2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Context-dependent dynamic weighting of information from multiple sensory modalities

Alexander Maye and Andreas K. Engel
University Medical Center Hamburg-Eppendorf
Dept. of Neurophysiology and Pathophysiology
Martinistr. 52, D-20246 Hamburg, Germany
Email: {a.maye,ak.engel}@uke.de

***Abstract*— A major problem for the application of sensorimotor approaches to robot control is the classification of states. The typically immense sizes of sensorimotor state spaces render it very unlikely that exactly the same states are visited by the robot several times. In order to learn about the consequences of alternative behaviors in these states, a classification of similar or related states is necessary. This requires a metric to measure similarity between states.**

Under the premise that the robot should maximize its fitness, we studied the correlations between sensory data in different modalities and fitness values. We found that this correlation structure can serve as a context-dependent weighting of the importance of individual sensory channels that allows to define such a metric. In a collision-avoidance scenario we demonstrate that this results in (i) faster learning of successful actions, (ii) an acquired differentiation between sensory modalities, (iii) the possibility to use the full sensors resolution without quantization or compression, and (iv) a means to enhance sensor failure resilience.

I. Introduction

The ability to generalize from previous experiences to new situations is one of the most important features of robotic control architectures. This requires to recognize if a given situation has been encountered before, so that the knowledge from previous experiences can be used to decide which action to take. In reinforcement learning [12], for example, the agent determines the state of the environment as a function of the sensory inputs. Hence the mapping between sensory inputs and states determines the agent's ability to respond to new inputs. However, sensory data are stochastic by nature, and not all changes in the state of the environment can be detected by sensors. This causes a considerable uncertainty in the state estimation that is tried to account for by probabilistic models (e.g., POMDP [1]).

In contrast to symbolic or representation-based approaches to robot control, sensorimotor approaches do not attempt to construct and maintain internal models of the environment such as maps, geometric models or scene descriptions. In the framework of Sensori-Motor Contingency Theory (SMCT, [8]) and its extension (eSMCs, [4]), states are constituted by sequences of sensory observations *and* the movements that caused these observations, hence allowing more reliable state estimation. A key assumption is that relations between movements and depending sensory changes are characteristic for the particular environmental conditions (spatial configurations, presence of objects etc.) and hence sufficient for the agent to adjust its behavior. The difficult problem of constructing high-level symbolic representations of the world state can be eschewed, which is what makes sensorimotor approaches an attractive alternative control schema.

At first the problem of generalization seems to be aggravated for eSMCs, as sensorimotor state space is dramatically expanded by considering movements and recent sequences thereof in addition to sensory data. Re-visiting the same state in this sensation-movement-time space is very unlikely unless some compression measures are taken, e.g. strong quantization or clustering of sensor values, small movement repertoires etc. The ability to generalize the experiences from similar eSMCs is required if learning should be fast, and this is where our motivation for this study originated. We will show that the context given by the set of currently exercised eSMCs allows to weight sensory inputs according to their relevance for the current situation. This effectively collapses sensorimotor state space along irrelevant dimensions and thereby facilitates generalization.

Mastery of eSMCs can be used to explain the different qualities of sensory experiences in humans [8]. According to SMCT, touching feels different from seeing not because the sensory organs are different, or different brain areas become activated, but because the laws that govern the changes of the sensory signals when the agent acts are different. Our study presents a computational model for how an autonomous agent can find out about the informativeness of sensory modalities in different situations and consequently differentiate its senses from an initially synesthetic state.

We study an experimental scenario where the robot learns to perform minimum-jerk locomotion in a rectangular confinement and avoid wall collisions (see Fig.1). The robot is able to move forward, backward, to the left and to the right. This is an extension of one of our previous scenarios where locomotion was only one-dimensional and fewer sensors were used [4], [5]. Control of the robot's locomotion works as follows: Each movement is internally evaluated with respect to the collision state, motor currents, and acceler-

The work was supported by the Sino-German research training group CINACS, DFG GRK 1247, www.cinacs.org, and by the EU project 'Extending sensorimotor contingencies to cognition - eSMCs', IST-270212, esmcs.eu

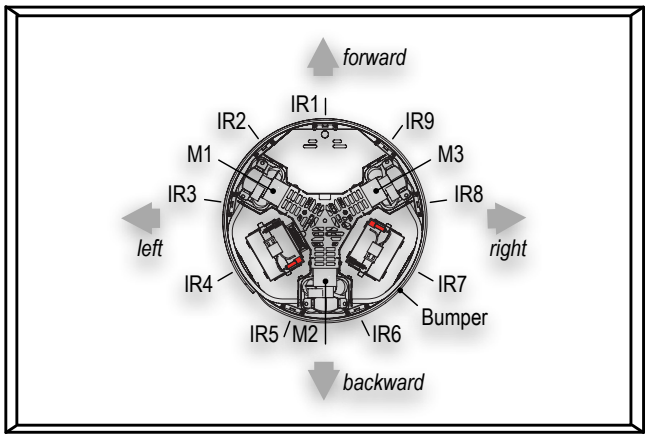

Fig. 1. Top view of the experimental scenario consisting of the Robotino® robot and the rectangular confinement (drawn to scale). The following sensor channels were used: IR1-9: infrared distance sensors, M1-3: motors (current consumption accessible to the robot), accelerometer (x-, y-, and z-direction, not shown), bumper (detects collisions, no directional information). IR1 is at the front.

ations, resulting in a fitness value or utility of the respective movement. Correlating the preprocessed sensor data with the utility magnitude yields information about the relevance of each sensory channel for the fitness of the robot. We used these correlations to weight the distance between the current sensorimotor context and all stored eSMCs. Differences in sensory channels with low correlations are neglected, and only eSMCs that match the current context in the relevant sensory channels are used to plan the next actions.

A simple example may illustrate the context-dependent sensory differentiation. When the robot moves forward, the correlation between its frontal distance sensor and the utility will be high, because low distance readings frequently go along with collisions, while high distance readings mostly result in a good (i.e., high) utility. For a rear distance sensor such a correlation does not exist, and the correlation coefficient will be low[1]. Weighting the Euclidean distance between the actual and stored feature vectors by the correlations for each distance sensor yields low distances for memories that match the front distance well, irrespective of the rear distance. This is a sensible behavior, because these memories are relevant for planning the next action. These eSMCs could be associated, for example, with previous experiences of a collision when the forward movement was continued, and an escape when a turn was taken on time. The interesting point is that without weighting the Euclidean distance by the correlation coefficients, differences in the front and rear sensors would have equivalent effects on the distance measure. For a given position and corresponding sensor readings the robot would consider memories with deviations in the rear distance as relevant as those with deviations in the front distance, and hence may not become aware of imminent collisions.

[1]When moving backwards the opposite is true: The rear distance sensor will have a high correlation, while the one of the front sensor will be low.

Our approach has a number of ramifications. The weighted matching we just described always yields those eSMCs that fit the current context best. This allows to use knowledge from situations that are similar to the current one, effectively constituting a method for generalizing across sensorimotor contexts. This generalization ability allows the exploitation of knowledge about the consequences of behavioral alternatives early on during the exploration phase. In addition, arbitrary resolutions of the sensory channels can be used: Even if the particular values are not matching exactly previous observations, the most similar eSMCs can be used to recognize the current context. And finally, for failing sensors the correlation between sensor readings and utilities becomes very low, what makes this sensor to be ignored henceforth in the search for matching eSMCs.

II. Methods

A. Overview of the eSMCs model

We extend the computational model of eSMCs that we introduced in [3] and used in previous studies with a one-dimensional action space [4], [5]. The model builds on discrete movements m and sensory observations in S sensory channels $\mathbf{o} = [o_1 \dots o_S]$ resulting from m, and considers sequences thereof over a finite history, $[mo(t)mo(t-1)\dots mo(t-h)]$. We call this an eSMCs of history or context length h. At each instant of discrete time t the robot records eSMCs of all history lengths from 0 up to a fixed horizon H. In addition, it computes the utility of its state at this instant as a function of some internal parameters, and stores it together with the set of eSMCs for all history lengths $h = 0 \dots H$. The utility function is a weighted sum of the state of the collision detector (bumper, $[0, 1]$), the average motor current ($[0, 2]$), the maximum increment in motor current ($[0, 2]$), and the acceleration change ($[-2, 2]$):

$$u = -bumper - \sum_{motors} 0.2 motor_{avg} - \sum_{motors} 0.2 motor_{inc} - 0.2 \max_{x,y,z}(|\Delta accel|)$$

This utility function has its global maximum in the trivial state of resting. Therefore the robot was not given the option to stop moving.

The robot stores for each eSMCs the movement $m(t+1)$ that was executed next and the vectors of sensory observations $\mathbf{o}(t+1)$ resulting from this movement. This allows prediction and action planning by forward chaining eSMCs. The utility of an eSMC reflects the aptness of the corresponding movement sequence for achieving a goal. This evaluation allows the agent to direct behaviors towards a goal, and we call such a movement sequence an action [6]. In our implementation eSMCs effectively constitute indexes into a memory of utilities and subsequent movement-observation pairs. For a detailed and more formal description we refer the reader to [3], [4], [5].

B. Sensor-utility correlations and a distance measure for eSMCs

For each movement sequence $\pi = [m(t) \dots m(t-h)], h \in [0, H]$, we compute the correlation coefficient between each sensory channel and the utility $u^\pi(t)$ by

$$r^\pi(s,\tau) = \frac{\langle (o_s(t-\tau) - \langle o_s(t-\tau)\rangle)(u^\pi(t) - \langle u^\pi\rangle)\rangle}{\sigma\left(o_s(t-\tau)\right)\sigma(u^\pi)}$$

Angular brackets denote expectation over repetitions of the same action sequence, and σ is the standard deviation.

These correlation coefficients allow the computation of an importance-weighted distance d between two eSMCs given by the same action sequence π and the sensory observation vectors $\mathbf{o}^i$ and $\mathbf{o}^j$ by

$$d(eSMC_i, eSMC_j) = \sum_{\tau=0}^{h}\sum_{s=1}^{S} \left\| r^\pi(s,\tau)\left(o_s^i(\tau) - o_s^j(\tau)\right)\right\|$$

Differences in sensory channels with a high correlation to the utility are weighted stronger and considered more important than those in channels that bear a low correlation.

In every iteration the distances d between the current set of eSMCs of all history lengths and all stored eSMCs with corresponding history length and movement sequence are computed. If a particular movement sequence π^h has not been carried out before, which is likely for larger history lengths h, the data structures for recording expectations and standard deviations are initialized, and the correlation coefficients are set to 1. Note that the range of sensory feature values o should be normalized if all channels should equally contribute to the distance metric. To allow for changes in the relevance of individual sensory channels over time, we compute all correlations in a sliding temporal window of 100 repetitions of the respective movement sequence.

C. Action selection

The distance measure allows to determine the eSMCs in the robot's memory that match the currently experienced sensorimotor context best. These eSMCs and their associated utilities are used to assess the outcomes of behavioral alternatives from previous experiences as described in [5]. The main idea is to use the information about the immediately following movement-observation pair that is stored for each eSMCs for generating predictions about sensorimotor events several steps ahead in time by forward chaining. Each prediction step is comprised of a movement, a set of possible sensory outcomes, and the associated utilities, and each sequence of prediction steps is considered a potential action. The cumulative utility of an action is given by the sum of the utilities of the individual movements,

$$u(\pi) = \sum_{\Delta t=1}^{T} u(m(t+\Delta t)),$$

with T the prediction horizon. The most promising action candidate is the one with the highest utility:

$$\hat{\pi} = \arg\max u(\pi)$$

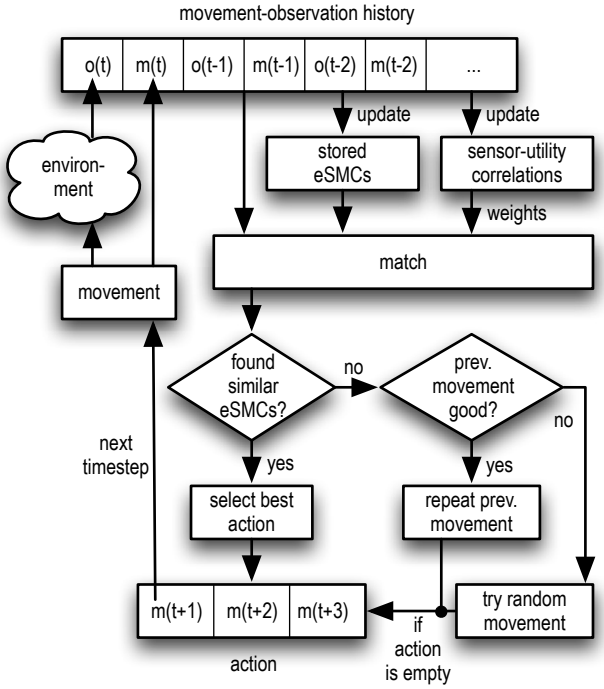

Fig. 2. Schematic overview of the eSMC-based robot control architecture.

The utility of the most promising action $u(\hat{\pi})$ is transformed into a probability $p = u(\hat{\pi})/T + 1$, and $\hat{\pi}$ gets into execution with this probability. Otherwise a new behavior is explored. In this manner actions of high utility are readily executed, while the likelihood for exploring alternatives is higher when the utility of the best known action is moderate. The robot follows the action plan until each movement was executed or the action is updated by new predictions. In order to curb the combinatorial explosion, only the 10 eSMCs with the closest match to current context are considered for generating predictions, and the prediction horizon is limited to $T = 2$.

We added one heuristic for the case when no action plan is available. This may happen because all movements of the last action plan have been carried out, due to lack of eSMCs that match the current context, or during a random switch to the exploration of new actions. This heuristic lets the robot continue the movement from the previous time step with a probability that is proportional to the utility that resulted from that movement. Otherwise a random movement is initiated. This lets the robot start with a momentum-driven behavior, which becomes gradually differentiated. The repetitive movements enable matches of larger context sizes already from the beginning of the learning phase with a notable effect on the learning rate. A schematic of the robot's control loop is shown in Fig. 2.

III. Results

We let the robot roam around the rectangular confinement, recording eSMCs and exploring the consequences of different behaviors. The trajectory in the left panel in Fig. 3 shows that during the first 1.5 minutes (180 epochs)

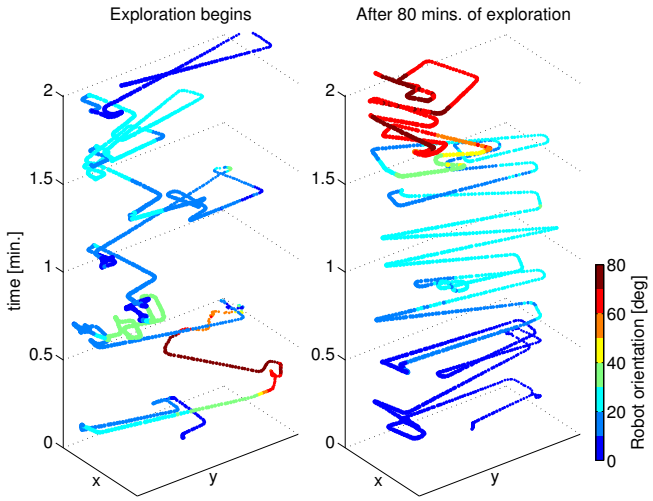

Fig. 3. Example trajectories of the robot at the begin of learning (left) and after substantial exploration (right panel). Z-axis is used to resolve movements over time, color represents orientation with regard to the main axis of the rectangular confinement.

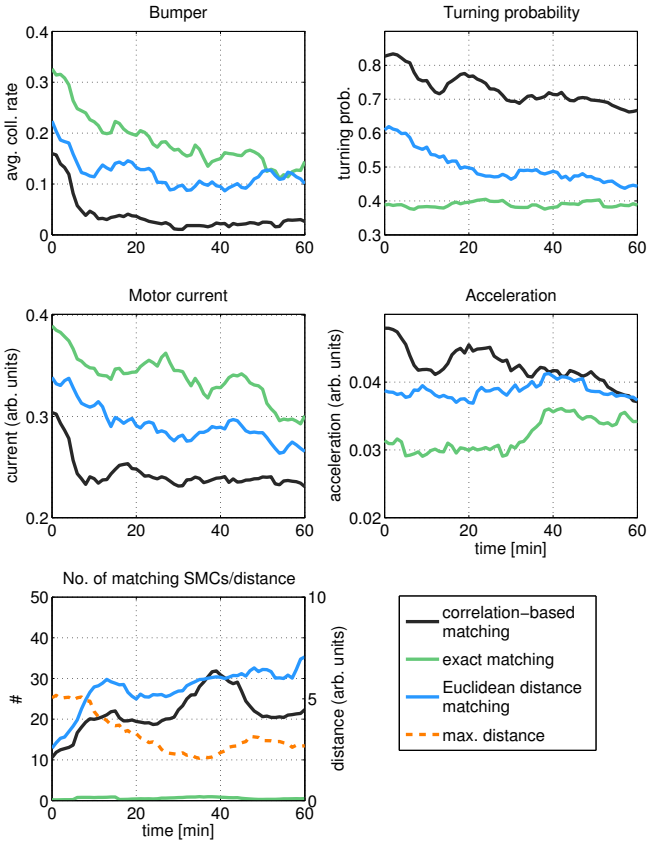

Fig. 4. Characterization of the robot's behavior when eSMCs for action planning are matched using the correlation-based distance metric (black curves), the Euclidean distance metric (blue curves), and exact matches (green curves). Curves show averages in a sliding window of 10 minutes

the robot spends most of the time exploring adequate reactions to wall collisions. Continued interaction with the walls also frequently changes the robot's orientation. The momentum heuristic in the action selection provides for straight movements between opposite walls. After about 1.5 minutes the robot increasingly becomes aware of imminent collisions and reacts with avoidance actions. The spiraling movements are not yet the energetically optimal solution, though. After about 70 minutes (8.400 epochs) the robot has learned energy-efficient collision-avoidance movements (see right panel of Fig. 3).

The color code in Fig. 3 shows that the orientation of the robot with regard to the main axis of the confinement changes by interacting with the walls. Pushing against a wall when the contact point between the circular periphery and the wall is not aligned with the force vector generates a torque that rotates the robot. Different orientations feature different sets of eSMCs, however. As a consequence, the robot's sensorimotor space has an orientation dimension, despite the fact that it does not dispose of actions that directly change orientation. This may explain the apparently long learning time until the robot can generally avoid collisions.

A. *Faster learning*

Strictly speaking the speed of learning is determined by the number of experiences or iterations per time the robot makes, and this is not changed by the distance metric we introduce. The metric provides however a means for generalization, allowing the robot to make use of experiences from similar situations and hence to show a behavior that makes the impression as if it learns faster indeed. Parameters that characterize the success of the robot's actions are shown in Fig. 4, and controllers using the correlation-based eSMCs matching, a simple Euclidean distance-based matching, and the exact matching are compared. The plots clearly show that the robot controlled by the correlation-based eSMCs matching can minimize its motor current and the collision rate already during the first 10 minutes of exploration. This required to switch the movement direction more frequently. Over time the robot learns to avoid unnecessary turns. The constantly low turning rate for the exactly matching eSMCs controller reflects the effect of the momentum heuristics in the action selection schema. Since this controller does not generalize, it does not switch movement directions in collision-prone situations that it has not experienced previously, resulting in the higher collision rate. The performance of the controller that uses the Euclidean distance to find similar eSMCs lies between the other two controllers. Depending on whether the non-matching sensory channels are relevant for the particular situation or not, the knowledge in the eSMCs with the lowest Euclidean distance may be appropriate or not for avoiding or escaping collisions.

The last panel in Fig. 4 shows how the number of similar eSMCs matching each situation grows over time. This reflects the growing sensorimotor knowledge which the robot acquires. The maximum correlation-based distance of similar eSMCs in each context does not seem to change much between the beginning and the end of the shown exploration period. The number of eSMCs, however, grows from about

10 in the beginning to more than 30 at the end, indicating that the action planning can draw on gradually more specific context knowledge. Using the Euclidean distance metric, approximately the same number of similar eSMCs are available. Since they may mismatch in sensory channels that are relevant in the current situation, the information about behavioral alternatives may be unreliable. Hence the agent is not able to select correct actions to the same extent as with the correlation-based similarity measure. In agreement with observations in our previous studies the number of exactly matching eSMCs grows significantly slower.

B. Differentiation between sensory modalities

Analysis of the correlations for the different sensory channels in the various action contexts after substantial exploration time reveals the relevance of each channel for selecting appropriate actions in each context. Fig. 5 shows examples of three representative movement sequences after about 80 minutes of learning eSMCs. When the robot has moved forward for 3 times steps (first row), the most relevant sensors predicting the utility of the current situation were the frontal distance sensors (IR1,2,9) two time steps ago (left panel). In the current time step their importance is lower (right panel), in favor of a higher importance of the collision detector and the motor currents (M1 and M3 are used for moving forward/backward). The relevance of these channels in the previous time step is in between (middle panel).

When moving backwards (second row), the most relevant distance sensors are the ones at the back (IR5,6), and the accelerometer signal (accelX) is reversed[2].

Closer inspection of the correlations for the distance sensors during these two actions shows non-zero, inverse correlation coefficients for the sensors opposite to the movement direction. This results from episodes when the robot's orientation was turned to about $90°$, i.e. like in Fig. 1. In this configuration the robot moved forward and backward along the short axis of the confinement, and both walls, at the front and at the back, were in range of the distance sensors all the time.

The last row in Fig. 5 shows the correlation structure when the robot changes movement from forward to right. As expected the most relevant distance sensors are the ones at the right periphery of the robot (IR7-9), and the correlation weighting shifts from the more frontal part to the lateral region over time.

C. Data resolution and the size of sensorimotor space

When matching the current sensorimotor context to stored eSMCs using the proposed distance metric, exact correspondence of the sensory feature values is not required. If the robot has explored a given action before, the correlation-based weighting schema yields the corresponding sensory experiences ordered by their similarity to the current sensory features. This should allow using arbitrary resolutions of the sensory features, because the distance-based matching will always find similar eSMCs among the stored eSMCs, irrespective of their absolute distances to the current context.

[2]The absence of any correlation for motor M2 on backward movements at times t and $t-1$ came as a surprise. We conjecture that the differing control of M2 in forward and backward movement commands is a peculiarity of the robot's odometry.

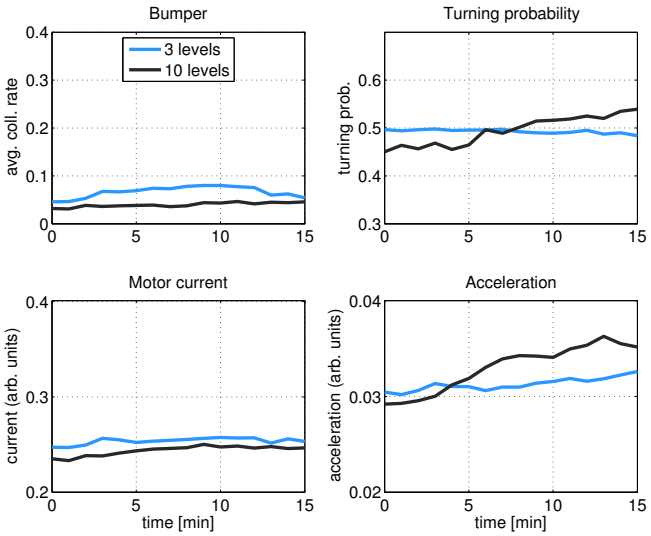

Fig. 6. Behavioral parameters (cf. Fig. 4) after learning eSMCs for 80 minutes when distance sensor readings were quantized to 3 (blue curve) and 10 levels (black curve). Despite the much larger sensorimotor space for 10 levels per distance sensor, the behavior is qualitatively the same, with an even somewhat lower collision rate.

We tested this hypothesis using 10 different levels for each of the 9 distance sensors instead of only 3. These high-resolution distance readings provide the robot with a much more precise estimate of its position with respect to nearby walls. At the same time it makes it extremely unlikely that the robot returns to exactly the same position associated with one of the stored distance patterns. The correlation-based matching yields eSMCs from places in the vicinity of the robot's current position. It is important to note that this does not mean ordinary Euclidean vicinity, but task- or context-dependent vicinity. When going forward, for example, the most similar eSMCs with regard to the frontal distance sensors will be considered, while the distances at the back are disregarded.

In Fig. 6 we show that the distance-based matching succeeded in identifying stored eSMCs with relevance for action planning even though exact matches could not be expected in the high-dimensional distance feature space (10^9 vs. 3^9).

D. Detecting and compensation sensor failures

In the trained robot a transient malfunction of one of the distance sensors was simulated by replacing the sensor readings with random values. We consider this type of failure much harder to detect than a simple shut-down of a sensor. Fig. 7 (left panel) shows that the correlations of the failing sensor with the utility of eSMCs of all history lengths drop to around zero, effectively disabling the use of this sensor for matching the current state with stored eSMCs. After the sensor resumed function, the correlations return

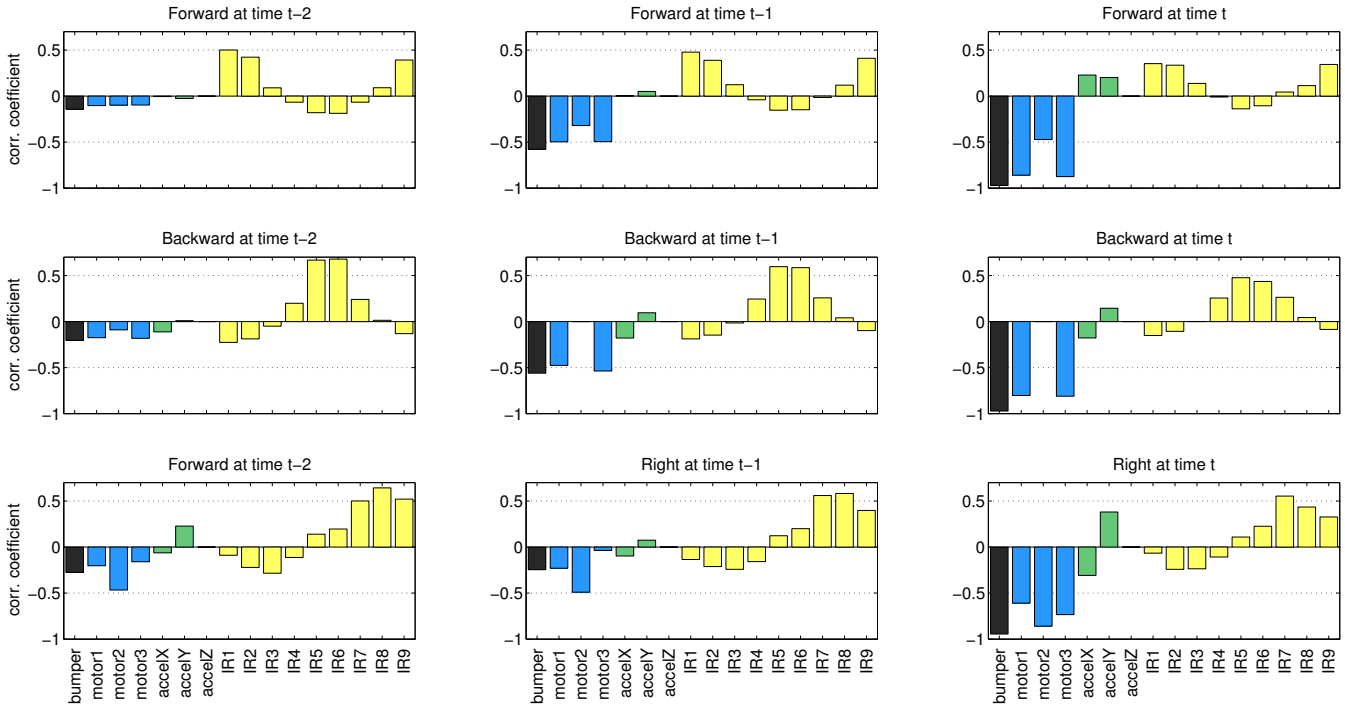

Fig. 5. Correlation structure of all sensory channels for three example movement sequences: moving forward (top row), moving backwards (middle row), and switching from moving forward to moving right (bottom row). For the sensor positions see Fig. 1.

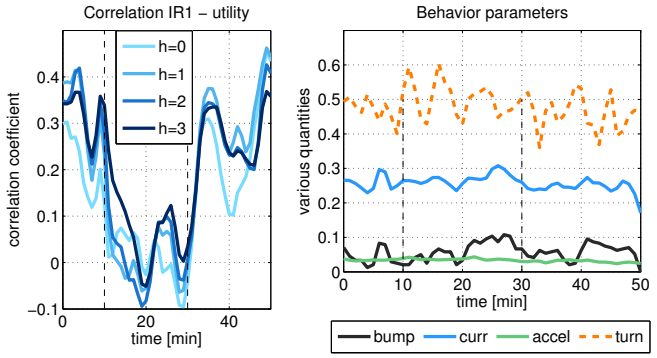

Fig. 7. After training, the robot was run for another 50 minutes with a simulated sensor malfunction in IR1 between 10 and 30 minutes. Shown are correlations between IR1 sensor features and utilities in eSMCs of different history lengths for moving forward (left panel) and time courses of parameters that reflect the robot's behavior (right panel).

to their previous levels. The simulated sensor malfunction transiently increases the robot's turning probability, but does not significantly affect the parameters that determine the fitness of the robot (right panel).

IV. Discussion and Related Work

In this study we presented a correlation-based metric for eSMCs and explored its application for autonomous robot control. The main contribution is a model for autonomous learning of sensory differentiation in robots. It allows to select a subset of sensory channels with the highest relevance in a given situation. This effectively constrains high-dimensional sensorimotor spaces to relevant subspaces and facilitates the search for previously made similar experiences that can be used for action planning.

Learning in high-dimensional sensorimotor spaces is treated in reinforcement learning as a value function approximation problem. In the majority of existing methods (see [12] for an overview) the action context is not sufficiently taken into account, though. Engaging context information in value function approximation methods bears some potential to speed up learning and to arrive at more context-specific value estimates.

A hierarchical neural network model that uses sensor-reward correlations for action selection is described in [9]. It uses the correlation between sensory features and rewards to bias the top-down propagation of states towards actions that promise good rewards. The correlations are not context-specific, though, and the structure across sensory channels is not analyzed. Our study shows that this correlation structure reflects an agent's knowledge about the importance of different sensory modalities in different tasks and that the agent can acquire this knowledge autonomously. This allows to employ the eSMCs-based controller in different robotic embodiments and hence makes it a kernel in the sense of [2]. The only parameter is the size of the temporal window over which the correlations are computed, controlling the robot's adaptation time to changes in the sensor properties.

An important feature in this respect is the possibility to handle sensor failures. In our approach failing sensors are not simply shut off. They are used to the extent they yield useful information and are re-enabled automatically if they resume function.

Apart from learning sensory differentiation our approach can be considered as a method for approximating a value function. Tile coding is a solution to this problem which is used in the context of reinforcement learning [11]. The value function is approximated by a weighted superposition across a set of tilings with different offsets, and the value of individual tiles is updated with samples that the agent generates. Adaptive tile coding does not require to specify number and resolution of tilings in advance and adjusts these parameters to the properties of individual value functions [13]. The main difference to our approach is that samples of the value function are used to update information in corresponding tiles, whereas our approach simply keeps all samples and determines the set of nearest neighbors at the selected location in sensorimotor space at the time the approximation is needed.

The results of our analysis of the correlation structure may be considered surprising in so far as they do not show a differentiation into physically distinct sensor classes, but that sensors are combined in a context-dependent manner. For moving forward, for example, the robot pays attention to the collision detector, the current consumption of the two driving motors, and the frontal distance sensors, while accelerations and signals from the rear distance sensors are ignored. The different combinations of sensors monitored in different contexts can be considered as "perceptual modalities", which may give a better account of human perception than explanations building on the anatomical or physiological differences of the senses [7].

Using the Pearson correlation coefficient in this study should be seen as just an example of measuring the relevance of sensory channels for the agent's fitness. This measure is certainly limited for it captures only linear relations. It may be appropriate for the sensor types that were used (distance, current, acceleration etc.), and that have mostly linear properties. The aptness for sensors with rather non-linear characteristics, e.g. pixels of a camera image, will be investigated in future studies. We expect that the correlation structure will not account for a detailed vision system, but may be useful for low resolution vision in simple environments, when single colors or brightness gradients bear sufficient information for the agent. Alternative measures like mutual information, transfer entropy, or causality measures may be better suited to capture more complex relations between sensors, actuators, and utilities. Transfer entropy, for example, can reveal directed, non-linear relationships in sensorimotor space that are mediated by the environment [10]. Developing our approach in this direction seems promising.

REFERENCES

[1] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101:99–134, 1998.

[2] F. Kaplan and P.-Y. Oudeyer. From Hardware and Software to Kernels and Envelopes: A Concept Shift for Robotics, Developmental Psychology and Brain Sciences. In J.L. Krichmar and H. Wagatsuma, editors, *Neuromorphic and Brain-Based Robots: Trends and Perspectives*, chapter 10, pages 217–250. Cambridge University Press, 2010.

[3] A. Maye and A.K. Engel. A discrete computational model of sensorimotor contingencies for object perception and control of behavior. In *2011 IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 3810–3815. IEEE, May 2011.

[4] A. Maye and A.K. Engel. Time scales of sensorimotor contingencies. In H. Zhang, editor, *Brain-Inspired Cognitive Systems 2012 (BICS 2012)*, volume 7366 of *LNAI*, pages 240–249. Springer-Verlag Berlin, Heidelberg, 2012.

[5] A. Maye and A.K. Engel. Using sensorimotor contingencies for prediction and action planning. In T. Ziemke, C. Balkenius, and J. Hallam, editors, *From Animals to Animats 12*, volume 7426 of *LNCS*, pages 106–116, Berlin Heidelberg, 2012. Springer.

[6] M. McGann. Enactive theorists do it on purpose: Toward an enactive account of goals and goal-directedness. *Phenomenology and the Cognitive Sciences*, 6(4):463–483, 2007.

[7] Marek McGann. Perceptual Modalities: Modes of Presentation or Modes of Interaction? *Journal of Consciousness Studies*, 17(1-2):72–94, 2010.

[8] J.K. O'Regan and A. Noë. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24:939–1031, 2001.

[9] D. Rawlinson and G. Kowadlo. Generating Adaptive Behaviour within a Memory-Prediction Framework. *PLoS ONE*, 7(1):e29264, 01 2012.

[10] N.M. Schmidt, M. Hoffmann, K. Nakajima, and R. Pfeifer. Bootstrapping perception using information theory: case studies in a quadruped robot running on different grounds. *Advances in Complex Systems*, 16(01n02), 2012.

[11] A.A. Sherstov and P. Stone. Function Approximation via Tile Coding: Automating Parameter Choice. In Jean-Daniel Zucker and Lorenza Saitta, editors, *Abstraction, Reformulation and Approximation*, volume 3607 of *Lecture Notes in Computer Science*, pages 194–205. Springer Berlin Heidelberg, 2005.

[12] R.S. Sutton and A.G. Barto. *Introduction to Reinforcement Learning*. MIT Press, Cambridge, MA, USA, 1st edition, 1998.

[13] S. Whiteson, M. E. Taylor, and P. Stone. Adaptive Tile Coding for Value Function Approximation. Technical Report AI - TR - 07 - 339, University of Texas at Austin, 2007.