

## 9. THE OCULOMOTOR SYSTEM: A BIOLOGICAL EXAMPLE

### 9.1 Oculomotor Control and Superior Colliculus

When reading, looking at a painting, or steering a car the eyes make numerous movements, most of which consist of short, jerky jumps called *saccades*. Saccades function to direct a small, slightly deepened region (the *fovea*) in the center of the retina towards particular locations in the visual field. An extraordinarily high number of light sensitive cells reside in the fovea, providing particularly high resolution of any object whose image resides there. As soon as the attention of an observer is caught by an object the image of which is outside of the fovea, a saccade moves the eyeball such that the image leaps into the fovea. When reading, for example, only a few letters can be simultaneously in focus, and even single words longer than two or three letters must be viewed piecewise by successive saccades. Most of the time, the saccade traverses an angle between four minutes and forty degrees ( Korn 1982). Larger changes in direction usually occur only when the eyes and head move simultaneously.

The control of eye movements, the oculomotor control, has often been the subject of neurophysiological investigations. The advantage one has in investigating the oculomotor system stems from the particularly close relation between eye motions and motor nerve signals. Due to the relatively small and constant mass of the eyeball and the capability of the muscle apparatus of the eye to react with a comparatively large and extremely rapid deployment of force, inertial effects play only a minor role, and the motions of the eyeball provide an accurate mapping of the nerve signals that control the muscle.

It is not precisely known where the decision is made concerning which object of the visual field should be in the primary focus. Clinical findings lead to the inference that parts of the *parietal lobe* play a vital role in this decision process ( Wurtz et al. 1986). Experiments have shown that the saccades are

triggered in the *superior colliculus*, a mounded, multilayered neuron sheet that is located in the upper region of the brain stem ( Sparks and Nelson 1987). The relation between the location of receptors on the retina which are, *e.g.*, excited by a small light point, and the place of neurons in the upper layer of the *superior colliculus* that are simultaneously excited, is continuous and topology conserving. This implies that a topographic *sensory map* from the retina to the upper layer of the *superior colliculus*, a so-called *retinotopic map*, is realized. In contrast to that, and essential for the saccadic control of eye movements, the lower layer provides an example of a *motor map* similar to the one described in Chapter 8. Locations in this layer correspond in an ordered way to saccadic changes in view direction that can be triggered by excitation of neurons at the corresponding locations. Such excitations can be artificially created by stimulation via inserted electrodes. With excitations thus invoked, the direction of the saccades turns out to be quite independent from the intensity of the stimulus; rather, their direction is mainly determined by the *location of the stimulus* in the layer.

There is an interesting relationship between the retinotopic map in the upper layer and the motor map in the lower layer. The layers lie against one another such that local excitations of the neurons in the lower layer trigger a saccade which moves the fovea to a location which was previously kept by the receptive fields of the corresponding top-lying neurons of the upper layer. In other words, if one transfers an excitation in the upper layer that was caused by a localized light stimulus on the retina to the directly underlying neurons of the motor map, then the result is an eye motion that leads the fovea to the light stimulus. This correspondence led to the formulation of the *fovealization hypothesis*: According to this hypothesis, the alignment of the sensory and the motor map in the *superior colliculus* serves to create saccades for the centering of images in the fovea (Robinson 1972).

The correct functioning of such a system demands that both maps correspond precisely to one another. This requires an exact, topographically ordered wiring from the retina to the sensory layer and also an exact, topographically organized assignment of saccadic motion vectors to the neurons of the motor layer. As experiments have also shown, the oculomotor system can adaptively follow changes in the interrelation between visual input and the saccades needed for centering. For example, test subjects were equipped with both contact lenses and eye glasses such that the corrections of the glasses and the contact lenses exactly cancelled each other. Because of the distance between glasses and contact lenses, such a combination acts like a (weak) “Galileic

telescope,” and previously correct saccades now miss their target by some degrees because of the combined devices. In the beginning the eye reacts with additional corrective motions after each saccade to compensate for the errors artificially created. This state of affairs does not remain, however; it has been experimentally determined that already after 9–14 minutes the oculomotor system has adapted the saccades so well that subsequent eye motions can no longer be distinguished from eye motions of subjects without the contact lenses/glasses combination. Larger corrective motions were no longer required (Henson 1977). This shows that beyond the formation of appropriate, topographically ordered connections, the oculomotor system can use the appearance of errors to adaptively change its saccades.

In the following we present a simple model which, on the basis of a few simple learning principles, can adaptively form a sensory map and corresponding motor map in order to control saccades. As before, we consider only a minimum of biological detail in order to motivate the following mathematical model. The starting point of our model is a lattice  $A$  of formal neurons, as they were described in Chapter 8. Each formal neuron  $\mathbf{r}$  corresponds to a receptive field centered on the retina at the location  $\mathbf{w}_{\mathbf{r}}$ , and the excitation of this neuron leads to a saccade which causes a translation of a visual object on the retina by a vector  $\mathbf{w}_{\mathbf{r}}^{(\text{out})}$ . The two layers of sensory and motor neurons are replaced by a single layer. The vectors  $\mathbf{w}_{\mathbf{r}}$  and  $\mathbf{w}_{\mathbf{r}}^{(\text{out})}$ , respectively, bring together the location of the receptive field of a sensory neuron  $\mathbf{r}$  and the corresponding saccade which is triggered when the underlying motor neuron  $\mathbf{r}$  is stimulated. Furthermore, we describe the correspondence between the visual stimulus and the resultant saccade simply by a pair of values  $(\mathbf{w}_{\mathbf{r}}, \mathbf{w}_{\mathbf{r}}^{(\text{out})})$  of the centrally localized neuron. In reality, the resultant saccade is determined by a *group* of excited neurons localized at  $\mathbf{r}$ . (In particular, in our model we do not imitate the continuous interpolation which is caused by the simultaneous activity of many neurons.) As long as saccades  $\mathbf{w}_{\mathbf{r}}^{(\text{out})}$  do not lead to a centering of the visual stimulus at the retina location  $\mathbf{w}_{\mathbf{r}}$ , corrective saccades are necessary. The following section will show how these corrective saccades can be gradually reduced by a simple learning method and how, simultaneous to that, the arrangement of the receptive fields, whose center points are determined by the vectors  $\mathbf{w}_{\mathbf{r}}$ , can organize itself topographically.

## 9.2 A Stepwise Method for Learning Saccades

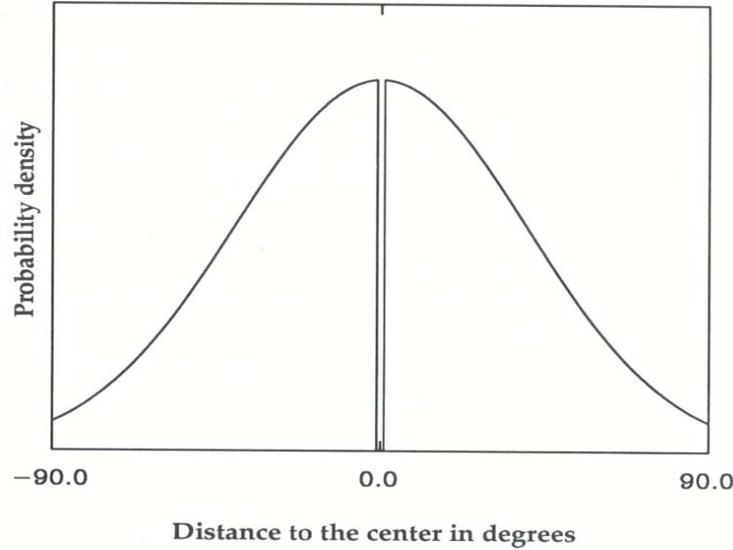
The learning algorithm of our model is suggested by the corrective saccades of the oculomotor system. If an object within the visual field draws one's attention, a corresponding saccade is triggered. This saccade might not lead the object's image precisely into the fovea in which case a second, corrective saccade would occur to reduce this error. If this correction step really brings an improvement, *i.e.*, the image moves closer to the fovea, the corrective saccade will be accepted for later eye movements. This means that a model neuron that has triggered a wrong saccade will next time trigger a saccade that is the sum of the wrong saccade plus the correction step. Only a single correction step is allowed in our model. After that, a new stimulus is presented, *i.e.*, the attention is shifted to a new, randomly chosen object in the visual field.

The model assumes that the new object of attention is randomly chosen each time, yet the choice is governed by a fixed probability density which qualitatively follows the natural distribution of receptors on the retina. The region in the fovea is unused because, there, no saccades become triggered. Figure 9.1 shows the chosen probability distribution of input stimuli as a function of the distance from the center of the retina. It corresponds to a Gaussian distribution with a width of  $\sigma_r = 40^\circ$  notched out at the center. The region of the fovea has a radius of  $1.0^\circ$  (Korn 1982).

The formation of the wiring between neurons and light receptors in the retina, *i.e.*, the sensory map, occurs in analogy to the formation of the somatotopic map of the hand (Chapter 7). For the simulation we use the substitute dynamics described in Chapter 7 as well as two-dimensional coordinate vectors  $\mathbf{w}_r$ . Again, an image point at the location  $\mathbf{v}$  on the retina selects that neuron  $s$  for which the distance  $\|\mathbf{w}_s - \mathbf{v}\|$  is smallest and causes an adaptation step (70) for the vectors  $\mathbf{w}_r$  which are the determining quantities of the "input wiring."

As a new feature, we add the learning of the output values, the saccade vectors  $\mathbf{w}_r^{(\text{out})}$ . The two-component vector  $\mathbf{w}_r^{(\text{out})}$  depicts the displacement of an image point on the retina that results from this saccade. In the ideal case this displacement leads into the center, *i.e.*, if one considers each vector  $\mathbf{w}_r^{(\text{out})}$  as "attached" to the receptor at  $\mathbf{w}_r$ , then all vectors must precisely end in the fovea.

These learning rules can now be mathematically formulated as follows. If  $\mathbf{v}$  is the distance vector on the retina from the fovea to an image point and



**Abb. 9.1:** The probability density of the input stimuli in our model as a function of the distance to the center of the retina. This probability density roughly corresponds to the receptor density on the retina excluding the region of the fovea.

$\mathbf{w}_s^{(\text{out})}$  the saccade of the neuron  $s$  which is most strongly excited by the image at  $\mathbf{v}$ , then the new retinal location of the image after executing the saccade  $\mathbf{w}_s^{(\text{out})}$  is given by  $\mathbf{v}' = \mathbf{v} + \mathbf{w}_s^{(\text{out})}$ . If  $\mathbf{v}'$  does not fall into the fovea, *i.e.*,  $\|\mathbf{v}'\| > R_{\text{fovea}}$ , where  $R_{\text{fovea}}$  corresponds to a diameter of  $1^\circ$ ,  $\mathbf{v}'$  selects, as  $\mathbf{v}$  did previously, a neuron  $s'$  which triggers another saccade  $\mathbf{w}_{s'}^{(\text{out})}$ , the *corrective saccade*. Every time this yields an improvement, *i.e.*, every time when  $\|\mathbf{v} + \mathbf{w}_s^{(\text{out})} + \mathbf{w}_{s'}^{(\text{out})}\| < \|\mathbf{v} + \mathbf{w}_s^{(\text{out})}\|$ , the original saccade  $\mathbf{w}_s^{(\text{out})}$  becomes improved by the corrective saccade  $\mathbf{w}_{s'}^{(\text{out})}$ :

$$\mathbf{u} = \mathbf{w}_s^{(\text{out,old})} + \mathbf{w}_{s'}^{(\text{out,correction})}. \quad (9.1)$$

Here again, we take advantage of the continuity of the mapping between retinal locations and saccade vectors by allowing the model neurons in the neighborhood to participate in the learning process of the output value  $\mathbf{w}_s^{(\text{out})}$ , in analogy with the learning step for  $\mathbf{w}_r$ . Just as in Chapter 8 we employ the formula

$$\mathbf{w}_r^{(\text{out,new})} = \mathbf{w}_r^{(\text{out,old})} + \epsilon' h'_{rs}(\mathbf{u} - \mathbf{w}_r^{(\text{out,old})}). \quad (9.2)$$

Here  $\mathbf{u}$  is the improved estimation for  $\mathbf{w}_s^{(\text{out})}$  defined in Eq. (9.1), and just as  $h_{rs}$  did previously,  $h'_{rs}$  depends only on the lattice distance  $d_{rs} = \|\mathbf{r} - \mathbf{s}\|$  between the neurons  $\mathbf{r}$  and  $\mathbf{s}$ . The parameter  $\epsilon'$  measures the learning step width.

Thus, our model for the oculomotor system can be summarized by the following steps:

0. Begin with a random assignment of the elements  $\mathbf{r}$  of the lattice  $A$  with receptive fields determined by the synaptic strengths  $\mathbf{w}_r$ , and with a random assignment of the saccades  $\mathbf{w}_r^{(\text{out})}$  to be triggered.
1. In accordance with the probability distribution  $P(\mathbf{v})$  shown in Fig. 9.1, choose a vector  $\mathbf{v}$  which represents a new “visual input.”  $\mathbf{v}$  points from the fovea to the retinal location of the new input.
2. Determine the center of excitation  $\mathbf{s}$  in the layer  $A$  of formal neurons by the condition

$$\|\mathbf{v} - \mathbf{w}_s\| \leq \|\mathbf{v} - \mathbf{w}_r\|, \quad \text{for all } \mathbf{r} \in A. \quad (9.3)$$

3. Perform a learning step

$$\mathbf{w}_r^{(\text{new})} = \mathbf{w}_r^{(\text{old})} + \epsilon h_{rs}(\mathbf{v} - \mathbf{w}_r^{(\text{old})}), \quad \text{for all } \mathbf{r} \in A \quad (9.4)$$

for the positions of the receptive fields.

4. Execute the saccade  $\mathbf{w}_s^{(\text{out})}$ , so that the position  $\mathbf{v}$  of the image is changed according to

$$\mathbf{v}' = \mathbf{v} + \mathbf{w}_s^{(\text{out})}.$$

- 5.a) If the visual object lies in the fovea, *i.e.*,  $\|\mathbf{v}'\| < R_{\text{fovea}}$ , then go back to step 1.
- 5.b) If the image does not lie in the fovea, *i.e.*,  $\|\mathbf{v}'\| \geq R_{\text{fovea}}$ , then determine the new center of excitation  $\mathbf{s}'$ , belonging to the retinal location  $\mathbf{v}'$  of the image, according to

$$\|\mathbf{v}' - \mathbf{w}_{s'}\| \leq \|\mathbf{v}' - \mathbf{w}_r\| \quad \text{for all } \mathbf{r} \in A. \quad (9.5)$$

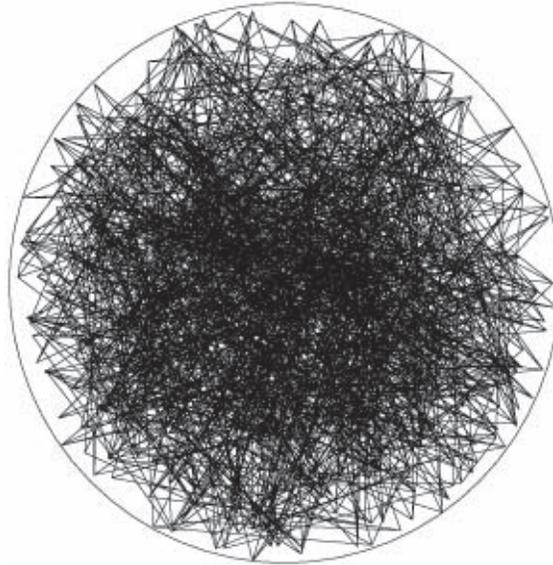
Execute a corrective saccade  $\mathbf{w}_{\mathbf{s}'}^{(\text{out})}$ . If the correction is an improvement, *i.e.*,  $\|\mathbf{v}' + \mathbf{w}_{\mathbf{s}'}^{(\text{out})}\| < \|\mathbf{v}'\|$ , perform a learning step for the saccades according to

$$\mathbf{u} = \mathbf{w}_{\mathbf{s}}^{(\text{out})} + \mathbf{w}_{\mathbf{s}'}^{(\text{out})} \quad (9.6)$$

$$\mathbf{w}_{\mathbf{r}}^{(\text{out,new})} = \mathbf{w}_{\mathbf{r}}^{(\text{out,old})} + \epsilon' h'_{\mathbf{rs}}(\mathbf{u} - \mathbf{w}_{\mathbf{r}}^{(\text{out,old})}) \quad \forall \mathbf{r} \in A \quad (9.7)$$

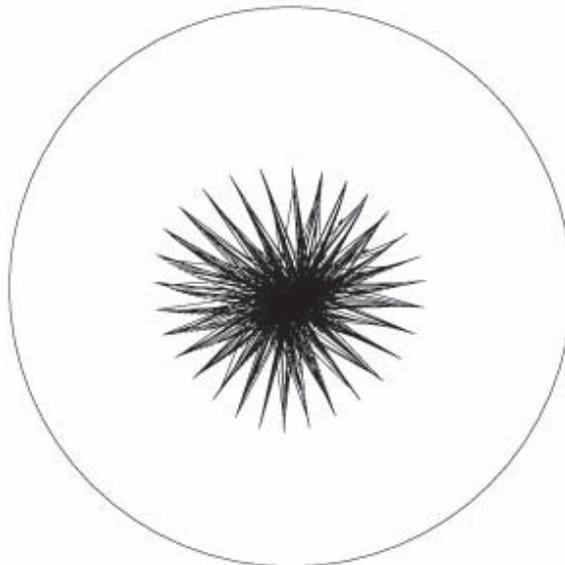
and go back to step 1. If the correction does not yield an improvement, *i.e.*,  $\|\mathbf{v}' + \mathbf{w}_{\mathbf{s}'}^{(\text{out})}\| \geq \|\mathbf{v}'\|$ , then omit the learning step and return to step 1.

Steps 1–3 constitute the algorithm for the formation of the topology-conserving map onto  $A$  as explained in Chap. 7. The newly added steps 4 and 5 change the assignment of  $A$  with output values  $\mathbf{w}_{\mathbf{r}}^{(\text{out})}$ . The change occurs according to the principle of *unsupervised learning* as mentioned in the previous chapter. The correct control actions must be discovered by the learning algorithm itself. This requires a search process in the space of possible values. In our algorithm this occurs by introducing corrective saccades. Again, a reward function is available only to tell how well the control has mastered the given task. In our case we employ the binary criterion “came closer to the fovea” versus “moved away from the fovea,” which decides between the alternatives “learn” versus “ignore.” By using the function  $h'_{\mathbf{rs}}$ , step 5 causes the neighboring neural units  $\mathbf{r}$  to participate in the learning process when the search of the neural unit  $\mathbf{s}$  is successful. This not only accelerates significantly the learning process but also contributes strongly to the convergence of the system to the desired state. Without the participation of neighboring neurons in the learning process, some of the neurons remain in a state in which saccades grossly deviate from the target value. This will be elucidated by the following presentation of simulation results.

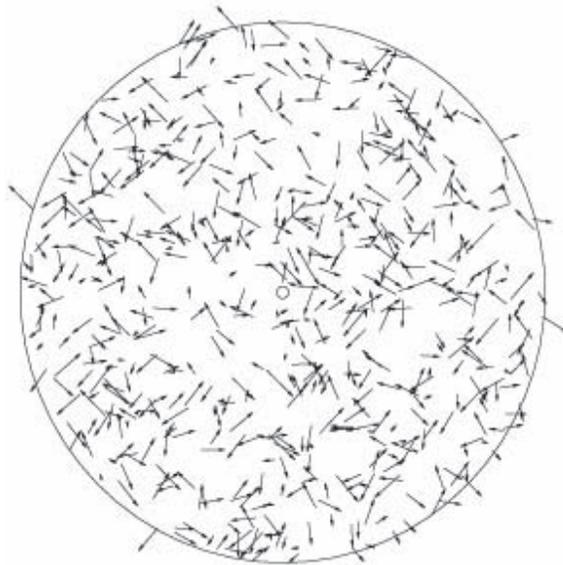


**Abb. 9.2:** Learning saccadic eye movements according to Eqs. (9.3) – (9.11). The figure shows the lattice in its initial configuration on the retina. The wiring between the receptors and the neural net is chosen completely at random.

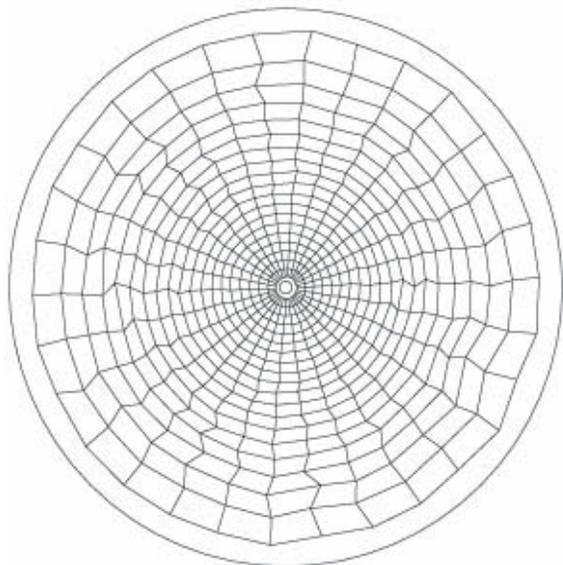
### 9.3 A Computer Simulation



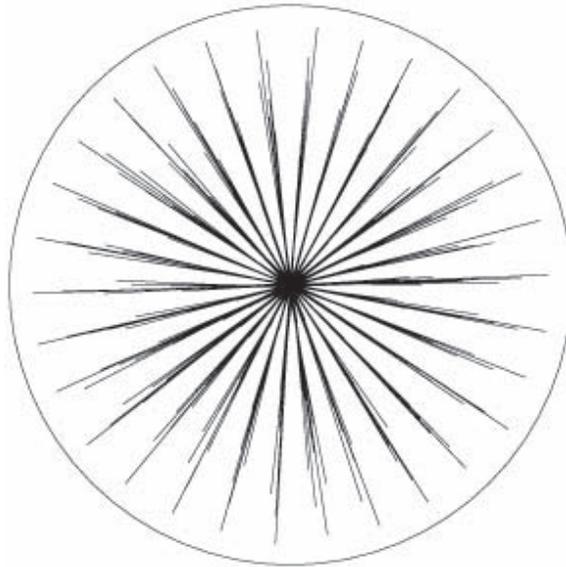
**Abb. 9.5:** After 20,000 learning steps the saccades have already become crudely ordered. All vectors point towards the center.



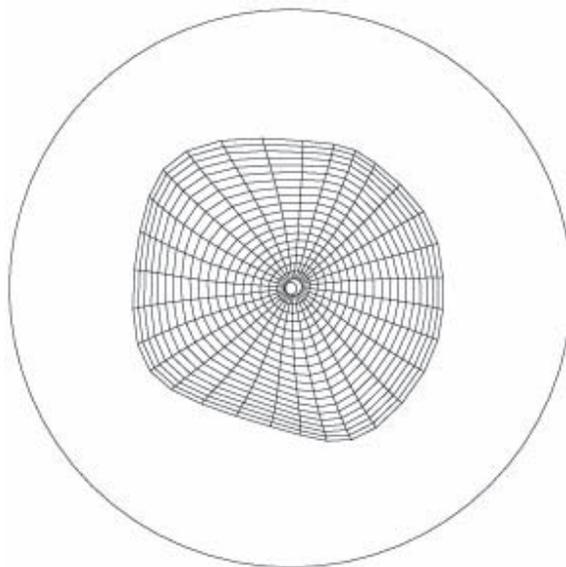
**Abb. 9.3:** The saccades associated with the lattice points at the start of the simulation. The direction and length of each vector are chosen at random. The variation in length corresponds to eye rotation angles between  $0^\circ$  through  $9^\circ$ .



**Abb. 9.6:** Learning saccadic eye movements according to Eqs. (9.3) – (9.11). The figure shows the state after 200,000 learning steps with an assignment between receptors and lattice points which is neighborhood-conserving and which follows the density  $P(\mathbf{v})$ .

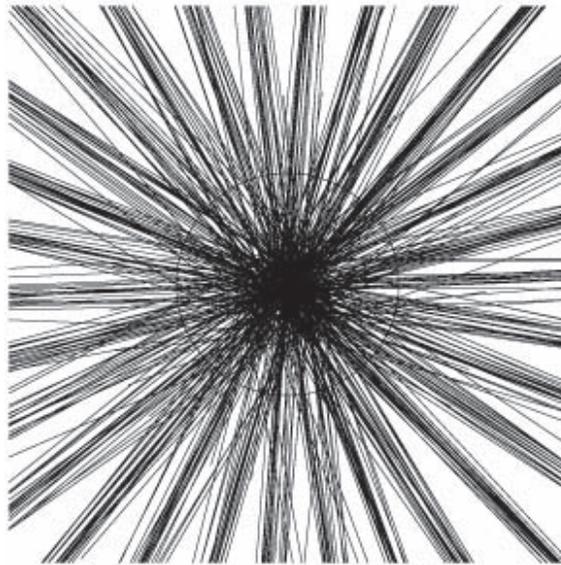


**Abb. 9.7:** After 200,000 learning steps the saccades point towards the center of the retina. A blowup of the foveal region (Fig. 9.5) makes visible the precise positions of the endpoints of the vectors  $\mathbf{w}_r^{\text{out}}$ .



**Abb. 9.4:** Learning saccadic eye movements according to Eqs. (9.3) – (9.11). The figure shows the lattice after 20,000 learning steps. At this time a recognizable ordering of the receptive fields has already taken place.

In a computer simulation of the above learning algorithm, we have chosen a ring-shaped lattice  $A$  with  $20 \times 30$  neurons. The simulation parameters were chosen as follows:  $\epsilon(t) = 1/(1 + 125t/t_{max})$ ,  $\sigma(t) = 10 \cdot \exp(-5t/t_{max})$ ,  $\epsilon'(t) = \exp(-5(t/t_{max})^2)$ , and  $\sigma'(t) = \exp(-5(t/t_{max})^2)$  with  $t_{max} = 200,000$ . Recorded at three different stages of the simulation, Figs. 9.2-9.4 show the assignment of  $A$  with receptive fields and the saccades to be executed.



**Abb. 9.8:** A blowup of Fig. 9.4b in the  $1^\circ$  foveal region which is indicated by the circle. One can see that all saccades actually lead into the foveal region and that the learning process has been successful.

The parameters for the learning of the receptive fields are chosen in such a way that they lose a large part of their plasticity at an early stage of the learning. This early freezing of the receptive fields is necessary because, when a receptive field is shifted, the target value of the saccade that is required at that location also changes. Without freezing, saccades that have already been correctly learned would become invalid under further changes of the receptive fields. For this reason the learning of the saccades requires the stabilization of the receptive fields.

Because of the rotational symmetry of the input stimulus distribution, the Kohonen net we employ is also rotationally symmetric. In the simulation the

net consists of twenty concentric rings with thirty neural units each. Every neural unit has two radial and two circumferential neighbors. The depiction of the net is done in an imaginary “projection onto the retina,” *i.e.*, each figure shows the retina, and for each neuron  $\mathbf{r}$  the center  $\mathbf{w}_{\mathbf{r}}$  of its receptive field is marked on the retina. To indicate the adjacency of neural units, we have connected the marks of lattice neighbors by lines.

The outer ring represents the whole visual field from  $-90^\circ$  to  $+90^\circ$ . The innermost ring encircles the fovea with a radius  $R_{fovea}$  which corresponds to an area of the visual field of  $1^\circ$ . The saccades  $\mathbf{w}_{\mathbf{r}}^{(out)}$  at each lattice point are drawn as arrows; they specify the shifting of an image on the retina when the saccade is executed.

Figure 9.2 displays the initial-state assignment with random synaptic strengths  $\mathbf{w}_{\mathbf{r}}$  and random saccades  $\mathbf{w}_{\mathbf{r}}^{(out)}$ . The magnitude of the saccades varies between  $0^\circ$  and  $9^\circ$ . Figure 9.3 shows the situation after 20,000 learning steps. At this stage a regular assignment between the retina and lattice points has already emerged and all the saccades are directed towards the center. In Fig. 9.4, after 200,000 learning steps, a well-ordered connectivity with the retina that is in accordance with the input stimulus distribution  $P(\mathbf{v})$  has become established. The receptive fields lie more densely in the region around the fovea than farther out, where the decrease of the Gaussian input distribution gives rise to a lower resolution. All the corresponding saccades now actually point towards the center of the retina. Because, in Fig. 9.4b, the positions of the endpoints of the vectors  $\mathbf{w}_{\mathbf{r}}^{(out)}$  cannot be clearly observed, Fig. 9.5 shows a blowup of the foveal region: one can notice that all saccades actually lead the targeted image into the fovea; the learning method has been successful.

The precise form of the distance measure between two neural units in the lattice is inconsequential for the organizational process to converge; however, sometimes a certain metric may fit a problem better than other distance measures. For example, in the above simulation we used the “Manhattan” rather than the Euclidean metric. (The “Manhattan-distance” between two lattice points  $\mathbf{r}$  and  $\mathbf{s}$  is the minimal number of lattice steps required to go from  $\mathbf{r}$  to  $\mathbf{s}$ .) Equal distances in the lattice, *i.e.*, in the neuron layer, can correspond to completely different distances between the corresponding receptive fields in the space of input signals. It is the distances between the centers of receptive fields which become visible in pictures such as Figs. 9.2–9.4. In contrast, it is the distances in the lattice which determine the spatial interaction between the neurons themselves and, thereby, determine the

distance-dependent adaptation steps in the model. This feature can be quite advantageous, as is particularly manifest in the vicinity of the fovea. There, receptors that are directly opposite to each other lie close together but have to learn saccades that differ as much as saccades that belong to receptive fields directly opposite and at the periphery of the retina. At both receptor pairs the required saccade directions of the partners differ by the same angle, namely  $180^\circ$ . Therefore, it makes sense to use the “Manhattan” metric which yields, for the foveal and peripheral pair, the same lattice distance between the diametrically opposite neural units, namely 15 lattice sites.

## 9.4 The Convergence of the Learning Process

In this section we demonstrate that under certain conditions the algorithm which was employed to learn the saccadic eye movements must converge. For this purpose we make two assumptions. First, we consider a stage at which a corrective saccade always gives rise to an improvement. This is valid when all vectors have oriented themselves towards the center which is, as one can see in Fig. 9.3, already the case after comparatively few learning steps (in most cases 10% of the total number is sufficient). In addition, we assume that the receptive fields on the retina lie dense enough that we may make the transition from discrete lattice points to a continuum of  $\mathbf{r}$ -values, as we did in Chapter 5 when we considered the representation of the ultrasonic spectrum on the bat’s auditory cortex. By this assumption the sets of values  $\mathbf{w}_{\mathbf{r}}$  and  $\mathbf{w}_{\mathbf{r}}^{(\text{out})}$  meld into continuous vector fields  $\mathbf{w}(\mathbf{r})$  and  $\mathbf{w}^{(\text{out})}(\mathbf{r})$ . Since we assume the lattice to be maximally ordered, the inverse  $\mathbf{r}(\mathbf{w})$  exists, and we can define

$$\mathbf{w}^{(\text{out})}(\mathbf{u}) = \mathbf{w}^{(\text{out})}(\mathbf{r}(\mathbf{u})). \quad (9.8)$$

$\mathbf{w}^{(\text{out})}(\mathbf{u})$  is the saccade which is triggered if the visual stimulus is at the retina location  $\mathbf{u}$ . In addition, a glance at the simulation data shows that  $\sigma'$ , the range of the interactions between neighbors, has a very small value from the start and decreases monotonically. We will see that from the time when all vectors are beginning to point towards the inner region, cooperation between neighbors is no longer necessary for convergence. If we set  $h'_{\mathbf{rs}} = \delta_{\mathbf{rs}}$ , then only the saccade at the lattice point  $\mathbf{s}$ , in whose receptive field the stimulus  $\mathbf{v}$  was located, experiences an adaptation step.

Under these conditions our learning algorithm can be mathematically formulated as follows: with an input stimulus  $\mathbf{v}$  the saccade  $\mathbf{w}^{(\text{out})}(\mathbf{v})$  is triggered

and leads to the retinal location  $\mathbf{v} + \mathbf{w}^{(\text{out})}(\mathbf{v})$ . The corrective saccade is then given by  $\mathbf{w}^{(\text{out})}(\mathbf{v} + \mathbf{w}^{(\text{out})}(\mathbf{v}))$ . Thus, the saccade at  $\mathbf{v}$  changes according to step 5 of our algorithm by

$$\Delta \mathbf{w}^{(\text{out})}(\mathbf{v}) = \epsilon' \mathbf{w}^{(\text{out})}(\mathbf{v} + \mathbf{w}^{(\text{out})}(\mathbf{v})). \quad (9.9)$$

It is beneficial to introduce the new variable

$$\mathbf{x}(\mathbf{u}) = \mathbf{u} + \mathbf{w}^{(\text{out})}(\mathbf{u}). \quad (9.10)$$

Here,  $\mathbf{x}(\mathbf{u})$  is the shift which the saccade  $\mathbf{w}^{(\text{out})}(\mathbf{u})$  still lacks to lead an image into the fovea. In our algorithm  $\mathbf{x}(\mathbf{u})$  should converge to zero since at the end it should be true that  $\mathbf{w}^{(\text{out})}(\mathbf{u}) = -\mathbf{u}$ . Equation (9.9) then can be written

$$\begin{aligned} \mathbf{x}^{\text{new}}(\mathbf{v}) &= \mathbf{x}^{\text{old}}(\mathbf{v}) + \epsilon' \mathbf{w}_{\text{old}}^{(\text{out})}(\mathbf{x}^{\text{old}}(\mathbf{v})) \\ &= (1 - \epsilon') \mathbf{x}^{\text{old}}(\mathbf{v}) + \epsilon' [\mathbf{x}^{\text{old}}(\mathbf{v}) + \mathbf{w}_{\text{old}}^{(\text{out})}(\mathbf{x}^{\text{old}}(\mathbf{v}))] \\ &= (1 - \epsilon') \mathbf{x}^{\text{old}}(\mathbf{v}) + \epsilon' \mathbf{x}^{\text{old}}(\mathbf{x}^{\text{old}}(\mathbf{v})). \end{aligned} \quad (9.11)$$

For estimation purposes we now want to make mathematically precise the condition that at some point in time all vectors have become oriented towards the center. “All vectors are oriented towards the center” shall mean

$$\frac{\|\mathbf{u}\| - \|\mathbf{u} + \mathbf{w}(\mathbf{u})\|}{\|\mathbf{u}\|} > \delta, \quad \text{for all } \mathbf{u} \in V, \quad \text{with } \delta > 0. \quad (9.12)$$

There should exist a fixed  $\delta > 0$  which satisfies Eq. (9.12) for all  $\mathbf{u} \in V$  simultaneously. Rearranging (9.12) yields

$$\|\mathbf{x}(\mathbf{u})\| < \|\mathbf{u}\|(1 - \delta), \quad \text{for all } \mathbf{u} \in V. \quad (9.13)$$

If we replace  $\mathbf{u}$  by  $\mathbf{x}(\mathbf{u})$  in Eq. (9.13), then

$$\|\mathbf{x}(\mathbf{x}(\mathbf{u}))\| < \|\mathbf{x}(\mathbf{u})\|(1 - \delta) \quad \text{for all } \mathbf{u} \in V \quad (9.14)$$

is also true. Through the triangle inequality (9.11) becomes

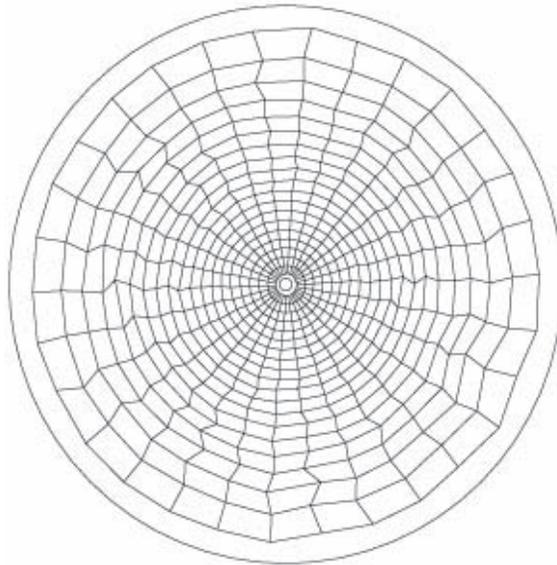
$$\begin{aligned} \|\mathbf{x}^{\text{new}}(\mathbf{v})\| &\leq (1 - \epsilon') \|\mathbf{x}^{\text{old}}(\mathbf{v})\| + \epsilon' \|\mathbf{x}^{\text{old}}(\mathbf{x}^{\text{old}}(\mathbf{v}))\| \\ &< (1 - \epsilon' \delta) \|\mathbf{x}^{\text{old}}(\mathbf{v})\|, \end{aligned} \quad (9.15)$$

since  $0 < \epsilon' < 1$  and, because of (9.12),  $0 < \delta \leq 1$ . Due to the nonexistent lateral interaction, the saccade of a lattice point is changed only if the stimulus falls into the lattice point's receptive field. But then, according to the

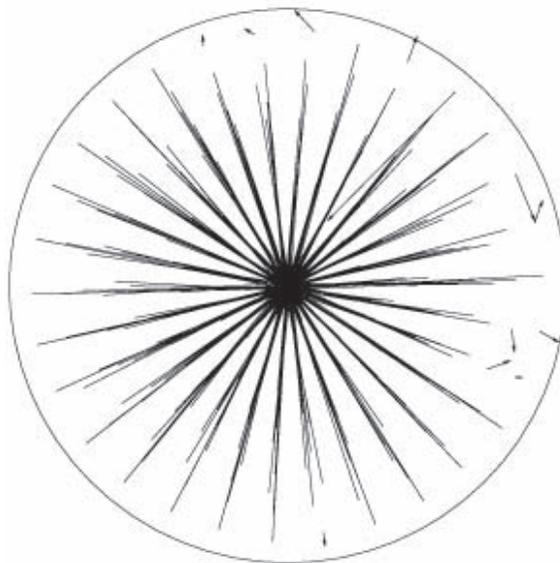
above inequality, the residual error is diminished by at least a factor  $(1 - \epsilon'\delta)$ . On average, after  $N$  learning steps the neural unit  $\mathbf{s}$  has been excited  $H_{\mathbf{s}}N$  times by a stimulus, where

$$H_{\mathbf{s}} = \int_{F_{\mathbf{s}}} P(\mathbf{v}) d\mathbf{v}. \quad (9.16)$$

$F_{\mathbf{s}}$  is the size of the receptive field of neural unit  $\mathbf{s}$  on the retina. Therefore, as the number of learning steps  $N$  increases, the error of the saccade at  $\mathbf{s}$  approaches zero on average faster than  $(1 - \epsilon'\delta)^{H_{\mathbf{s}}N}$ , resulting in the convergence of our algorithm under the above conditions.



**Abb. 9.9:** Learning saccadic eye movements according to Eqs.9.2 – 9.11, but without cooperative learning as described through (9.11), i.e., with  $h_{\mathbf{rs}} = \delta_{\mathbf{rs}}$ . After 200,000 learning steps the rotationally symmetric Kohonen net again displays the same assignment between receptors and lattice points.



**Abb. 9.10:** Without cooperation between neighbors, not all of the saccades learn the targeted value. In particular a few saccades in the outer region of the retina give rise to totally wrong directions.

If after only about 10% of the total number of simulation steps a state is reached where the learning algorithm for the saccades, even without interacting neighbors, safely converges towards the targeted values, then the question arises whether we need the cooperation between neighbors at all, especially since the range was very small from the start as we can see from the simulation parameters. An answer is apparent in Fig 9.6 where a simulation result is shown with the same simulation parameters as before, except with the cooperation between neighbors “turned off.” There we recognize that a few saccades in the outer region of the retina deviate completely from their targeted output values and even point away from the center. This is due to those vectors, which after being acted upon by a stimulus, do not find an appropriate corrective saccade leading the stimulus closer to the center and, thereby, yielding a learning step (9.11). By an initial random assignment there are always a few saccades which evolve by the learning algorithm in a way that, at some point in time, they point into the receptive field of a neuron whose corrective saccade does not give rise to an improvement. These badly learned saccades appear mainly in the retina’s outer region because

there they often find only themselves for corrective saccades. In the learning algorithm without cooperation between neighbors these saccades have no longer the opportunity to rotate into the correct direction. Furthermore, we realize that the overall convergence of the system is slightly worse than in the case of cooperative learning. In particular, the end points of all vectors do not yet lead into the fovea even after the simulation has been terminated after 200,000 steps. This arises simply from the reduced rate of convergence. A reduced convergence rate is to be expected since without cooperation a saccade only changes if the center of a stimulation lies precisely in the saccade's receptive field, and neurons no longer profit from their neighbors. In principle, however, there is no reason why with further learning steps, except for some "runaways" in the outer region, the same desired state should not be reached.

At the beginning of the learning process, the cooperation between neighbors is essential. At this stage, it has the task of rotating all vectors towards the inner region. This cooperation is indispensable for a successful, overall convergence. The isolated saccades which continuously point towards the outside and which at the end of the simulation without cooperative learning would not have changed their direction are now shifted towards the center by their neighbors with more favorable starting values. Indeed, with cooperation between neighbors, all saccades point towards the center after only 10% of the learning steps (Fig. 9.3), creating a basis for the desired convergence of the total system.

## 9.5 Measurements on Human Subjects

In the model for the learning of oculomotor control presented above, we employed, with the introduction of corrective saccades, a very simple learning principle. Perhaps too simple, since measurements by Becker and Fuchs (1969) show that the simulation results of the algorithm do not quite agree with experimental observations. The experiments of Becker and Fuchs with human subjects show that the saccades almost never lead directly into the fovea. One can infer, however, that the "mistakes" occur by design since the errors of the first saccades are not randomly distributed around the fovea. The first saccades are, with few exceptions, too short (*undershoot*). In fact, it is believed that the intentional use of a first saccade which is too short gives rise to advantages, possibly in a better planning and easier execution

of subsequent eye movements. Compared to those observations, our model learns its saccadic eye motions much too “well” because at the end of the learning process all of our saccades precisely lead in the fovea. In our simple example we have omitted the complicated aspect of planning and restricted ourselves to the images of immobile objects. An extension of the learning model must deal with the tracking of objects, where planning ahead will play an important role. Perhaps the intentional use of *undershoots* could provide advantages for a learning method extended in such a way.

Through the pole-balancing and saccadic-control problems, we have seen how self-organizing, topology-conserving maps can be used in a natural way for the learning of input/output relations in the form of adaptively organized “look-up” tables. The topology-conserving feature of the map makes possible the cooperation between neighboring neural units, which strongly contributes to the method’s rapid convergence. This will be developed further in the following chapters. The adaptive capabilities of the map make possible an automatic optimization of the choice of value pairs represented in the table. Nonetheless, for mappings between higher dimensional spaces, a very high number of value pairs must be stored. This difficulty can be somewhat eased by the use of locally valid linear mappings instead of value pairs. With linear mappings, more complex control tasks can be handled, such as those that arise in the motion control of robots. In Chapters 10, 11 and 12 this will be demonstrated with a neural network which learns the control of a triple-jointed robot arm and one which learns to control a robot arm with redundant degrees of freedom.