

Neural Architecture for Sensorimotor Anticipation

Horst-Michael Gross, Volker Stephan, Torsten Seiler

Technical University of Ilmenau
Department of Neuroinformatics
98684 Ilmenau , Germany
email: homi@informatik.tu-ilmenau.de

Abstract

The anticipatory approach to perception and sensorimotor control presented here is based on the prediction of sensory consequences of hypothetically executed actions or action sequences. Based on this anticipatory approach, we present a neural architecture that is able to anticipate and evaluate hypothetically sensorimotor sequences and to evolve an initially reactive behavior into a planning and forecasting one. We demonstrate the performance of this kind of generative predictive perception in the light of a local navigation task of a KHEPERA robot. Results of vision-based sensorimotor experiments performed with the internally simulating control architecture are presented.

1 Introduction and Basic Idea

With our *anticipatory approach* to perception and sensorimotor control [8; 9; 10] we want to avoid the common separation of *perception* and *generation of behavior* and fuse both aspects into one neural process. Therefore, we consider perception, especially perception of space and shape, as internal simulation of a number of hypothetical actions (or action sequences) and anticipation of their sensory consequences. In this respect, perception is regarded to be an active process of generating sensorimotor hypotheses based on experienced and learned sensorimotor relations. This point of view emphasizes the generative and predictive character of perception considering both *sensory* and *motor* aspects of the action-perception-cycle. On the one hand, all hypothetical actions and their expected sensory consequences '*describe*' the current sensory situation. On the other hand, from this set of '*descriptive*' actions those can be *selected for execution in reality* which result in a positive effect concerning the goal of the system. From this point of view, perception of a visual scenery, for example in a navigation task, can be characterized as follows: an object in a certain distance may entail visual impressions (e.g. the optical flow, see below) that enable

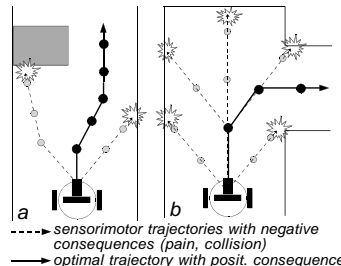


Fig. 1: Local navigation as sensorimotor controlled internal simulation of hypothetical actions and anticipation of their sensory consequences - simulated and evaluated sensorimotor trajectories.

a prediction of the expected visual and tactile consequences in case of performing typical actions. Based on the real visual impressions, a number of hypothetical actions are simulated internally (see exemplary trajectories in Fig. 1 a, b): those actions having a *positive* consequence (e.g. a collision-free movement) are preferred for execution (black trajectories in Fig. 1 a, b). All simulated actions that entail *negative* effects (collision, pain) will not be executed in reality, but, nevertheless, they contribute to the description of the present situation, too. So, it is possible to characterize the visual scenery immediately in categories of behavior, i.e. by a set of actions which describe possible methods of local navigation.

2 Sensorimotor Control Task

We have investigated the applicability of our anticipatory approach and the developed neural control architecture (see section 3) in the light of a local navigation task in order to demonstrate the efficiency and soundness of our concept. In this task, a mobile robot has to anticipate the sensorimotor consequences of its possible actions in order to navigate successfully, especially in critical regions of the environment, for example, in front of obstacles, on intersections, etc. (see Fig. 1 and 8 - left). Instead of a simulator, we used our mobile miniature-robot KHEPERA to learn to navigate in a real-world environment. If a mobile robot is to interact in a useful way with its surroundings, it is essential that it has sensory systems that are able to determine the basic 3D structure of the nearby environment around it. A possible way is provided by the *optical flow*, especially the *optical flow field*. An optical flow field is determined by three things that are of direct importance for our anticipatory approach: the robot egomotion, the scene motion, and the 3D structure

*Supported by Deutsche Forschungsgemeinschaft, DFG-Grants Gr1378, 1-2 (SEMRINT-Project) and Gr1378, 2-1 (SEMRINT-Project)

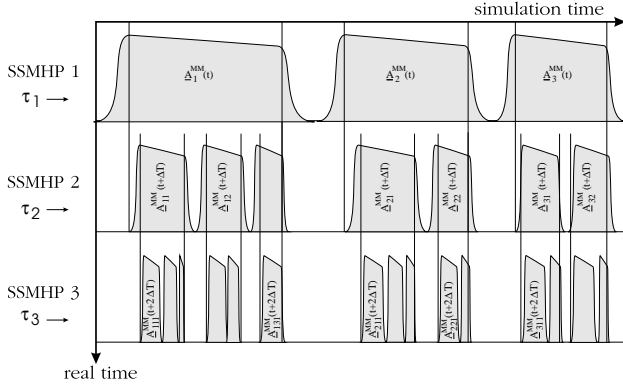


Fig. 2: Time behavior of hypothesis activities generated in three concatenated Prediction Modules operating at staggered time-scales of a time hierarchy. Dynamic approach to a combined depth and width search for the best evaluated action sequence within the sensorimotor space.

of the scene. In other words, the optical flow field yields implicit information about spatial distances of the robot to objects in the environment considering the actions carried out immediately before. Regardless of the typical optical flow problems, DUCHON [3], KRÖSE [6], and others could demonstrate practically that it is possible to navigate a mobile robot just using the optical flow field. To estimate the optical flow, we utilize the region-based correlation approach by CAMUS [2]. Each movement of our robot KHEPERA is bipartioned in translatory and rotatory components, which are executed sequentially. During each translatory movement, a sequence of consecutive images is taken with the robot’s camera in order to stabilize the estimation process by averaging of microstep flow fields.

3 Neural Architecture for Sensorimotor Anticipation

In this section, we present our neural control architecture that learns to predict and evaluate the sensory consequences of hypothetically executed actions. For that purpose, the architecture must be able to simulate alternative sensorimotor sequences, select the best evaluated one, and start to execute the sequence in reality. Thus, it shows planning behavior at the sensorimotor, i.e. subsymbolic level. Since the intrinsic dynamics of the internal simulation process is much faster than the dynamics of the action-perception-cycle in reality, it is possible to simulate and evaluate several hypothetical action-perception-cycles before final selection and execution.

3.1 Parallel-sequential Search in Sensorimotor Space at Different Time Scales

Sequences of hypothetical actions and corresponding sensory predictions are generated in our architecture by chaining of several sensorimotor Prediction Modules (PM) (see Fig. 3) that operate at staggered time-scales of a time hierarchy (see Fig. 2). The first Prediction Module (PM1) operates on the real senso-

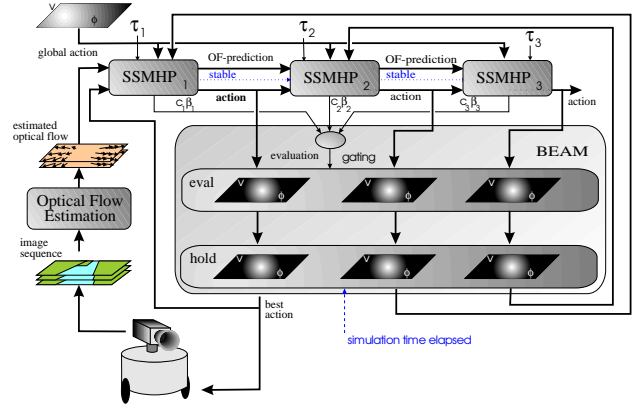


Fig. 3: Chaining of PMs (SSMHPs) operating at staggered time-scales for internal simulation of sequences of hypothetical actions and expected sensory consequences. The architecture realizes a dynamic search in depth and width and selects the best evaluated sensorimotor sequence from all simulated ones. Depth and width of the search are limited both by resources (number of modules) and time (available time for relaxation to alternative solutions, and reaction cycle length of the robot).

ry input, simulates a set of alternative actions, and predicts hypothetical sensory consequences of these actions. The second module PM2 just operates on the predicted sensory consequences of the simulated PM1-actions, generates itself a set of alternative actions, and anticipates their sensory consequences. This internal simulation and prediction process on hypothetical data is repeated in the subsequent modules. Because the PMs operate at predefined, staggered time scales of that time-hierarchy, each of the modules is only given that limited time to dynamically generate alternative hypotheses during which the output of its predecessor remains stable (see Fig. 2). To reduce the complexity of search in the sensorimotor space, we use a dynamic approach, which only selects the most promising sensorimotor transitions. For that purpose, each PM is able to generate a set of evaluated action hypotheses that are selected in descending sequence beginning with the best evaluated one. This is a dynamical version of a *search in width* with the advantage that the order of search can be continuously reshuffled due to adaptation of the action values or a global modulation of all action hypotheses.

Subsequently, we explain the essential neural sub-systems of our anticipating control architecture and describe an internal simulation cycle.

3.2 Sensorimotor Prediction Module

Fig. 4 depicts the internal architecture of a single PM. Each PM consists of a submodule for generating topologically coded action sets (*action suggestion*), a submodule for sequential selection of single actions within this map (*action selection*), a submodule for prediction of the sensory consequences of the selected action (*optical flow prediction*), and an evaluation module for the selected action (*hypothesis evaluation*). A real sensory input $\underline{x}^{OF}(t)$ and the last real

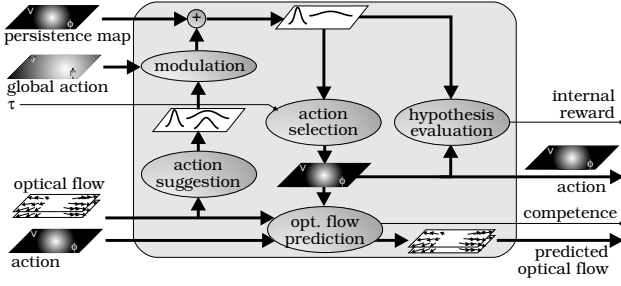


Fig. 4: Architecture of one Prediction Module (PM)

action $\underline{a}^{real}(t)$ caused this input are processed in two pathways: a) the *sensorimotor mapping* with the subsystems action suggestion, modulation, action selection, and hypotheses evaluation and b) the *motorsensory prediction* with the optical flow prediction subsystem. In these pathways, hypothetical actions and their expected sensory consequences can be generated that ‘describe’ the current sensory situation. What action will be selected from this set of ‘descriptive’ actions for execution depends on the internal evaluation of the generated sensorimotor sequences as a whole.

Sensorimotor mapping: In this pathway, a topologically organized motor map that codes a set of alternative actions is generated in the *action suggestion* subsystem. A location in this motor map is corresponding to a specific action within a quasi-continuous action space (e.g. velocity and steering angle of a navigation movement), while the activation of the corresponding neuron in the field represents the value of this action learned in previous “action-perception-cycles” (see Fig. 8-right). To generate the motor map, this subsystem is realized as a neural function approximator ($X^{OF}(t) \rightarrow A(t)$) that is based on an adaptive vector quantization technique combined with a topologically distributed reinforcement learning mechanism [4]. The activity distribution in this motor map can be directly modulated by an external input, e.g. a global map (see Fig. 4), to influence the subsequent action selection process. This *modulation* is a simple control mechanism that allows it to integrate a single PM into a multi-modular control architecture (see Fig. 3). To select alternative actions in the motor map, a neural field dynamics with sequential selection behavior is used in the *action selection* subsystem. The aim of this dynamics is the sequential selection of motor hypotheses in form of activity blobs in the motor map. Since we are interested in alternative motor hypothe-

ses, the selected blob needs to be stabilized only for a certain time. After that time, the selected region has to be inactivated to give the next hypothesis the chance to win the competition. Our selection mechanism is based on a non-linear AMARI-field dynamics [1]. The field dynamics just selects the local region with the highest energy within the map and suppresses all other neurons in the field. So, the localization of the winner blob codes the action hypothesis with the highest evaluation for the current sensory situation. In order to generate a sequence of alternative action hypotheses, we extended the AMARI-dynamics by a temporal self-inhibition of the currently selected winner-blob. A detailed description of this self-inhibition mechanism is given in [11]. With this extension, alternative action hypotheses can be found and selected sequentially taking their competition energy that corresponds to the input-specific “value” of the suggested actions into consideration. This way, the field dynamics selects the most promising action with the highest evaluation first and later less promising ones, if the available time for relaxation to alternative solutions allows this (Fig. 5). Result of the sequential search is a spatio-temporal sequence of temporarily stable action suggestions (see Fig. 2-first row, too) that are fed into the *motorsensory prediction* pathway, where they modulate the *optical flow prediction*-subsystem (see below). Consequently, the parameters for the internal field dynamics directly control the breadth of search.

The *hypothesis evaluation* subsystem computes an evaluation signal that is based on the selected action and the action suggestion. This way, the sequential selection dynamics controls the *width of search* in each PM. The maximum *depth of search* depends on the number of replicated PMs organized in the time hierarchy (see section 3.3).

Motorsensory prediction: One of the main ideas of our anticipatory approach is that the subsequent flow field just depends on the previous one and the executed action, if the environment is static. If the robot is moving forward with a slight turn to the left, for instance, the flow field shifts to the right and vice versa. Therefore, it is assumed that the *optical flow prediction* subsystem can learn to anticipate the sensory consequences of selected actions. Without a successful prediction at the sensory level it is impossible to chain several PMs in order to generate hypothetical sensorimotor sequences. To predict a hypothetical flow field of a hypothetical action, the prediction network has to approximate the following mapping, which is usually non-linear because of the distortions of the optical system: $OF^{real}(t) \times A^{hyp}(t) \rightarrow OF^{hyp}(t+1)$. We assume, that this motor controlled prediction can be approximated by a piecewise linear function approximation. We investigated several neural architectures for the sensorimotor based optical flow prediction, among other things a distributed architecture of action-specific, linear function approximators (see Fig. 6). This approach realizes a divide-and-conquer strategy that is somewhat akin to that of the “Adaptive Mixtures of Local Experts” of [5]. Our flow pre-

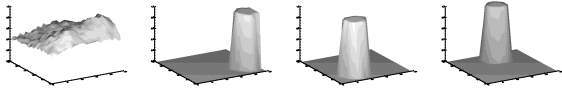


Fig. 5: Extended AMARI-Dynamics for sequential generation of motor hypotheses. Starting from an stable input that codes the evaluated action suggestions for a given sensory situation (left), the dynamics successively selects alternative motor hypotheses in the field.

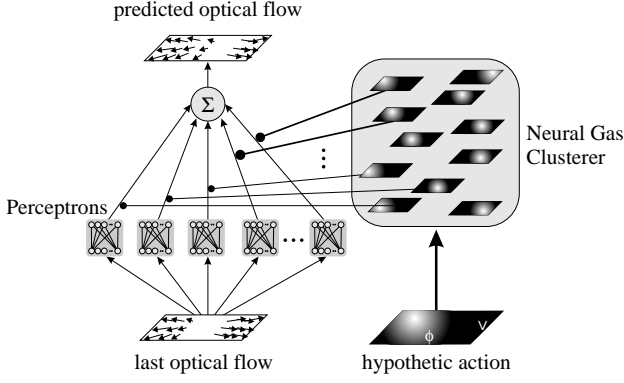


Fig. 6: Schematic diagram of our modular optical flow prediction network with n expert networks (linear perceptrons) and a gating network.

dictors are corresponding to the local experts, and our action map clusterer to the selector network that assigns tasks of different motor context to the individual experts. The mapping of the experts is controlled by a neural vector quantization technique, a Neural Gas [7], that is gating the different flow predicting perceptrons in dependence on the action selected by the *sensorimotor mapping* pathway. The final prediction of the succeeding flow field $\underline{\hat{x}}^{OF}(t+1)$ results from a weighted superposition of the outputs of the local experts that is controlled by the activation y_i^M of the gating Neural Gas neurons.

$$\underline{\hat{x}}^{OF}(t+1) = \frac{\sum_{r=0}^k y_{i(r)}^M \cdot \underline{W}_{i(r)}^P \cdot \underline{x}^{OF}(t)}{\sum_{r=0}^k y_{i(r)}^M}$$

with $y_{i(r)}^M = e^{-r/\sigma}$

The index r results from a “neighborhood ranking” of the reference vectors \underline{w}_i^M for the current motor input $\underline{a}^{real}(t)$ ($r=0$ stands for the best-matching neuron i with the lowest Euclidean distance d_i), k controls the influence of the other Neural Gas neurons.

3.3 Multi-modular Architecture

A single *PM* can sequentially generate hypothetical actions and estimate the resulting sensory consequences together with corresponding evaluations. Sequences of hypothetical actions and corresponding sensory expectations are generated by chaining of several *PMs* operating at staggered time-scales (see Fig. 2). As mentioned above, the first *PM* operates on the real sensory input, the second one on the prediction of the first, and so on (see Fig. 3). The actions proposed by the *PMs* and all evaluations of a stable action sequence are fed into a short-term memory called *Best Evaluated Action Memory - BEAM* (Figure 3). A newly hypothesized action sequence overwrites a stored sequence if the cumulative evaluation of the new one is higher than that of the stored. The *BEAM* is inactivated if the input for the first *PM* is destabilized, which means that a new, real sensorimotor situation has been achieved. This strategy allows the internal simulation of more than one action sequences, but just

the first action of the action sequence that has been evaluated best is executed in reality. Moreover, the sequence of hypothetical action maps that is stored in the short-time memory of *BEAM* is used to prefer the selected sequence in the next simulation cycle as a kind of persistence (Fig. 3). This is realized by a specific delayed feedback of the memorized action maps to the *PMs*, where they modulate the subsequent action selection processes. Our dynamic hierarchy with staggered time-scales guarantees that, after a short setup-time, the internal simulation process can be interrupted at any time. Then, that action sequence which has been evaluated best up to this moment is selected and executed. This way, it is possible to find an at least good solution at any time. From the robotics, this is well-known as *anytime behavior*. If there is enough time, our architecture can find the best evaluated action sequence on the basis of those motorsensory relations that have been experienced, evaluated, and learned in previous action-perception-cycles.

3.4 Learning

The learning within the *PMs* is performed after execution of a real action by comparing the real and the predicted sensory situation ($\underline{x}^{OF}(t+1)$, $\underline{\hat{x}}^{OF}(t+1)$) considering the reinforcement signals received from the environment. It is to remark, that the learning occurs independently in both pathways. Hence, a non-optimal selected action is not a problem for the learning of the sensory prediction. On the other hand, an erroneous prediction has no influence on the reinforcement learning in the *sensorimotor mapping* pathway.

Sensorimotor Mapping: The generation of evaluated motor maps that consider all experienced actions is based on a self-organizing sensorimotor function approximation on the basis of reinforcement learning (RL). A detailed explanation of this kind of neural field based action value learning is presented in [4]. In a first implementation, the action values of the *action suggestion* subsystem are adapted by the following RL-rule:

$$\Delta w_{ri}(t) = \eta \cdot y_i^S(t) \cdot \underline{a}_r^{real}(t) [R - \underline{a}_r^{hyp}(t)]$$

where r is the position in the motor map, $\underline{a}^{hyp}(t)$ is the suggested set of alternative actions, $\underline{a}^{real}(t)$ is the actually executed action that has been selected by the neural field dynamics as winner-blob, y_i^S is the actual activation of the Neural Gas neurons quantifying the sensory input $\underline{x}^{OF}(t)$, η determines the learning rate and $R \in [0..1]$ is the immediate reward for the action carried out. The reward is based on a predefined reward function that considers the tactile feedback after execution of $\underline{a}^{real}(t)$.

Motorsensory prediction: In this pathway, the weights \underline{w}_i^M of the neurons i in the vector quantizer, which clusters the *action space*, are adapted according to their activation y_i^M and the executed action $\underline{a}^{real}(t)$ by the following equation:

$$\Delta \underline{w}_i^M = \eta^M \cdot y_i^M \cdot [\underline{a}^{real}(t) - \underline{w}_i^M]$$

The adaptation of the local experts for optical flow prediction is started automatically when a relative-

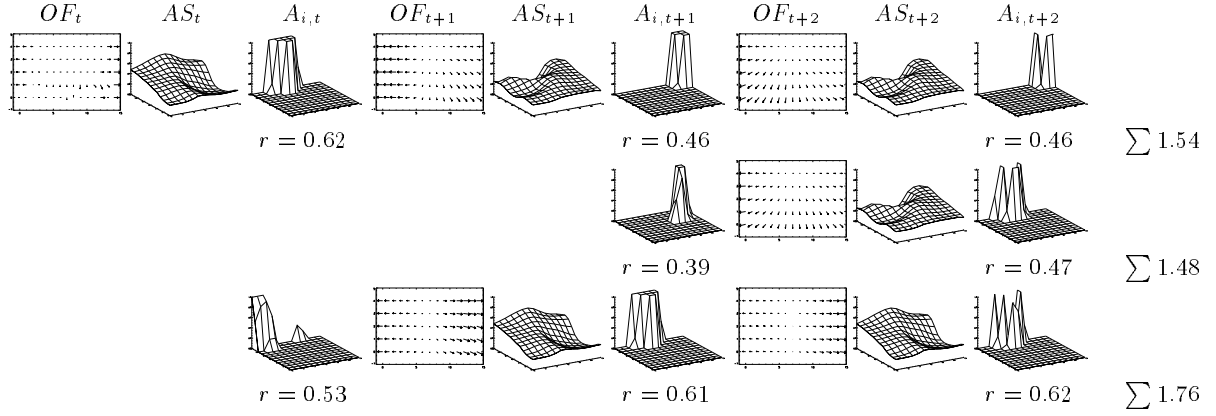


Fig. 7: Parallel-sequential search in sensorimotor space at different time-scales: generation and evaluation of alternative sensorimotor sequences in three chained PMs operating at staggered time-scales. The architecture realizes a dynamic search in depth (x-direction) and width (y-direction) and selects the best evaluated sensorimotor sequence from all simulated ones. The action selected first of all in PM1 (top row left: $r = 0.62$) is the best evaluated action, but it does not consider possible sensorimotor states in the future (reactive behavior). With internal simulation (planning behavior), the motorsensory sequence simulated last in PM1-PM3 (bottom row) can accumulate an higher expected total reward (1.76) than the first one (1.54). Therefore, the first action of the last sequence will be executed in reality. This approach shows anytime behavior.

ly stable clustering of the action space has been reached. To speedup this learning process, the optical flow mapping is initialized as a noisy identity function. This way, the mapping network can more rapidly learn the sensory effects of actions considering the non-linearities of the optical system. The learning of the perceptrons of our mixtures-of-experts approach is based on the delta rule that is controlled by the gating network, the vector quantizer:

$$\Delta \underline{W}_i^P = \eta^P y_i^M [\underline{x}(t+1) - \underline{W}_i^P \underline{x}(t)] \underline{x}^T(t)$$

The quality of optical flow prediction is stored in a position specific competence signal $c_{\underline{r}}$ that is determined by a leaky integration of the local prediction errors $E_{\underline{r}}$, where \underline{r} is the position in the flow field $\underline{x}(t+1)$, and η^c a control parameter:

$$c_{\underline{r}}(t+1) = (1 - \eta^c) \cdot c_{\underline{r}}(t) + \eta^c \cdot 1/(1 + E_{\underline{r}})$$

This competence signal will only increase if the prediction error decreases. The aim of a competence controlled prediction is to prevent an ineffective prediction, because it is not useful to investigate hypothetical situations if the competence for their prediction is too low. This allows a very efficient use of limited computational resources, especially of time, for internal simulation of alternative actions. For example, a selected hypothetical action (winner-blob) is immediately inhibited in the *action selection* subsystem, if the sensory prediction competence for that specific action is too low to proceed with this sequence.

4 Results

First of all, some typical sensorimotor situations, the corresponding optical flow fields, and the suggested actions for the next action-perception-cycle are presented in Fig. 8. Figure 7 shows several anticipated sensorimotor sequences and illustrates the character of the internal simulation and selection process. By

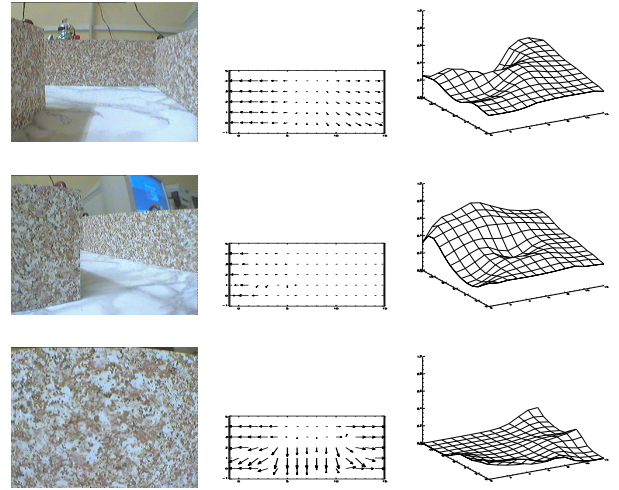


Fig. 8: Results of optical flow estimation and action suggestion: typical sensory situations (left), corresponding, sparse-coded optical flow fields (middle), and suggested action value maps in the action suggestion subsystem (right). Positively evaluated actions generate higher, poor evaluated ones lower activities in these maps. Velocity is coded in y-direction, steering angle in x-direction. In the upper situation the robot avoids left turns (because of the wall), in the middle one it can move straight ahead (free space). The last situation represents an obstacle in front of the robot, therefore no “move forward” hypotheses are generated, merely slow turn left or turn right movements.

means of that, we can demonstrate the differences between reactive and planning behavior in our architecture. Characteristically for reactive behavior is that the action selection only depends on the actual situation and the immediately expected reward. In contrast to this, our anticipating system can consider the future rewards of hypothetical sensorimotor situations, too. By means of internal simulation, it can look ahead

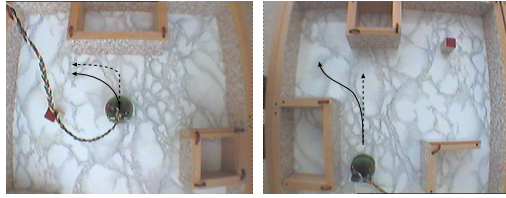


Fig. 9: (Left) Visually guided local navigation as result of anticipation of sensory consequences. The anticipatory agent (full line) moves earlier to the left than the reactive one (dashed line). (Right) The full trajectory shows a local navigation maneuver modulated by a global influence that prefers a movement to the left. This influence can modulate the local search dynamics only, when the obstacle on the left disappears from the field of view. Then the robot moves to the left and follows the global intention. The dashed trajectory is an example for a local navigation without a global top-down modulation. The robot is moving straight ahead until an obstacle is blocking its way.

and select that action sequence, out of all simulated sequences, which yields the highest total reward in the future. Exemplarily, we can show that an anticipating system is able to avoid an obstacle in the environment earlier than a reactive one (Fig. 9-left). Another important aspect, the influence of a global modulation upon the anticipation process, is illustrated in Fig. 9 (right).

5 Outlook

The anticipatory approach presented here leaves room for further development. Although far from fully evaluated yet, it can be considered as an interesting route for further work towards the synthesis of adaptive behavior in truly anticipating neural agents. For this reason, we will investigate in detail, how our neural architecture can evolve an initially reactive behavior into a planning and anticipating one. Basis for this behavior is the ability of our architecture to compose and evaluate unknown sensorimotor sequences, which have never been experienced as a whole before, from experienced and evaluated subsequences, or transitions. Based on this ability to link dynamically, it can be expected, that our control architecture is very flexible in environments that require flexible system goals. For example, for a strong reinforcement learning method, like Q-learning [12], it is very time-expensive to relearn all action values if the reward function is changed, since all possible state-action sequences must be experienced again as whole to determine the new cumulative reinforcements for these sequences. Instead, our approach has to relearn the immediate reinforcements only, new sensorimotor sequences can be composed from known transitions and subsequences which can be interpreted as an *elementary sensorimotor alphabet*. Naturally, this alphabet is highly specific for each agent.

The predictability of sensory consequences of hypothetical actions makes it possible to detect differences between real and predicted flow fields. This enables it to direct the internally visual attention to the contradictory regions in the input, for example, to get more

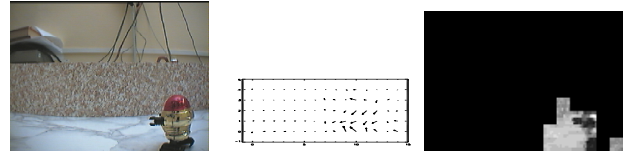


Fig. 10: Performance of the anticipating architecture to find dynamic objects in a dynamic environment (left) by comparing predicted and real optical flow fields. The detected differences (middle) correspond to the dynamic obstacle, the moving small robot. The last figure (right) shows the result of a gating process that transfers the image regions with a high prediction error to subsequent processing levels, e.g. to control selective attention.

detailed information to disambiguate the sensory situation. We use this prediction error for recognition of dynamic obstacles that actively move in the field of view. Fig. 10 shows very first results of this approach to a prediction based control of selective attention.

An outstanding detailed comparison between our anticipating approach and a model-free, strong reinforcement learning, like, for example, the Q-learning [12], is to demonstrate the superiority of generative predictive perception for sensorimotor control.

References

- [1] S. AMARI. Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields, *Biol. Cybern.*, **27** (1977) 77-87.
- [2] T. CAMUS. Real-Time Optical Flow, *PhD thesis*, Brown University, Dept. CS, 1994
- [3] A.P. DUCHON, W.H. WARREN & L. PACK KAEHLING. Ecological robotics: Controlling behaviour with optical flow, 1994
- [4] H.-M. GROSS, V. STEPHAN & M. KRABBES. A Neural Field Approach to Distributed Reinforcement Learning in Continuous Action Spaces, to appear in *Proc. IJCNN'98, Anchorage*, IEEE Press 1998
- [5] R.A. JACOBS, M.I. JORDAN, S.J. NOWLAN, AND G.E. HINTON. Adaptive Mixtures of Local Experts, *Neural Computation*, **3** (1991) 79-87.
- [6] B. KRÖSE, A. DEV, X. BENAVENT & F. GROEN. Vehicle Navigation on Optical Flow, *Proc. of RWC Symp. 1997*, RWC Tech. Rep., TR - 96001, 89-95
- [7] TH. MARTINETZ & K. SCHULTEN. A Neural Gas Network learns Topologies, *Proc. ICANN'91*, 397-402, Elsevier 1991
- [8] R. MOELLER & H.-M. GROSS. Perception through Anticipation, in *Proc. PerAc'94*, 408-411, IEEE Computer Society Press 1994
- [9] R. MOELLER Wahrnehmung durch Vorhersage - Eine Konzeption der handlungsorientierten Wahrnehmung, *PhD thesis*, TU Ilmenau, Dept. CS, 1996
- [10] T. SEILER, V. STEPHAN, A. HEINZE & H.-M. GROSS. Handlungsauswahl durch Antizipation sensomotorischer Konsequenzen, *Proc. SOAVE'97*, Fortschritt-ber. VDI, Reihe 8/663: 109-118, VDI-Verl. 1997
- [11] V. STEPHAN & H.-M. GROSS. Formerhaltende sequentielle visuelle Aufmerksamkeit in columnar organisierten neuronalen Feldern, in: *Proc. DAGM'97*, 411-418, Informatik aktuell, Springer 1997
- [12] C. WATKINS & P. DAYAN. Q-learning, *Machine Learning*, **8** (1992) 279-292