# Learning to Look and Looking to Remember: A Neural-Dynamic Embodied Model for Generation of Saccadic Gaze Shifts and Memory Formation

Yulia Sandamirskaya and Tobias Storck

**Abstract.** Looking is one of the basic sensorimotor behaviours, which entails representation of the visually perceived target and transformation of this representation in a motor signal, which moves the eye to center the target object in the field of view. Looking facilitates memory formation, bringing objects into the portion of the retinal space with a higher resolution. It also helps to align the internal representations of space with the physical environment. In this chapter, we present a neural-dynamic architecture, which integrates several processes, involved in looking, such as target selection, generation of motor signal, adaptation of gaze shift's amplitude, memory formation, scene exploration, and the coordinate transformations. We demonstrate the functioning of the architecture on a simulated robotic agent and provide a discussion of its implications in terms of neural-dynamic and cognitive modelling.

## 1 Introduction

When we look around a room, we don't even notice our own fast and frequent gaze shifts - saccades, - which scan the environment around us, bringing potentially interesting portions of visual input into our fovea for detailed examination. We, instead, have an illusion of a continuous perception of the room and objects in it, ready to be acted upon. We can direct actions at objects around us based on our memory, collected in the fixation periods between discrete saccadic eye movements. How does our neural system accomplish this task? This question has been asked for decades now [2, 12]. How may we build an artificial looking system with similar properties?

Let us follow the processes involved in looking and which our neural system uses to derive a useful representation of the body's surroundings from a sequence of saccades. First, the luminance of light, reflected from objects around us, induces

Yulia Sandamirskaya · Tobias Storck
Insititut für Neuroinformatik, Ruhr-Universität Bochum, 44780 Bochum, Germany
e-mail: {yulia.sandamirskaya,tobias.storck}@ini.rub.de



P. Koprinkova-Hristova et al. (eds.), *Artificial Neural Networks*,
Springer Series in Bio-/Neuroinformatics 4, DOI: 10.1007/978-3-319-09903-3_9

activity patterns on the retina of our eyes. This activity is projected onto the visual cortex and subcortical structures, in which the neuronal attentional mechanisms select a single target for the next saccade. Next, the precise and fast eye movement is generated and brings the interesting part of the visual input into the fovea. Falling onto the fovea, the features of the visual patch may be examined at a better resolution and the identity of the object in this location may be recognised. Moreover, at this moment, the location of the visual patch has to be stored in a useful way, i.e. so that it may be fused with information, collected from previous and upcoming fixations.

Several parts of this process have been examined experimentally, and both neurologically and behaviourally realistic models were developed. Despite the seeming simplicity of the looking system[1], the neural circuitry involved in saccades generation has an immense complexity. For instance, the selection of the saccade's target involves attentional processes, which include both bottom-up, saliency-based computation [17], and top-down influence of context, presence of distractors and alternative targets, memory, or task cues. The generation of a saccade towards the selected target, in its turn, is not trivial, since the variability in the neuromuscular system of the eye calls for a permanently running calibration process between the retinal frame of reference and the motor system [5, 15, 24]. Such calibration processes related to the control of gaze-shifts were found in cerebellar cortex and are hypothesised to be modulated by reward-related basal ganglia loops. Finally, the reference frame transformations between retinotopic, gaze-centred, head-, and body-centred coordinates are needed to fuse information between saccades, to generate saccades to remembered targets if intermediate saccades are involved (double-step saccades), as well as to directed, e.g., arm movements towards visually perceived targets. These processes are hypothesised to be located in the prefrontal cortex [9].

As summarized by Girard and Berthoz [13], several cortical and subcortical regions are involved in generation of the saccadic eye movements (or gaze shifts) in humans and primates. Each of these regions, in its turn, has a complex structure and the modelling and experimental work on untangling these structures – taking into account the behavioural and neurophisiological data – is far from being complete yet.

In this paper, we show how a neural-dynamic framework, based on Dynamic Field Theory, allows to implement an embodied, dynamic, autonomous, and adaptive model for looking behavior. We demonstrate the properties and function of the model by implementing the neural-dynamic architecture on an embodied (here, simulated) robotic agent, connecting its sensors (camera) and motors to the architecture. We introduce visual scenes to the robotic camera and observe the behaviour – both of the network and of the simulated robotic hardware. This approach enables a tight integration between modelling and behaviour analysis, linking the neural, architectural level with the behavioural level.

Our model integrates several functional components, which evolve around looking behaviour. In particular, the agent is able to calibrate itself and learns to look at objects based on a vision-based error-estimation module. The agent is able to

[1] Looking is probably the most basic behaviour, which arises on the intersection between the sensory input and motor control; compare it with arm movements, for instance.

adapt to abrupt changes in the sensorimotor plant or in the environment, as demonstrated in adaptation experiments. Moreover, the robot builds a memory of the observed scene, which integrates the feature information about the objects in the scene with spatial information. A transformation to a body-centred reference frame is performed to make memory independent of the current gaze direction of the camera. We deal with exploration of a scene and inhibition of return, which allows to scan the scene with several objects with different saliency. We demonstrate memory and double-step saccades, as well as simulate an experiment on saccadic adaptation.

Certainly, the model does not address all issues and does not tackle all problems, involved in understanding the saccade generating circuitry. However, it makes an important step towards this understanding, by, first, testing a framework, in which neuronal mechanisms may result in real world behaviour, bridging the mechanistic and behavioural levels. Critically, in this framework, the modelled neural circuitry may be simulated continuously in time and may be linked to noisy and varying real-world sensory inputs and effectors. Second, the framework allows to integrate many functional modules and thus study the saccadic system as a whole, in an integrative framework, instead of looking at its subsystems in isolation. Such a 'wholistic' approach provides additional constraints for modelling, which are set by considering the system as a whole, instead of looking at it as a sum of its components. Finally, the functionally-driven approach allows to look beyond a mechanistic devision of the saccadic system according to neurophysiological units, found in the brain, and allows to consider solutions, in which particular function is served by several such units and is not localised to a particular brain structure. We believe that such functional models will ultimately shed light on how the brain solves the problem of generating precise saccades.

The model, presented here, is not claimed to be complete and we also didn't aim at relating its elements to the neuronal structures, although such mapping will be exemplarily done in the discussion, but is an initial presentation of the neural-dynamic modelling framework with, hopefully persuasive, demonstration of its power to understand the neural circuitry in behavioral terms. The architecture may also have technical application, being a proof-of-concept demonstration of a self-calibrating robotic system, similar to a model presented in [3], but with an increased autonomy, which learns to look at objects in its environment and autonomously update the involved internal sensorimotor mappings when needed.

## 2 Methods

The architecture for generation of adaptive looking behaviour, which we present in this chapter, is built within Dynamic Field Theory (DFT) – a mathematical and conceptual framework for modelling cognition based on population-level descriptions of neuronal dynamics [1, 33]. In DFT, continuous activation functions, called Dynamic Neural Fields (DNFs), represent different perceptual and motor variables, critical for behaviour of an embodied agent. The DNFs may be coupled through mappings, analogous to weight matrixes of neural networks, which are subject for

adaptation and learning. Critically, DNFs follow the dynamics, which allows for localised attractor patterns – peaks of activity, which make DNF architectures behaviourally robust – the behaviourally relevant states of the dynamics correspond to attractors and thus the correct behaviour may be guaranteed.

Here, we present the main functional and structural units of DFT, used in the architecture, whereas a more thorough discussion of the DFT framework may be found in [32, 29].

## 2.1 Dynamic Neural Fields

Dynamic Neural Fields describe activity of populations of biological neurons in terms of continuous functions, defined over behavioural parameters, to which the neurons are sensitive. These continuous functions are descriptions, which abstract away the discrete nature of individual neurons, as well as the spiking character of their activity. A differential equation (1) specifies how the activation functions evolve in time in DFT, as analysed by [42, 14, 1].

$$\tau \dot{u}(x,t) = -u(x,t) + h + \int f(u(x',t))\omega(|x'-x|)dx' + I(x,t). \tag{1}$$

In Eq. (1), $u(x,t)$ is the activation of a dynamic neural field (DNF) at time $t$; $x$ is one or several behavioral parameters (e.g., color, pitch, space, or velocity), over which the DNF is spanned; $\tau$ is the relaxation time-constant of the dynamics; $h$ is the negative resting level, which defines the activation threshold of the field; $f(\cdot)$ is the sigmoidal non-linearity shaping the output of the neural field when it is connected to other fields or self-connected; the self-connections of DNFs are shaped by a Mexican hat lateral interaction kernel, $\omega(|x'-x|)$, with a short-range excitation and a long-range inhibition; $I(x,t)$ is the external input to the DNF from the sensory systems or other DNFs.

Equation (1) defines an attractor for the activation function, which is determined by the external input, $I(x,t)$, the resting level of the field, $h$, and the lateral interactions, specified by the kernel, $\omega(|x-x'|)$. Critically in the DFT framework, a distinctive type attractor of a DNF is a localized activity peak, which may be "pulled up" by the lateral interactions from a distributed input with inhomogeneities. Such peaks of activation are units of representation in Dynamic Field Theory [33]. Because of the stability and attractor properties of the DNF dynamics, cognitive models formulated in DFT may be coupled to real robotic motors and sensors and were shown to generate cognitive behavior in autonomous robots [32, 6]. In particular, activity peaks stabilise decisions about detection of a salient object in the visual input and selection among alternatives, the stabilised representations are critical to linking to motor control and performing cognitive operations on representations [29].

## 2.2 Autonomous Control in DNF Architectures

A single DNF converges to an attractor – one or several localised activity peak(s) – in response to its inputs and would stay in this state unless inputs change, bringing about a different attractor. In this "passive" mode, a DNF could serve as a detection mechanisms, but wouldn't be able to autonomously control actions of an agent. To overcome this limitation, the DFT framework has been extended recently with dynamical structures, which enable activation and deactivation of DNFs depending not only on the sensory inputs, but also on the cognitive task, which the agent faces [31, 27]. These dynamical structures are zero-dimensional DNFs, or dynamical nodes, with an activation following the Equation (2).

$$\tau \dot{n}(t) = -n(t) + h + cf(n(t)) + I(t). \tag{2}$$

The dynamics of activation function $n(t)$ of a dynamical node is equivalent to dynamics of a zero-dimensional DNF. This dynamics is bi-stable: the node can be in an active (i.e., activation is above the threshold, defined by the sigmoidal output function $f$) or in an inactive state. An active state may be sustained event if the external input, $I(t)$, which initially activated the node, decreases below the activating levels. The self-excitatory term with a strength $c$ accounts for this behavior. If the self-excitation parameter is high enough, the node may stay active even if the initial input, which caused its activation, ceases completely. In this case, an external inhibitory input is needed to deactivate the node.

For autonomous control of DNFs, two such nodes are introduced for every elementary (cognitive) behaviour in the agent's repertory: an *intention node* activates DNFs, which bring about a particular (motor or intrinsic) action, and a *condition of satisfaction node* is activated when the desired action outcome is perceived to be achieved and the intention can be inhibited [30].

In our looking architecture, these nodes play an important role in starting and finishing different phases of the gaze shift – the saccade initiation and termination, activation of the fixation, learning, and memory formation processes.

## 2.3 Learning with DNFs

The basic learning mechanism in the DFT is the formation of memory traces of positive activation of a DNF [41]. The memory trace – called *preshape* in DFT – is a dynamical layer, which receives input from the respective DNF and projects its output back to this DNF. The memory trace projection facilitates activation of the DNF at previously activated locations (positive preshape), or inhibits DNF at previously activated location (negative preshape, which accounts for habituation and exploration). The preshape layer follows the equation (3), [29].

$$\begin{aligned}\tau_l \dot{P}(x,t) = \lambda_{build}\Big(-P(x,t) + f\big(u(x,t)\big)\Big) f\big(u(x,t)\big) - \\ -\lambda_{decay} P(x,t)\Big(1 - f\big(u(x,t)\big)\Big).\end{aligned} \tag{3}$$

Here, $P(x,t)$ is the strength of the memory trace at site $x$ of the DNF with activity $u(x,t)$ and output $f\big(u(x,t)\big)$, $\lambda_{build}$ and $\lambda_{decay}$ are the rates of build-up and decay of the memory trace. The build-up of the memory trace is active on the sites with a high positive output $f\big(u(x,t)\big)$, the decay is active on the sites with a low output. The memory trace $P(x,t)$ is an additive input to the DNF dynamics.

Two DNFs may be coupled through a higher-dimensional memory structure, similar to a weight matrix in the standard neural networks. In DFT, such weight matrix is adapted through the mechanism of memory trace formation, which is in this case equivalent to a Hebbian learning process. The coupling is strengthen between locations in two DNFs, which are activated simultaneously, according to Equation (4).

$$\tau\dot{W}(x,y,t) = \varepsilon(t)\Big(-W(x,y,t) + f(u_1(x,t)) \times f(u_2(y,t))\Big) \cdot \Big(f(u_1(x,t)) \times f(u_2(y,t))\Big). \tag{4}$$

Here, the weights function, W(x, y, t), which couples two DNFs, $u_1(x,t)$ and $u_2(y,t)$ has an attractor at the intersection between positive outputs of the DNFs. The intersection is computed as a sum between the output of $u_1$, expanded along the dimensions of the $u_2$, and the output of the $u_2$, expanded in the dimensions of the $u_1$, augmented with a sigmoidal threshold function (this neural-dynamic equivalent to the Kronecker product is denoted by the $\times$ symbol). Learning is only active at the intersection of the active regions in fields $u_1(x,t)$ and $u_2(y,t)$ to prevent spontaneous forgetting of previously learned associations. The shunting term $\varepsilon(t)$ limits learning to time intervals, specified by the autonomous control of the overall architecture, as will be exemplified in our model.

The learning process is functionally robust if the coupling is updated only when the two input DNFs are in a correct, behaviourally meaningful, state. In the looking architecture, presented here, we combine the elements of intentionality with learning dynamics to demonstrate how the sensorimotor mapping, involved in looking behaviour, may be autonomously learned.

## *2.4 DFT in Modelling Saccade Generation*

In our work, we rely on several properties of DFT, which already have been probed in modelling certain aspects of saccades, in particular the target selection and time course of decision making in saccade preparation. The following models are particularly relevant to our work, since they use the same mathematical model for the layer, which performs selection of saccadic targets and thus our model "inherits" the properties, established for this layer.

The first model, which used a dynamic neural field as a layer for target selection in a saccades generation system was introduced by Kopecz and colleagues [18]. The planned eye movement was represented in a neural field with visual and task-related (pre-) input converging on this field. The model could account for a transition from averaging between two presented targets to precise saccades to one of the two targets

depending on the distance between the targets and the strength of the memory for the two locations from preceding saccades. The architecture features an active fixation system, similar to the one we use in our model.

The effects of lateral interactions in superior colliculus (SC) on the saccade reaction time were studied by Trappenberg and colleagues [38] in their work, which provides a link between behavioural studies of saccade generation and the underlying neural substrate. This model considers integration of different input sources for the selection of a single target in SC with a dynamic neural field and may account for experimentally observed delays in saccade initiation. Behavior of the model closely resembles activity of neurons and the model can account for many observations related to saccade initiation.

Willimzig et al [40] has also studied the time-course of saccadic decision making within dynamic field theory. To initiate a saccade, the system has to overcome fixation, which may be more or less difficult depending on the fixated object and the overall scene. The transition from fast averaging saccades to more time-lagged selective saccades is again demonstrated here using DNF layer for target selection. A similar DNF framework for movement representation was used in modelling preparation of arm movements [7].

These previous models have demonstrated the power of DFT in preparation and planning of saccades. Here, we extend these models to an architecture, which may actually realise the planned saccades, check for their accuracy, and adapt saccades' amplitude if needed.

## 3 The Model

### *3.1 The Overall Architecture for Looking*

Figure 1 depicts the overall architecture for looking. It is a fairly complex network, since it accomplishes several functions apart from generating gaze shifts towards visually perceived targets. The architecture may be described in terms of the following interconnected modules. The *perception* module integrates over time and stabilises the visual input from a simulated robotic camera, as well as the proprioceptive input from the motor system of the simulated robotic agent. The precise saccades to the visually perceived targets are generated in the *saccade generator system*. The amplitude of the saccades for every retinal position and current gaze angle is learned in a set of gain maps within the *learning* module. There is one set of gain maps for each of the two motors of the robot. Next, the *fixation system* tracks the object between successive saccades and triggers memory processes in the *memory system*, which, on the one hand, steers exploration of the scene, decreasing the competitive advantage of those object in the scene, which are already put to memory, and, on the other hand, creates an allocentric representation of the objects in the scene, which can be used to generate saccades from memory, double step saccades (using the *planning saccades system*), or arm movements towards the target. Next, we will walk through the most important parts of this architecture, explaining how different functions of the network are brought about.

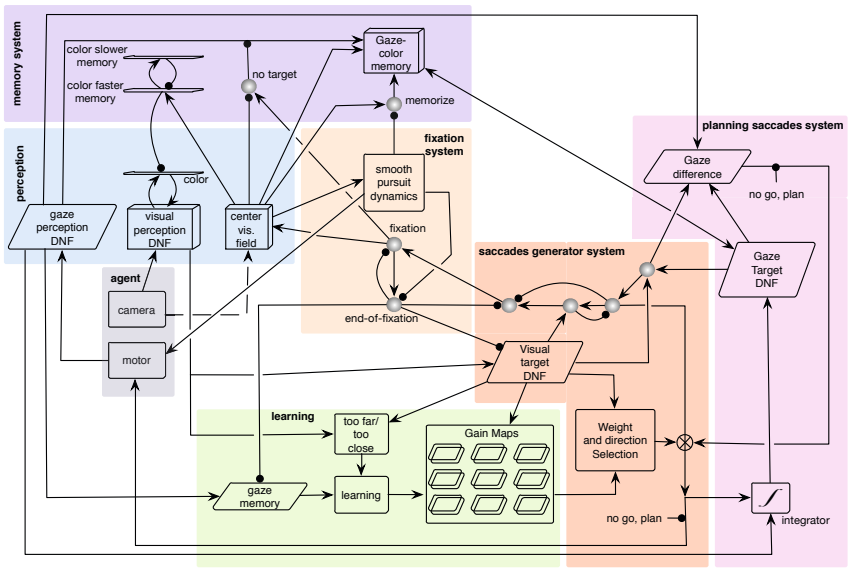



**Fig. 1** The overall architecture

## *3.2 Perception*

The perceptual system consists of a three-dimensional *visual perception DNF*, which spans the dimensions of color and the retinotopic space (modelled to be a cartesian space here, but a polar version with foveal expansion is possible as well). The RGB output of the robotic camera is split into three color channels – hue, saturation, and value. The saturation channel is used to perform the basic figure-ground segregation and create a course saliency map, which highlights regions in the visual space, for which the hue value is extracted and input to the visual perception DNF. The hue (color) dimension is additionally stabilised by a coupled one-dimensional color DNF. The visual perception DNF requires this supporting input to form a stabilised activity peak over a selected object in the scene. This support from the color field is suppressed for objects, which are already stored in memory. Thus, such objects have a disadvantage in the competition to be selected for the next looking act[2].

The *center of visual field* is another three-dimensional DNF, which receives input from the central portion of the camera image. This field may only build activity peaks from the camera input when the *fixation node* is active and provides an additional boost to this DNF (i.e., raises its resting level), signalling that a saccadic gaze shift has been finished. An activity peak in the center of visual field DNF activates

[2] This 'habituation' happens along color dimension, but habituation along retinal space is also possible [36], both processes have to be balanced to account for human looking data. Here, we keep this system simple since accounting for experimental data is not the focus of the work reported here.

two processes in the architecture: memory formation and smooth pursuit dynamics, briefly described in the next section.

The *gaze perception DNF* receives input from the motor system of the robotic agent and represents the current proprioceptive state of the motor system.

### 3.3 Motor Control

The motor system of the agent consists of two motorised joints – the pan and tilt joints of the camera-head unit. The pan joint corresponds to rotation of the camera head around the vertical axis between the two cameras of our robot (only one of the cameras was used in this work; the camera's optical axis is not crossing the axis of the pan joint). The tilt joint corresponds to the incline of the camera – a rotation around the horizontal axis (which is also not aligned with the image plane of the cameras). Note that the arrangement of the motor rotation axes and the camera image provides for a non-linear mapping between the reference frame of the camera image ('retinal' reference frame) and the reference frame of the motor system.

The two motors – the pan and tilt motors – are servo motors, which may receive both position and rotation speed commands. We have used speed control in this work to make the motor system somewhat confirm with the biological motor systems (still, buying into significant simplifications compared to muscle control). The two motors coarsely correspond to the horizontal and vertical control of the eye. The rotation of the eye-ball was not modelled here.

During saccades, the saccade generator system sends velocity commands directly to the two motors of the camera head. During smooth pursuit movement, a dynamical system, which has an attractor at the visually perceived target, provides the velocity input for the two motors and performs visual servoying around the target object. The visual servoying towards the target is not possible during saccadic gaze shifts, since they are performed too fast for the visual processing to influence their course.

Next, we will describe the system, which may generate correct velocity commands based on the visual perception of the target in the retinal frame of reference.

### 3.4 Saccade Generator

Since the saccades (in humans and primates) are very fast movements[3], the neural system needs to generate the complete velocity profile, which will bring the eye's fovea onto the target. In our architecture, this velocity profile is generated by a neural oscillator, as described next. In particular, the neural oscillator is part of a central control unit, which generates the saccades gaze shifts and controls the temporal dynamics of the overall architecture. This unit consists of six dynamical nodes, depicted in Figure 2.

[3] Which ecologically makes sense, to minimise the time when the eye moves and both vision and calibration with the outside world are disturbed.

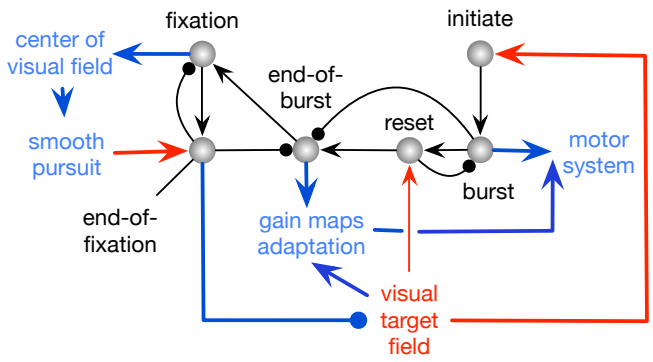



**Fig. 2** A neural circuit, which generates the gaze-shift velocity profile. Grey shaded circles denote the six central nodes, which control gaze shift generation. Black arrows show excitatory connections between the nodes, lines with filled circles – the inhibitory connections. Blue arrows are outputs of the system to other parts of the architecture, red arrows are inputs to the nodes structure.

The *burst* and *reset* nodes constitute the neural oscillator. If a constant input is provided to the burst node, this pair of nodes gets activated and deactivated in alternation: the burst node activates the reset node, which in its turn deactivates the burst node and consequently looses its own activation, and the cycle repeats. In our setting, however, this system, generates a single "oscillation" in the following way. The reset node is lifted to be self-sustained by the input from the *visual target field* and only looses its activation when the whole gaze-shift action is finished (i.e. when the visual target field is inhibited by the *end-of-fixation node*). The active reset node keeps the burst node inhibited until a new target is selected for the gaze shift.

Activation of the burst node drives the motor system of the agent by sending velocity commands to both the pan and tilt motors. The amplitude of the burst specifies the peak velocity of the gaze shift and, implicitly, its amplitude. The busts's amplitude is set by the adaptive *gain maps* – one for the horizontal movement, generated by the pan motor, and one for the vertical movement, generated by the tilt motor. The adaptive dynamics of the gain maps will be presented in Section 3.5.

The burst node is activated by the *initiate gaze shift node*, which is effectively the intention node of the gaze shift behaviour. The *end-of-burst node* detects when one burst is finished and activates the *fixation node*, which, eventually, drives the fixation and the smooth pursuit systems of the architecture. When the fixation system brings the object in the center of the visual field and the memory for this object is updated, the *end-of-fixation node* is activated and resets the gaze shift elementary behaviour by inhibiting the visual target field, which provided the initial input to this system.

Figure 3 shows the time-course of activation of the five nodes from Figure 2 for two subsequent gaze shifts, demonstrating how the autonomous organisation of the architecture works – like a clock, turned on by the target field when a new target to

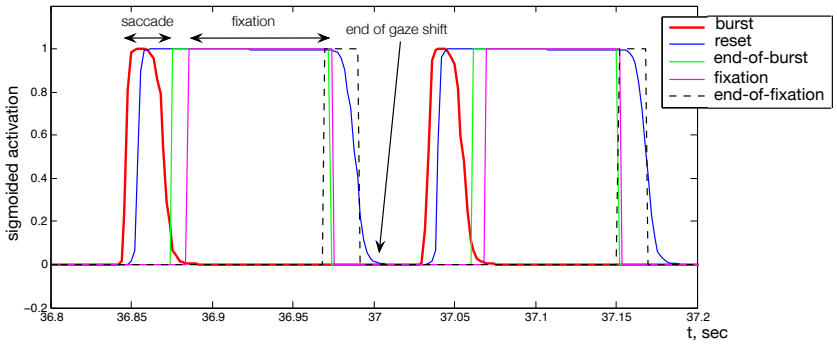



**Fig. 3** A sequence of two gaze shifts, each consisting of a saccadic part, driven by the Burst node (red line shows sigmoided activation of the Burst node), and a fixation part, in which the target object is tracked, memory is formed, and adaptation is performed if needed (magenta line shows activation of the Fixation node), the gaze shift is finished when the End-of-fixation node is activated (dashed black line) and the Reset node (green line) is inhibited, releasing the Burst node from inhibition, which may now generate the next saccade.

look at is detected and turned off by the fixation dynamics, when the target is centred in the camera's view and its location and features (color) are stored in memory.

## *3.5 Saccade Amplitude and Gain Maps*

The peak velocity of the saccadic gaze shift is defined by the values, stored in the *gain maps*, which are initially learned and are constantly adapted in a learning process after each gaze shift. Each gain map is defined over the same space as the visual target DNF – the retinotopic (here, camera image) space. There is one gain map for each starting gaze angle, i.e. for every possible state of the eye (here, of the camera motors). This means that the gain map is a four-dimensional structure, with two visual (retinal) dimensions and two motor dimensions. Further, there is one such map associated with each of the motors of the system (the pan and tilt motors).

The maps control the precise vector of the saccadic gaze shift as follows. The output of the visual target field is multiplied with a slice of the four-dimensional gain map, selected by the currently perceived (before the gaze shift) gaze position (the pan and tilt values stored in the *gaze memory DNF*). When a peak is built in the visual target field, this multiplication effectively selects a region on the gain map, which is specified by the location of the activity peak in the visual target field. Thus, the gain maps function as synaptic weights between the visual target field and the circuitry, which generates motor commands. Moreover, these synaptic weights are modulated by the input, which specifies the gaze angle before the saccade. Formally, the gain maps constitute a tuneable (or steerable) map between the retinal and the motor frames of reference [9]. The result of the multiplication of the output of the visual target field with the gain map is integrated and is connected to the output of

the *saccade generator system* (Figure 2), amplifying the amplitude of the burst of the neural oscillator. Figure 4 illustrates this process.

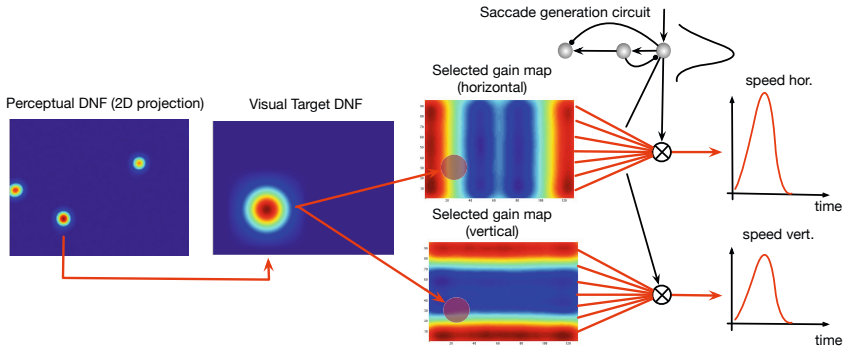



**Fig. 4** Circuitry to generate saccades with a precise amplitude

The gain maps are initially homogeneous – all values in them are set to ones. When a visually perceived target appears in this state, a saccade is generated with such an amplitude in both motors that does not bring the target object into the central portion of the camera image (retina). Thus, the fixation system does not get engaged, but the error estimation module is activated and estimates whether the saccade was too long or too short (*too far / to close* module in Figure 1) in each of the four direction in the image: left, right, up, and down (corresponding to four direction of the eye movements' 'synergies'). The learning mechanism of Equation (5) updates the gain map at the position, which corresponds to the active region of the visual target field and the gaze angle before the saccade. The direction of adaptation (increase of decrease of the values in the selected region of the gain map) is defined by the output of the error estimation module.

$$\tau_l \dot{G}^{h,v}(x,y,k,l,t) = \varepsilon^{h,v}(t) f(u_{EoS}(t)) \Big( f(u_m(k,l,t)) \times f(u_{tar}(x,y,t)) \Big). \qquad (5)$$

Here, $G^{h,v}(x,y,k,l,t)$ are two sets of gain maps (for the '*h*orizontal' and '*v*ertical' components of movement). Each of the $k \times l$ gain maps in the two sets is defined over the dimensions of the visual target DNF, $u_{tar}(x,y,t)$. Each set spans $k \times l$ different initial motor states (the pan and tilt joint angles in our setup). The gains change in the map(s), which are selected by the output of the motor DNF, $f(u_m(k,l,t))$, at the locations, which are set by the activity peak in the target DNF. $f(u_{EoS}(t))$ is the output of the end-of-saccade node, which is required to be positive (saccade finished) for learning to become active. $\varepsilon^{h,v}(t)$ is the error in each of the movement components, $\tau_l$ is the learning rate.

Thus, after each unsuccessful saccade the gain maps are corrected slightly in a localised region. After sufficient experience with looking at visual targets in

different locations in the image and from different starting gaze angles, the complete gain maps are learned and the system is able to perform precise saccades from any configuration. The maps are updated locally over a few gaze-shifts if an unexpected change in the sensorimotor plant happens.

### *3.6 Memory Formation and Exploration*

After a precise saccade, the target object falls into the central part of the visual field, as detected by the *center of visual field DNF*. The object is being tracked now by the *smooth pursuit dynamics*, which sets an attractor at the visually perceived target and visually servoys the camera at the objects. When the object is centred in the visual field, its memory is formed in the three-dimensional *color-gaze memory field*, which is critical for creation of an allocentric memory for the object and for generation of memory saccades. The part of the architecture, which forms memory during fixation and biases the perceptual system during exploration is shown in Figure 5.

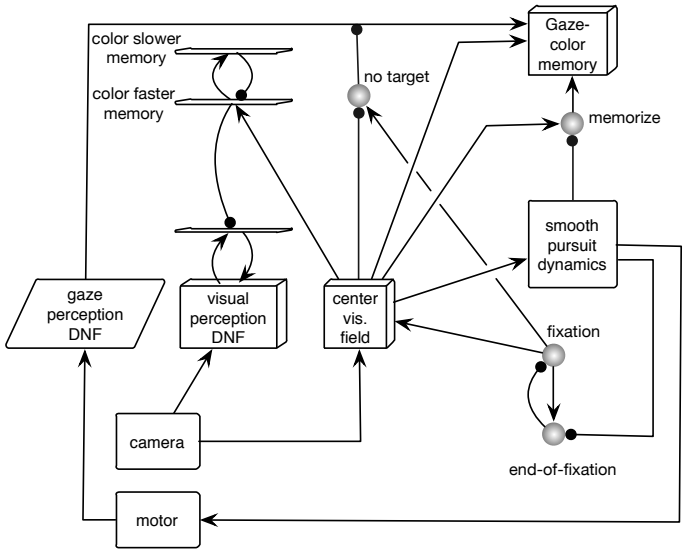



**Fig. 5** Part of the architecture responsible for memory formation and exploration (habituation)

Here, the activated *center of visual field DNF* activates the smooth pursuit dynamics, which fixates the object in the center of visual field, even if the object moves. As long as movement is generated by the smooth pursuit dynamics – i.e. the object is not yet centred in the visual field and its location in gaze coordinates is not stable, – the memory formation process is suppressed by a lacking boosting input from the *memorise node*. When the object is centred in the visual field of view, the memory

formation process is activated in the three-dimensional *color-gaze memory DNF*, which forms an allocentric memory of the objects in the visual scene, represented in gaze-coordinates. If there is no object perceived in the center of visual field (the *no target node* is activated), the corresponding location in the color-gaze memory DNF is inhibited. If there was an object memory stored at this location, its memory representation ceases.

On the other hand, the *center of visual field DNF* boosts the *color faster memory field*, which builds a memory representation of the color of the currently observed object. This representation inhibits the visual perception DNF along color dimension for the subsequent saccades. Thus, when the next saccade target is selected, the colors that are already in memory have less chances to induce a peak in the perception DNF. The *color slower memory* builds up preshape activity hills on a more slower time scale. These preshape hills eventually delete the peaks from the faster color memory field, releasing the respective color to participate in the target selection during exploration.

### *3.7 Prediction and Memory-Driven Saccades*

On the right side of Figure 1, the planning saccade system is depicted. In this system, the saccades' targets are represented in the body-centred gaze coordinates. This gaze-target DNF receives input either from the gaze-memory DNF, if saccade is to be performed to a memorised object, or from an integrator, which predicts the gaze coordinate of a visually perceived target without performing a saccade. The latter path allows the system to perform double-step saccades, when two targets are presented to the system and the agent has to saccade to them in a sequence only when both targets are no longer visible. In this case, the system "simulates" two saccades from the fixation point and stores the predicted gaze angle after each such virtual saccade. These stored gaze-representations of the targets' locations can then be used to generate eye movement towards both targets, even though the retinal representation of the second target would shift after the first saccade.

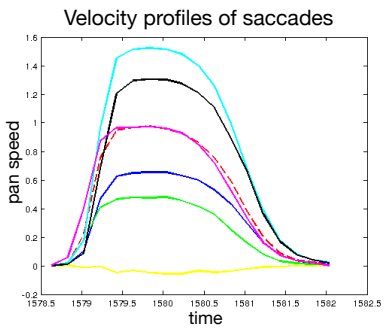
## 4 Results of Simulated Experiments

The architecture, sketched in Figure 1, was implemented using an open-source C++ framework *cedar*, `www.cedar.ini.rub.de` [20], which allows to build neural-dynamic architectures using a graphical user interface and simulate them efficiently on a conventional computer. The simulation basically consists in solving the coupled differential equations (the core equations, which form the building units of the architecture, were presented in Section 2 of this chapter) using Euler method. The *cedar* framework also offers an interface to robotic hardware, including cameras. The interface may be used both with real and simulated robots. In this work, we have used a simulated CoRa robot [16], in particular its camera, mounted on a motorised pan-tilt unit. Although *cedar* offers plotting routines to visualise the architecture during its function for monitoring purposes, we have collected the data

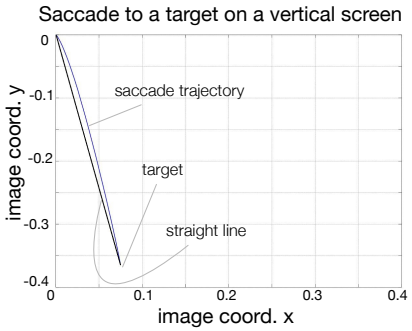
from simulation and have analysed and plotted it offline, using standard MATLAB routines. The data consists of matrices of activation of the dynamic neural fields and nodes of the architecture. Here, we exemplify some of the architecture's functions based on this collected data from simulated experiments.

## *4.1 Gaze Shifts Generation*

Saccades, generated by humans or primates, have particular properties, which are well-studied in a controlled experimental settings. Here, we demonstrate that the basic properties of real saccades may be replicated in our system. Note that the architecture was not tuned to model the dynamic properties of human saccades, but still the trajectories and velocities of saccades resemble the respective behavioural plots. Figure 6a shows examples of velocity profiles, generated by the neural-dynamic gaze-shift generator in the simulated robot for gaze shifts of different amplitudes. Note the constant duration of the gaze shift, as typically observed in eye movement studies, and the varying peak velocity for saccades of different aptitudes. This property is part of the 'main sequence', postulated in studies of the primates' saccade generating system.




**(a)** Generation of saccades of different amplitudes. Demonstration of the relation between the peak velocity and amplitude of saccadic gaze shifts.




**(b)** Trajectory of an oblique saccade.

**Fig. 6** Illustration of the basic properties of the saccade generating system

Figure 6b shows the trajectory of an eye movement (projection of the gaze direction on a vertical plane, in which the target is presented) towards a target, which requires activation of both the vertical and horizontal movement systems. The trajectory is very close to a straight line, as observed experimentally. In our architecture, this property is achieved by a mechanisms, similar to the mechanism proposed by [35] for generation of two-dimensional saccades: both horizontal and vertical

movements are produced by the same neural burst generator, which output is scaled differently for the two components of the movement. This results in almost straight oblique saccades, when both the target and the gaze direction are projected on a vertical plane. The shape of saccade trajectories in 3D, when targets are located on a horizontal plane at different distances, is yet to be established experimentally and can be simulated in our framework.

### *4.2 Scene Exploration and Memory Formation*

Figure 7 demonstrates how a visual scene (here, consisting of three coloured objects) is explored by the system.

On the left, four snapshots from the robot simulator show the simulated camera image and the visual scene in front of the robotic camera head. In the first snapshot, the robot observes the scene with three objects. The red object induces the largest blob in the camera image (since it is closer to the robot on the table) and is selected by the visual perception DNF as the saccade's target. The robot performs a saccadic gaze shift followed by a fixation dynamics towards the red object (second snapshot). Now, only the yellow object is visible in the camera image and is selected by the robot for the next gaze shift. When the yellow object is centred in the camera, both blue and red objects are visible (third snapshot). Although the red object is more salient in the camera image again, the blue object is selected for the next saccade, because of the inhibitory influence of the memory on the selection dynamics in the visual perception DNF. Finally, the blue object is fixated by the system (forth snapshot) and its representation is stored in the gaze-based scene representation.

On the right of Figure 7, the summed activations of the saliency in the camera image and the visual target DNF are shown. Note that during exploration, many locations have significant saliency in the camera image (in fact, more than shown in the figure – these are only the regions, which where active for longer periods of time). In the visual target DNF, to the contrary, only the locations selected for the saccade targets leave traces. The third plot shows the gaze-based (body-centred) representation of the visual scene, built-up during scene exploration: after each successful saccade, the gaze angles of the camera head are stored, which correspond to the object in the scene, i.e. which bring the object in the center of camera field of view (if the body of the robot does not move relative to the scene). The forth plot shows the projection of line of sight of the camera on the table surface during the experiment, as viewed from above. These projections show the scan paths of the camera head over the scene.

In this demonstration, it may be seen how a visual scene triggers a sequence of saccades in the system. Each object is fixated by the camera in a succession and the gaze angle of the robot during fixation is stored as a self-sustained peak in the memory field. The resulting memory representation may be used to direct saccades to memorised, but currently not observed objects, as well as used for control of movements, generated by other effectors of the robot (e.g., reaching), even if the robot looks away from the object (e.g., to look at the arm or the next object in a longer sequence of actions).

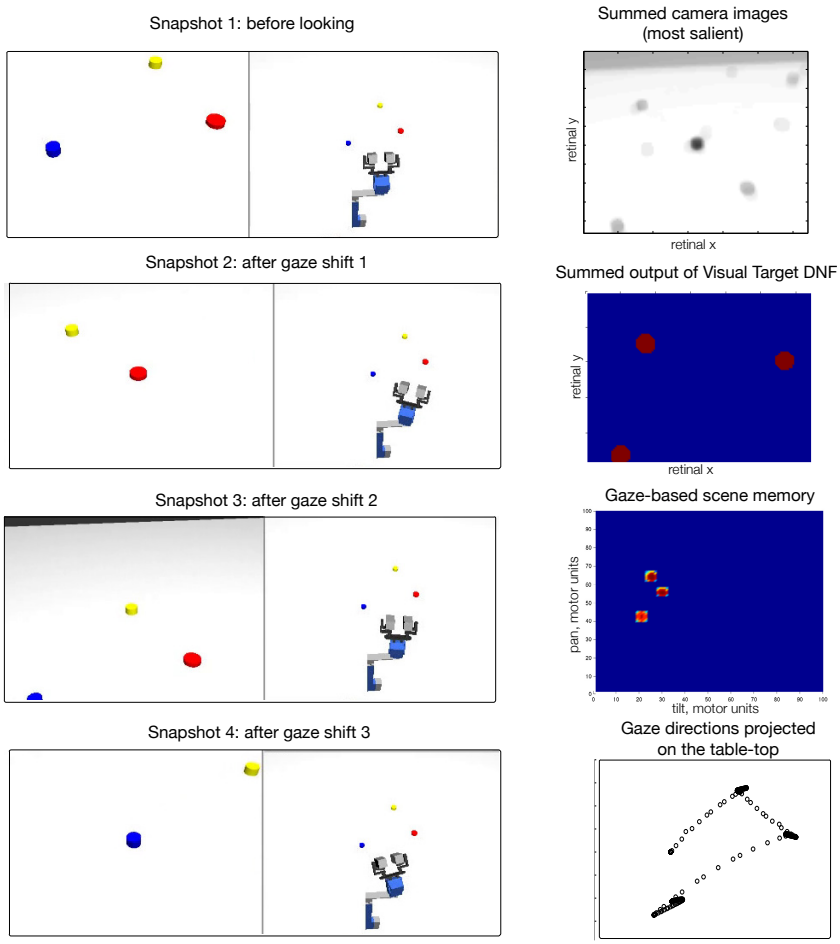



**Fig. 7 Left**: Simulated robot exploring the scene. **Right**: the summed visual input to the architecture, the summed activity of the visual target DNF, gaze-based memory of the scene at the end of exploration, and the gaze-trajectory, projected on the table-top (sampled at 100 frames per second).

## *4.3 Gain Maps Learning*

The precise gaze shifts, demonstrated in the previous experiment, are the result of a learning processes, in which the gain maps, which specify the saccades' amplitudes, are adapted. Figure 8 shows the convergence of the learning process for a single location in one of the gain maps, in which the initial error of more than 3 cm is reduced over a few saccades (five here) to values below 0.5 cm.

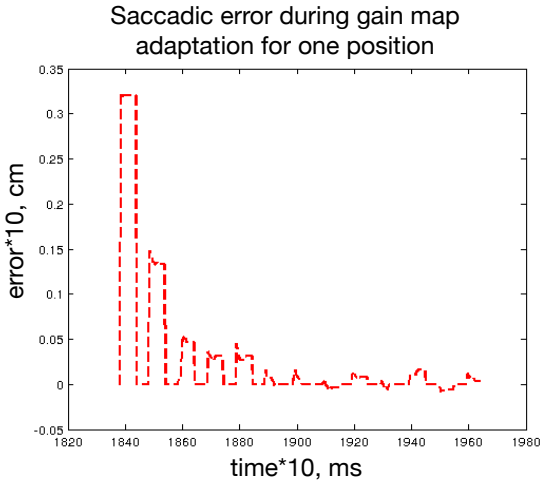



**Fig. 8** Convergence of gain map for a single location

The two four-dimensional gain maps, learned for the horizontal and vertical movements are shown in Figure 9. The maps reflect the geometry of the robot and implicitly encode the amplitudes of the pan and tilt shifts for different shifts in the retinal frame of reference. Note that the mapping between the image and motor coordinates is non-linear here and changes significantly with initial gaze angle of the camera head. Especially with changing initial tilt-configuration, the amplitude of the motor signal changes for the same shift in the retinotopic coordinates. The pan-configurations play a lesser role in our robotic architecture, since changes in initial pan do not change the mapping between shifts in retinal and motor coordinate frames much.

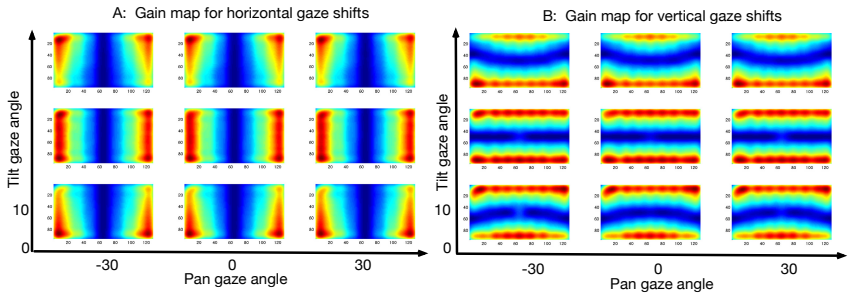
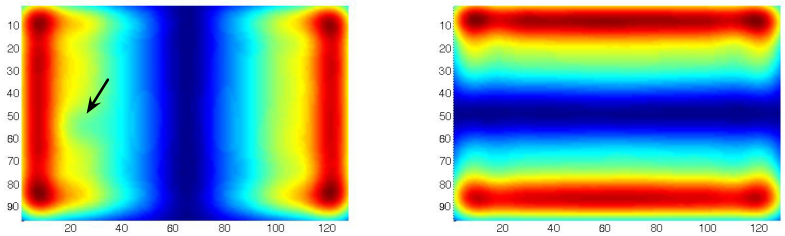



**Fig. 9** Gain maps learned by the system. Nine slices are shown for three selected pan and tilt values of the initial pose (gaze angle) of the camera head.

### *4.4 Modelling Adaptation Experiments*

Adaptation of the amplitude of the saccadic gaze shift, as demonstrated in [24] is exemplified in Figure 10. Here, the robot first learned the complete gain maps and was able to perform precise saccadic gaze shifts from any starting configuration. In a scenario, which simulated the adaptation experiment, the target object was shifted horizontally during the saccadic gaze shift, so that the saccade landed (in the case, shown here) behind the target, i.e. the saccade was too long. The perceived error was used to updated the gain map, similarly as during the initial learning process, so that the new, adapted, gain map generated a saccade of the amplitude, which brought the shifted target in the fovea.

The adaptation is only effective for a localised region both in terms of retinal location of the target and the gaze angle prior to the saccade. In combination, the adaptation generalises to a region in the allocentric (here meaning gaze-angle independent) reference frame. This result is conform with recent experimental studies [44].




**(a)** Map for horizontal gaze shift **(b)** Map for vertical gaze shift

**Fig. 10** Gain maps adapted at one location. The adapted region is marked with the arrow.

The time-course of the simulated adaptation experiment is shown in Figure 11. This figure shows the increase of saccadic error at the offset of the adaptation experiment (seventh saccade shown in the figure) and a gradual decrease of the error back to the optimal level in the course of several saccades.

## 5 Discussion

This paper has introduced a computational framework and a neural-dynamic architecture for generation of adaptive looking behaviour in an embodied agent. The behaviour and the computational network share several characteristics with the human looking system.

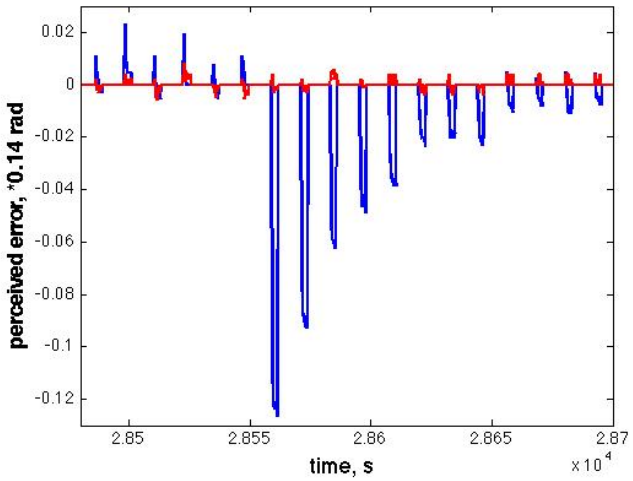



**Fig. 11** Time course of adaptation: error magnitude for the adapted location over the course of adaptation experiment

## *5.1 Strengths and Limitations of the Architecture*

The power of the framework we used in our modelling is in its dynamics. Thus, the introduced architecture is a process model and allows to model not only the structure, but also the dynamics of neural processes within this structure. Moreover, this dynamics is autonomous and embodied, which means that it may be connected to real sensors and motors and produce behaviour in real time.

The dynamic fields theory provides for stability of the building units of the architecture and enables coordination between different subsystems.The stability property of the dynamic neural fields has been established theoretically [1, 14, 42] and leads to robust, controllable, behaviour. The behaviourally relevant states of the neural system are represented in this framework as attractors, which persist long enough to have impact on the downstream structures. Transitions between attractors are instabilities and their course is autonomously controlled by an interconnected set of dynamical nodes, which organise behaviour of the architecture in time.

In the architecture, many different functional subsystems are integrated, some of them were developed in recent years in the DFT framework, others are introduced here for the first time (e.g. the adaptive weights coupled to a neural oscillator, the motor-based scene memory, the error-detection network, the coordination between smooth pursuit dynamics and saccade generation). Several functions of the looking system are implemented in our model, such as memory formation, formation of allocentric scene representation, habituation, scene exploration, adaptation, and learning. Although not all properties of looking behaviour in humans and primates

have been accounted for in our current system, the framework has a potential to be extended and refined while keeping the behavioural stability of the overall network.

Another important characteristics of our architecture is that it is a function-based model. This allows to avoid narrowing modelling too much on single brain areas, as has been advocated early on in studies of the saccadic system [39, 13]. Thus, the architecture presented here is a behavior-based, functional model, which may be mapped on the substrate of neural circuits, involved in generation of eye movements. This mapping onto neuronal substrate, however, even for a single functional module of our architecture requires considering interactions among different neural circuits. This avoids a simplifying view, when one cortical or subcortical region is made responsible for a single function and every function is assigned to a single brain region. Since our architecture was inspired by the neural and behavioural findings about biological saccadic systems, we will provide a brief discussion of the neural plausibility and relevance of our model in Section 5.2.

Several properties of the saccadic system have not been modelled here. Thus, our current system cannot produce saccades with different durations. A more flexible in this respect neural oscillator could solve this problem, or substituting the oscillator with a resettable integrator, as in classical saccade generation models [28]. This modification will also solve the problem that an interrupted saccade, currently, will be resumed, but won't end at the correct pose.

## *5.2 Discussion of the Architecture in Relation to the Neural Mechanisms of Saccades Generation*

The currently most widely accepted picture of the saccade generating circuitry includes the following neuronal structures [13]: superior colliculus (SC), saccadic burst generators in the reticular formation, cerebellum, basal ganglia, and cortical structures. Here, we review briefly how each of these brain areas is reflected in our architecture.

### 5.2.1 Superior Colliculus

The SC is considered to be responsible for representation of the amplitudes and directions of saccades in retinal coordinates. In our architecture, the visual (retinotopic) target field accomplishes this role, since activity peaks in this DNF encode the saccades targets in retinal coordinates. But also the spatial component of the perceptual DNF probably corresponds to one of the SC layers. This field selects the saccadic target, but does not keep this representation fixed during the eye movement, but tracks the visual input to some extent. The burst, reset, and fixate nodes, roughly correspond to the three different types of saccade-related neurons, found in SC [21, 43]. The nodes perform the spatial-to-temporal transformation, which converts the location of the activated region on the visual target map into a temporal signal encoding the desired speed of the eyes, similar to other neural models, e.g. [26]. The nodes of the saccade generating circuit may be seen in close connection to

the visual target DNF, and a more precise model of SC would have a number of layers with the topology of the visual target DNF and interconnected among each other like the nodes in Figure 2. We are not modelling the log-polar topology of the SC here – our visual target DNF has a cartesian structure, but the log-polar organisation of this field can also be used [38].

### 5.2.2 Saccadic Burst Generators

Our saccade-generating circuit is similar to the burst, buildup, and fixate neurons, proposed in several models for saccade generation [13]. This circuit provides for the temporal properties and dynamics of saccades, in particular the relation between velocity, amplitude, and duration of the saccade, time of its initiation, and reaction to perturbations. The mechanisms we used to produce two-dimensional saccades are closely related to mechanisms proposed in earlier models [8, 10], which include a shared burst generator driving the two components of the eye movement through different gain factors. This setup results in straight oblique saccades. Experimentally, it has not yet been decided on the nature and substrate of the saccade generating motor signal [13], but an alternative to our solution would be a classical model by [28], in which the motor command is integrated to achieve the target pose during saccade generation. This solution may be realised in our architecture, but would require the transformation from the retinal representation of the target to its motor representation to be learned before this integration leads to saccades of the correct amplitude.

### 5.2.3 Cerebellum

Fine-tuning of the saccadic amplitude, as well as corrections for changes in motor plant are found to happen in the cerebellum [26, 22]. This structure is also considered to be responsible for taking the starting position of the eye into account and compensating for the non-linearities of the motor system. In our architecture, the gain maps correspond to this function of the cerebellar structures, in particular, they attenuate saccades and provide for adaptation in the saccadic circuitry. The gain maps also reflect to models of the cerebellum as the locus of supervised learning, responsible for long-term calibration and adaptation of saccades' gains. E.g., a similar, but more abstract and relying on different error signals model is discussed in [4] . Gain adaptation in our framework is similar to adaptation process described in [11, 5]. Although in several models [19, 26, 23] the control of saccades' accuracy is controlled by the cerebellum only, cortical structures are probably also involved in this process [13]. Which, again, shows that the correspondence of an established function to the neural substrate might not be unequivocal.

### 5.2.4 Basal Ganglia

Error signal, generated by the 'too-far, too-close' module in our architecture, could be neurally associated with inferior olive, as in the model of [34]. The error in our

model is determined directly based on the visual input after the saccades, whereas in their model the error is defined based on proprioception, using the memory of the saccade signal, the non-confirmation of the correct cascade by the fixation system, and the motor signal of the corrective saccade. In our framework, the error estimation module takes input from the visual memory of the saccadic target and the perceived visual representation of the target after the saccade and estimates whether the target after the saccade is on the same or on the different side of the midline compared to the original location of the target on the retina. This very basic operation delivers meaningful results, which drive learning in the correct direction, even when the corrective saccade also cannot be performed in the correct direction yet, since the system is completely uncalibrated. An elaborated visual processing mechanism to estimate visual error after saccade, which drives adaptation, has been discussed in relation to experimental work on saccadic adaptation [24].

Since basal ganglia are probably also involved in overall temporal coordination of saccade generation, our gaze shift generator nodes circuitry, which is responsible for temporal coordination of the gaze shifts, fixation, adaptation, and memory formation, could partially reside in this region.

#### 5.2.5 Cortical Structures: Adaptation and Spatiotopic Visual Maps

The location of adapted region in our gain maps depends on a combination of the retinal position of the target and the position of the eye before the saccade. Consequently, adaptation effectively influences the target location in the allocentric, body centred frame of reference. This is conform with a recent experimental finding [44], which establishes that adaptation influences saccade targeting for the same position in the allocentric space. The spatial spread of the adapted region in our architecture corresponds to the experimental findings as well [15]. Our gaze-based memory representation of the visual scene resembles recent evidence, which shows that the motor representation is what is updated in the double-step paradigm and is probably used to plan multiple saccades [25]. Both these functions – learning of the saccade's amplitude and formation of memory for the visually observed scene – at least partially are solved by cortical structures [37].

## 6 Conclusions and Outlook

The architecture presented here demonstrates how looking behaviour may be learned autonomously and lead to formation of memory in body-centered coordinates, which may be used to direct actions at objects around us. The model demonstrates what it takes to create an illusion of the stable perception from a sequence of fast eye movements, in particular the critical role in this process of adaptation and learning, the temporal coordination between different processes and sensorimotor structures, and interplay between memory formation and exploration. Considering the integrated, dynamical system of functional modules allowed us to reveal how closely interconnected different functions of the looking system may be. There are different

directions, in which this architecture may be developed. To add neural plausibility and account for neural and behavioural data is one of them, whereas extending the architecture towards a system, capable of learning to generate arm movements towards visually observed targets and thus creating a more complex self-calibrating embodied agent is another possible direction.

**Acknowledgements.** The authors gratefully acknowledge the financial support of DFG SPP *Autonomous Learning*, within Priority Program 1567.

## References

1. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. Biological Cybernetics 27, 77–87 (1977)
2. Aslin, R.N.: Perception of visual direction in human infants. In: Visual Perception and Cognition in Infancy, pp. 91–119 (1993)
3. Chao, F., Lee, M.H., Lee, J.J.: A developmental algorithm for ocularmotor coordination. Robotics and Autonomous Systems 58(3), 239–248 (2010)
4. Dean, P., Mayhew, J.E., Langdon, P.: Learning and maintaining saccadic accuracy: a model of brainstem-cerebellar interactions. Journal of Cognitive Neuroscience 6(2), 117–138 (1994)
5. Ebadzadeh, M., Darlot, C.: Cerebellar learning of bio-mechanical functions of extra-ocular muscles: modeling by artificial neural networks. Neuroscience 122(4), 941–966 (2003)
6. Erlhagen, W., Bicho, E.: The dynamic neural field approach to cognitive robotics. Journal of Neural Engineering 3(3), R36–R54 (2006)
7. Erlhagen, W., Schöner, G.: Dynamic field theory of movement preparation. Psychological Review 109, 545–572 (2002)
8. Fuchs, A.F., Kaneko, C.R.S., Scudder, C.A.: Brainstem control of saccadic eye movements. Annual Review of Neuroscience 8(1), 307–337 (1985)
9. Gail, A., Andersen, R.: Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor transformations. The Journal of Neuroscience 26(37), 9376–9384 (2006)
10. Gancarz, G., Grossberg, S.: A neural model of the saccade generator in the reticular formation. Neural Networks (1998)
11. Gancarz, G., Grossberg, S.: A neural model of saccadic eye movement control explains task-specific adaptation. Vision Research 39(18), 3123–3143 (1999)
12. Gibson, J.J.: The perception of the visual world (1950)
13. Girard, B., Berthoz, A.: From brainstem to cortex: computational models of saccade generation circuitry. Progress in Neurobiology 77(4), 215–251 (2005)
14. Grossberg, S.: Nonlinear neural networks: Principles, mechanisms, and architectures. Neural Networks 1, 17–61 (1988)
15. Hopp, J.J., Fuchs, A.F.: The characteristics and neuronal substrate of saccadic eye movement plasticity. Progress in Neurobiology 72(1), 27–53 (2004)
16. Iossifdis, I., Bruckhoff, C., Theis, C., Grote, C., Faubel, C., Schöner, G.: CORA:An Anthropomorphic Robot Assistant for Human Environment. In: Proceedings of the 2002 IEEE Int. Workshop on Robot and Human Interactive Communication, Berlin, Germany, September 25-27, pp. 392–398 (2002)

17. Itti, L., Koch, C.: Computational modeling of visual attention. Nature Reviews Neuroscience 2, 1–11 (2001)
18. Kopecz, K., Schöner, G.: Saccadic motor planning by integrating visual information and pre-information on neural, dynamic fields. Biological Cybernetics 73, 49–60 (1995)
19. Lefèvre, P., Quaia, C., Optican, L.M.: Distributed model of control of saccades by superior colliculus and cerebellum. Neural Networks 11 (1998)
20. Lomp, O., Zibner, S.K.U., Richter, M., Rañó, I., Schöner, G.: A Software Framework for Cognition, Embodiment, Dynamics, and Autonomy in Robotics: cedar. In: Mladenov, V., Koprinkova-Hristova, P., Palm, G., Villa, A.E.P., Appollini, B., Kasabov, N. (eds.) ICANN 2013. LNCS, vol. 8131, pp. 475–482. Springer, Heidelberg (2013)
21. Munoz, D.P., Wurtz, R.H.: Saccade-related activity in monkey superior colliculus. I. Characteristics of burst and buildup cells. Journal of Neurophysiology 73(6), 2313–2333 (1995)
22. Optican, L.M.: Sensorimotor transformation for visually guided saccades. Annals of the New York Academy of Sciences 1039, 132–148 (2005)
23. Optican, L.M., Quaia, C.: Distributed Model of Collicular and Cerebellar Function during Saccades. Annals of the New York Academy of Science 956, 164–177 (2002)
24. Pelisson, D., Alahyane, N.: Sensorimotor adaptation of saccadic eye movements. Neuroscience & Biobehavioral Reviews 34, 1103–1120 (2010)
25. Quaia, C., Joiner, W.M., FitzGibbon, E.J., Optican, L.M., Smith, M.A.: Eye movement sequence generation in humans: Motor or goal updating? Journal of Vision 10(14) (2010)
26. Quaia, C., Lefèvre, P., Optican, L.M.: Model of the control of saccades by superior colliculus and cerebellum. Journal of Neurophysiology 82(2), 999–1018 (1999)
27. Richter, M., Sandamirskaya, Y., Schöner, G.: A robotic architecture for action selection and behavioral organization inspired by human cognition. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS (2012)
28. Robinson, D.A.: Oculomotor control signals. In: Lennerstrand, G., Bach-y Rita, P. (eds.) Basic Mechanisms of Ocular Motility and Their Clinical Implications, pp. 337–374. Pergamon Press, Oxford (1975)
29. Sandamirskaya, Y.: Dynamic Neural Fields as a Step Towards Cognitive Neuromorphic Architectures. Frontiers in Neuroscience 7, 276 (2013)
30. Sandamirskaya, Y., Richter, M., Schöner, G.: A neural-dynamic architecture for behavioral organization of an embodied agent. In: IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL EPIROB 2011) (2011)
31. Sandamirskaya, Y., Schöner, G.: An Embodied Account of Serial Order: How Instabilities Drive Sequence Generation. Neural Netw. 23(10), 1164–1179 (2010)
32. Sandamirskaya, Y., Zibner, S.K.U., Schneegans, S., Schöner, G.: Using Dynamic Field Theory to extend the embodiment stance toward higher cognition. New Ideas in Psychology 31(3), 322–339 (2013)
33. Schöner, G.: Dynamical Systems Approaches to Cognition. In: Sun, R. (ed.) Cambridge Handbook of Computational Cognitive Modeling, pp. 101–126. Cambridge University Press, Cambridge (2008)
34. Schweighofer, N., Arbib, M.A., Dominey, P.F.: A model of the cerebellum in adaptive control of saccadic gain. Biological Cybernetics 75(1), 19–28 (1996)
35. Scudder, C.A.: A new local feedback model of the saccadic burst generator. Journal of Neurophysiology 59(5), 1455–1475 (1988)
36. Spencer, J.P., Schöner, G.: Embodied Approach to Cognitive Systems: A Dynamic Neural Field Theory of Spatial Working Memory. In: . . . Annual Conference of the Cognitive . . . , pp. 2180–2185 (2006)

37. Steve, N.G., Charles, T., Benoît, G.: Saccade learning with concurrent cortical and subcortical basal ganglia loops. arXiv preprint arXiv:1312.5212, 1–34 (2013)
38. Trappenberg, T.P., Dorris, M.C., Munoz, D.P., Klein, R.M.: A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. Journal of Cognitive Neuroscience 13(2), 256–271 (2001)
39. Tweed, D., Vilis, T.: A two dimensional model for saccade generation. Biol. Cybern. 52, 219–227 (1985)
40. Wilimzig, C., Schneider, S., Schöner, G.: The time course of saccadic decision making: dynamic field theory. Neural Networks: the Official Journal of the International Neural Network Society 19(8), 1059–1074 (2006)
41. Wilimzig, C., Schöner, G.: The Emergence of Stimulus-Response Associations from Neural Activation Fields: Dynamic Field Theory. In: Proceedings of the Twenty-Seventh Annual Cognitive Science Society, pp. 2359–2364. Cognitive Science Society, Stresa (2005)
42. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13, 55–80 (1973)
43. Wurtz, R.H., Optican, L.M.: Superior colliculus cell types and models of saccade generation. Current Opinion in Neurobiology 4, 857–861 (1994)
44. Zimmermann, E., Burr, D., Morrone, M.C.: Spatiotopic Visual Maps Revealed by Saccadic Adaptation in Humans. Current Biology (2011)