

Extension of the ALVINN–Architecture for Robust Visual Guidance of a Miniature Robot*

M. Krabbes, H.–J. Böhme, V. Stephan, H.–M. Gross
Technische Universität Ilmenau
Fachgebiet Neuroinformatik
D-98684 Ilmenau, Postfach 100565
Markus.Krabbes@Informatik.TU-Ilmenau.DE

Abstract

Extensions of the ALVINN–Architecture are introduced for a KHEPERA–miniature robot to navigate visually robust in a labyrinth. The reimplementantion of the ALVINN–approach demonstrates, that also in indoor–environments a complex visual robot navigation is achievable using a direct input–output–mapping with a multilayer perceptron network, which is trained by expert–cloning. With the extensions it succeeds to overcome the restrictions of the small visual field of the camera by completing the input vector with history–components, intrduction of the velocity dimension and evaluation of the network’s output by a dynamic neural field. This creates the prerequisites to take turns which are no longer visible in the actual image and so make use of several alternatives of actions (f.e. at crossings).

1. Introduction and scenario

Topic of the project GESTIK is to develop a neurally based control architecture for a mobile robot to navigate visually while maintaining "eye contact" with an operator to follow his (gesture based) orders [1]. Static and dynamic obstacles on the route are to be avoided suitably. For the intended performance it is significant, that only local behavior is feasible, because the real environment never appears unambiguously with respect to the global position in the operating field (no unambiguous landmarks). This is no restriction in context of the project’s topics, because through cooperation of the operator and the vehicle the desired performance is achievable completely and unequivocally. This interactivity is achieved by a heterarchic structure of agents to represent the complete situation–specific action–space in a separable manner. These several agents are physically identical as described in detail in this paper, but trained with different intentions [2].

Target system is the robot MILVA ([http://](http://cortex.informatik.tu-ilmenau.de/technik.html)

cortex.informatik.tu-ilmenau.de/technik.html), which is equipped with a triocular vision system and on-board-PC. Presently the miniature robot KHEPERA (round, $\phi=55$ mm; central color camera) serves as experimental platform for the investigations introduced here, whose practical application proceeds remarkably unproblematic and can substitute simulations well.

The idea of ALVINN [3] is to use direct feed forward processing of a camera–picture to a steering angle by a two layer multilayer perceptron (MLP) with a small number of hidden units (about 4) to steer a street–vehicle on different kinds of roads. Therefore the images of a car–mounted camera were recorded together with the corresponding steering manipulations of the driver on extended trips to teach the network this relation by backpropagation ('expert–cloning'). The analogous value of the appropriate steering angle was presented by the activations of a whole vector of output neurons in topological coding instead of coding by a single neuron.

2. Realization of the ALVINN–approach on KHEPERA

For the control architecture of KHEPERA a two–layer–perceptron in exclusive feed–forward structure is used as network too. Preprocessed and subsampled images form the input vector of the network, read from the KHEPERA–PAL–camera with a framegrabber.

Due to the wooden walls on a light blue ground, the labyrinth appears within sharp blue–yellow contrast. Therefore it is suggested to perform a conversion of the original picture into the blue–yellow–activations of the physiological color space according to [4]. A suitable sigmoid function spreads the values to a range $(-1 \dots 1)$. (fig. 1) The network has to map the situation–adequate steering angle in topological coding. Because KHEPERA with its two driven wheels has no explicit control parameter "steering angle", it is to be calculated from the speed difference of the two wheels ($LS=LeftSpeed,RS=RightSpeed$) by

*These studies are part of the project GESTIK supported by the Thuringian Department of Science, Research and Culture.



Figure 1. The steps of image preprocessing for KHEPERA: the original picture (left), the blue-yellow-activations (2nd f. left), after threefold subsampling (2nd f. right), and dynamic adaptation (right).

normalization to the total velocity. The following equation is applied to map this angle onto a m -dimensional output vector representing a Gaussian of M Neurons width (fig. 2):

$$y_i = e^{-\frac{(i - \frac{m+1}{2} \cdot (1 + \frac{LS-RS}{LS+RS}))^2}{M \cdot \frac{LS+RS}{LS+RS}}} \quad i \in [1 \dots m]; \quad (1)$$

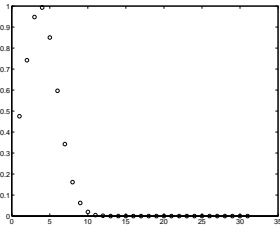


Figure 2. One-dimensional topologic coding of the steering angle: Graph for $m=31$, $M=10$, $RS=5 \times LS$

The multilayer perceptron is trained with a database containing 2500 samples of input - output patterns, recorded from the exemplar behavior of a human trainer ("expert") and preprocessed as described above. It has to be pointed out that the navigation while recording proceeds exclusively based on the video-images of KHEPERA's on-board-camera to guarantee consistent data sources in training and recall. Therefore the expert steers the robot using a video head-set.

During the network-teaching the whole database is repeatedly presented in mixed order. In this way it is possible to use direct learning instead of batch learning without disadvantages in respect to both robustness and learning speed. As number of hidden units, 6 proved to be suitable. Finally, the proper quantity of hidden units should be determined empirically. An unnecessarily high number can only be identified by the development of quasi identical weight patterns in the weight diagram of the network.

As a useful number of cycles to repeat the training set, only a range between 10 and 50 can be specified. Although a considerable minimization of the error, integrated across the training set, becomes obvious in this range, there is no crucial improvement of the subjective behavior with extended training. Because of the considerable expenditure to

create training examples, the data base for validation is limited and therefore not able to completely avoid overfitting. It should be noted here that no benchmark was developed yet to compare different training sets, teaching cycles, network configurations and so on. Perspectively, a test must be realized, which is based on recording the real behavior combined with an evaluation of yet to be defined criteria. Figures 3 and 6 are first results of these procedures, visualizing a tracking of a colored label on top of the KHEPERA.

To use the network, firstly the centre of gravity of the output activation was calculated to project it to a relation of wheel speeds, which corresponds to the maximum of the Gaussian at this position. This approach is simplified as compared to the methods of [3] (best correlating Gaussian) and [5] (only a window for the neighbourhood of maximum) but yields sufficient results. As a constant basic speed, a usual value from the teaching phase was chosen. KHEPERA masters all good-natured situations, even in an unknown labyrinth, largely free of collisions and shows the expected behavior. It lines up centrally in alleys; if only one wall is visible, it attempts to follow the wall with a lateral distance typical for alley passages (fig. 3). It is necessary to keep parts of the floor in the visible range of the camera permanently. While doing this, the restricted area of view appears very limiting (fig. 4).

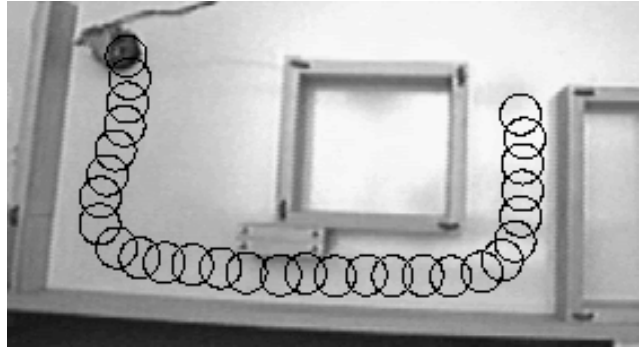


Figure 3. The achievable behavior of KHEPERA with the original ALVINN-approach (circle=KHEPERA-base).

It is possible, that the vehicle crashes into obstacles which it does not see at this moment. That's why the expert uses special behavior while creating the training set. Typical for that is to go straight ahead as far as possible and to turn away only if the way in front of the vehicle is blocked. With the necessity to keep a part of the floor always in the image, the radius must not fall below 10 cm (outer wheel) in bending alleys. With that there is no exit from blind alleys closer than 20 cm. It turns out, that also the vertically limited visible field has a negative effect because the estimation

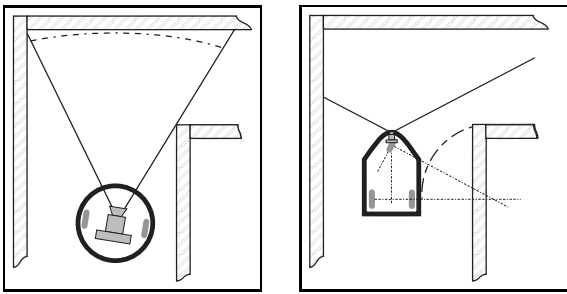


Figure 4. The restricted visible field of the KHEPERA-camera limits considerably the achievable behavior (left), comparable to conditions for the MILVA-robot (right).

of distances to obstacles is based on recognition of the position on the floor.

3. Extensions of the ALVINN-approach

3.1. Path window

Because the same problems will appear for the MILVA-robot with its traktix from 3-wheel-kinematics in a comparable manner (fig. 4), structures are to be developed already in the KHEPERA-scenario to create adequate behavior through a sensomotory projection. Since the previous images of the path history have equal importance for the navigation behavior, a path window was introduced. This seems more appropriate compared with the use of recurrent networks or dynamic neurons, respectively (fig. 5). This structure is a simplified application of the "Sliding-Window"-technique in TDNN according to [6]. In this path window, previous input vectors (images) are presented in addition to the actual one. For such an expansion of the network input layer, it is important to use a constant distance for the images presented simultaneously. To realize this also with variable speed, the input vectors are saved up to a determined horizon together with the step distance belonging to them. From this buffer distance-equalized input vectors are chosen. To teach the network with such an input history, a database of training sets was created, which consists of a continuous sequence of training patterns. In this database, a training sample at any position can be pointed to by reading of a section according to the dimension of the path window.

It was interesting to observe that even a human expert does not succeed in turning around a corner with the current picture only. The action-reaction-feedback from *target of motion* – *KHEPERA-movement* – *change of the picture* is so mediate, that only extended practice leads to some routine. That's why it was necessary to complete the senses of

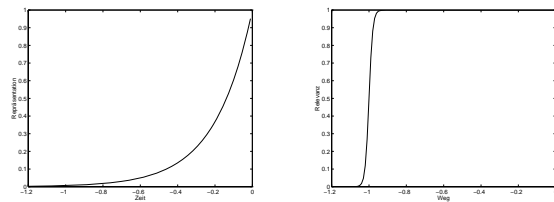


Figure 5. Whereas recurrent network structures represent past information comparable to PT1-behavior (left) the relevance of picture information keeps on constant high level up to a path corresponding to the vehicle's dimension and completely collapses after that (right).

the driver by 2 "senses of touch" in addition to the visual field. The distance value of the 2 outer lateral IR sensors of KHEPERA are passed on to the driver by headphones as stereo noise. But this procedure contains a latent danger; that the behavior of the teacher is dominated too much by this additional, for the network not usable modality (IR-sensors) leading to inconsistencies between training and network recall.

Application of this extension shows the desired success not only in lateral positioning to obstacles or changing alley widths, even those which are no longer visible in the actual image (Figure 6), but also in reduction of ambiguous situations. Now decisions how to move are possible also in positions close to a wall, based on the information in the path history (Figure 7). Moreover, now there are no problems to lose the "eye"-contact to the floor temporarily in sharp curves. As a number of input vectors in the path window, 3 proved suitable in a distance of about 3 cm (conforms to half KHEPERA-diameter). More history layers increase the number of adaptive weights in the network too much, higher distance of the vectors causes too much variety to be retraced by the network in a sensomotrical coincidence.

Generally the introduction of the path window is a leap in quality compared to the original architecture because the sensory basis is created to make use of alternatives of actions (of driving forward) possible, f.e. driving along turns which are not longer visible in the actual image. So it becomes feasible to create several training sets based on different action incentives to achieve agents with different intentions of behavior.

3.2. Dimension of speed

Simultaneously to the extensions described above, a second dimension was integrated in the action space to represent the vehicle speed according to behavior of the expert. Therefore, the previous m -dimensional output vector was converted to a $m \times n$ -dimensional one to code steering angle

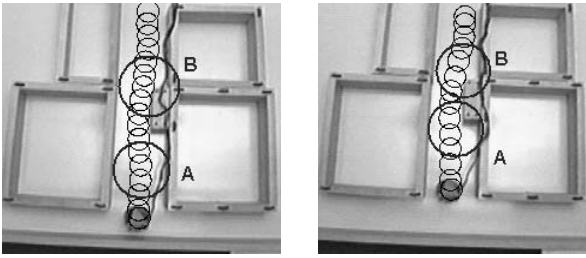


Figure 6. The recording of the vehicle-behavior near a lateral obstacle in the labyrinth shows the achieved progresses using the path window: The original network is forced to evade very early (A) and reverts to the alley centre prematurely (B) (left). With the additional information from the path window both evading (A) and returning (B) take place with the expected distance to the obstacle (right).

and speed in a two-dimensional (about $M \times N$ dimensioned) Gaussian (eq. 3, fig. 8).

$$y_{ij} = e^{-\frac{(i - \frac{m+1}{2} + (1 + \frac{LS-RS}{LS+RS}))^2}{M}} \cdot e^{-\frac{(j - \frac{N}{2} - \frac{LS+RS}{LS_{max}+RS_{max}}(n-N))^2}{N}} \quad (2)$$

$i \in [1 \dots m]; j \in [1 \dots n]$

The afferent (in each case top) and efferent (in each case bottom) weights of 6 hidden units of a trained network with 2-dimensional output coding without path window are depicted in the weight diagram fig. 9. In the weight patterns of the input layer the structure of the input data appears clearly, the output layer shows strongly excitatory (light) and inhibitory (dark) weights in the typical regions of the action space (stop at $v=0, \phi=0$: in the action map central at bottom).

3.3. Focus of activation

From a methodological point of view the two-dimensional representation offers crucial advantages compared to a conceivable double one-dimensional mapping of steering angle and speed, because multiple action suggestions are **separable** across the activity distribution of output neurons. With the resulting output activation, a dynamic neural field according to [7] and [8] can be appropriately stimulated. Hence, the desired effects can be achieved, which are included in a "intelligent" interpretation of the network's output:

1. Mechanisms of selection of maxima focus on the region of highest intensity. For this the region has to

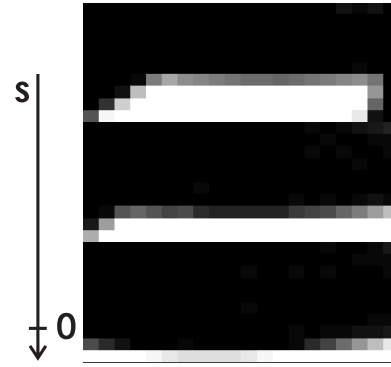


Figure 7. Path window with 3 input vectors (3 x 10 image lines): While the actual vector (bottom) represents an ambiguous situation, a turn to the right is recognizable with the additional information from the vectors past.

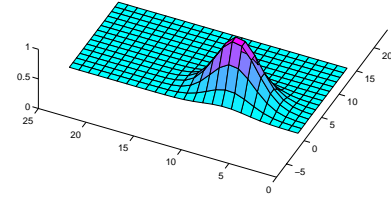


Figure 8. Two-dimensional topologic action coding for a slow turn right in a 10 x 5 dimensional Gaussian at a 25 x 15-dimensional output vector (as mesh-plot).

exceed its neighborhood both in spatial expansion and activation. The activation of all non-supported regions is inhibited.

2. After selection of a local maximum it will be focused on, even if its activity falls below that of others, up to a certain limit. Such cases appear when the actual situation, represented in the input vector, was not or hardly contained in the training set of the network, caused by noise effects, dynamic obstacles or momentary loss of the ground in the actual view. This hysteresis quality leads to a kind of perseverance avoiding the loss of focus during short-term crashes of local activity.
3. Even during a movement of the local activation its centre will still be matched. With this, the never stationary "blobs" on the output layer can be tracked permanently.

The neural field consists of a two-dimensional layer of first order dynamic neurons (time constant τ), whose activities

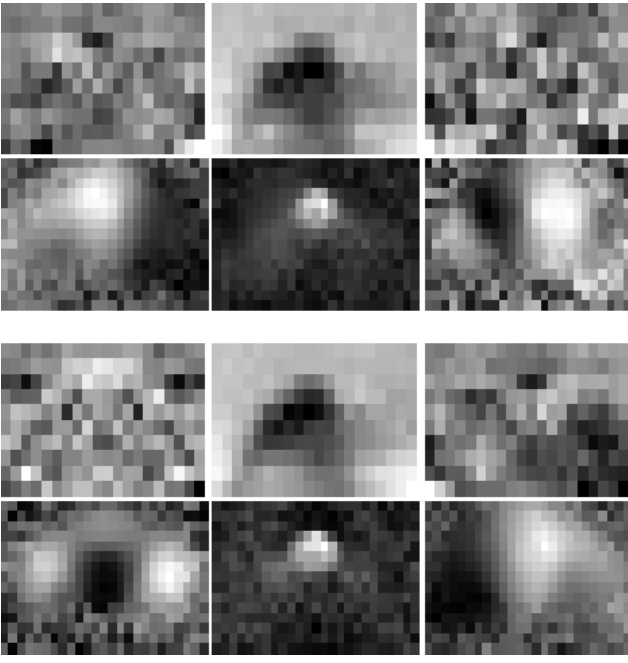


Figure 9. The weight diagram of a 200-6-375–network, see text for explanation.

are changed according to the following differential equation:

$$\tau \frac{d}{dt} u(\vec{i}, t) = -u(\vec{i}, t) + h + x(\vec{i}, t) + \int_R w(\vec{i} - \vec{i}') S[u(\vec{i}', t)] d^2 i' \quad (3)$$

The change of the state of a neuron at position \vec{i} in the assembly is a function of its previous state $u(\vec{i}, t)$, the global inhibition $h < 0(H_0)$, the sum of all neighbouring neurons weighted by the distance function $w(\vec{i} - \vec{i}')$, the threshold function $S(\cdot)$ and last but not least the input pattern $x(\vec{i}, t)$. For simulation the differential equation is suitably approximated with the one step technique according to EULER and CAUCHY using step width ΔT in the following time discrete equation for the state activities:

$$z_{ij}(k+1) = (1 - \alpha)z_{ij}(k) + \alpha \left\{ w^I x_{ij} + \int_R w(\vec{i} - \vec{i}') S[u(\vec{i}', t)] d^2 i' \right\} \quad (4)$$

with:

$$\alpha = \frac{\Delta T}{\tau} \quad (5)$$

$$w(\vec{i} - \vec{i}') = C \cdot e^{-\frac{\|\vec{i} - \vec{i}'\|^2}{2\sigma^2}} - H_0 \quad (6)$$

$$S(z) = \frac{1}{1 - e^{-\frac{z-0.5}{0.5}}} \quad (7)$$

Equation 6 defines the lateral distance function $w(\cdot)$ with local excitatory and global inhibitory effect such that only one compact cluster can succeed. As threshold functions serves a suitable sigmoid function (eq. 7).

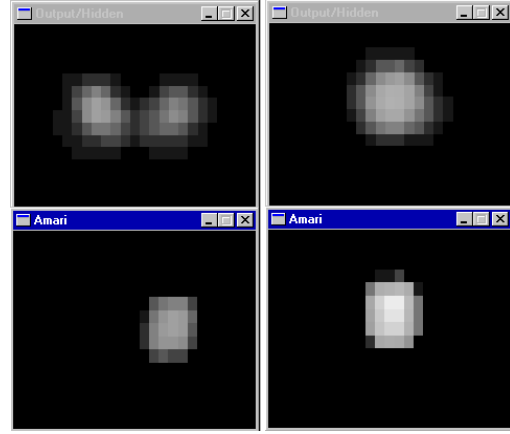


Figure 10. The activation of the output layer of the net (each on top) and the neural field operating on it: a conflict situation on the left and a non-ambiguous situation on the right.

Figure 10 demonstrates the selective effect of the neural dynamics. But the static image can hardly convey the real behavior of this neural layer, as especially its time series of activation is symptomatic for the global behavior. On the basis of the activation of the dynamic neural field, a drive command could be generated without mistakes by determination of the centre of gravity. However, the best results were achieved by gating a local region of the original output activation through suppressing all other neurons, whose corresponding units in the dynamic layer were inhibited. Because of the high resolution in the output coding the position of the maximum activation in this gated window can be used directly with suitable low pass filtering afterwards.

4. Outlook

The used dynamic neural field will gain further significance for the fusion of intentionally different taught agents to represent the complete situation-conform repertoire of actions (see section 1.). The fusion proceeds in that their gated output activations are overlaid appropriately (maximal activation in easiest case) to stimulate a similar neural layer to focus in a energetic centre for selecting the finally executed action. A decomposed multi-agent-structure, as proposed here, offers the opportunity of a local nonspecific modulation of the several agents by a hierarchically higher

decision level to produce an interactive complete architecture, as draft in fig. 11:

1. The individual agents project the visual data stream (distance-normalized vector sequence) onto a situation-adequate action (suggested actions) coded topologically in two-dimensional maps. They receive their intentional orientation by training with different training sets (aimed at "turning left" / "going straight ahead" / "turning right") created by an expert (human pretraining).
2. The action suggestions of the agents operating on identical data streams are superposed (+) and can be biased by instructions of the operators intentionally (human manipulation), as well as with regard to an action (desirable action). Because of the purely modulating manipulation it is certain, that the vehicle will act according to the external instructions, but only with its own repertoire of behavior in conformance to the actual situation.
3. For further adaption of the individual agents the finally selected (action selection) and executed action is returned to the agents to assign them the achieved success. The basis for evaluation is the sensory experience of collisions, which leads as internal reinforcement the entire system to select another (better) action in the next comparable situation. The module human evaluation represents an evaluation by an expert with external reinforcement.

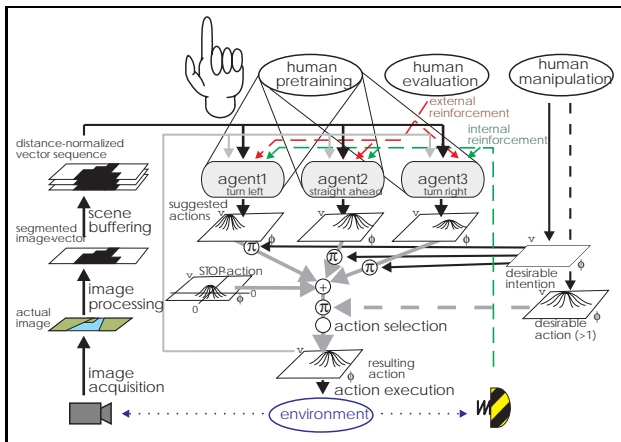


Figure 11. Draft of a total architecture with active learning, intentional agents.

As mentioned above, an adaptive design of the several agents should be achieved through completing the MLP-networks with an active learning structure, which in addition to the situation-specific action learns an estimation

of the network's competence in this situation. This structure is comparable to the AHC (adaptive heuristic critic) – approach according to [9]: the additional network functions as adaptive critic element (ACE) and learns the value function of the fixed policy of the MLP-network. In contrast to the MLP working as function approximator, for this ACE it is necessary, that its architecture is not based on a statistical learning of the input-output-mapping, but able to grasp critical exceptions in one step (one-shot-learning) (fig. 12).

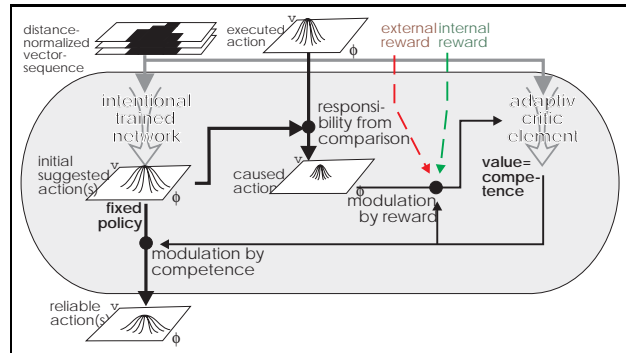


Figure 12. The functional diagram of a complete agent in AHC-architecture.

As noted, the actions available for the execution may differ from those suggested by an individual agent. Therefore, a value of responsibility must be computed by comparison of suggested and finally executed action. The ACE learns a prediction of the reward (0...1) in the actual situation, which is used to modulate the output activation of the MLP-network.

References

- [1] Böhme, H.-J., Brakensiek, A., Braumann, U.-D., Krabbes, M., and Groß, H.-M., "Neural architecture for gesture-based human-machine-interaction," in *Gesture-Workshop Bielefeld*, September 1997.
- [2] Krabbes, M., Böhme, H.-J., Stephan, V., and Groß, H.-M., "Handlungsorganisation mittels intentionaler neuronaler agenten," in *13. Fachgespräch Autonome Mobile Systeme*, Stuttgart, Oktober 1997.
- [3] D. A. Pomerleau, *Neural Network Perception for Mobile Robot Guidance*, Kluwer Academic Publishers, Boston/Dordrecht/London, 1993.
- [4] Pomierski, T. and Groß, H.-M., "Biological neural architecture for chromatic adaptation resulting in constant color sensations," in *Proc. of the ICNN-96, Washington DC*. 1996, pp. 734–739, IEEE-Press.

- [5] Fäustle, P., Daxwanger, W., and Schmidt, G., “Steuerung lokaler fahrmanöver durch direkte kopplung abbildender sensorik an künstliches neuronales netz,” in *10. Fachgespräch Autonome Mobile Systeme*. 1994, pp. 214 – 225, Springer Verlag.
- [6] Waibel, A. and Hampshire, J., “Building blocks for speech,” *Byte*, vol. 8, pp. 235 – 242, 1989.
- [7] Shun-Ichi Amari, “Dynamics of pattern formation in lateral-inhibition type neural fields,” *Biological Cybernetics*, vol. 27, pp. 77 – 87, 1977.
- [8] K. Kopecz, “Neural field dynamics provide robust control of attentional resources,” in *Aktives Sehen in technischen und biologischen Systemen*. 1996, pp. 137 – 144, Proceedings of the German Society of Computer Science (GI) Workshop, Hamburg.
- [9] Barto, A. G., Sutton, R. S., and Anderson, C.W., “Neuronlike adaptive elements that can solve difficult learning control problems,” in *IEEE Transactions on Systems, Man, Cybernetics SMC*, 1983, vol. 13, pp. 834 – 846.