Progressive Learning of Sensory-Motor Maps through Spatiotemporal Predictors

Erhard Wieser and Gordon Cheng

Abstract— **Developmental robotics suggests that the forward and inverse kinematics should be learned through a sensory-motor mapping, instead of being programmed in advance. Motor babbling and goal babbling are two common approaches to generate training samples used to acquire a sensory-motor mapping. Motor babbling typically needs a considerable amount of training data and time to acquire a sufficient mapping, while goal babbling poses difficulties on how to select appropriate goals. In this paper, we propose a neurobiologically-inspired system to progressively learn a sensory-motor mapping bootstrapped from a simple constrained DOF exploration, which generates much less training data than motor babbling. Our proposed system is designed according to two neurobiologically-inspired paradigms: spatiotemporal prediction and uniformity. The spatiotemporal prediction capability facilitates the acquisition of sensory-motor mappings with less amount of training data on the one hand, and facilitates robust behaviour on the other hand. The uniform system design structure is the foundation for building a scalable architecture for cognitive development. We use an improved version of our predictive action selector (*PAS*) as building block of our system. We validate a PAS on a 2 DOF robot head where the robot learns object tracking and evading. Then we validate a second PAS on a 5 DOF arm where it learns reaching.**

Keywords: development of motor and cognitive skills, behaviour bootstrapping, neurorobotics

I. INTRODUCTION

A. Learning the Sensory-Motor Mapping as Prerequisite for Robot Behaviour

In this paper, we take on the developmental perspective and solve a basic problem in robotics, which is the mapping between the visual space and the joint space, *i.e.* the mapping between the visual representation of a task-relevant feature (corresponding to the robot limb or an object of interest) and the degrees of freedom (*DOF*) of a robot limb. Developmental studies such as [1], [2], [3], [4] point out that the acquisition of this visuo-proprioceptive mapping is the prerequisite for learning subsequent motor and cognitive skills, particularly related to hand-eye coordination. In order to solve this visuo-proprioceptive mapping problem on a real humanoid robot, we use our *predictive action selector* (*PAS*) proposed in [5], where our PAS controlled only the 2 DOF head of the humanoid robot NAO and facilitated the emergence of meaningful behaviour like object tracking and evading. However, it was not clear whether the original PAS [5] can be scaled up to deal with more than 2 DOF. In this paper, we explain the neurobiological grounding of

Erhard Wieser and Gordon Cheng are with the Institute for Cognitive Systems, Technische Universität München, München, Germany, http://www.ics.ei.tum.de. E-mail: erhard.wieser@tum.de, gordon@tum.de. Video to this paper: http://web.ics.ei.tum.de/~erhard/videos/SMLearn.mp4
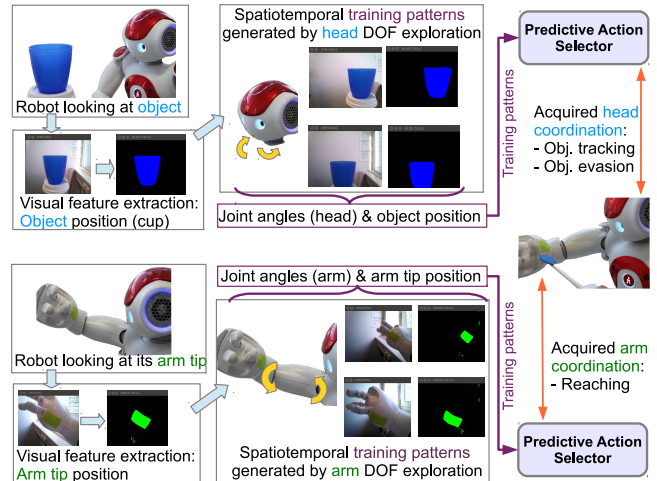
Fig. 1. Progressive learning of sensory-motor maps from scratch: The top half shows a humanoid robot who is facing an object of interest (cup). Our system starts with a constrained DOF exploration of the head joints (one DOF moving at a time, only once). During exploration, sensory-motor data is sampled and fed into our predictive action selector (PAS) [5] which we have improved in this paper. The PAS bootstraps head coordination (tracking behaviour and evading behaviour) from the training data. The bottom half shows the same concept re-used for the arm. Again, the system performs a constrained DOF exploration (one DOF moving at a time, only once) while observing the arm tip. A second PAS bootstraps arm coordination (reaching behaviour) from the training data.

our PAS, and improve its algorithm in order to make it more scalable and more accurate. Thus, our contributions are as follows: We show that our improved PAS can be viewed as an architectural building block for learning sensory-motor maps. Instead of using common methods like motor babbling or goal babbling to generate training data, we use a constrained DOF exploration. By doing so, we emulate the pre-structuring of the human biological motor system in an early developmental stage, where the motions are constrained to simplify the learning [6], [1]. Compared to our original version [5], our improved PAS algorithm is more scalable and it has an increased accuracy in visual tracking. Compared to the related works, our proposed system is a neurobiologically-grounded approach for the progressive learning of sensory-motor maps. It offers the advantage of requiring less training samples and exploiting spatiotemporal prediction for the bootstrapping of robust behaviour. The prediction capability is applied to a multitude of different visual features (*e.g.* object of interest, arm tip, *etc.*). Figure 1 shows the conceptual diagram of our approach.

B. Related Works

In robotics, a common method for the sensory-motor mapping is the forward kinematics as well as the inverse kinematics, also known as the forward model and the inverse model, respectively. Kinematic models are often pre-programmed, for example by using the Denavit–Hartenberg method or the screw-based method [7]. Despite the accuracy of *a priori* kinematic models, they cannot cope with sudden changes in the robot's morphology. For example, the robot might loose (parts of) one limb or one DOF during an accident, but is still supposed to re-calibrate its kinematics and continue to be operational. The developmental perspective offers a solution which enables the robot to learn this sensory-motor mapping by itself in a self-exploratory manner. The developmental perspective of learning sensory-motor schemes by self-exploration, in particular learning visuo-motor coordination, has been widely investigated from different approaches. We identify two main approaches: motor (or body) babbling on the one hand, and goal babbling (including goal-directed exploration) on the other hand. In the following, we review each of these approaches.

In motor babbling (or body babbling) [8], [9], [10], [11], random DOF motions are generated. Kuniyoshi *et al.* [8] realize self-exploration of the embodiment on a simulated baby robot. Its limbs are driven randomly in simulated water, and spatiotemporal correlation of the sampled tactile and motor signals is used to acquire a body map. Schillaci [9] applies motor babbling to realize the generation and sampling of rich proprioceptive and visual data, followed by a k-NN implementing the forward as well as the inverse model, all together with the aim of learning a body map. Takahashi *et al.* [10] use babbling to provide enough amount of data for training dynamic neural networks to acquire a body image. Kajić *et al.* [11] apply babbling to gather training data for self-organizing maps (SOMs) [12]. In [11], they use 2D SOMs in order to realize the learning of hand-eye coordination. The works [8], [9], [10], [11] use motor babbling for the acquisition of the sensory-motor mapping. Babbling generates a rich amount of training data. Sometimes, a big amount of training samples is desirable, for example when a deep neural network is being trained, like in [10]. Mapping models using SOMs also require a considerable amount of training samples. For example, in [11], Kajić *et al.* mention that 74143 input vectors are collected and used for training the SOMs, and they also mention that their random body babbling lasted approximately 40 minutes.

Goal babbling [13], [14], [15], [16] offers a more efficient approach towards learning the sensory-motor mapping. Rolf, Steil, and Gienger [13] demonstrate that goal babbling allows to learn inverse kinematics on humanoid robots (up to 50 DOF). Their work has been inspired by the insight that goal-directed actions occur early in infant development [17]. Rolf and Steil also show the success of their approach on different robot morphologies such as a bionic elephant trunk [15]. Baranes and Oudeyer [14] show that for the learning of inverse models, an exploration of goals in the task

space is faster than an exploration in the motor space. They use an exploration metric ("measure of interestingness") in order to guide the learning by selecting appropriate goals. Schmerling, Schillaci, and Hafner [16] apply goal babbling to learn coordinated control of the head and arm of a NAO robot. They use the Explauto framework [18] as a toolbox to select the goals in a curiosity-driven manner.

In the context of a reaching task on a simulated robot arm, Stulp and Oudeyer [19] validated a method of selectively activating or freezing individual DOF. They report that their algorithm runs a uniform body babbling procedure before it starts to freeze individual DOF.

C. Our Approach

Compared to the related works, our approach is more plausible from the neurobiological point of view. Instead of motor babbling or goal babbling, our system uses a constrained exploration of DOF, where each DOF is simply moved back and forth only once, with one DOF moving at a time. The benefit of this constrained DOF exploration is the drastic reduction of the movement space and the number of training samples. Numerous studies, such as [20], [21], [6], and [1], provide evidence that constraining the movement space is a property of the biological motor system in an early developmental stage. By constraining the movement space, the training data is reduced and less time is spent to acquire an early sensory-motor mapping, which in turn can be refined later on to learn hand-eye coordination. Compared to [19], our exploration algorithm freezes all DOF except for an active one, switching from one active DOF to the next along the kinematic chain. Our action selection algorithm then bootstraps meaningful behaviour, such as reaching, by scalable prediction mechanisms. Prediction is known to be an essential mechanism of the neocortex [22], [23], and it is beneficial for bootstrapping sensory-motor coordination. An example for sensory-motor coordination is the reaching ability, since it requires robustness to temporal occlusions of the target object and a mechanism to bring the arm tip back into the robot's field of view [24].

Our approach is visualized in fig. 1. Compared to sophisticated biologically-inspired control models such as [25], we make an abstraction in order to use pixels and joint angles as the only physical quantities. In order to learn hand-eye coordination, our system uses two PAS modules: one for head control, and the other for arm control. Whether they are used for one or the other, both PAS modules operate in the same manner internally. A PAS does *not* contain any Jacobian or (inverse) kinematic model. Instead, it learns the sensory-motor associations from scratch, based on a simple constrained DOF exploration. In [5], we have grounded our PAS model in a cognitive science point of view. Here, we also ground our PAS model in a neurobiological point of view. For this purpose, we need to recall that the PAS itself contains two important sub-modules: the self-motion predictor (*SMP*), and the task-relevant (visual) feature predictor (*TRFP*). Our SMP encodes the predicted effects of the robot's own motion (ego-motion or self-motion) on a perceived

visual feature. Encoding of self-motion is also present in the neocortex [26]. Our TRFP encodes predicted features of an external object relevant for the robot's interaction with the world, for example the position of an object of interest. The perception of self-motion as well as the perception of object motion can each be mapped to specific cortical regions [26], [27]. The SMP and the TRFP both utilize our version of the multiple timescale recurrent neural network (*MTRNN*) for the learning and prediction of spatiotemporal patterns. We re-implemented the original MTRNN [28] and modified it to work with sigmoid neurons only, in order to make it more uniform [29] and thus closer to the neocortex from a neurobiological point of view. The dorsal premotor cortex (PMd) is involved in (action) response selection [30]. This response selection property is modelled by our proposed action selection algorithm.

II. SYSTEM DESCRIPTION

A. Definitions

We start with the explanations of important terms used to describe our system.

Sensory-motor pattern: A vector encoding sensory-motor features (*e.g.* the perceived position of an object in the image plane, joint angles) at a particular moment in time. A sensory-motor pattern $\mathbf{x}_i$ at time i consists of a visual pattern vector $\mathbf{v}_i$ and a proprioceptive pattern vector $\mathbf{p}_i$, eq. (1).

$$\mathbf{x}_i^T = \begin{bmatrix} \mathbf{v}_i^T & \mathbf{p}_i^T \end{bmatrix} \quad (1)$$

The visual pattern encodes the position of a visual feature of interest in the camera image plane (robot's field of view). This can be the position of an external object, or the position of the robot's hand, or a desired goal position. The proprioceptive pattern encodes the DOF positions of a particular limb, *e.g.* the head or the arm. The visual feature positions and DOF positions are normalized such that each pattern contains values between 0.0 and 1.0. Note that there are two types of visual patterns: $\mathbf{v}_f$ encoding the *feature of interest*, and $\mathbf{v}_g$ encoding a pre-specified *goal*.

Sensory-motor sequence: A sequence S_c of length $m + 1$ is defined as a series of these sensory-motor patterns, *i.e.* $S_c = (\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m)$. The subscript c denotes the identity index of a sequence. We term c as *context* because our PAS learns sequences by its recurrent neural network, and each learned sequence can be represented and recalled by the initial context states of the network. The length m is termed as prediction length, it determines how many time steps a sequence is predicted forward. Another term for sequence is spatiotemporal pattern.

B. Input and Output of a Predictive Action Selector

Consider a PAS module: Its input data consists of $\mathbf{v}_g$, $\mathbf{v}_f$, and the current $\mathbf{p}_{in}$ of the limb to be controlled. Its output data is $\mathbf{p}_{out}$, the updated positions sent to the robot limb. Both $\mathbf{v}_g$ and $\mathbf{v}_f$ are in the robot's field of view (*FOV*).

For learning head control, $\mathbf{v}_f$ encodes the perceived position of an external object of interest, *e.g.* a cup. The goal $\mathbf{v}_g$ is constant and set to $[0.5 \quad 0.5]^T$, representing the middle of

the FOV. For learning arm control, $\mathbf{v}_f$ encodes the perceived position of the robot's arm tip. The goal $\mathbf{v}_g$ is set to be the position of the target object when attempting to reach.

C. Constrained DOF Exploration

Whichever limb the robot learns to control (*i.e.* head or an arm), the constrained DOF exploration phase stays the same. The content encoded by $\mathbf{v}_f$ is pre-specified such that $\mathbf{v}_f$ encodes the cup position in case of head control, or the arm tip position in case of arm control. The robot explores the DOF of its limb while observing $\mathbf{v}_f$. Only one DOF is moved at a time, then followed by the next DOF, and so on until all DOF of the limb are explored. Starting from the home position h, each DOF is moved to an upper limit $h + l$ and back to home position, and then moved to a lower limit $h - l$ and finally back to home position. DOF positions are recorded only when starting from home position, and then going to either upper or lower limit. Thus, our system records two simple visuo-proprioceptive sequences for each DOF separately. One additional sequence is recorded which is the *idle* sequence where all DOF of the limb are in home position. Thus, the number of recorded sequences is $2 \cdot N + 1$ where N is the number of DOF of the limb to be controlled. Each DOF is only moved by a few degrees. The explored range is the same for every DOF of a limb.

D. Improved Prediction Model

For the scope of this paper, we focus on the main algorithm of the predictive action selector. The PAS algorithm uses the predicted patterns delivered by the SMP and TRFP. However, our algorithm presented in [5] needs to be improved regarding its scalability, in order to deal with more than 2 DOF and to increase its accuracy. Here, we improve the foundations in [5] without changing the original idea behind it.

Although it would be sufficient for our system to use a simple recurrent neural network (with one timescale for the context neurons), we still decided to adopt and modify the MTRNN [28] because it offers the possibility to learn more complex spatiotemporal patterns (which may be utilized in our future work).

1) Predictive Minimization: The PAS minimizes a value function V through a process of internal prediction [5]. We explain how this method can be used in two different scenarios, one is for the control of the head, and the other is for the control of the arm:

$$\min V(c, m) = \min \| \mathbf{v}_g + a_f \cdot \mathbf{v}_f + a_s \cdot \Delta\mathbf{v}_s(c, m) \| \quad (2)$$

The scalars a_f and a_s are called alteration parameters. They support an integration of multiple sensory input [31] and they bias the robot's behaviour, *e.g.* changing from object tracking to object evasion [5]. The content of $\mathbf{v}_g$ and $\mathbf{v}_f$ depends on the controlled limb, see section II-B. Note that the pattern $\mathbf{v}_f$ will not be observable all the time due to occlusions. In such a case, a prediction of $\mathbf{v}_f$ is delivered by the TRFP of the PAS. The difference pattern $\Delta\mathbf{v}_s(c, m)$ is of key importance: It describes the *predicted change* of the visual feature, this change is caused by *self-motion* only. In case of head control,

$\Delta\mathbf{v}_s$ is the predicted change of the external object caused by the robot's own head motion. In case of arm control, $\Delta\mathbf{v}_s$ is the predicted change of the arm tip marker caused by the robot's own arm motion. The difference pattern depends on the selected sequence c and on the prediction length m, and it is computed by:

$$\Delta\mathbf{v}_s(c,m) = \mathbf{v}_f(c,m) - \mathbf{v}_f(c,0) \qquad (3)$$

The pattern $\mathbf{v}_f(c,m)$ is $\mathbf{v}_f$ of the sequence c at time step m. The pattern $\mathbf{v}_f(c,0)$ is $\mathbf{v}_f$ of the sequence c at time step 0. The patterns $\mathbf{v}_f(c,m)$ and $\mathbf{v}_f(c,0)$ are delivered by the SMP of the PAS. The biological analogy is that $\mathbf{v}_f$ and $\Delta\mathbf{v}_s$ correspond to input from the visual pathway (MST region), especially $\Delta\mathbf{v}_s$ describing the predicted perception of self-motion relative to a feature of interest [26].

2) Processing Steps: We use the method of predictive minimization as part of our improved action selection algorithm. In contrast to [5], the improvements encompass the calculation of a new position for each DOF separately and the automatic adjustment of the prediction length m depending on the perceived visual error. Benefits are the scalability regarding the number of DOF and the loss of one model parameter. By the following processing steps, we model the response selection property of the dorsal premotor cortex (PMd) and the computation of motor signals in the primary motor cortex (M1).

For each DOF of a limb, the PAS algorithm does the following steps 4 to 10:

It first computes the current visual error.

$$e_v = \|\mathbf{v}_g - \mathbf{v}_f\| \qquad (4)$$

Based on this visual error, the algorithm determines the prediction length m as:

$$m = \lfloor \kappa \cdot e_v \rfloor \qquad (5)$$

where κ is a constant termed prediction length factor. By using m, the algorithm computes the self-motion effects $\Delta\mathbf{v}_s$ and selects the visuo-proprioceptive sequence c_w (winner sequence) which minimizes V.

$$\min V(c,m) \quad \text{given a set of learned sequences} \qquad (6)$$

The sequences needed for step 6 have been generated by the constrained DOF exploration before and have been learned by the SMP. Once the algorithm has determined c_w, it computes the new DOF position by

$$\Delta p(c_w,m) = p(c_w,m) - p(c_w,0) \qquad (7)$$
$$p_{out} = p_{in} + \Delta p(c_w,m) \qquad (8)$$

where $p(c_w,m)$ is the predicted DOF position extracted from the visuo-proprioceptive winner sequence c_w at length m. After determining p_{out}, the algorithm updates the visual feature, in order to consider the predicted change $\Delta\mathbf{v}_s(c_w,m)$ which will occur once p_{out} is commanded to the DOF:

$$\mathbf{v}_f^* = \mathbf{v}_f + \Delta\mathbf{v}_s(c_w,m) \qquad (9)$$

Finally, the algorithm updates the input visual feature by

$$\mathbf{v}_f \leftarrow \mathbf{v}_f^* \qquad (10)$$

III. EXPERIMENTS

A. Setup

We used a NAO robot to evaluate our system. The NAO was in sitting position throughout all experiments. We used a green colour marker for a colour-based detection of the robot's arm tip in its FOV. The colour marker was attached on the NAO's right arm tip, located between its wrist joint and its fingers. Besides the colour information, the contour of the marker in the image frame is extracted and serves as a distance measure. The smaller the contour size, the greater the distance of the marker (*i.e.* the arm tip) with respect to the top camera frame. Since the NAO does not support stereo vision, we used the size of the extracted contours of the arm tip marker and the object in order to extract the distance information relative to the robot's top camera. Figure 2 shows the NAO reaching for an object (a blue marker attached on a chopstick). The sampling rate of

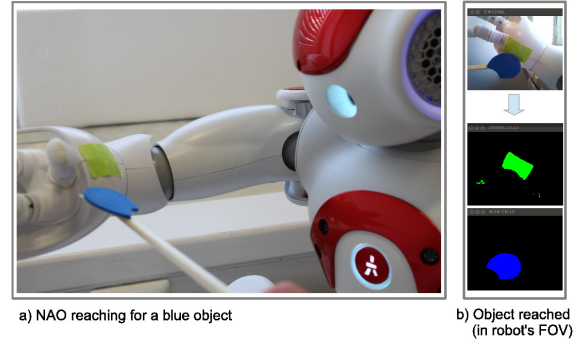a) NAO reaching for a blue object   b) Object reached (in robot's FOV)

Fig. 2. Experimental setup: The left part (a) shows a NAO (in sitting position). A green marker is attached on its right wrist and serves as the arm tip feature. A blue marker is attached on a chopstick which is moved by a human. The blue marker serves as the object feature in case of the reaching experiment. The right part (b) shows the extraction of the arm tip and object feature in the robot's FOV. In contrast to the left part, the camera image in the right top part shows the best situation possible in case of the reaching experiment, where the robot has moved its arm tip to the object as close as possible (corresponding to a normalized visual error of ca. 0.2 in the robot's FOV).

the joints was 20 Hz, which was also the update rate of a PAS. During the head DOF exploration phase for learning head control, our system collected 48 sample vectors in total. During the arm DOF exploration phase for learning arm control, our system collected 126 sample vectors in total. In each case, these sample vectors are the training data for the SMP of the PAS controlling the particular limb.

B. Neural Network Parameters

We employ an one-to-one mapping, *i.e.* each dimension of a normalized pattern vector is mapped to one neuron of the input-output group of the neural network. In the MTRNN of the SMP, the first part of input-output neurons represents $\mathbf{v}_f$, while the second part of input-output neurons represents $\mathbf{p}$. In the MTRNN of the TRFP, the input-output neurons represent $\mathbf{v}_f$. The MTRNN of the SMP was set to single timescale mode because the sequences gathered during DOF exploration are short enough. In the single timescale

mode, all context neurons have the same timescale and all context neurons have connections to input-output neurons and vice versa, whereas in the multiple timescale mode, the slow context neurons are disconnected from the input-output neurons and vice versa. The MTRNN of the TRFP was also set to single timescale mode. We use the following symbols for the neural network parameters: N_{IO}, N_{FC}, N_{SC} denote the number of input-output, fast context, and slow context neurons, respectively. Parameters τ_{IO}, τ_{FC}, τ_{SC} denote the timescale of input-output, fast context, and slow context neurons, respectively. Table I summarizes our network parameter settings for the PAS learning head control and for the PAS learning arm control. Note that the TRFP of a PAS predicts

TABLE I
NEURAL NETWORK PARAMETER FOR THE PAS LEARNING HEAD
CONTROL AND FOR THE PAS LEARNING ARM CONTROL

MTRNN param.	N_{IO}	N_{FC}	N_{SC}	τ_{IO}	τ_{FC}	τ_{SC}
SMP_{Head}	4	5	5	7	20	20
TRFP_{Head}	2	5	5	7	20	20
SMP_{Arm}	8	15	5	7	20	20

the motion of an external object. TRFP_{Arm} was not needed in the reaching experiment, since the prediction of the arm tip is already done as part of the action selection process (see steps 9 and 10 of section II-D.2).

C. Acquired Head Control: Tracking an Object of Interest

We validated our improved PAS model on the head, especially to investigate the improvements in accuracy compared to our earlier model. The parameter settings of the SMP and TRFP of the head PAS stayed the same as during the learning phase. A cup was moved in front of the robot and the tracking skill was validated. The visual goal pattern was constant and set to $\mathbf{v}_g = [0.5 \quad 0.5]^T$, corresponding to the center of the robot's FOV. During this interaction, the cup was moved behind another object, causing a short-term occlusion. Compared to our previous work [5], we could observe that our improved PAS model has at least the same or even a better accuracy in terms of the remaining visual error. Once the object comes to rest, the remaining visual error is below 4 - 5 %, as reported in [5]. Nevertheless, our improved PAS model often achieves up to 1 - 2 % remaining visual error. For the head PAS, we set $\kappa = 15$. The improved PAS also supports object evasion behaviour, like shown in [5]. The results of the head control are shown in fig. 3.

D. Acquired Arm Control: Reaching for an Object of Interest

We validated whether our improved PAS model has successfully learned to control the right arm of the NAO. This experiment demonstrates the scalability of our model from 2 to 5 DOF. For now, we validated the arm coordination when the head was in a resting position facing the arm. In case of arm control, the visual goal pattern was set to be the object position (blue marker on the chopstick, see fig. 2). For the arm PAS, we set $\kappa = 8$. The parameter settings of the SMP of the arm PAS stayed the same as during the learning
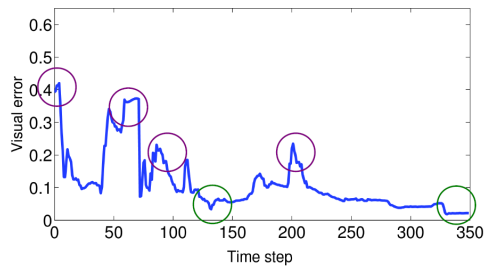

Fig. 3. Visual error in the robot's FOV while our PAS is controlling the head to track an object. At the onset (left purple circle), the object appears in the corner of the robot's FOV, causing a large error which the robot compensates for by moving accordingly. The error is also high between the time steps 60 and 72 due to an occlusion of the object. Nevertheless, the robot's behaviour is still enough robust in terms of a continuous tracking made possible by using the prediction of the object location. Additional peaks in the error around time steps 85 and 200 arise from sudden motions of the cup. Once the cup is only slightly moved or not moved any more (green circles), the robot fixates its attention on it with a remaining visual error of 1 - 2 %.

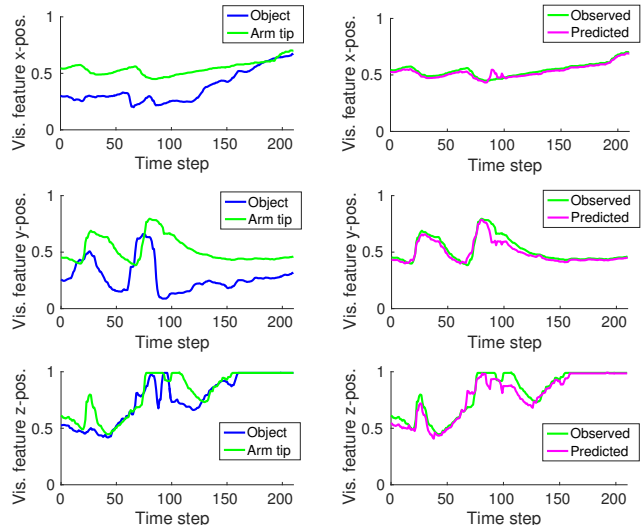phase. We summarize the results in fig. 4. In this reaching

Fig. 4. Observed and predicted visual features during the reaching. The left side shows the target object (blue marker on a chopstick end, represented by the blue signals) which was moved randomly in front of the robot, with variations in the distance. The arm tip (green signals) was approaching the object. The robot reached the object around time step 160. A residual error remained in the y-position, which is acceptable in this scenario because the markers of the arm tip and chopstick are physically next to each other. Then the measured normalized visual error is about 0.2 or less, meaning that the object is within the robot's grasp. The right side shows the observed arm tip feature (green) and its prediction (magenta) during the reaching.

experiment, the visual goal was not predicted, although this is possible by using the signal delivered by TRFP_{Head}. The arm PAS predicts the arm tip feature and relies on this predicted signal when the arm tip is not visible. The best prediction was 150 $msec$ ahead in time. This arm tip prediction is shown in the right side of fig. 4. The reaching ability shows that our improved PAS model can deal with more than 2 DOF. In our conducted reaching experiment, the elbow was the most active joint which especially compensated for variations in

the distance between object and robot.

IV. CONCLUSION

We presented an improved version of our predictive action selector (PAS) which can be viewed as a crucial building block for the construction of architectures for cognitive development. Our PAS is a neurobiologically-inspired model for the progressive learning of sensory-motor maps based on a constrained DOF exploration. Progressive learning is realized by exploring the DOF of one limb and learning the mapping, and then repeat this procedure for another limb. By this approach, one PAS learned to control the head of the robot. Then, a second PAS learned to control the arm. Compared to our previous work, we achieved an improvement in the object tracking accuracy (1 - 2 % remaining visual error as the minimum error measured) when controlling the head. In case of arm control, we demonstrated the acquisition of an early reaching skill. Our proposed PAS model can integrate both visual and proprioceptive information in a very rapid way with considerably less training data compared to approaches using motor babbling. For the learning of object tracking via head motions, our system collected 48 sample vectors. For the learning of reaching, our system collected 126 sample vectors. This relative small amount of training data was enough to bootstrap tracking and reaching at an early developmental stage. When predicting the arm tip position in the field of view, the best prediction achieved was 150 $msec$ ahead in time at a sampling rate of 20 Hz. Since the PAS computes predictions milliseconds ahead in time, potential applications encompass motion correction and behaviour switching.

REFERENCES

[1] D. Vernon, C. von Hofsten, and L. Fadiga, *A roadmap for cognitive development in humanoid robots*, ser. Cognitive Systems Monographs, R. Dillmann, Y. Nakamura, S. Schaal, and D. Vernon, Eds. Springer-Verlag Berlin Heidelberg, 2010, vol. 11.

[2] M. Asada, K. F. MacDorman, H. Ishiguro, and Y. Kuniyoshi, "Cognitive developmental robotics as a new paradigm for the design of humanoid robots," *Robotics and Autonomous Systems*, vol. 37, no. 2–3, pp. 185–193, 2001.

[3] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: a survey," *IEEE Transactions on Autonomous Mental Development*, vol. 1, no. 1, pp. 12–34, 2009.

[4] Y. Kuniyoshi, Y. Yorozu, S. Suzuki, S. Sangawa, Y. Ohmura, K. Terada, and A. Nagakubo, "Emergence and development of embodied cognition: a constructivist approach using robots," *Progress in Brain Research*, vol. 164, pp. 425–445, 2007.

[5] E. Wieser and G. Cheng, "Predictive action selector for generating meaningful robot behaviour from minimum amount of samples," in *Proceedings of the IEEE International Conference on Development and Learning and Epigenetic Robotics*, 2014, pp. 139–145.

[6] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: a survey," *Connection Science*, vol. 15, no. 4, pp. 151–190, 2003.

[7] C. R. Rocha, C. P. Tonetto, and A. Dias, "A comparison between the denavit–hartenberg and the screw-based methods used in kinematic modeling of robot manipulators," *Robotics and Computer-Integrated Manufacturing*, vol. 27, no. 4, pp. 723–728, 2011.

[8] Y. Kuniyoshi, Y. Yorozu, Y. Ohmura, K. Terada, T. Otani, A. Nagakubo, and T. Yamamoto, "From humanoid embodiment to theory of mind," in *Embodied Artificial Intelligence, LNAI 3139*, F. Iida, R. Pfeifer, L. Steels, and Y. Kuniyoshi, Eds. Springer-Verlag Berlin Heidelberg, 2004, pp. 202–218.

[9] G. Schillaci, "Sensorimotor learning and simulation of experience as a basis for the development of cognition in robotics," Ph.D. dissertation, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2014.

[10] K. Takahashi, T. Ogata, H. Tjandra, Y. Yamaguchi, and S. Sugano, "Tool-body assimilation model based on body babbling and neurodynamical system," *Mathematical Problems in Engineering*, vol. 2015, pp. 1–15, 2015.

[11] I. Kajić, G. Schillaci, S. Bodiroža, and V. V. Hafner, "Learning hand-eye coordination for a humanoid robot using soms," in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, 2014, pp. 192–193.

[12] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[13] M. Rolf, J. J. Steil, and M. Gienger, "Goal babbling permits direct learning of inverse kinematics," *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 3, pp. 216–229, 2010.

[14] A. Baranes and P.-Y. Oudeyer, "Active learning of inverse models with intrinsically motivated goal exploration in robots," *Robotics and Autonomous Systems*, vol. 61, no. 1, pp. 49–73, 2013.

[15] M. Rolf and J. J. Steil, "Efficient exploratory learning of inverse kinematics on a bionic elephant trunk," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 6, pp. 1147–1160, 2014.

[16] M. Schmerling, G. Schillaci, and V. V. Hafner, "Goal-directed learning of hand-eye coordination in a humanoid robot," in *Proceedings of the IEEE International Conference on Development and Learning and Epigenetic Robotics*, 2015, pp. 168–175.

[17] C. von Hofsten, "An action perspective on motor development," *Trends in Cognitive Sciences*, vol. 8, no. 6, pp. 266–272, 2004.

[18] C. Moulin-Frier, P. Rouanet, and P.-Y. Oudeyer, "Explauto: an open-source python library to study autonomous exploration in developmental robotics," in *Proceedings of the IEEE International Conference on Development and Learning and Epigenetic Robotics*, 2014, pp. 171–172.

[19] F. Stulp and P.-Y. Oudeyer, "Emergent proximo-distal maturation through adaptive exploration," in *Proceedings of the IEEE International Conference on Development and Learning and Epigenetic Robotics*, 2012, pp. 1–6.

[20] N. A. Bernstein, *The co-ordination and regulation of movements*. Pergamon Press Ltd., 1967.

[21] E. C. Goldfield, *Emergent forms: origins and early development of human action and perception.* Oxford University Press, 1995.

[22] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[23] ——, "Hierarchical models in the brain," *PLoS Computational Biology*, vol. 4, no. 11, pp. e1 000 211, 1–24, 2008.

[24] C. Gaskett and G. Cheng, "Online learning of a motor map for humanoid robot reaching," in *Proceedings of the 2nd International Conference on Computational Intelligence, Robotics and Autonomous Systems*, 2003, pp. 1–6.

[25] E. Todorov, "Direct cortical control of muscle activation in voluntary arm movements: a model," *Nature Neuroscience*, vol. 3, no. 4, pp. 391–398, 2000.

[26] R. S. Zemel and T. J. Sejnowski, "A model for encoding multiple object motions and self-motion in area mst of primate visual cortex," *The Journal of Neuroscience*, vol. 18, no. 1, pp. 531–547, 1998.

[27] G. Wiest, M.-A. Amorim, D. Mayer, S. Schick, L. Deecke, and W. Lang, "Cortical responses to object-motion and visually-induced self-motion perception," *Cognitive Brain Research*, vol. 12, no. 1, pp. 167–170, 2001.

[28] Y. Yamashita and J. Tani, "Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment," *PLoS Computational Biology*, vol. 4, no. 11, pp. e1 000 220, 1–18, 2008.

[29] M. H. Johnson, *Developmental cognitive neuroscience.* John Wiley & Sons, 1997.

[30] D. Badre and M. D'Esposito, "Is the rostro-caudal axis of the frontal lobe hierarchical?" *Nature Reviews Neuroscience*, vol. 10, no. 9, pp. 659–669, 2009.

[31] G. Cheng and Y. Kuniyoshi, "Complex continuous meaningful humanoid interaction: a multi sensory-cue based approach," in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 3, 2000, pp. 2235–2242.