ_j together define the prediction function $\mathbf{P}(\mathbf{a})$ defined over the entire action space. Because the function λ_j models the receptive field of a CDN j on the underlying population of motor neurons, we follow our previous assumption and implement each λ_j as a multivariate Gaussian like

$$\lambda_j(\mathbf{a}) = e^{-\frac{1}{2}(\mathbf{a}-\boldsymbol{\mu}_j)^\top \boldsymbol{\Sigma}_j^{-1}(\mathbf{a}-\boldsymbol{\mu}_j)}, \tag{7}$$

where $\boldsymbol{\Sigma}_j$ is the covariance matrix and $\boldsymbol{\mu}_j$ is the location of the receptive field of CDN j. Measurements motivating a Gaussian model for movement fields in the SC-MD-FEF pathway are presented in [39]. Additionally to the argument for biological plausibility given in this section, we will present in Sect. V results which reveal the actual required shape for this receptive field (see also Fig. 10).

In conclusion, the free parameters of our model are $\mathbf{P}_j$, $\boldsymbol{\Sigma}_j$, and $\boldsymbol{\mu}_j$. These parameters define the plasticity of the modeled corollary discharge circuit. While $\mathbf{P}_j$ directly defines a prediction operator based on feedforward connection weights, $\boldsymbol{\Sigma}_j$ and $\boldsymbol{\mu}_j$ code for topological plasticity in the CDN layer allowing for changes in position and shape of each field with respect to the underlying motor area. Note, a natural constraint for the entries of $\mathbf{P}_j$ is to require them to be greater or equal to zero. Negative values would not make sense in the described scenario. Also note, given the equivalence of the biologically inspired interpretation of the corollary discharge circuit as described in the previous section and the mathematical model proposed in this section, we use in the reminder of this article the terms “feedforward connection activation” and “prediction matrix entry” interchangeably as they have the same meaning to us.

C. Plasticity

To learn the free parameters of our model, the agent executes a number of actions to experience and cover a given action space in the sense of reafference exploration or motor

babbling described in II-C. During this exploration phase, triplets $(\mathbf{a}^a, \mathbf{o}_s^a, \mathbf{o}_{s+1}^a)$ consisting of before and after stimuli for a given action are sampled. Given such stimulus samples, we choose the adaptation process of the model to follow a minimization on the prediction error like

$$\begin{aligned} &\left(\mathbf{P}_j^*, \mathbf{\Sigma}_j^*, \boldsymbol{\mu}_j^*\right) = \\ \text{argmin} \quad &\sum_a \left\| \sum_j \lambda_j\left(\mathbf{a}^a\right) \mathbf{P}_j \ \mathbf{o}_s^a - \mathbf{o}_{s+1}^a \right\|^2 \\ \text{s.t.} \quad &\mathbf{P}_j \geq \mathbf{0} \end{aligned} \tag{8}$$

This optimization problem can be addressed by a number of methods. For example, to find a (locally optimal) solution, different gradient descent methods are applicable and readily available in both batch and online versions. In a batch version, the optimization is solved given a number of collected $(\mathbf{a}^a, \mathbf{o}_s^a, \mathbf{o}_{s+1}^a)$ triplets, while in the online case, sample pairs are sequentially becoming available as experienced. A particular approach to solve Eq. 8 is presented in the next section.

IV. Method

To find $\mathbf{P}_j^*$, $\mathbf{\Sigma}_j^*$ and $\boldsymbol{\mu}_j^*$, we first collect a batch of triplets $(\mathbf{a}^a, \mathbf{o}_s^a, \mathbf{o}_{s+1}^a)$ sampled from the considered motor space and a given environment. Each triplet is acquired as follows: First, the agent is placed at a randomly chosen position at which it records stimulus $\mathbf{o}_s^a$. Then, a random action $\mathbf{a}^a$ is chosen and the agent is displaced in a discrete step to a new position at which it records the stimulus $\mathbf{o}_{s+1}^a$. The number of triplets used to train a setup of the model as presented in Sect. V is 5000. We then solve Eq. 8 with the Levenberg-Marquardt algorithm which we found to have nice convergence properties while being relatively simple to implement, see for example [54]. Based on empirical evidence presented in Sect. V-D, we reduce $\mathbf{\Sigma}_j$ to a diagonal matrix thereby constraining the receptive field functions of corollary discharge neurons λ_j to be axis-aligned Gaussians. The constraint $\mathbf{P}_j \geq \mathbf{0}$ was implemented by adding an exponential penalty function to the optimization. While it is no problem to find a solution for $\mathbf{P}_j^*$, $\mathbf{\Sigma}_j^*$ and $\boldsymbol{\mu}_j^*$ with an online method, convergence is much slower, we therefore choose here the batch approach for practical reasons. However, we note that under different circumstances an online implementation might be preferable, e.g. for a purely biologically inspired implementation in a robot with stronger memory constraints and a longer exploration phase.

The experiments presented in Sect. V were initialized as follows: The locations of CDN receptive fields $\boldsymbol{\mu}_j$ were set according to a uniform random distribution. The sizes $\mathbf{\Sigma}_j$ of CDN receptive fields were set to a fixed value. The prediction matrices (feedforward connections) were initialized to zero. It is important to note that with a randomized initialization, nothing prevents the adaptation process from converging to a locally optimal solution. However, from a biological perspective the initialization of the CDN layer corresponds to a topology generated by ontogenetic development (compare Sect. II-C). Hence, we can in fact expect a coarse structure to be present before ontogentic adaptation starts. We conjecture here that an initialization provided by morphogenetic development might be relatively close to the final solution.

In experiments which include actions leading to a dilation of the stimulus, the action is encoded like $\zeta = \log(z)$ where z can be seen as a zoom factor while ζ can be interpreted as the distance of the visual sensor to the observed scene. For example, in a setup as introduced in the next section and illustrated in Fig. 5, an action ζ means moving the sensor along the vertical direction changing the distance to the observed image i_0. With this choice, we obtain a situation-independent, and at the same time energetically plausible representation of an action dilating the stimulus. Situation independence is achieved in the sense that an action composition like $+\zeta_a - \zeta_a$ leaves the sensor stimulus $\mathbf{o}_0$ invariant. This means the agent's current state does not influence the effect of an action ζ_a. Energetically plausible refers to the fact that in a physical setup, ζ might directly relate to voltage or current applied to an actuator moving the agent for example towards or away from an observed scene. Thus, encoding dilation as $\zeta = \log(z)$ appears reasonable, as moving away $(-\zeta_a)$ or towards $(+\zeta_a)$ a scene requires the same amount of energy which would not be reflected by the zoom factor z.

V. Results

We will now consider a specific instance of the model introduced above. An agent with four degrees of freedom and a given sensor topology observes a 2-dimensional environment given as a grayscale image $i_0(x)$. The agent can modify the current observation $\mathbf{o}_s$ by executing actions from a 4-dimensional action space spanned by translations (x- and y-direction), rotations, and changes in distance to i_0 (dilation). The setup and the four available degrees of freedom are illustrated in Fig. 5. The sensor can be seen as moving over i_0 which in Eq. 2 is expressed as moving from i_s to i_{s+1}.

Note, the oculomotor system of primates and other animals with binocular vision does not directly implement actions leading to stimulus rotation and dilation. However, confirming the versatility of the proposed model, we are going to present in this section an experiment which also covers self-induced stimulus rotation and dilation. Such actions are in particular relevant for the development of artificial systems as the related sensorimotor interaction patterns do play an important role e.g. for visual processing during locomotion or for visual perception during object manipulation.

In what follows, we explore two specific sensor layouts as shown in Fig. 6. The foveal layout shown in Fig. 6(b) was generated according to a logarithmic spiral. Retinotopic layouts found in living organisms with binocular vision follow closely such a density distribution up to a small area in the very center which deviates from this law. In [55] an approximation of this deviation is formulated, however we do not consider this area. We investigate each layout under a subspace of the full 4-dimensional action space. For the grid layout, we choose to analyze the adaptation process of the proposed model under translational actions (subsequently called *grid/translation* setup). The foveal layout is used to train the model for rotation and dilation actions (subsequently called *fovea/rotation-dilation* setup). A more in depth discussion on the choice of this pairing of sensor topologies and action spaces is given in Sect. VI.

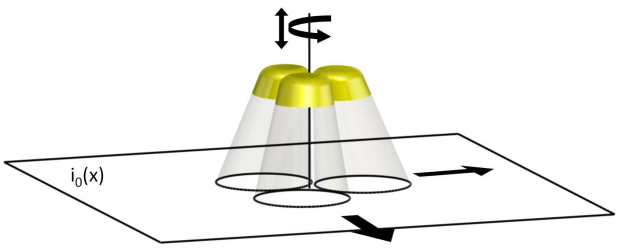
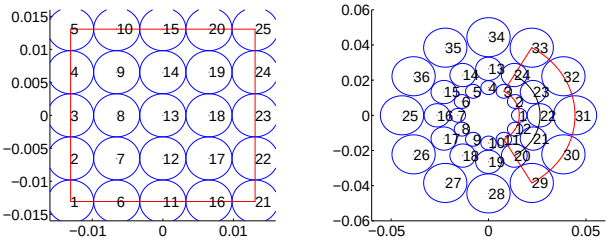


Fig. 5. Model instantiation considered in Sect. V. The agent, a rigid sensor with a given configuration of a number of receptive fields (yellow), observes a 2-dimensional world $i_0(x)$. The action model is implemented as sensor translation actions in x- and y-directions, rotation around the axis perpendicular to i_0, and changes in distance to i_0.

A. Visualization

The results for the investigated topologies and action spaces are summarized in Fig. 7 and Fig. 8. On the left of Fig. 7 and Fig. 8 the motor layer of the respective setup is shown. For the *grid/translation* setup shown in Fig. 7, the motor layer covers the shown range of actions inducing horizontal and vertical stimulus shifts. For the *fovea/rotation-dilation* setup shown in Fig. 8 the motor layer covers the actions leading to stimulus rotations and dilations in a range as denoted. Each point on the motor layer represents an action relative to the sensor's original position. Thus, in Fig. 7 each point on the motor layer corresponds to a motor signal leading to a shift in horizontal or vertical direction of the visual stimulus. In Fig. 8 each point on the motor layer corresponds to a motor signal triggering an action inducing a rotation or dilation of the visual stimulus. Comparing with the graphical model introduced in Sect. II-B, the selection of a particular location in the motor layers shown on the left in Fig. 7 and Fig. 8 is analogous to determining the resulting motor signal for a population coded motor layer by summing active motor neurons as illustrated in red in Fig. 3 layer (a); hence a point in the shown motor layers is the analogon to a peak in activation in the red layer in Fig. 3. Points in light gray depict the 5000 random actions taken while training the model. Black ellipses depict the location and variance of the multivariate Gaussian used to model the corollary discharge neurons (CDNs) on the



(a) Grid Layout (b) Log-spiral Layout

Fig. 6. The two sensor topologies considered for the presented results. Circles represent the standard deviation of the Gaussian receptive fields. Axis units refer to the size of the 2-dimensional world on which the sensor moves and which extends over a range of $[-1, +1]$ in x- and y-direction. Outlined areas denote action sampling ranges. a) Uniform grid-like 5×5 layout with 25 receptive fields. b) Foveal layout parametrized according to a growth spiral with 12 branches each with 3 receptive fields following $\rho = 0.0063e^{0.33\phi}$.

motor layer as described in Sect. III and by Eq. 7. Comparing with Fig. 3 these ellipses represent the gray layer (b). The center of each ellipse is determined by the corresponding $\boldsymbol{\mu}_j$ of the depicted CDN, the size and shape of each ellipse is drawn according to the corresponding $\boldsymbol{\Sigma}_j$. In both figures, Fig 7 and Fig. 8, the sub-figures (a) illustrate the initial configuration from which the optimization was started: CDN locations $\boldsymbol{\mu}_j$ were initialized randomly and all $\boldsymbol{\Sigma}_j$ were set to a default value. Sub-figures (b) illustrate CDNs in their converged configuration as $\boldsymbol{\mu}_j^*$ and $\boldsymbol{\Sigma}_j^*$, optimized according to Eq. 8. Note, we have to be aware of boundary effects when inspecting the results of the motor layer organization. This is due to the fact that we inevitably have to rely on a finite range for action sampling. Thus, we expect to observe some disturbances for corollary discharge neurons located at the sampling border. Sub-figures on the right side of Fig. 7 and Fig. 8 show the activation of feedforward connections in matrix format for nine selected CDNs. Comparing with the graphical model of the proposed corollary discharge circuit as shown in Fig. 3, each depicted matrix describes the activation of feedforward connections linking one CDN in layer (b) to the visual area (d). The activation of all connections coded by a matrix $\mathbf{P}_j$ is initially set to zero (black). In sub-figures (b) of Fig 7 and Fig. 8, matrices $\mathbf{P}_j$ are plotted after one iteration of maximizing Eq. 8. Sub-figures (d) show the same matrices but now in converged configuration $\mathbf{P}_j^*$. As described in Sect. III, an activation matrix $\mathbf{P}_j$ is in mathematical terms a linear prediction operator which can be used to predict a future stimulus $\mathbf{o}_{s+1}$, given the current stimulus $\mathbf{o}_s$ and an action $\mathbf{a}$ as shown in Eq. 6. Hence, the 9 prediction operators shown in Fig. 7 (d) and Fig. 8 (d) are the prediction operators valid for the 9 central nodes and their influence areas learned as $\boldsymbol{\Sigma}_j^*$ and $\boldsymbol{\mu}_j^*$ in Fig. 7 (c) and Fig. 8 (c). Their values are shown grayscale coded.

In summary, sub-figures (c) and (d) in Fig 7 and Fig. 8 illustrate the visual stimulus prediction function as learned by the proposed model. Sub-figure (d) depicts the activation of feedforward connections or the prediction operator $\mathbf{P}_j^*$ learned for each CDN. Sub-figure (c) illustrate the topological organization and influence area of each CDN as defined by $\boldsymbol{\mu}_j^*$ and $\boldsymbol{\Sigma}_j^*$. The result is a smooth, non-linear prediction function constructed as the combination of linear prediction operators with overlapping influence areas.

How this model and the learned function can be used to predict visual stimuli for any action contained in the covered motor space is illustrated in Fig. 9. In Fig. 9 (a) the converged configuration for the *grid/translation* setup is shown and a randomly selected motor action – not contained in the set of actions used to train the model – is marked bold red. Indices in this sub-figure enumerate corollary discharge neurons. According to the previously learned CDN topology shown in Fig. 9 (a), CDN 24 and CDN 21 are the corollary discharge neurons most activated by the chosen action. Their activations are given by Eq. 7 according to their receptive fields yielding $\lambda_{24} = 0.71$ and $\lambda_{21} = 0.24$ for the chosen

action.[3] Fig. 9 (b) shows the prediction operators $\mathbf{P}_{24}$ and $\mathbf{P}_{21}$. They are linearly combined according to Eq. 6 using λ_{24} and λ_{21} to obtain the prediction operator optimal for the chosen action. In the upper half of Fig. 9 (c) an example input stimulus is shown. Circles correspond to visual receptive fields as shown in layer (d) in Fig. 3. The grayscale color-code of each circle filling represents the recorded input stimulus for each visual receptive field. At the bottom of Fig. 9 (c) the predicted stimulus is shown as obtained from the given input stimulus and the prediction operator assembled according to the triggered action. Note, the action selected in sub-figure (a) corresponds to a shift of the visual area by one visual receptive field distance to the right and a bit more than half a visual receptive field distance upwards. Therefore, the predicted visual stimulus is shifted one visual receptive field distance to the left and approximately half a visual receptive field distance down. The downward shift results in a blur due to the resolution of the visual system. Also note, indices in sub-figures (b) and (c) are related and denote visual receptive fields.

B. Results Grid Layout

When training the model using the grid layout under translational actions, we find that the motor layer converges to a configuration where corollary discharge neurons are distributed in the action space on a regular grid. This is visible in Fig. 8(a), in particular for non-boundary CDNs. Notably, their locations in action space coincide with the spacing of receptive fields in the sensor layout. We note that the prediction matrix of the CDN with index 14 converged to zero and has no impact on the final cost function. From our privileged perspective, we can see that CDN 14 should have been placed somewhere between CDN 16 and 4 to improve the present solution. In the shown case, the algorithm converged to a local optimal solution where CDN 14 has no contribution and the area between CDN 16 and 4 is covered by slightly more outstretched neighbor CDN receptive fields. This slightly increases the prediction error in this neighborhood of the action space but has no severe impact on the prediction ability.

In Fig. 8(b), the learned prediction matrices of nine selected CDNs are shown. For a graphical interpretation of the shown prediction matrices, read an entry q in row r of a prediction matrix for action a as the activation of the receptive field connection between receptive field q and r. For example, CDN number 15 located in the action space at $(0, 0)$, not surprisingly shows a single diagonal of non-zero entries. Other $\mathbf{P}_j$ show non-zero entries with according offsets. For example CDN number 25 covers actions where the values of the first column of visual receptive fields (indices 1 to 5) contribute to predict future values of the second column (indices 6 to 10 in Fig. 6(a)). Interestingly, the entries for the unpredictable receptive fields 1 to 5 converged to be non-zero in the diagonal. This is due to the fact that for natural images with low spatial frequency and small sensor translation distances, the future activation of a receptive field with unpredictable input is best described by values of previously close receptive fields. Despite the fact that we expect $\mathbf{P}_j$ to be sparse in general due to spatiotemporal relationships between visual receptive fields, we find the converged configuration to feature prediction matrices with an exceptionally small number of non-zero entries. We will get back to this observation in more detail in Sect. VI.

To address questions regarding the global optimal solution, we ran a number of optimizations where each run started from a different randomly initialized configuration. Measuring for each converged configuration the total prediction error over all sampled actions, we are able to confirm with a high degree of certainty that the globally optimal configuration is the one where all 25 CDNs are arranged on a regular 5×5 grid. We were unable to find another configuration with a smaller overall prediction error.

C. Results Foveal Layout

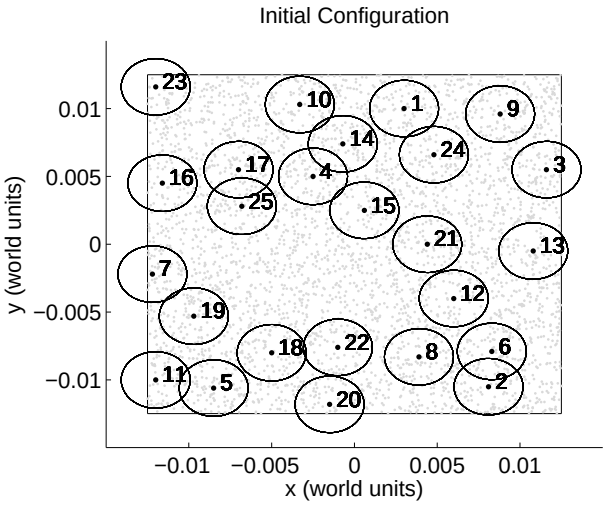
When training the model using the foveal layout with actions leading to stimulus rotation and dilation, we find that the CDN layer converges to a configuration where CDN receptive fields are regularly distributed on concentric circles. This configuration can be seen in Fig. 8(c), where the location and size of CDNs is plotted with respect to ζ as described in Sect. IV (x-axis), and the angle of rotation (y-axis). Note, as the vertical axis in Fig. 8 (a) and (c) denotes rotation, vertical lines in these plots describe circular arcs, and vertically aligned CDNs lie on concentric circles. Thus, as for the grid layout, the organization of CDN receptive fields in the action space happens to reflect the spatial layout of the visual receptive fields in the sensor area. In Fig. 8(d) the learned prediction matrices of nine selected CDNs are shown. As for the grid layout, visual receptive fields which cannot be predicted accurately happen to be approximated by their own previous value (diagonal entries). And again, we observe that all $\mathbf{P}_j$ are exceptionally sparse.

Unlike for the grid-like setup, the results presented for the foveal setup represent what we suspect to be the global optimal solution. This assumption is supported again by the fact that no other solution found led to a smaller overall prediction error. We therefore have strong reasons to believe that the positioning of the CDNs as shown in Fig. 8(c) corresponds to the globally optimal one.

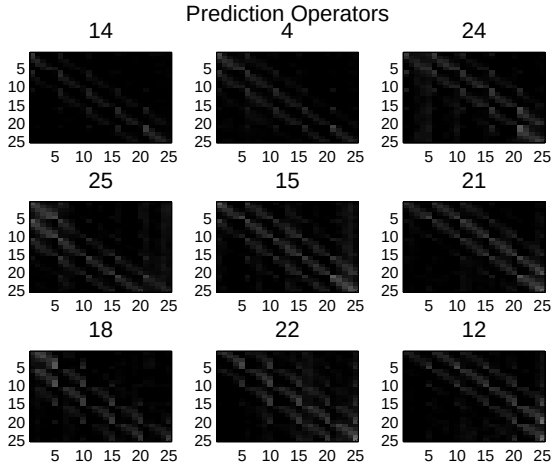
D. Result Validation

From the above presented results it is not directly visible why the organization of the motor layer converges to the described configurations. To get a notion of the driving force behind the organizing process, it is useful to first inspect the underlying function $\mathbf{P}(\mathbf{a})$. This is difficult as $\mathbf{P}(\mathbf{a})$ defines for every action a matrix $\mathbf{P}^a$ of size $N \times N$. Nonetheless, to get an impression of what the trained model is actually approximating, we visualize a particular entry of this matrix for a number of random actions. In Fig. 10(a) we plot the selected entry using prediction matrices $\mathbf{P}^a$ learned by linear regression from multiple samples for each action. For a comparison, Fig. 10(b) shows the same matrix entry obtained from our

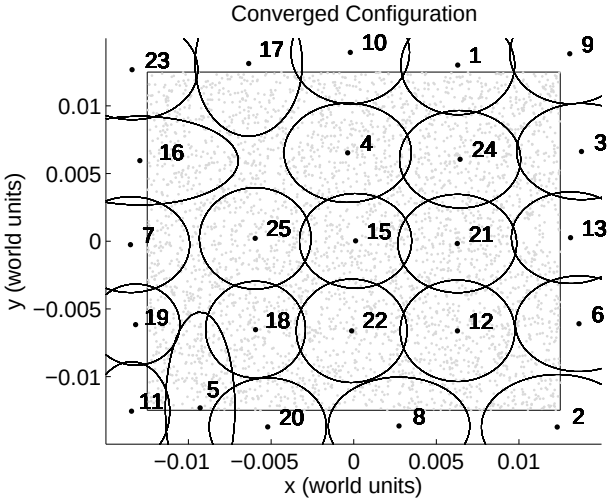
[3] In favour of a comprehensive illustration, we discard in this example the contribution of all other less activated CDNs.



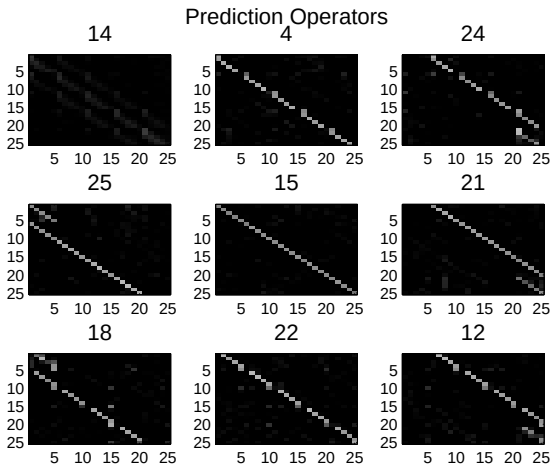
(a) Initial configuration of the receptive fields of each corollary discharge neuron on the motor space (iteration 0).



(b) Prediction matrices after one iteration of the Levenberg–Marquardt optimization algorithm. Each matrix was initialized to zero at iteration 0 (black).



(c) Final configuration of the receptive fields of each corollary discharge neuron (CDN) on the motor space (iteration 1000). Note, CDN 14 was suppressed during the optimization, see also Sect V-B.



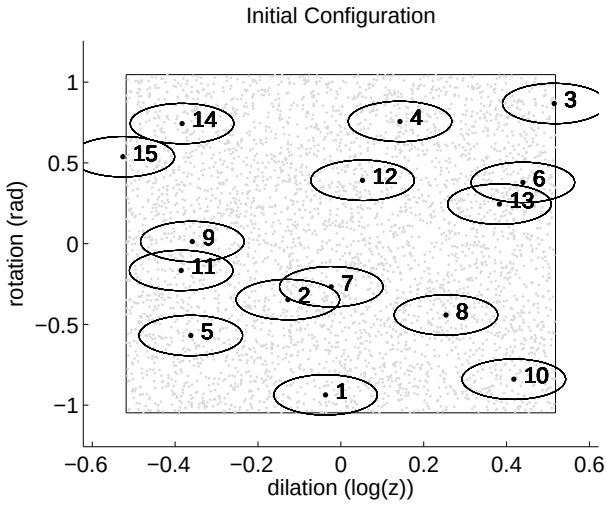
(d) Prediction matrices after 1000 iterations of the Levenberg–Marquardt optimization algorithm. The prediction matrices associated to the corollary discharge neurons covering the center area of the motor space are shown.

Fig. 7. Optimization of the proposed model for a visual sensor area receiving input according to a retinotopic mapping as shown in Fig. 6(a) and a motor space covering translational actions. Left: Representation of the visual motor space for translations where each point represents a shift relative to the sensor's original position. In grey, the sampled displacements used to train the model (5000). In black, the receptive fields (visualized as ellipses) of each corollary discharge neuron ($\mathbf{\Sigma}, \boldsymbol{\mu}$). Right: Feedforward connection weights of nine corollary discharge neurons displayed in the format of prediction matrices $\mathbf{P}$ as described in Sect. III. Each matrix $\mathbf{P}_j$ is shown as a table of size 25×25 with matrix entries color-coded in grayscale (black $= 0$, white $= 1$). Note, without any specific assumptions, the receptive fields of corollary discharge neurons converged to locations in the motor space which correspond to translational actions which match exact shifts of visual receptive fields, and which allow for a particularly small number of feedforward connections (i.e. particularly sparse prediction operators $\mathbf{P}_j$).

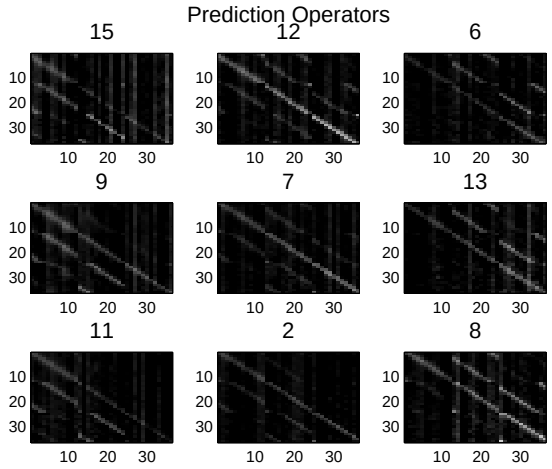
model $\sum_j \lambda_j(\mathbf{a})\mathbf{P}_j$ using the parameters learned in Sect. V-C. Comparing the two plots, two things become apparent: first of all, the values plotted in Fig. 10(a) resemble closely a multivariate Gaussian and are therefore well approximated by the linear interpolation shown in Fig. 10(b); secondly, even though we sampled a selected prediction matrix entry for the non-uniform sensor layout and the rotation-dilation action space, the resulting distribution resembles an axis-aligned Gaussian. The second observation justifies the previous decision to restrict $\mathbf{\Sigma}_j$ to be diagonal matrices.

VI. Discussion

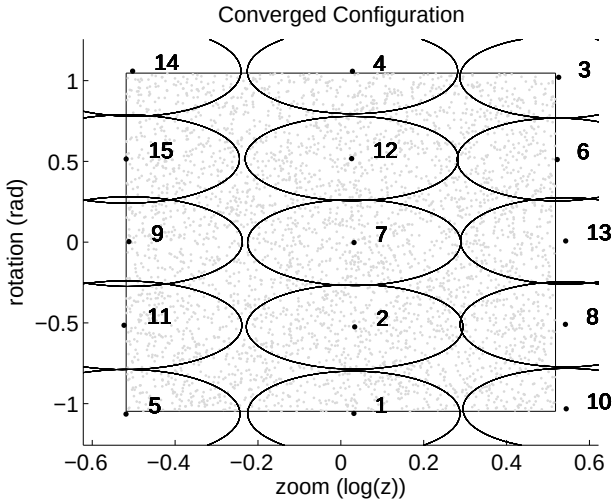
We have proposed an adaptive model of a visual corollary discharge circuit (CDC) following state of the art functional understanding of a particular visual CDC in the primate brain. The mathematical formulation of the proposed model is directly deduced from a biologically inspired representation, where we translated neurons and neural feedforward connections to a weighted linear combination of prediction matrices. In particular, we were interested in modelling the capability of a CDC to adapt during ontogenetic development in order to optimize the prediction of visual stimuli for previously unknown sensor topologies and movement behaviors. Inspired



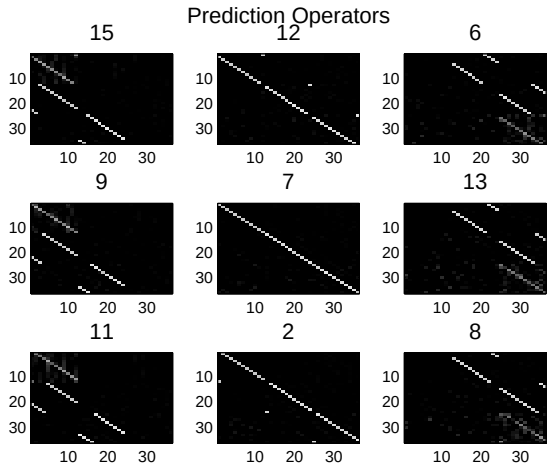
(a) Initial configuration of the receptive fields of each corollary discharge neuron on the motor space (iteration 0).



(b) Prediction matrices after one iteration of the Levenberg–Marquardt optimization algorithm. Each matrix was initialized to zero at iteration 0 (black).



(c) Final configuration of the receptive fields of each corollary discharge neuron on the motor space (iteration 600).

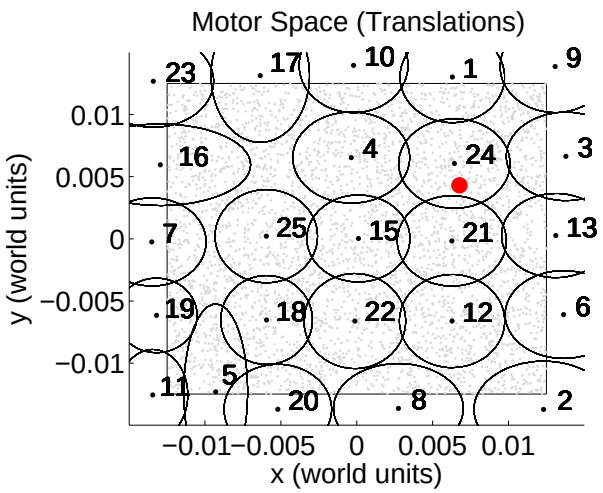


(d) Prediction matrices after 600 iterations of the Levenberg–Marquardt optimization algorithm. The prediction matrices associated to the corollary discharge neurons covering the center area of the motor space are shown.

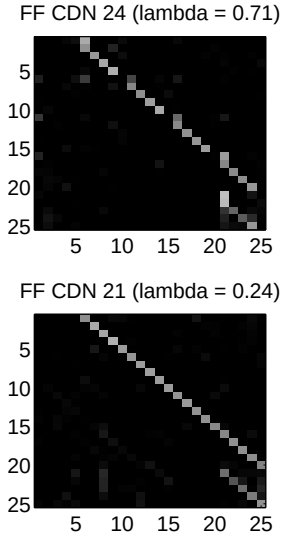
Fig. 8. Optimization of the proposed model for a visual sensor area receiving input according to a retinotopic mapping as shown in Fig. 6(b) and a motor space covering rotational actions and dilation. Left: Representation of the visual motor space for rotations and dilations where each point represents a rotation and change in distance z relative to the sensor's original position. In grey, the sampled displacements used to train the model (5000). In black, the receptive fields (visualized as ellipses) of each corollary discharge neuron ($\mathbf{\Sigma}, \boldsymbol{\mu}$). Right: Feedforward connection weights of nine corollary discharge neurons displayed in the format of prediction matrices $\mathbf{P}$ as described in Sect. III. Each matrix $\mathbf{P}_j$ is shown as a table of size 36×36 with matrix entries color-coded in grayscale (black $= 0$, white $= 1$). Note, without any specific assumptions, the receptive fields of corollary discharge neurons converged to locations in the motor space which correspond to rotational actions and dilations which match exact shifts of visual receptive fields, and which allow for a particularly small number of feedforward connections (i.e. particularly sparse prediction operators $\mathbf{P}_j$).

by neuroscientific evidence of neural plasticity, we introduced an adaptive intermediate layer of corollary discharge neurons (CDNs) able to adjust to an unknown sensor layout and movement behavior by changing location and size of their receptive fields with respect to the underlying layer of motor neurons. We presented results which demonstrate the resulting adaptation process which is driven by the minimization of the stimulus prediction error. Notably, we deliberately kept at all time the layout of the sensor and the effects of executed actions as a black box implementation of which we knew nothing about. Considering a physical agent acting in a 3-dimensional world, the proposed model is applicable respecting the constraint described in Appendix A. For a sensor following the typical pinhole camera model, the imposed constraint is fulfilled for example for a spatially fixed camera allowed to rotate along all axes, or for a camera observing a planar scene and moving such that it is always facing the plane. The requirement is violated if a given action in combination with the observed environment leads to motion parallax.

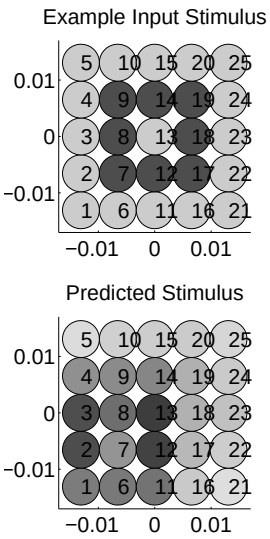
On the basis of the obtained results, we observed that optimal configurations of the proposed model feature corollary discharge neurons with very sparsely activated feedforward connections, or in other words, have particularly sparse prediction matrices associated. Hence, the optimization discovers and takes advantage of locations in the action space where visual stimulus prediction can be done with an especially



(a) Motor space with a particular action marked red. Receptive fields of corollary discharge neurons are shown according to the final configuration shown in Fig. 7.



(b) Feedforward connections of corollary discharge neurons 24 and 21 (in the format of prediction matrices $\mathbf{P}_{24}$ and $\mathbf{P}_{21}$).



(c) Input stimulus (top), and predicted stimulus (bottom), where the latter was obtained using a prediction operator $\mathbf{P}$ according to Eq. 6.

Fig. 9. Example usage of the obtained corollary discharge circuit as presented in Fig. 7: a) The agent selects an action in the available motor space; the action is marked red and corresponds to a shift of the visual area with a horizontal component to the right and a vertical component upwards. Indices enumerate CDNs. b) Corollary discharge neurons (CDNs) are activated according to their receptive fields. For the action shown in red, the two CDNs with highest contribution are 24 and 21 with activations $\lambda_{24} = 0.71$ and $\lambda_{21} = 0.24$ (according to Eq. 7). c) Using feedforward connections $\mathbf{P}_j$ activated according to Eq. 6, the future stimulus can be predicted from a given example stimulus. Note, the selected location in the motor space corresponds to a shift of the visual area by one visual receptive field distance to the right and a bit more than half a visual receptive field distance upwards. Therefore, the predicted visual stimulus is shifted one visual receptive field distance to the left and approximately half a visual receptive field distance down. The downward shift results in a blur due to the resolution of the visual system. Note, indices in sub-figures (b) and (c) are related and denote visual receptive fields. In sub-figure (c), the activation of visual receptive fields is color-coded in grayscale.

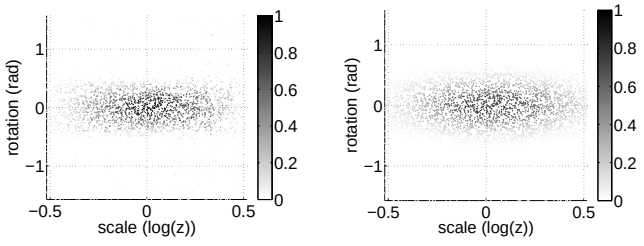
simple prediction model. Even though we expect the activation of feedforward connections to be sparse in general due to spatiotemporal constraints between visual receptive fields, the number of active (non-zero) feedforward connections in the found solutions is sparser than expected. Of course, actions which fall in between such locations still require a more complex prediction operator, however, with the proposed model, appropriate prediction networks are accurately composed for such locations by linearly interpolating the result of several CDNs as described by Eq. 6. Inspecting Fig. 7 and Fig. 8, it also becomes apparent that the organization of CDNs on the motor space follows the discretization of the visual area. This essentially means, the density distribution of receptive fields in the visual area is projected onto the motor space, which, according to the above stated observation, proves to be an efficient mapping for a corollary discharge circuit.

The reason for the observed tendency towards solutions with a sparser feedforward network can be understood starting from observations made in Sect. V-D. As shown in Fig. 10(a), the activation of a single entry of a linear prediction matrix plotted over the action space resembles closely a multivariate Gaussian. Therefore, to best approximate such an activation function, a CDN with a Gaussian receptive field has to be located at the center of this distribution adapting its receptive field size according to the given shape of the activation area. Configurations for which this can be best achieved for all receptive field connections of the CDN's prediction matrix approximate best the actually required activation of receptive field connections and are at the same time particularly sparse. However, the existence of actions which allow for such configurations is defined by a particular sensor/actions pairing (SAP). In [16], we investigated the relationship between sensor topology, sensor movements and linear stimulus predictors by presenting a measure which – given a particular sensor layout and behavior – qualifies locations in the action space according to the complexity of the prediction model required at that location. The application of this measure to the grid and foveal sensor layout under translation, rotation and dilation showed that, while the SAP grid/translation and fovea/rotation-dilation define a clear set of actions for which particularly simple prediction models exist, the remaining combinations, grid/rotation-dilation and fovea/translation, do not define a similarly clear defined set of actions suitable for simple prediction models. In the present work, we presented results for SAPs which are known to define a set of actions with particularly sparse prediction matrices. We observed that this choice facilitates a clear and unique organization of the CDN layer. For other pairings, the adaptive organization of the introduced model finds just as well a solution, but optimal configurations are not as well defined as for the chosen pairings. This means, for the grid/rotation-dilation and the fovea/translation pairings we could not identify solutions which we reproducibly and conclusively encountered to be

globally optimal configurations.

In conclusion, particularly sparse CDCs have a number of advantages. First of all, considering a physical implementation, the feedforward network of a sparse CDC can be built with a small number of neural connections. With respect to the visualization of the model in Fig. 3 this means, a majority of the feedforward connections (c) can be completely removed. Furthermore – in a computational sense – the generation of the predicted stimulus is facilitated because a smaller number of operations is required due to the reduced number of connections.

Apart from direct physiological and computational implications, a second characteristic of the obtained solutions was observed. We find that the tendency of the CDN layer to organize as a sparsely activated feedforward network has parallels with a concept termed *self-similarity* introduced by [1]. They showed in an inventive work the following: a set of points randomly distributed on a disk converges to a stable configuration with a highly regular structure if, i) the points are conjointly transformed by rotations, dilations and translations applied according to a given probability distribution; and ii) when after each iteration each point is moved towards transformed points lying closest. Interestingly, under the introduced rules, certain action probability distributions induce point distributions which resemble closely receptive field distributions as found in foveal sensor layouts. The action probability distributions which lead to such configurations are composed of rotation and dilation actions uniformly distributed over an arbitrary range, combined with translation actions distributed over a *limited* range. For a visualization, the reader might refer to Fig. 11 in [1]. With respect to our work, the organization of such a set of visual receptors – represented as simple points – can be seen in a duality with the adaptation of the CDN layer. Receptors distributed on a disk are in our situation spatially extended receptive fields of corollary discharge neurons integrating input from the motor space. Self-similarity, measured by Clippingdale and Wilson as the average distance of transformed receptors to closest previous receptors, is expressed in our case by the sparsity of prediction matrices $\mathbf{P}_j$ specifying how many previous receptors influence a transformed receptor. The following listing summarizes



(a) Sampled (b) Interpolated

Fig. 10. Activation of the receptive field connection $(18, 18)$ in the foveal layout plotted over the rotation-dilation action space. The left plot shows $\mathbf{P}_{(18,18)}$, where for each location $\mathbf{P}$ was explicitly computed by linear regression. The right plot shows $\mathbf{P}_{(18,18)}$ approximated by $\sum_j \lambda_j(\mathbf{a})\mathbf{P}_{j(18,18)}$ with parameters learned as shown in Fig. 8(c) and 8(d).Other entries in $\mathbf{P}$ than $(18, 18)$ show similar activation distributions centered at different locations.

correspondences for the two scenarios:

Clippingdale	Our work
receptors	CDNs
receptor distances	non-zero $\mathbf{P}$ entries
actions are fixed	topology of visual area fixed
finds opt. receptor conf.	finds optimal actions

According to these relations, the duality between both situations can be sketched as follows: while Clippingdale and Wilson consider a given set of actions and find a receptor configuration which has optimal self-similarity in terms of point distances, we consider a given sensor layout and find transformation actions that transform this layout in such a way that the activation of transformed receptors can be predicted using input from a minimum number of previous receptors. Note, in our case we observe this tendency even though we do not explicitly optimize for self-similarity.

In summary, the introduced model of a corollary discharge circuit successfully learns to predict visual stimuli while at the same time optimizes its topological structure as to increase the sparsity of feedforward connections. Hence, the adaptation process implicitly not only optimizes prediction ability of the network but also discovers – if they exist – locations in the given action space which allow for a particularly simple prediction model with the given sensor layout. Applications exploiting this relationship remain to be explored in future work. For example, if we are interested in generating behaviors supporting simplified corollary discharge signals, CDNs with a particularly sparse feedforward network are good candidates to compose corresponding action sequences. Or conversely, in search for a good sensor/actions pairing, a good combination might be found by tracing the presence of corollary discharge neurons with a particularly sparse feedforward network.

Appendix A
Relations between States and Observations

The introduced model of a corollary discharge circuit assumes that actions can be modeled by Eq. 2. Here we revisit this point and provide more detail as to how this model links with real world agents.

When modeling agents one usually considers a state space $\mathcal{S}$ describing not just the agent but the whole world. An agent action is represented by a function $h : \mathcal{S} \to \mathcal{S}$ which changes this world state. In this work, the agent is allowed to observe the world using a visual sensor which works in two steps. First it is assumed that there is a surface onto which light is projected represented as a function $g : \mathcal{S} \to \mathcal{I}$. Here $\mathcal{I}$ is a function space where each element $i : \mathbb{R}^2 \to \mathbb{R}$ is a function returning the projected intensity at each point on the surface. As a second step, this surface has several receptive fields which are able to integrate the projected intensities on a particular area, each producing a stimulus to the agent. In the diagram shown in Fig. 11, the space of these stimulus is called $\mathcal{O}$, and is captured by the observation model presented in the text as Eq. 1.

Considering the establishment of the proposed CD circuit, a constraint is posed on the agent action h, requiring that it induces a transformation $a : \mathcal{I} \to \mathcal{I}$. This means that the

projection on the sensor surface after an action h must be perfectly reconstructable solely from the previous projection. If one constrains the set of actions to have this property, then for all agent purposes an action can be fully described as the function a instead of considering the full agent state model acted on by functions h. Unfortunately this requirement is too strict to satisfy exactly for most general applications. One particular exception arising in biology is the considered case of eye movement actions. In this case the surface onto which the world is projected is a sphere and the eye movement rotates the projection on the sphere.

This work focuses not on representing a but instead in predicting the observed stimulus after an action is taken, solely from the previously observed stimulus. Notice that the existence of a does not guarantee that the observed stimuli are predictable. For this to happen, the action must be such that the integrating receptive fields line up before and after the action is taken, corresponding to a permutation of the observed stimulus in line with the concept of self-similarity introduced by [1]. We emphasize that it is this interrelation which we explore in this work and which allows for a particularly sparse feedforward network. The fact that the organization of such a network is implicitly dependent on a well concerted pairing of sensor topology and action space is consistent with observations made for living organisms. Animals typically feature a highly specific pairing of behavior and sensor structure which is favorable from this point of view, see also [16].

The model presented in Sect. III-A also imposes that a be linear for technical reasons. Note that since a is actually an operator acting from a function space to a function space, this is not as limiting as it might seem at first glance. Consider the example $a(i)(x) = a_s(x)i(a_p(x))$ where $a_s(x) : \mathbb{R}^2 \to \mathbb{R}$ and $a_p(x) : \mathbb{R}^2 \to \mathbb{R}^2$ are any linear or nonlinear functions. The corresponding operator a is linear as can be quickly checked for all x:

$$\begin{aligned}(\alpha i_1 + \beta i_2)(x) &= a_s(x)(\alpha i_1 + \beta i_2)(a_p(x)) \\ &= \alpha a_s(x)\, i_1(a_p(x)) + \beta a_s(x)\, i_2(a_p(x)) \\ &= \alpha a(i_1)(x) + \beta a(i_2)(x).\end{aligned}$$

Appendix B
Proof of Linearity

Considering actions which lead to perfectly predictable changes in observations, the argument for linearity is a direct consequence of the world, action and sensor model. First note that the observation function (1) is linear in i. Then, with perfectly predictable observations, equation (4) is perfectly satisfied, meaning that each receptive field value satisfies

$$\begin{aligned}&\mathbf{o}_{s+1} = p^a(\mathbf{o}_s) \\ \iff &\begin{bmatrix} o_1(a(i)) \\ o_2(a(i)) \\ \vdots \\ o_N(a(i)) \end{bmatrix} = p^a\left(\begin{bmatrix} o_1(i) \\ o_2(i) \\ \vdots \\ o_N(i) \end{bmatrix}\right).\end{aligned} \tag{9}$$

Since a and o_i are linear, given any two images i_1 and i_2 and any two scale factors, α and β, the previous satisfies

$$\begin{bmatrix} o_1(a(\alpha i_1 + \beta i_2)) \\ o_2(a(\alpha i_1 + \beta i_2)) \\ \vdots \\ o_N(a(\alpha i_1 + \beta i_2)) \end{bmatrix} = p^a\left(\begin{bmatrix} o_1(\alpha i_1 + \beta i_2) \\ o_2(\alpha i_1 + \beta i_2) \\ \vdots \\ o_N(\alpha i_1 + \beta i_2) \end{bmatrix}\right)$$

$$\alpha \begin{bmatrix} o_1(a(i_1)) \\ o_2(a(i_1)) \\ \vdots \\ o_N(a(i_1)) \end{bmatrix} + \beta \begin{bmatrix} o_1(a(i_2)) \\ o_2(a(i_2)) \\ \vdots \\ o_N(a(i_2)) \end{bmatrix} = p^a\left(\alpha \begin{bmatrix} o_1(i_1) \\ o_2(i_1) \\ \vdots \\ o_N(i_1) \end{bmatrix} + \beta \begin{bmatrix} o_1(i_2) \\ o_2(i_2) \\ \vdots \\ o_N(i_2) \end{bmatrix}\right)$$

which, when equation (9) is replaced on the left hand side

$$\alpha p^a\left(\underbrace{\begin{bmatrix} o_1(i_1) \\ o_2(i_1) \\ \vdots \\ o_N(i_1) \end{bmatrix}}_{\mathbf{x}}\right) + \beta p^a\left(\underbrace{\begin{bmatrix} o_1(i_2) \\ o_2(i_2) \\ \vdots \\ o_N(i_2) \end{bmatrix}}_{\mathbf{y}}\right) = p^a\left(\alpha \underbrace{\begin{bmatrix} o_1(i_1) \\ o_2(i_1) \\ \vdots \\ o_N(i_1) \end{bmatrix}}_{\mathbf{x}} + \beta \underbrace{\begin{bmatrix} o_1(i_2) \\ o_2(i_2) \\ \vdots \\ o_N(i_2) \end{bmatrix}}_{\mathbf{y}}\right)$$

$$\alpha p^a(\mathbf{x}) + \beta p^a(\mathbf{y}) = p^a\left(\alpha \mathbf{x} + \beta \mathbf{y}\right),$$

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

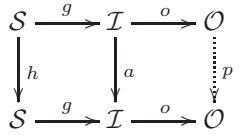


Fig. 11. Relationships and constraints between the state space $\mathcal{S}$, the sensor surface $\mathcal{I}$, and the space of sensor stimuli $\mathcal{O}$. Functions h, a and p denote action transitions. Functions g and o describe how $\mathcal{I}$ is generated from $\mathcal{S}$ and sampled in $\mathcal{O}$.

Acknowledgment

This work was supported by the European Commission Project FP7-ICT-248366 RoboSoM, by the Portuguese Government – Fundação para a Ciência e Tecnologia (FCT) project PEst-OE/EEI/LA0009/2011, project DCCAL PTDC/EEA-CRO/105413/2008, and FCT grant SFRH/BD/44649/2008.

References

[1] S. Clippingdale and R. Wilson, "Self-similar neural networks based on a kohonen learning rule," *Neural Netw*, vol. 9, no. 5, pp. 747–763, 1996.
[2] J. Elman, E. Bates, A. Karmiloff-Smith, M. Johnson, D. Parisi, and K. Plunkett, *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press, 1996.
[3] D. M. Wolpert and J. R. Flanagan, "Forward models," in *Oxford Companion to Consciousness*, T. Bayne, A. Cleeremans, and P. Wilken, Eds. Oxford University Press, 2009, pp. 294–296.
[4] T. B. Crapse and M. A. Sommer, "Corollary discharge across the animal kingdom," *Nat. Rev. Neuroscience*, vol. 9, pp. 587–600, 2008.
[5] ——, "Corollary discharge circuits in the primate brain," *Curr. Opin. Neurobiology*, vol. 18, pp. 552–557, 2008.
[6] L. L. E. Massone, "Sensorimotor learning," in *The Handbook of Brain Theory and Neural Networks*, M. A. Arbib, Ed. The MIT Press, 1995, pp. 860–864.
[7] A. Pouget and L. Snyder, "Computational approaches to sensorimotor transformations," *Nature Neuroscience*, vol. 3, pp. 1192–1198, 2000.
[8] B. Girard and A. Berthoz, "From brainstem to cortex: Computational models of saccade generation circuitry," *Prog. Neurobiol.*, vol. 77, pp. 215–251, 2005.
[9] J. D. Crawford, W. P. Medendorp, and J. J. Marotta, "Spatial transformations for eye-hand coordination," *J. Neurophysiology*, vol. 92, pp. 10–19, 2004.
[10] D. M. Wolpert, J. Diedrichsen, and J. R. Flanagan, "Principles of sensorimotor learning," *Nature Reviews Neuroscience*, vol. 12, pp. 739–751, 2011.

[11] C. Laschi, G. Asuni, E. Guglielmelli, G. Teti, R. Johansson, H. Konosu, Z. Wasik, M. C. Carrozza, and P. Dario, "A bio-inspired predictive sensory-motor coordination scheme for robot reaching and preshaping," *Auton. Robots*, vol. 25, pp. 85–101, 2008.

[12] E. Falotico, M. Taiana, D. Zambrano, A. Bernardino, J. A. Santos-Victor, P. Dario, and C. Laschi, "Predictive tracking across occlusions in the icub robot," in *Proc. of the 9th Int. Conf. on Humanoid Robots*, 2009, pp. 486 –491.

[13] C. Quaia, L. M. Optican, and M. E. Goldberg, "The maintenance of spatial accuracy by the perisaccadic remapping of visual receptive fields," *Neural Networks*, vol. 11, pp. 1229–1240, 1998.

[14] M. Lungarella and O. Sporns, "Information self-structuring: key principle for learning and development," in *Proc. of the 4th Int. Conf. on Development and Learning*. IEEE, 2005, pp. 25–30.

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

[16] J. Ruesch, R. Ferreira, and A. Bernardino, "A measure of good motor actions for active visual perception," in *Proceedings of the 10th IEEE International Conference on Development and Learning (ICDL-2011), to appear*, 2011.

[17] E. von Holst and H. Mittelstaedt, "Das reafferenzprinzip (wechselwirkung zwischen zentralnervensystem und peripherie)," *Naturwissenschaften*, vol. 37, pp. 464–476, 1950.

[18] R. Sperry, "Neural basis of the spontaneous optokinetic response produced by visual inversion," *J Comp Physiol Psychol*, vol. 43, pp. 482–489, 1950.

[19] B. Webb, "Neural mechanisms for prediction: Do insects have forward models?" *Trends in Neuroscience*, vol. 27, pp. 278–282, 2004.

[20] J. F. Poulet and B. Hedwig, "A corollary discharge mechanism modulates central auditory processing in singing crickets," *J. Neurophysiology*, vol. 89, pp. 1528–1540, 2003.

[21] F. Delcomyn, "Corollary discharge to cockroach giant interneurons," *Nature*, vol. 269, pp. 160–162, 1977.

[22] D. Edwards, W. Heitler, and F. Krasne, "Fifty years of a command neuron: the neurobiology of escape behavior in the crayfish," *Trens Neurosci.*, vol. 22, pp. 153–161, 1999.

[23] E. Ahissar and D. Kleinfeld, "Closed-loop neuronal computations: focus on vibrissa somatosensation in rat," *Cerebral Cortex*, vol. 13, pp. 53–62, 2003.

[24] C. F. Moss and S. R. Sinha, "Neurobiology of echolocation in bats," *Curr. Opin. Neurobiol.*, vol. 13, pp. 751–758, 2003.

[25] M. M. Umeno and M. E. Goldberg, "Spatial processing in the monkey frontal eye field. i. predicitive visual responses," *J. Neurophysiology*, vol. 78, pp. 1373–1383, 1997.

[26] R. C. Miall and D. M. Wolpert, "Forward models for physiological motor control," *Neural Networks*, vol. 9, pp. 1265–1279, 1996.

[27] S. Vaziri, J. Diedrichsen, and R. Shadmehr, "Why does the brain predict sensory consequences of oculomotor commands? optimal integration of the predicted and the actual sensory feedback," *J. Neuroscience*, vol. 26, pp. 4188–4197, 2006.

[28] J. F. Poulet and B. Hedwig, "New insights into corollary discharges mediated by identified neural pathways," *Trends in Neuroscience*, vol. 30, pp. 14–21, 2007.

[29] C. C. Bell, V. Han, and N. B. Sawtell, "Cerebellum-like structures and their implications for cerebellar function," *Ann. Rev. Neuroscience*, vol. 31, pp. 1–24, 2008.

[30] M. C. Bushnell, M. E. Goldberg, and D. L. Robinson, "Behavioral enhancement of visual responses in monkey cerebral cortex: I. modulation in posterior parietal cortex related to selective visual attention," *J. Neurophysiology*, vol. 46, pp. 755–772, 1981.

[31] M. E. Goldberg and M. C. Bushnell, "Behavioral enhancement of visual responses in monkey cerebral cortex: Ii. modulation in frontal eye fields specifically related to saccades," *J. Neurophysiology*, vol. 46, pp. 773–787, 1981.

[32] W. C. Hall and A. K. Moschovakis, *The Superior Colliculus*. Taylor & Francis Inc., 2003.

[33] D. A. Robinson, "Eye movements evoked by collicular stimulation in the alert monkey," *Vision Research*, vol. 12, pp. 1795–1808, 1972.

[34] C. Lee, W. H. Rohrer, and D. L. Sparks, "Population coding of saccadic eye movements by neurons in the superior colliculus," *Nature*, vol. 332, pp. 357–360, 1988.

[35] A. K. Moschovakis, "The superior colliculus and eye movmement control," *Curr. Opin. Neurobiology*, vol. 6, pp. 811–816, 1996.

[36] J. R. Duhamel, C. L. Colby, and M. E. Goldberg, "The updating of the representation of visual space in parietal cortex by intended eye movements," *Science*, vol. 255, pp. 90–92, 1992.

[37] C. L. Colby and M. E. Goldberg, "Space and attention in parietal cortex," *Ann. Rev. Neuroscience*, vol. 22, pp. 319–349, 1999.

[38] M. A. Sommer and R. H. Wurtz, "Influence of the thalamus on spatial visual processing in frontal cortex," *Nature*, vol. 444, pp. 374–377, 2006.

[39] ——, "What the brain stem tells the frontal cortex. i. oculomotor signals sent from superior colliculus to frontal eye field via mediodorsal thalamus," *J. Neurophysilogy*, vol. 91, pp. 1381–1402, 2004.

[40] ——, "Brain circuits for the internal monitoring of movements," *Ann. Rev. Neuroscience*, vol. 31, pp. 317–338, 2008.

[41] M. M. Umeno and M. E. Goldberg, "Spatial processing in the monkey frontal eye field. ii. memory responses," *J. Neurophysiology*, vol. 86, pp. 2344–2352, 2001.

[42] A. Murthy, S. Ray, S. M. Shorter, E. G. Priddy, J. D. Schall, and K. G. Thompson, "Frontal eye field contributions to rapid corrective saccades," *J. Neurophysiology*, vol. 97, pp. 1457–1469, 2007.

[43] A. P. Batista, C. A. Buneo, L. H. Snyder, and R. A. Andersen, "Reach plans in eye-centered coordinates," *Science*, vol. 285, pp. 257–260, 1999.

[44] C. L. Colby, J. R. Duhamel, and M. E. Goldberg, "Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area," *J. Neurophysiology*, vol. 76, pp. 2841–2852, 1996.

[45] K. Nakamura and C. L. Colby, "Updating of the visual representation in monkey striate and extrastriate cortex during saccades," *Proc. Natl. Acad. Sci. USA*, vol. 99, pp. 4026–4031, 2002.

[46] A. S. Tolias, T. Moore, S. M. Smirnakis, E. J. Tehovnik, A. G. Siapas, and P. H. Schiller, "Eye movements modulate visual receptive fields of v4 neurons," *Neuron*, vol. 29, pp. 757–767, 2001.

[47] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *Physiology*, vol. 160, pp. 106–154, 1962.

[48] ——, "Receptive fields and functional architecture of monkey striate cortex," *Physiology*, vol. 195, pp. 215–243, 1968.

[49] D. H. Hubel, T. N. Wiesel, and S. LeVay, "Plasticity of ocular dominance columns in monkey striate cortex," *Phil. Trans. R. Soc. Lond. B*, vol. 278, pp. 377–409, 1977.

[50] R. Pinaud, L. A. Tremere, and P. Weerd, Eds., *Plasticity in the Visual System: From Genes to Circuits*. Springer US, 2006.

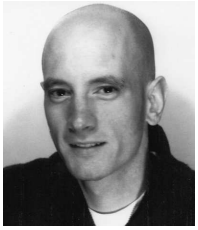
[51] A. J. King, M. E. Hutchings, D. R. Moore, and C. Blakemore, "Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus," *Nature*, vol. 332, pp. 73–76, 1988.

[52] T. D. Mrsic-Flogel, S. B. Hofer, C. Creutzfeldt, I. Cloëz-Tayarani, J. P. Changeux, T. Bonhoeffer, and M. Hbener, "Altered map of visual space in the superior colliculus of mice lacking early retinal waves," *J. Neuroscience*, vol. 25, pp. 6921–=6928, 2005.

[53] J. Schouenborg, "Learning in sensorimotor circuits," *Curr. Opin. Neurobiology*, vol. 14, no. 6, pp. 693–697, 2004.

[54] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

[55] 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.



Jonas Ruesch received the M.Sc. degree in Mechanical Engineering in 2005 from the Swiss Federal Institute of Technology, Zurich. He is currently working towards the Ph.D. degree in Electrical and Computer Engineering at the Computer and Robot Vision Laboratory at Instituto Superior Técnico, Lisbon. His current research focuses on the co-development of sensory and motor structures for active visual perception in embodied artificial agents.

Ricardo Ferreira graduated in Electrical and Computer Engineering in 2004 at Instituto Superior Técnico (IST). In 2006 he received his M.Sc. degree studying underwater stereo reconstructions of 3D scenes when observed through an air-water interface. His Ph.D. studies focused on reconstructing paper-like surfaces from multiple camera images, and were concluded in 2010. Both his M.Sc. and Ph.D. degrees were obtained at IST. Currently his research interests focus on manifold optimization and general geometric problems arising in robotics and computer vision. He is currently involved in the European research project RoboSoM.

Alexandre Bernardino received the Ph.D. degree in Electrical and Computer Engineering in 2004 from Instituto Superior Técnico (IST), Lisbon. He is an Assistant Professor at IST and Researcher at the Institute for Systems and Robotics (ISR-Lisboa) in the Computer Vision Laboratory (VisLab). He has participated in several national and international research projects and published research papers in top journals and conferences in the field of robotics, vision and cognitive systems. He has been associate editor in large robotics conferences and reviewer for multiple journals and conferences. His main research interests focus on the application of computer vision, cognitive science and control theory to advanced robotic and surveillance systems.