Authorization to lend and reproduce this thesis

As the sole author of this thesis, I authorize Brown university to lend it to other institutions or individuals for the purpose of scholarly research.

....
Hans-Martin Rudolf Arnoldi

....
(Date)

I further authorize Brown University to reproduce this thesis by photocopying or other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

....
Hans-Martin Rudolf Arnoldi

....
(Date)

Development of a Coordinate Transformation for the Superior Colliculus

by
Hans-Martin Rudolf Arnoldi
Diplom-Informatiker, Technische Universität München, 1990

Thesis
Submitted in partial fulfillment of the requirements for
the Degree of Master of Science
in the Department of Cognitive and Linguistic Sciences at Brown University

May 1993

© Copyright
by
Hans-Martin Rudolf Arnoldi
1993

This thesis by Hans-Martin Rudolf Arnoldi is accepted in its present form by
the Department of Cognitive and Linguistic Sciences as satisfying the
thesis requirement for the degree of
Master of Science

Date ....

Heinrich H. Bülthoff

Approved by the Graduate Council

Date ....

Acknowledgments

This thesis would not have been possible without the reciprocal scholarship of Brown University and the German Academic Exchange Service (DAAD). I would like to thank both institutions for their financial support and their trust in my abilities.

The work also benefits from software tools for neural network simulations developed at Kratzer Automatisierung GmbH, F.R.G.

I am indebted to my advisor Heinrich Bülthoff, and to Jim Anderson who initiated my contact with Brown University. Nathan Intrator, the third member of my committee, provided important guidance at a very critical stage of this work.

I also wish to thank all faculty, students, and staff of the Department of Cognitive and Linguistic Sciences. They have been very supportive during a difficult time.

My deepest gratitude goes to Michael Scharf, who taught me how to express my complicated ideas and thoughts in fluent and readable English.

Contents

List of Figures

(Hecht-Nielson, 1987; Hecht-Nielson, 1988) (Meredith and Clemo, 1989) (Wise and Irvine, 1983) (Hirsch et al., 1985) (King and Palmer, 1985; King and Palmer, 1983) (Jay and Sparks, 1987a; Jay and Sparks, 1984; Lee et al., 1988) (Poeppel, 1973; Poeppel and Logothetis, 1986) (Meredith and Stein, 1986b; Meredith and Stein, 1986a; Meredith et al., 1987; Stein et al., 1989; Meredith and Stein, 1990) (Zipser and Andersen, 1988b; Goodman and Andersen, 1989) (Kohonen, 1982; Hertz et al., 1991; Grossberg, 1988)

Chapter 1

Introduction

Fundamental questions arise when trying to understand a cognitive system: How does it combine different inputs and modalities in order to perceive the environment? How does the system use information from different sources yet still yield the most appropriate behavior, regardless of whether the incoming information is contradictory, or has no interaction at all? The answers relate to every domain within cognitive science. For example, in natural language understanding the interaction of the meaning of a word and its grammatical role is a problem of combining information from different sources. A similar problem in vision is the interaction of different cues for depth perception, such as binocular stereo, shading, texture, and motion.

Using an engineering approach on a restricted class of problems, one could construct an automaton or an algorithm which achieves the most accurate input-output mapping for a given system. But this cannot be an entire solution to the problems raised above. Providing a mechanism of how the brain might combine different sources of informations is one thing; but if it is static, it is prone to fail when the environment is changed. Under this assumption, the focus of attention shifts : How can a cognitive system develop the capabilities to combine different inputs from the environment? The classic answer is that it must use a highly adaptive mechanism, one which learns through experience.

In order to approach these rather abstract questions, this thesis concentrates on a well-established brain structure which accomplishes a multisensory integration: the mammalian superior colliculus (SC). In the deeper layers of the SC visual, auditory, and somatosensory modalities are active in order to guide saccadic eye movements and therefore devote visual attention to certain objects in the environment. The descending efferents of the SC are well known for the initiation of saccades. These neurons orient the fovea as the focus of visual attention towards an "interesting" object. Towards this purpose, the SC receives input from the retina, the cochlea, and somatosensory cells. In mice, it also integrates sensory inputs from the whiskers. The work presented here is restricted to the sensory fusion of auditory and visual information (Fig. 1.1).

There is a limited amount of literature on auditory processing in the SC. One of the more prominent papers (Jay and Sparks, 1987b) showed that the auditory receptive fields in the SC shift in accordance with the position of the eyes. This led to a very recent model (Groh and Sparks, 1992) of the transformation of auditory input into a motor error coordinate system. A "shifter circuit" which controls information flow dynamically (Anderson and van Essen, 1987) and a neural network for reorganization of spatial relationships (Trehub, 1987) could also be utilized by the SC. But all these models fail to account for the experimental

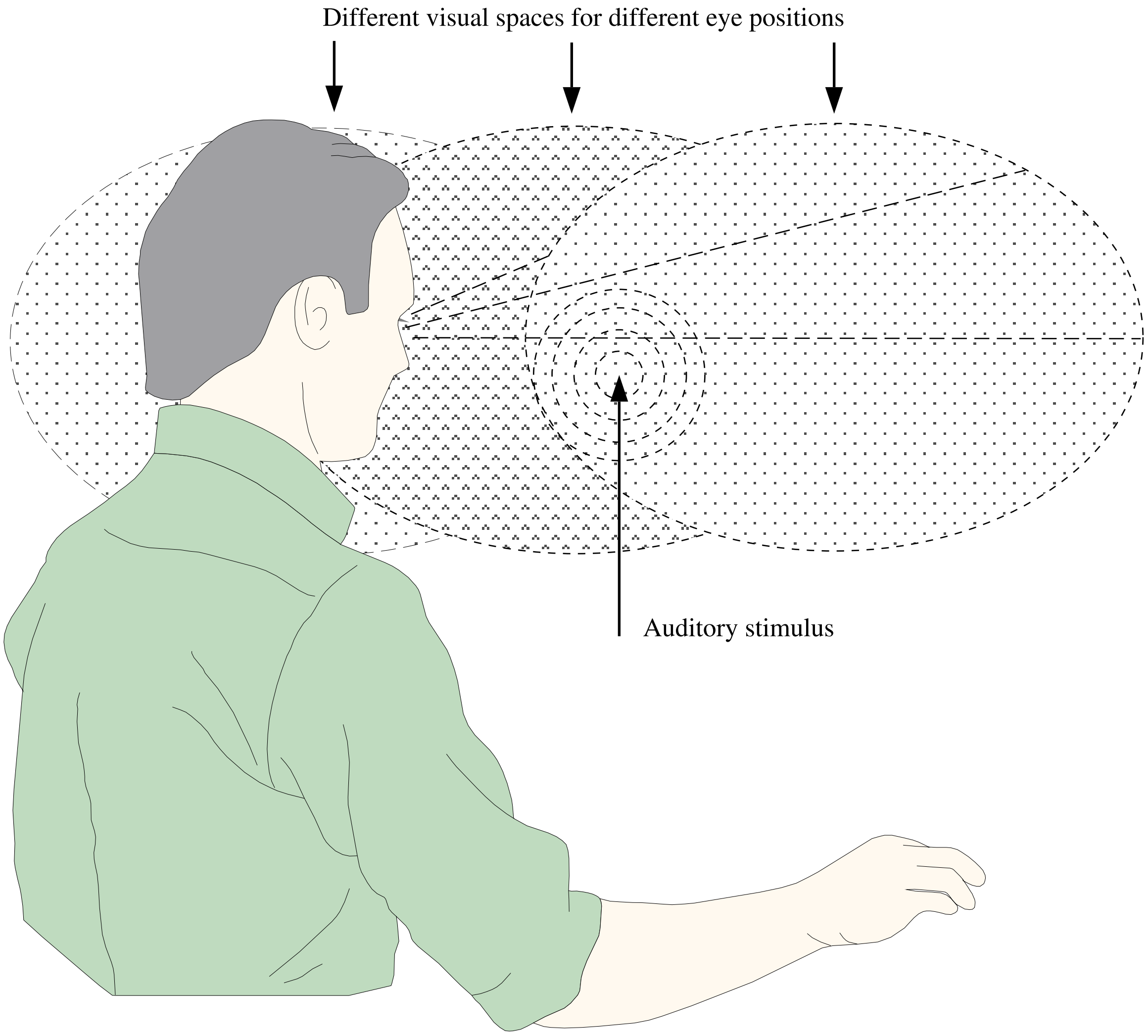

Figure 1.1: The problem: an auditory stimulus changes location within a visual space that moves with the eye.

findings that the auditory receptive fields in the SC change with regard to experience of the environment (Withington-Wray et al., 1990b). In the tradition of Hubel and Wiesel's deprivation experiments on receptive fields in visual cortex (Hubel and Wiesel, 1962) it was demonstrated that the collicular system must be highly adaptive to the environment. A variety of other papers also stress this point (King and Moore, 1991). For the optical tectum of barn owls, an equivalent to the mammalian SC, the same plasticity of the auditory map was found (Knudsen and Brainard, 1991). This thesis attempts to account for these facts, and focuses on a system which develops the transformation of auditory input through visual and auditory experience.

The development of a coordinate transformation was previously addressed in the neural network literature. Zipser and Andersen (1988a) introduced a network which learns the transformation from visual stimuli on the retina into a head-centered coordinate system. This task appears to be the exact opposite of what the model of the SC should do. In the SC head-centered stimuli must be transformed into a coordinate system based on the motor error on the retina. The major drawback of Zipser's and Andersen's network is their use of back propagation, a biologically questionable learning rule. Besides, their transformation happens in a different brain structure, the parietal cortex. Unlike the pathway from auditory cortex to SC the parietal cortex characteristically has very strong feedback projection which could be used to send an error signal. There are also subtle differences concerning the receptive fields employed in parietal cortex and SC.

The back propagation network was also described in supplementary papers (Zipser and Andersen, 1988b; Goodman and Andersen, 1989) which proved to be very valuable for this work, as they describe precisely how the eye position signal influences the process of transformation. In a more recent simulation Mazzoni, Andersen, and Jordan (1991) used a different and biologically more plausible learning rule but produced the same interesting result that artificially learned receptive fields resemble the receptive fields found in the parietal cortex.

The neural computations needed to transform a head-centered auditory stimulus into an oculocentric coordinate system require information about eye position. Without knowledge of how the retina relates to head-centered space, it is extremely complicated to align a visual stimulus with a sound in the environment. A computational approximation similar to Marr and Poggio's algorithm for stereo vision (Marr and Poggio, 1976) could match both sensory inputs without eye position information. A solution for the superior colliculus (SC), however, requires the mechanism to compensate large misalignments within a very brief period of time. Initial saccades to an auditory target have a latency of 300 msec (Lueck et al., 1990) and do not allow for a time-consuming relaxation process. Information about eye position is necessary to allow fast processing in the coordinate transformation.

In the course of this work a great amount of time was devoted to analyzing the influence of tonically discharging neurons. These neurons provide the eye position signal in the simulations, and model neuronal activity in the parietal cortex which is related exclusively to eye position. The neurons discharge linearly with the orbital angle of the eye. The slopes and the offsets change, but the overall behavior of these tonic neurons remains the same. They either increase or decrease their firing rate linearly with regard to the position of the eye in orbit. Therefore, these neurons compute their activation through the following formula:

$$\text{neuron activation} = \text{slope} * \text{eye position} + \text{offset}$$

The incorporation of eye position signal into the system substantially complicates the transformation process. The tonic discharge of eye position neurons represents the orbital angle of the eye using a frequency code. This information must be combined with the place code for the auditory stimulus. The mechanism that accomplishes this combination is complex and poorly understood. A new organization is suggested here in order to take information about eye position into account. In this organization the tonic neurons form a topological map on which a neuron with an on-center-off-surround receptive field can encode the actual eye position. The y-cells in the retina have a receptive field of this type, and therefore this same distribution of synaptic weights might also be found in those auditory brain structures which project to the deeper layers of the SC.

In this thesis a network is proposed with which the coordinate transformation in the SC can be learned through the experience of coinciding visual and auditory stimuli. It is the working hypothesis that on average a visual and an auditory stimulus will have the same source in the environment. In neural network terms one assumes that the association to be learned is provided by an external teacher. In a first attempt the different inputs were associated with a two layer network. This proved to be unsuccessful because similar inputs have to be mapped onto different outputs, according to the combination in which the inputs occur. A two layer network cannot cope with the task of associating similar inputs with outputs which are far apart.

Instead of using back-propagation as the dominant learning rule for a multiple layered network another approach was chosen : With unsupervised Hebbian learning an internal representation of auditory space evolves in an intermediate layer. With the use of lateral inhibition a new state space develops in which input combinations are mapped into internal representations. These internal representations are modulated through the tonic discharge of neurons coding the eye positions.

The internal representations which evolved in the intermediate layer can then be associated with an activity pattern in the output layer. In latter versions of the network these internal representations were pre-defined by the pattern of connectivity between input and internal layer. This resulted in a less adaptive but smaller network which could process the information more efficiently. The exact appearance of the output activity is provided as a teaching input. This output layer models one of the deeper layers in the SC which receives teaching input as it ascends from the retina.

Chapter 2

Biological Basis

Among the first to investigate the interactions of different modalities in the deep layers of the superior colliculus (SC) were Dräger and Hubel in the mid 70's. They reported superimposed maps in the SC of mice for visual, auditory, and somatosensory stimulus sources in the environment. Their work is not concerned with change in eye position because the "mouse moves its eyes very little" (Draeger and Hubel, 1975). Either the inferior colliculus or the auditory cortex were assumed to be possible sources for auditory input.

Wise and Irvine (1983) focused on the auditory input to the deep layers of cat superior colliculus. They found a predominance of monaural excitatory input from the contralateral ear which could be inhibited by an ipsilateral stimulation. Thus, they doubt the existence of restricted spatial fields. This lack of tonotopicity is confirmed by Hirsch et al.(1985) for intermediate and deep layers of the cat superior colliculus. King and Palmer's results (1983) indicate a map of auditory space in the SC of guinea pigs. They report receptive fields of visual and auditory inputs to be coincident over a wide range. But they also found severe discrepancies at some locations which they suspect to result from abnormal eye position under anaesthesia. All authors report a first spike latency of 6 to 16 ms before auditory stimuli show a response in the SC.

2.1 Shifts of auditory receptive fields

Pöppel (1973) first proposed that either the visual or the auditory map in the SC would have to move with the eyes in order to preserve an alignment. The interactions of eye positions and auditory receptive fields were addressed experimentally by Jay and Sparks (1984). In their experiments, they establish that the auditory receptive fields in primate superior colliculus are not organized in head-centered coordinates. An extended description of their results was given in subsequent papers (Jay and Sparks, 1987a; 1987b). Their experiments on monkeys demonstrate that visual and auditory stimuli use a common coordinate system and a common motor pathway to initiate visual saccades. They present experimental results which prove that the "response to auditory stimuli was significantly affected by variations in eye position" (Jay and Sparks, 1987b). It is from within this context they speak of a shift of auditory receptive fields. Although their conclusions about receptive fields evoked skepticism because they contradicted other data in cat superior colliculus (Harris et al., 1980), they are now generally accepted for the primate superior colliculus. A recent publication also confirms their findings for cats (Hartline et al., 1989) as auditory receptive fields in the superior colliculus were found which shift with voluntary eye movements. Jay and

Sparks (1987a,1987b) speculate that the auditory signal is mapped onto a spatial framework encoded in motor error coordinates. In this framework, the activation of a cell specifies the change in eye position required to focus a target. Therefore, the projection of auditory input on the SC must be dynamic; with a change in eye position, the same auditory stimulus must activate different locations in the SC.

Very recently two neural network models were proposed which accomplish this exact shift of activation (Groh and Sparks, 1992). They present a biologically plausible explanation of how a brain-like structure could solve the task, transforming signals from a head-centered coordinate system into an oculocentric one. They examine how to construct a minimal neural circuitry for the task. The present work extends this approach, so that this transformation is learned through the experience of the environment. Shifter circuits that could be utilized by the SC have been proposed earlier (Anderson and van Essen, 1987; Trehub, 1987). But none of these proposals can account for the flexibility and adaptation of the mechanism developed in the brain. The previously proposed networks appear to be engineered simply to fulfill their task. They do not include even a faint hint of how the structure could have evolved and adapted through experience and learning in the environment.

2.2 Development of auditory receptive fields

A variety of sources in the literature establish that the auditory receptive fields in the SC are shaped by visual experience. This plasticity was first demonstrated in the optic tectum of barn owls, which is the homologues of the mammalian superior colliculus (Knudsen et al., 1984a). In their experiments the researchers used ear plugs to create abnormal binaural cues, and to change the temporal relationships of frequency components. They found a sensitive period correlated to the growth of head and ears during which the tectal auditory space was re-calibrated. Visual instruction in the tectum was discovered to be responsible for the alignment of visual and auditory maps (Knudsen and Brainard, 1991). The process involved is highly accurate, as it copes with changes in interaural intensity and interaural time differences, (both cues in the order of decibels or microseconds). It also seems to be highly adaptive : young barn owls can reorganize the artificially induced alignment after removal of the ear plug (Knudsen et al., 1984b). The tested owls managed to develop the same sensory maps as barn owls raised under normal conditions.

Similar evidence for plasticity of auditory maps was found in the SC of ferrets (King et al., 1988). The authors displaced the visual topography in the SC and investigated the effects on the development of the auditory map. The effect of strabismus was compensated by a shift in the auditory map, but this adaptive mechanism seemed to be exhausted with severe eye rotations; the auditory map in the SC was completely disorganized. The authors admit that the difference found between strabismus and eye rotation could also be due to a disturbed eye position signal. However, they prefer to explain it as a limit in the capability to reorganize the auditory map. In a recent publication (King and Moore, 1991) it is still acknowledged that an impaired eye position signal could be responsible for this limited re-alignment.

The SC of the guinea pig was examined in a series of papers (Withington-Wray et al., 1990b) lately. Although the central auditory pathways of the guinea pig are functionally developed before birth, an ordered map of auditory space in the superior colliculus is not evident until 32 days after birth. Different phases can be distinguished in the development of this map. The transition to a topographic order occurred remarkably rapid, in the third phase from the 31st to the 32nd day. It is assumed that the development of a topographic map is delayed until the head is fully maturated. Increasing interaural distance while the head is still growing would mean a constant adjustment of the auditory map.

In two successive papers (Withington-Wray et al., 1990a; 1990d) the authors analyze the influence of visual and auditory experience on the development of the auditory map. Both dark rearing and auditory deprivation prevent the emergence of a topographic order of auditory space in the SC. Neither the cochlea, primary auditory pathways, or inferior colliculus revealed any signs of impairment. Effects are limited to the auditory receptive fields in the SC.

Visual and auditory experience crucially shape the auditory map in the SC of the Guinea pig during a very short developmental period from the 26th to the 30th day after birth (Withington-Wray et al., 1990c). Animals developed a normal representation of auditory space in the SC after receiving visual and auditory input only in this four day period. Auditory receptive fields became further selective after the crucial period, as ongoing head growth must be compensated. But without the initial settings within the four day crucial period, the animal fails to construct a topographically organized map of auditory space. In their latest publication (Withington-Wray et al., 1991) the authors investigate the role of the visual cortex in the developmental process of the auditory map. They ascertain that the visual cortico-tectal pathway has no developmental influence on the construction of the auditory map in the superior colliculus. Therefore, they conclude that intra-collicular visual-auditory interactions are responsible for the formation of the auditory map.

2.3 Multisensory integration

The transformation of auditory space into a collicular coordinate system is constrained by visual and auditory interactions within the SC, where stimuli from each sensory modality converge in deep layer cells. Meredith and Stein (1986a,1986b) found two different classes of multisensory interactions, generally characterized as response enhancement and response depression respectively. They report these interactions take place in deep laminae cells of the SC which receive input from visual, auditory, and somatosensory stimuli. The different modalities enhance each other's effect on the cell when originating from the same stimulus source in the world; the cell's response is depressed when the inputs have different causes in the environment. King and Palmer (1985) confirm these findings, and explain that although some stimuli could not produce spike responses, they were able to modify the responses to other modalities. These findings might explain the strange mode of interaction which Meredith and Stein call "multiplicative" in nature. Normally neurons aggregate their synaptic inputs without higher order interaction like multiplication. A neuron activated to just below threshold by one input might respond vigorously with an additional input. Thus, this interaction appears to be nonlinear compared to the amount of incremental input.

In a subsequent paper, Meredith et al. (1987) explored the effect of temporal disparity of a stimulus combination on a single neuron. Essentially they discovered a functional link. Auditory and visual stimuli mutually enhance their response as long as they "belong" to the same cause with respect to temporal coincidence. But the same stimulus combination also produces a response depression when presented with a large temporal disparity. Different response latencies for auditory stimuli and visual stimuli complicate the task of determining whether two stimuli coincide. On average, a visual stimulus takes 82 msec to produce a response in the SC; unlike auditory stimuli which have a latency of merely 19 msec.

Differences in the transmission speed of sound and light change this relationship with respect to the distance of the stimulus origin in the environment. Meredith et al. (1987) found very long interactive periods (up to 400 msec) and suggest that long "discharge trains" in the cells overcome different latencies of the superior colliculus. Enhancing and depressing influences occur during these interactive periods. Pöppel and Logothetis (1986) propose a mechanism using neuronal oscillations either to align stimuli belonging to each other, or to separate uncorrelated stimuli.

For the spatial component of multisensory integration in the SC, Meredith and Stein (1986a) extract the following rule : "coincident multisensory stimuli produce enhancement, whereas disparate stimuli produce either depression or no interaction". Stein et al. (1989) present additional behavioral parallels to these neurophysiological findings: Cats were trained to react to a particular visual cue. Their performance improved considerably when an auditory stimulus appeared at the same location. A disparate auditory stimulus depressed responses to the visual cue. This mode of behavioral interaction was also validated by psychophysical experiments (Lueck et al., 1990) which concentrated on initial saccades to auditory stimuli. When a visual stimulus occurred with an auditory stimulus in the same hemifield, a "center-of-gravity" effect was reported for amplitude and latency of the saccade. A visual stimulus in the opposite hemifield had an inhibitory influence on the same parameters.

The visual receptive fields of cells in the deep layers of the superior colliculus have a significantly larger diameter in comparison to the visual receptive fields in superficial layers. Non-visual receptive fields of deeper laminae cells are also very large. Meredith and Stein (1990) conclude the process of multisensory integration might compensate different eye positions using the great overlap of receptive fields. They hypothesize that "receptive-field overlap is easier to maintain between 2 large receptive fields than 2 small ones". They notice a tendency in animals with significant freedom to move their eyes and pinae to have larger visual receptive fields than those incapable of significant displacements of different modalities. Though aware of the results regarding shifting auditory receptive fields by Jay and Sparks (1987b), they concentrate on explaining the integration of somatosensory input for which no compensatory shifts have yet been found.

2.4 Distributed representations

The question of spatial representation in the SC was discussed in a variety of publications. It is generally accepted now that in the deep layer of the SC, a spatial distributed neural code represents the motor error signal to focus a stimulus location. The distribution of activity over a whole population of cells initiates a saccade to move the eyes to a target (McIlwain, 1976). The premotor areas of the brain stem which orient the eyes are not excited by the local and strong activation of few neurons. Several different models of how the distributed and widespread activation of cells translates into the appropriate saccadic eye movement have been proposed (McIlwain, 1991). In all models, the discharge of a single neuron in the SC represents a movement vector of particular size and direction. The vector summation model assumes that a patch of activity represents an eye movement, which is the sum of all those mini-vectors encoded as cell activities (Fig. 2.1). Lee et al. (1988) question the contradictory results of this vector summation, and propose a population averaging model. In their model, the final eye movement is computed as the weighted average of vectors represented by the active population (Fig. 2.2). The results by Lee et al. (1988)

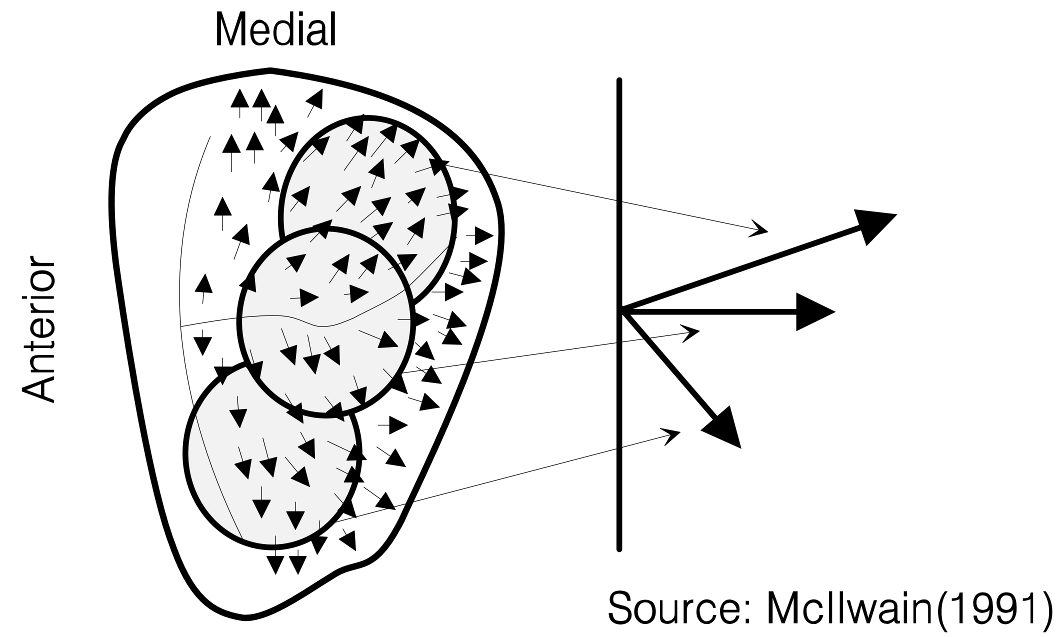

Figure 2.1: A patch of "mini-vectors" represents the movement vector of the eye.

also demonstrate that it is not the peak of activity which encodes the saccade. Although they only deactivated a peripheral subset of the active population (not the most active region), the evoked eye movement changed. Therefore even these peripheral activations must influence the initiation of a saccade.

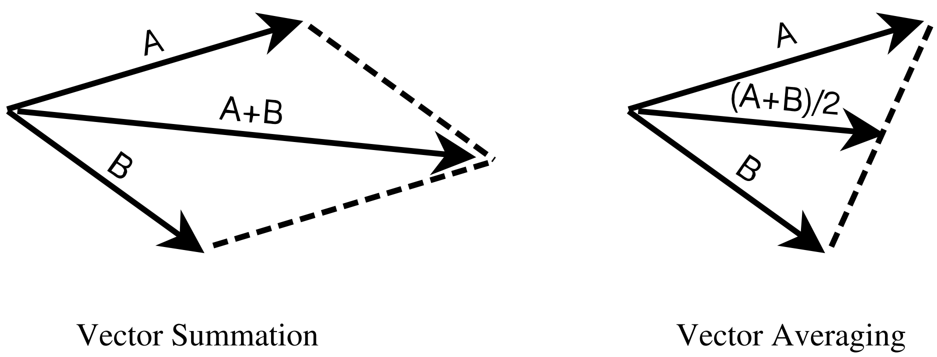

Figure 2.2: A comparison between vector summation and vector averaging model.

2.5 Sources of auditory input

None of the above work addresses an important issue: where does the auditory input to multisensory cells in the deeper layers of the SC come from? By tracing back the auditory input source, a modeler could carefully and explicitly limit the search parameters for the transformation process. The inferior colliculus and auditory cortex are repeatedly mentioned as possible sources. In the auditory cortex, Meredith and Clemo (1989) identified the anterior ectosylvian sulcus (AES) as the major and only region to project robustly into the SC. However, this projection is extremely divergent/convergent, diffuse, and without a clear topographic relationship. It is an exclusively excitatory projection and the authors conclude that it merely helps superior colliculus cells to exceed threshold. No direct pathway from the superior colliculus back to AES was found by which an error signal could be transmitted. The AES itself was not tested in respect to effects of eye position.

According to Groh and Sparks (1992), the inferior colliculus cannot embody the primary auditory input source. Average response latencies for auditory stimuli in the deeper laminae of the SC around 44.8 ms deviate too much from those latencies in the inferior colliculus around 28 ms. The inferior colliculus neither seems to form a topographic map of auditory space, nor does it have cells which are affected by eye position (Jay and Sparks, 1987b). But these data result from examining just five cells particular in the inferior colliculus. More experimental work is needed to provide a more thorough picture of the biological basis of auditory input sources to the SC.

Chapter 3

The Model

The general framework provided in chapter 2 of "Parallel Distributed Processing" (Rumelhardt et al., 1986) is used here in order to list all the assumptions and simplifications made in the simulations.

3.1 The set of processing units

Throughout this thesis, the processing units will be called neurons as they were inspired by neurophysiological data. They have no more sophisticated means of input-output processing than the neurons found in brains. In fact they have less sophisticated mechanisms, because adaptation within the neuron was not modeled. "Real neurons" rarely fire constantly over a long period of time. Even with stronger input, the neuron will exhaust itself and fire less frequently. On the other hand, a neuron usually becomes more sensitive to any input when it has not fired for a long time. This adaptation and sensitization are easily included in a model when an internal memory is added for each neuron. In computer science a similar extension of the finite state machine exists in the concept of a push down automaton. The push down automaton allows a class of problems previously thought insolvable to be resolved, but it also adds to the complexity of the algorithms. In neural network simulations the results are so hard to analyze and the behavior is so erratic in terms of predictability that in most of the models, this feature is not included. It has not been included in these simulations for the same reasons.

The neurons themselves do not represent higher level concepts as in more psychologically inspired simulations. The point at which a stimulus from the retina or in auditory space is perceived is represented as a gaussian activity on a topologically ordered map of input neurons. The neurons coding the eye position are activated with different slopes and offsets, but they always respond linearly with the eccentricity of the eye. They form a redundant representation of the eye position. The internal layers, which do not receive input from outside, form their own representations called "internal" representations. Because of the lateral inhibition in the internal layers, these representations are also ordered topologically but their meaning is not defined in advance.

The superior colliculus (SC) is known to represent information with a distributed activation of cells (McIlwain, 1991). Thus the output of the network was assumed to form a population code or averaging code of the motor error coordinate. In the simulations the network always learned to generate a response with exactly this type of distributed activity.

3.2 The state of activation

The neurons in a network can have different states of activation. Biologically it would be plausible to have two states for the neuron: firing or not firing. Such a model still lacks a refractory period, and therefore departs again from a real neurophysiological simulation. Using differential equations might overcome this problem but in the case of tonic input, such an approach is far too slow to yield any interesting results in a acceptable period of time. A neuron discharging tonically does not carry information in a single spike. The frequency of firing contains the actual information and therefore the spikes must be integrated over time. It might be possible to simulate three to five neurons in a competitive cluster with very short time steps or differential equations, but for a bigger network involving layers of neurons the biological processes must be somehow simplified.

In the simulations the activation state of a neuron ranges from 0 to 32767. This number is used to aggregate all spikes in a time interval and determine the activation state from that sum total. In the first series of simulations an activation value in the interval [0,255] was used. This turned out to be insufficient to accurately simulate a lateral inhibition. After each propagation step in a lateral inhibition network, the new state of activation of the inferior neuron must be smaller than before; otherwise the lateral inhibition halts at a median level. With only 256 discrete activation states this often occurred, especially when the inhibition is kept small to prevent oscillations. Therefore the number of possible activation states had to be increased. Still, the lateral inhibition is not guaranteed to work properly for all possible input combination. For a perfect winner-take-all, the activation states must be extended to a doubled or even higher precision. Essentially for every precision and every winner-take-all network, one can construct an input combination which cannot be resolved to have one active neuron only.

3.3 The output function

The state of activation of a neuron is propagated to all postsynaptic neurons. As the activation function encodes the number of spikes in a hypothetical time interval this aggregate of spikes is used as input to all postsynaptic neurons. In PDP-terms (Rumelhardt et al., 1986) the output function is the identity function.

3.4 The pattern of connectivity

All neurons are not interconnected, as the notion of a receptive field becomes meaningless when every neuron is connected to all other neurons in the network. Instead, the connections between the input, internal and output neurons are chosen to reflect differing brain structures. These connections will be called synapses, again indicating the intent to stay close to neurophysiology. Most of the synapses are initialized randomly, and modified during the course of processing and learning. Thus the existence of a synapsis allows the possibility of learning certain influences.

Synapses are not symmetric. In most cases the postsynaptic neuron does not project to the presynaptic neuron at all. The existence of feedback projections in some brain structures suggests a different learning rule than it is present in other brain structures with strict feed-forward synapses. In the parietal cortex, for example, many synapses have been found to project back, which could indicate the existence of an error signal. However, this does not hold for the pathway from auditory cortex to SC. In fact, the deeper layers of the SC do not project to any cortical area. They project to the thalamic nuclei which in turn project to the cortex (Sparks and Hartwich-Young, 1989).

In the model the strength of a synapses is assumed to be in the interval [-32767,32767]. A synapses can be defined between any two neurons in the network.

3.5 The rule of propagation

In the model, all neuron activities are updated simultaneously. It is here where the model departs radically from neurophysiological processes. The brain lacks an external timing pulse after which the new activity state of each neuron is determined. A synchronous update procedure has many implications for simulations. A winner-take-all competition is revealed to be more challenging to implement because of oscillations which easily occur with synchronous updating. But the frequency of these oscillations is arbitrary, determined only by the updating process. The neuronal oscillations of Gray, König, Engel, and Singer (1989) are an entirely different matter, determined by a continuous, physical process. Again, synchronous updating is not capable of a precise simulation. Asynchronous updating also cannot be the answer. Even if it changes the speed of the timing pulse and helps to prevent oscillation, it does not better mirror the biological process. Also, asynchronous updating needs extra cpu-time to determine which neuron is to be updated and to assure that all units have been updated after each cycle. Again, using differential equations seems to be the proper approach, but from the practical point of view (CPU-time), it is an unacceptable way to actually simulate a large number of neurons.

3.6 The activation rule

A standard clipping function (Anderson et al., 1977) was used to transfer the weighted input into a state of activation. With two parameters, (an upper limit and an lower limit) the net input to a neuron was compressed into the interval [0,32767]. A net input less than the lower limit did not yield any activation, and a net input higher than the upper limit saturates at the maximum activation state. In between the input was mapped by a linear function onto the state of possible activations. A linear function computes quickly but again is not realistic. In biological systems such a transfer function should have a sigmoid form as in figure 3.1. When storing a look-up table for the sigmoid activation function of each neuron, one quickly reaches the limits of memory capacity. In order to work with a function which is neither CPU-intensive nor memory exhausting, one might neglect the difference between a linear and a sigmoid function, but one should also be aware of the consequences. Especially at the points where the clipping function is not differentiable, lateral inhibition will oscillate more often. Two neurons which are above the upper limit have no way of knowing who is supposed to be more strongly activated. This might be resolved after a couple of iterations but it could also lead to an oscillation.

The activation rule is a very powerful instrument in network simulations. Future research should investigate how to realize it more dynamically. Adaptation and sensitization of a single neuron could be easily modeled through a change in the activation function. A lot of time in network simulations is wasted in adjusting the activation function to the input. Therefore, an adaptive activation function would be very helpful. A neural system which adjusts itself to the inputs could be a real step forward in neural network research. A

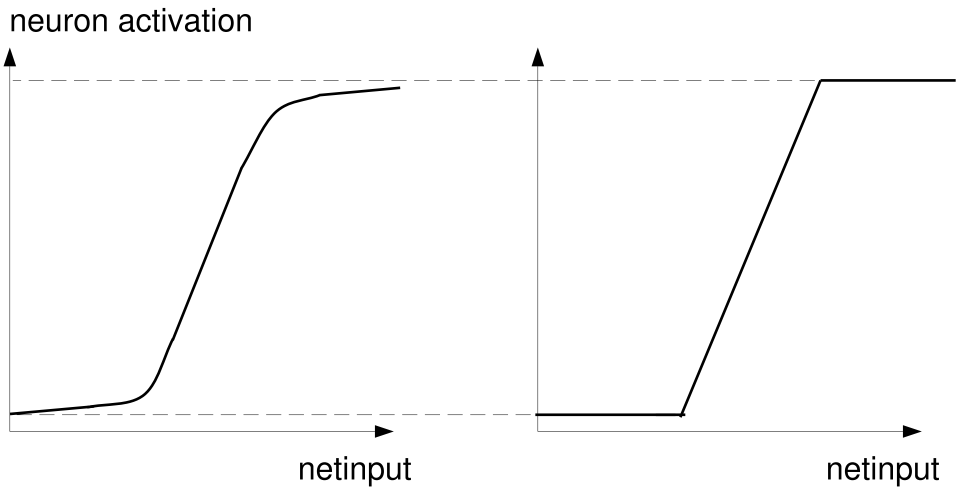

Figure 3.1: A sigmoid activation function compared to a clipping function

common upper and lower limit of the activation function should evolve in the course of processing with the weighted input to every neuron in between. But this must be achieved by a local process without communication among the neurons or a maximum search over all neuron activities. For such a local process, a neuron must evaluate its inputs statistically over time and be able to store the result of that evaluation.

3.7 The learning rule

In the simulations two different learning paradigms were used : Unsupervised and supervised learning. Both learning rules do not digress significantly from Hebbian learning (Hebb, 1949). They use only information which is available locally at each synapsis.

In the unsupervised learning paradigm the algorithm distinguishes whether the postsynaptic neuron is active or inactive. An inactive postsynaptic neuron does not alter its synaptic weights at all. There is a possibility of subtracting a decay term for inactive postsynaptic neurons but it was not utilized in the simulations. When the postsynaptic neuron is active a weight change Δw is computed by applying the following formula:

$$\Delta w = \left(c * \frac{PresynapticActivity}{MaximumActivity} - x\% * w\right) * \frac{PostsynapticActivity}{MaximumActivity}$$

where	w	=	current strength of the synapsis
and	x%	=	adjustable percentage
and	c	=	learning constant

Subtracting a percentage of the weight with every learning step proved to be a very efficient way of coping with a couple of problems. First of all it regulates the weights and prevents an unbounded increase of synaptic strength which is typical of pure Hebbian learning. The weights are decreased smoothly and they approach zero values without overshooting into negative values. This asymptotic behavior on the upper and the lower limit of the possible weight scale resembles the saturation effects found in biological systems.

Both decrease and increase of synaptic weight can be simulated with the same mechanism when it decreases the weights in accordance with its current strength. In some models the learning algorithm distinguishes between the case when the presynaptic activity is above or below a learning threshold. By that comparison the algorithm determines whether a learning constant will be added or subtracted. This learning threshold is a superfluous, artificial parameter redundant with the learning formula suggested here.

The main reason, though, to choose this special learning rule was the fact that in a competitive cluster one neuron started to take over, respond to every input, and to suppress all other neuron activity in its cluster. Once one neuron had an advantage over the others it would always win the competition and block any development of selectivity within the cluster. The subtraction of a percentage of the weight bans this behavior. A neuron trying to increase its synaptic weight too excessively will be inhibited by a subtraction relative to the current strength of the weights. In practice selectivity developed very easily and smoothly with that kind of learning.

For the supervised learning paradigm different modified versions of the "Widrow-Hoff" error correction procedure (Widrow and Hoff, 1960) were used. Hebbian learning which does not take the current error into account did not converge into a satisfying association. The internal representation had too much overlap to be associated with the teaching input by a simple Hebbian procedure.

The sign of the error was adequate for learning the association between internal and output layer. The learning rule employed the following formula:

$$\Delta w = \left(c * \frac{PostsynapticActivity}{MaximumActivity} * \frac{PresynapticActivity}{MaximumActivity}\right) * signum(error)$$

where	c	=	learning constant
and	error	=	Teaching input − postsynaptic activity.

Of course a learning rule which has more information available also establishes the correct association. Applying the degree of error to the formula, changes it to

$$\Delta w = \left(c * \frac{TeachingInput - PostsynapticActivity}{MaximumActivity} * \frac{PresynapticActivity}{MaximumActivity}\right)$$

3.8 The environment

In the learning phase the input patterns were presented at random order. Eye position and auditory input had a uniform random number distribution. In a natural environment additional head movements might yield a normal distribution of eye positions centered around the zero degree orbital angle. This would affect the distribution of teaching inputs in the output layer. For simplicity reason these factors were omitted in the model but they would rather improve performance than deteriorate it.

Chapter 4

A two layer network

A researcher should always start with the simplest model possible, especially when he aims to simulate a complicated system. For neural networks the most simple model consists of two layers: an input layer and an output layer. The synaptic weights between the two layers are modified throughout the process of learning. These weights store the association between input and output layer.

Following the above paradigm, a model of the coordinate transformation in the superior colliculus (SC) was approached with a two layer network. In most neural networks every neuron in the input layer is connected to every neuron in the output layer. As the notion of a receptive field becomes meaningless in a fully connected network, a different pattern of connectivity was adopted here. The input layer (and with it the scheme to determine the receptive fields) was further divided into three different subunits. These three different "sublayers" should be understood as a simulated retina, an auditory map, and a representation of eye position respectively.

4.1 Input and output representation

A patch of neuron activity was used to represent visual input on the retina as well as auditory input on the auditory map. For eye position, a different coding scheme was applied. Neurophysiological studies report that eye position effects produce a tonic background activity in the neurons (Andersen et al., 1985; Andersen and Gnadt, 1989). Thus, in this work the neurons which encode the eye position are called tonic neurons. Most models (Andersen and Zipser, 1988; Pouget et al., 1992) with an eye position signal use two different categories of tonic neurons. One class encodes the horizontal orbital angle of the eye, the other encodes vertical eye deviation. The tonic signal is supposed to originate from the motor command to the muscle fibers which hold the eye in its position. Considering the configuration of six muscles which control eye movements, the coordinate system for eye position signals should be complicated in nature. However, as the fovea moves in a two-dimensional plane and as the system is much easier to analyze in a orthogonal coordinate system, a more sophisticated model of feedback from the eye muscle command was sacrificed for simplicity. In the present model the eye position was represented by two sets of neurons, each coding one axis of two-dimensional space. In each set, a neuron either increased or decreased its activity linearly with respect to the orbital angle of the eye (Fig. 4.1). Different slopes and offsets in a set provided a redundant but easily distinguishable representation of eye position.

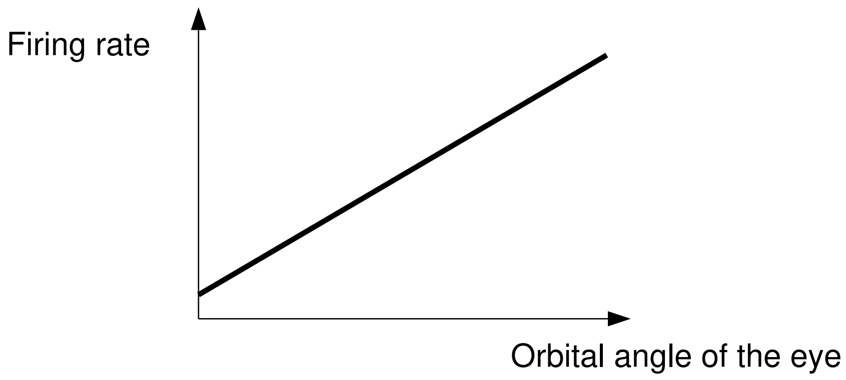

Figure 4.1: The firing rate of a neuron encoding the position of the eye in orbit.

The output layer represents one of the deeper laminae in the SC. The central idea of these simulations is to exploit the overlap of point images from neurons in the different input layers in order that the network learns and forms a particular representation for each input combination in the output layer (Fig.4.2). The learning algorithm also has the advantage of a consistent learning environment. Every auditory stimulus must have a cause which can be perceived as a coincident visual stimulus. Therefore, the two stimuli are correlated and can be associated with the same output. Together with an eye position signal, this combination constitutes an unique input pattern. An equivalent network scheme was suggested by Meredith and Stein (1990) to explain the considerable increment of deep laminae visual receptive fields compared with the receptive fields found in the superficial layer.

A modeler could incorporate a retina consisting of the fovea, a region with high density of receptors in the center, and peripheral regions with a less dense distribution of photoreceptors. By the same token, he could also simulate other subtle particularities. For example, the afferent projections from the retina to the superficial layers of the SC can be described by a complex logarithmic function (Ottes et al., 1986). The above issues are beyond the scope of the current project. In order to be able to concentrate on the actual task of integrating different modalities it was assumed that the input patterns are preprocessed, and that the input layers project homogeneously to the output layers. As input patterns consist of activity patches, the question of peripheral input did not appear relevant.

4.2 Pattern of connectivity

Central attention must be devoted to specifying which input locations produce a common activity in the output layer. Only if appropriate regions in the input layers project to the same region in the output layer, a new activation pattern will develop there. The projections from different modalities to the deeper layers of the SC are not mapped very well by neurophysiological experiments. Further experiments could define regions in auditory space which project to the same deeper layer area as the fovea. McIlwain (1976) answered similar questions with his notion of a “point image”. According to his findings, a stimulation of cell in the retina affects the same number of neurons in the SC regardless of whether the stimulation occurred in the fovea or in the periphery. The number of neurons influenced by the stimulations was revealed to be equal for a each cell on the retina.

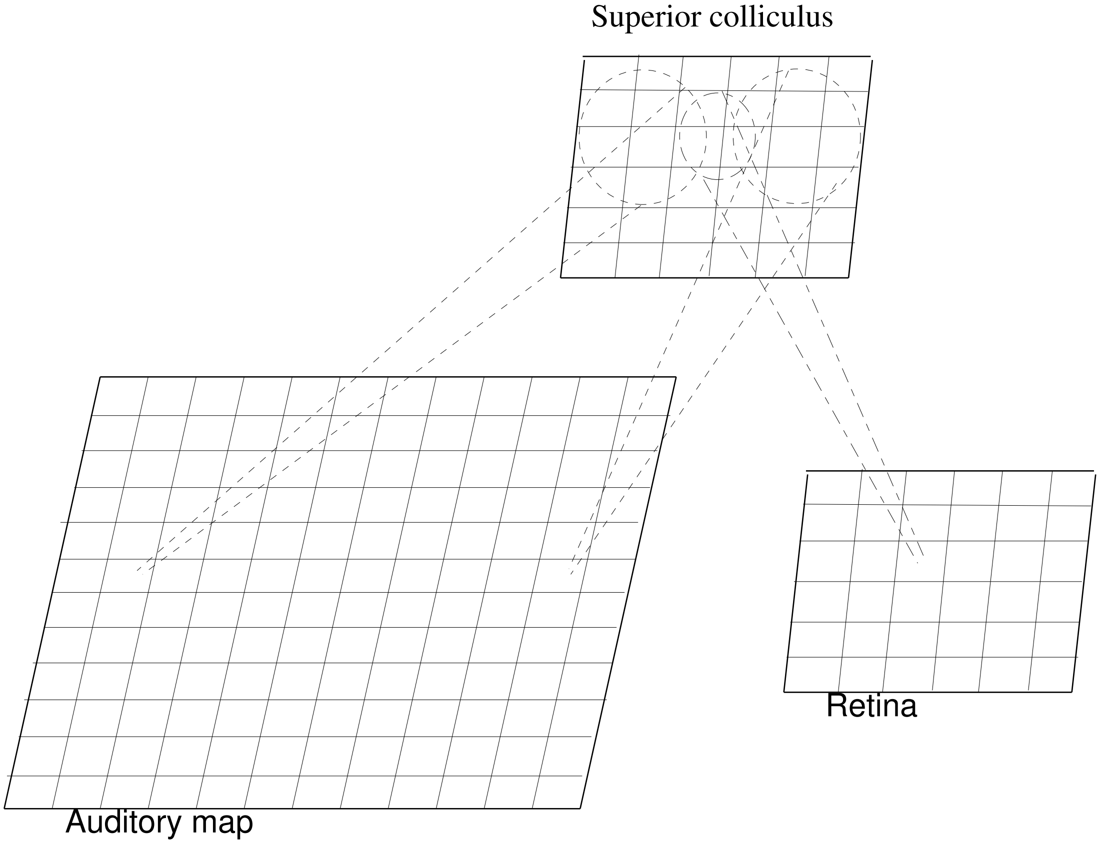

Figure 4.2: Point images of auditory and visual input neurons in the SC

In the model the size of auditory space is related to the size of the retina as in figure 4.3. An auditory stimulus can be perceived at every position in the three-dimensional space around the head. However, that part of auditory space is delimited which might coincide with the retina under every possible orbital angle of the eye. If the retina receives input from a space which is x degrees wide, and if horizontal eye movements range from y degrees to the left to y degrees to the right, the limited auditory space will be x + 2y degrees wide to cover every point the retina can see with extreme orbital angles of the eye.

If a region on the visual input map and a region on the auditory input map can be activated simultaneously by the same stimulus source, they will have to project to the same area in the output layer in order to form an association. Thus, the receptive fields on the auditory map in a two layer network must be bigger than those on the visual map. They must include all possible eye movements.

Neurons in the deeper layers of the SC have very large receptive fields on the retina. Meredith and Stein (1990) report a mean diameter of 66.9 degrees. Thus, the auditory receptive fields must be enormous in a two layer network. In the simulations, a network with a retina consisting of 40 x 40 pixels and an auditory map consisting of the equivalent of 80 x 80 pixels needed 3 million connections. This indicates that some kind of preprocessing must occur. It supports a claim that auditory space must be transformed into an internal representation which then projects to the SC.

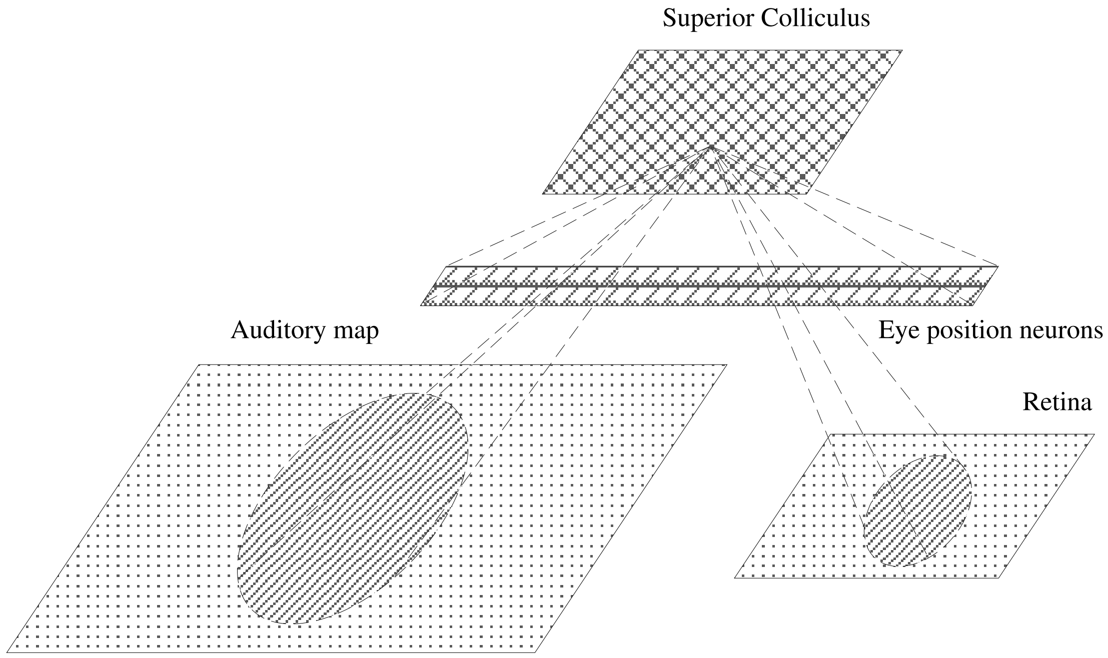

Figure 4.3: The network for the simulations

4.3 Results

In the simulations it was impossible to associate each input combination to a unique output pattern. The network developed different strategies to cope with output patterns which depend on the combination of input stimuli, and with the fact that the same stimulus in one input layer produces different output activities. In all simulations neurons in the output layer concentrated their synaptic weights on a single input source. At first, the tonic neurons were the main target. Tonic background activity was always present in every output neuron, while auditory and visual input could be either on or off. Thus, the network specialized to respond to the tonic neurons. All the other synaptic weights set to visual or auditory stimuli decayed in the course of learning as they were not constantly reinforced.

The influence of tonic neurons must always remain background activity in the output neurons. Whenever the tonic input activates an output neuron to its maximum level, this output neuron is useless for its actual task. The output neuron's activity cannot be used to detect the coincidence of auditory and visual input. Therefore the tonic influence must be delimited. However, the tonic input also has to create an impact in the output layer because the eye position signal discriminates between different input combinations. In the simulations the tonic input should determine a center of activity in the overlap of point images from auditory and visual stimuli (Fig. 4.2). The tonic influence must be fairly balanced which implies the synaptic weights to tonic neurons must not be modified.

In successive simulations the synaptic weights to the tonic neurons were not learned. The neurons in the output layer specialized on the visual input. They decayed all synapses in their auditory receptive fields and responded exclusively to visual input. The same number of active neurons represented visual and auditory stimulus. But the probability of a neuron to be active on a visual map consisting of 1600 neurons is much higher than for a neuron on the auditory map of 6400 neurons. The ratio of active to inactive neurons is important.

A neuron in auditory space is statistically less often active than a neuron in visual space. As learning is proportional to activity, the auditory influence is more likely to decay. The visual input is dominant.

A modality can vary its impact on the output neurons with the number of active neurons coding a stimulus. The output layer neurons reacted with a new strategy as soon as the influences of the visual and auditory stimuli were controlled for equality. The output neurons decay all synaptic weights no matter where they connect. This seems to be a conflicting result. How can a neuron reduce its response activity when learning itself depends on this activity? The combinations of stimulus locations in the input layers are distributed equally. In each learning trial the synaptic weights for one stimulus combination are reinforced and all other synaptic weights decay. Without a statistical structure in the stimulus combinations the output neurons cannot extract any patterns. Therefore the decay term steadily decreases all weights. At a certain point in the process of learning the neurons in the output layer cease to respond no matter what input combination is presented.

A change in the relationship between decay and reinforcement of synaptic weights can address this kind of "fading away". A higher learning constant guarantees that the increase of synaptic weights will be equivalent to the loss of synaptic weights in each learning trial. In the simulations a constant and cyclic re-coding resulted in receptive fields which reflect the last few stimulus combinations.

These simulations suggest that the coordinate transformation cannot be learned in a two layer network with the given input representations. A two layer network is forced to neglect an input which seems to bear no statistical information. The task would be more suitable for this network architecture if the input configuration obeyed a non-uniform statistical distribution. In the SC no preference for a certain configuration of input stimuli can be present. All combinations have the same probability in a natural environment.

4.4 The Brain-State-in-a-Box Model

An association may be realized in a supervised paradigm even if it cannot be formed with unsupervised learning. In order to examine this possibility the "Widrow-Hoff" error correction procedure (Widrow and Hoff, 1960) was applied to the coordinate transformation task. Eye position, auditory, and visual stimulus were auto-associated in a fully connected network. This network merely tests whether a stable association can evolve. A brain-state-in-a-box model (Anderson et al., 1977) provides the network parameters. These simulations simplify the input pattern to a great extent as one dimension in space was extracted and learned. Visual and auditory stimuli, as well as eye position were represented as Gaussian activations on a one-dimensional scale.

The results of these simulations were as discouraging as the results for preceding models with an unsupervised paradigm. Some input combinations would be auto-associated but others never managed to form stable pattern. The association process also required a large number of learning trials to show any effect. Some input combinations were totally ignored in the course of learning. In the final test a combination of input stimuli was chosen randomly and used as a starting activity in the network. The activity was iterated for a number of propagation cycles and most of the time the network deteriorated with this input combination. The network activity changes arbitrarily, and settles on a meaningless activity pattern. These results indicate that a two layer architecture does not support the coordinate transformation task.

Chapter 5

Back propagation

Neural networks with three layers have greater capabilities than two layer networks for building appropriate association. Whenever a task cannot be solved with simple association in a two layer network, a modeler is forced either to change the representation used so far or to continue the simulations with a three layer network. In order to model the transformation of auditory stimuli into an oculocentric coordinate system the options among representations are restricted. The representations of the eye position signal as well as auditory and visual stimuli are defined by neurophysiological results. An approach with a three layer network offers an alternative, especially if it self-organizes a new representation in the middle layer.

The introduction of a learning rule for three layer networks has revived the field of neural computation in the last few years. It was the first learning rule able to overcome the limitations of a two layer network. This learning rule is generally known as back propagation (Rumelhardt et al., 1986) and it can train arbitrary input-output associations. However, back propagation is not as consistent with neurophysiological data as Hebbian learning (Hebb, 1949). It does not provide a sound theory of how the brain could develop its particular properties, and thus is usually rejected by those researchers working in the field of neuroscience.

In a back propagation network the network activity is first propagated forward from the input units over the hidden units to the output units. Then, an error signal (the difference between actual and desired output) must be propagated backward in reverse direction through the same synaptic weights which created the error. In reality, a synapse never functions as a pathway in both directions. The passage of error information backward through a synapse is an inherent flaw in the architecture of back propagation networks, one which cannot be remedied by any biologically plausible mechanism.

A network resulting from back propagation learning sheds no light on the workings of its mechanisms. The meaning and function of units in the hidden layer is dubious, and these hidden units do not offer any clues as to why and how the network solved its task. However back propagation learning does establish the required associations in a way which seems to exploit the same statistically optimal solution the brain uses. But, again, it does not help to unravel these solutions; in a way, back propagation further obscures the problem. A researcher is always inclined to interpret simulation results in accordance with his theory. Back propagation results can easily be tuned to neatly fit almost any theory.

Implementations of back propagation can also be criticized for normalizing all network activity. Further, the idea of gradient descent in back propagation learning presumes a network where every neuron in one layer is connected to every other neuron in the next layer. Such a fully connected network seems rather irritating when working with brain structures and receptive fields.

Back propagation is a very powerful method for finding optimal solutions. Thus, it is appealing to engineers who apply it to problems in their domain with encouraging results and increasing popularity. Researchers in psychology (Rumelhardt and McClelland, 1986) can use it to simulate cognitive processes, as a model of higher level cognition seems to be less dependent on actual physical implementation. But in order to explore the brain on the neurophysiological level of cells, synapses, and receptive fields, back propagation does not appear to be a very helpful tool.

5.1 Back propagation in the posterior parietal cortex

Back propagation was applied to a task similar to coordinate transformation in the superior colliculus (SC). A three layer network (Zipser and Andersen, 1988a; Andersen and Zipser, 1988) was trained to transform a stimulus on the retina into a head-centered coordinate system where stimulus location was encoded independent of eye position. A gaussian pattern of activity on a simulated retina of 8 x 8 pixels represented an input stimulus. Neurons coding eye position formed a second input to the network. The activity of one class of eye position neurons was computed by linear functions of horizontal orbital angle of the eye. The other class of eye position neurons varied their activity linearly with vertical eye position. For each neuron a different function with either a positive or a negative slope mapped eye position into a neuron activity. These functions had different intercepts, but they always increased or decreased their output monotonically with eye position.

The authors used two different formats in order to encode the output stimulus into a head-centered coordinate system : a gaussian activity pattern on a two dimensional map and a monotonic format equivalent to the representation of eye position. Basically, both methods generated the same results. Within 1000 trials, the network learned to establish the correct association between input and output. The mainly appeal of the simulations was that the units in the hidden layer combine eye position signals and input from the retina in a way analogous to neurons in the parietal cortex. This kind of combination was characterized by "planar gain fields" (Andersen et al., 1985; Andersen and Zipser, 1988). The entire next chapter is devoted to this issue. The retinotopic receptive fields generated in the hidden layer also correspond to the retinotopic visual receptive fields resulting from recording data in the parietal cortex. However, the wide variety of large and complex receptive fields prevented a quantitative comparison between model and experiment.

The monotonic output format encodes an eye position under which the stimulus is focused on the fovea. It is a head-centered signal which might be utilized by the brain to initiate a saccade to the stimulus. Goodman and Andersen (1989) investigate this possibility and saturate the activity of hidden layer units artificially. They report the same eye-position-dependent saccadic behavior for the network's output as was recorded in brain stimulations of area 7a of the parietal cortex. Zipser and Andersen (1988b) also examine other formats for output representation. Back propagation learned the association for all of the output formats. However, the most analogous relationship between the neurons in area 7a of parietal cortex and the hidden layer units was found with the combination of the original gaussian/monotonic format.

Coordinate transformation in area 7a of parietal cortex defines a simple problem. As displayed in figure 5.1 a two-dimensional vector for the stimulus on the retina is added to a second two-dimensional vector which encodes the eye position. This results in a two-dimensional vector for a head-centered stimulus. Different representations for eye position and stimuli complicate the computations for the back propagation model. In the model with a gaussian output format (Zipser and Andersen, 1988a) the input and output layer have the same size. The training set must be severely limited to exclude those input combinations which would produce a stimulus location outside the output layer. However, back propagation associates every input-output combination regardless of edge effects.

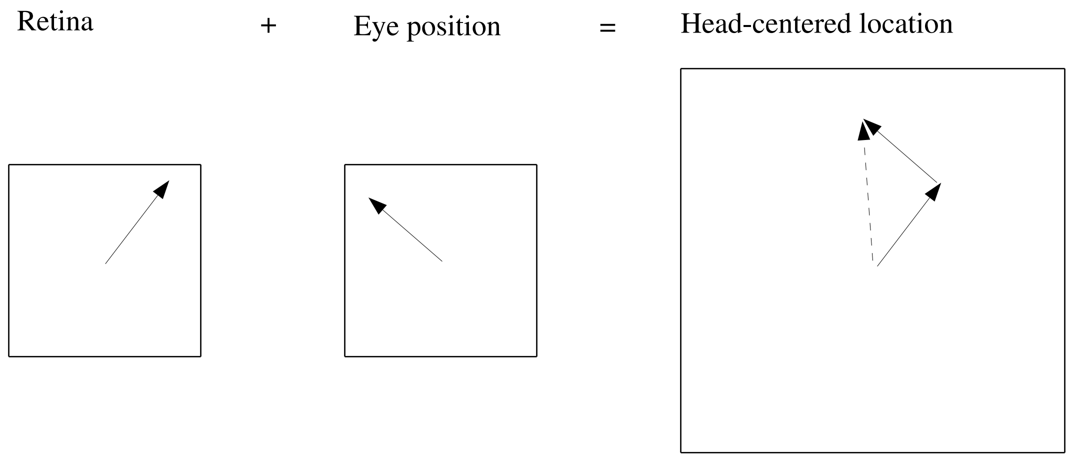

Figure 5.1: A stimulus on the retina is converted into a head-centered position

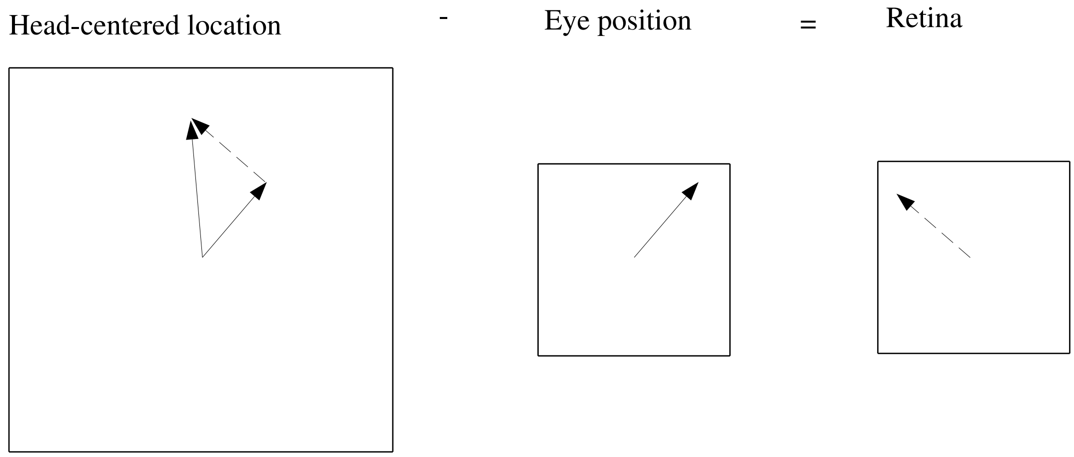

Figure 5.2: A stimulus in head-centered auditory space is converted into an oculocentric coordinate system.

Coordinate transformation in the SC seems to be the exact opposite of the task of the parietal cortex. The eye position signal must be subtracted from a head-centered auditory stimulus in order to compute the stimulus in oculocentric coordinates (Fig. 5.2). However, the similarity fades for a neuronal realization of the transformations. Obviously the receptive fields of neurons in the intermediate layer of the transformation must be chosen more carefully by the superior colliculus than the parietal cortex. A receptive field might cover the retina as in the model of the parietal cortex, but the auditory input map consists of all the space which might coincide with the retina under every orbital angle of the eye

(Fig. 5.3). How, then, could a receptive field cover that entire space? In the SC, large but restricted receptive fields exist after birth. They show a "marked preference for stimuli in the contralateral hemifield" (Withington-Wray et al., 1990b) and develop a very distinctive, sharply pointed shape. It seems highly unlikely that development or learning starts with unrestricted receptive fields as it does for back propagation.

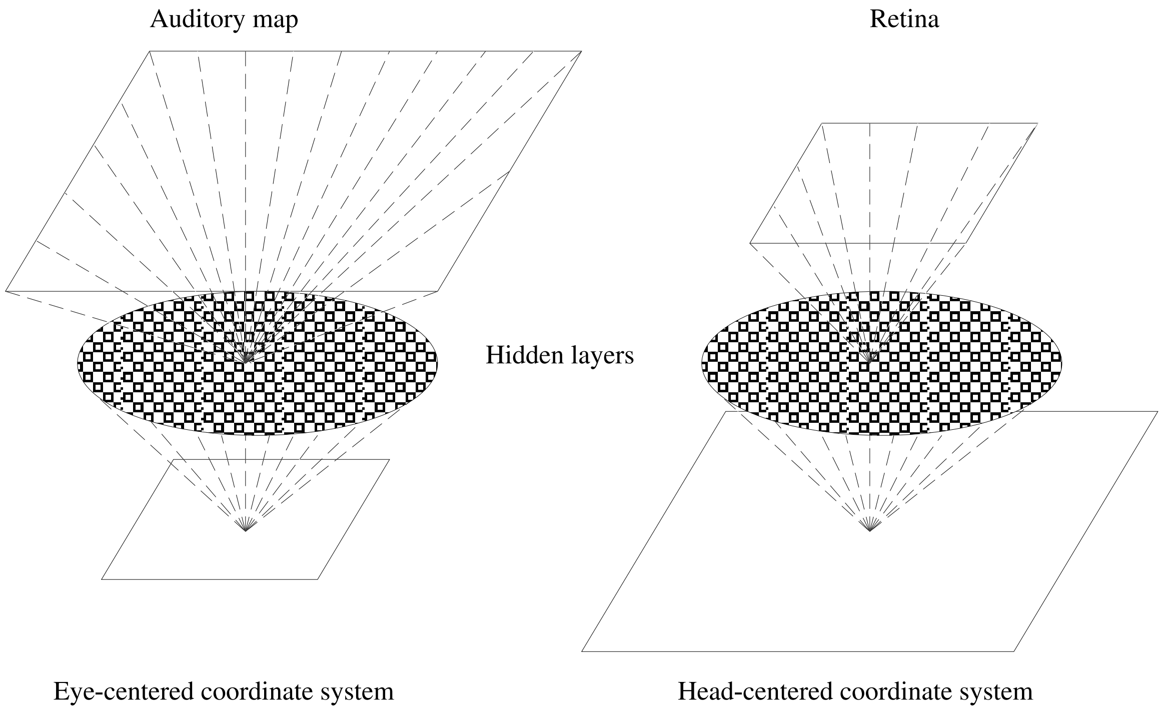

Figure 5.3: The two problems in a neuronal realization.

In a variation of the back propagation algorithm, the weights are shared within the hidden layer (LeCun et al., 1990), thereby utilizing restricted receptive fields. This algorithm was applied to a hierarchical transformation of space (Pouget et al., 1992), a similar task to that of the parietal cortex. The learning rule was not further pursued here, because it seemed to have even less biological foundation than back propagation.

5.2 Back propagation in the superior colliculus

In the present study, a back propagation network was trained to accomplish coordinate transformation in the SC. The motivation for these simulations was not to model neural processes, but to understand the underlying mechanism of the transformation. The statistically optimal solution which back propagation learning generates might provide a general idea of how to implement the transformation process in a more biologically based framework. My back propagation simulations were never intended as simulations of the "real" processes in the SC.

All simulations were performed with the PlaNet simulator version 5.6, developed at the University of Colorado at Boulder. A back propagation network with the same size capabilities as the previous two layer model requires too many synaptic connections to be simulated on a moderate speed computer. In a back propagation network, every unit in the input layer is connected to every unit in the hidden layer. Thus every unit in the hidden layer needs 6400 synaptic connections to an auditory input map with an equivalent size of 80 x 80 pixels on the retina. The network size fairly well explodes with regard to the number of hidden units.

Instead of decreasing the size of the auditory map, the vertical component of input and output was omitted in the model. The horizontal component of eye position signal as well as visual and auditory stimuli was extracted. The auditory input stimulus formed a gaussian activity on a one-dimensional map of horizontal auditory space. Current research in auditory perception also suggests a separate processing of the vertical and horizontal components of an auditory stimulus. Eye position neurons only represented the horizontal orbital angle of the eye; vertical eye position was disregarded. The output of the network consisted of a gaussian activity encoded on a one-dimensional map. It represented the horizontal component of an eye-centered stimulus. As this model was intended to study the mechanisms of coordinate transformation, these simplifications did not devalue the simulations.

Different configurations for back propagation networks varied in terms of the number of hidden units. All of the networks learned the association of eye position and auditory stimulus to an eye-centered stimulus with a root mean square error of below 0.3%. Back propagation learning did not have any difficulty in establishing the input-output associations. However, the simulation did not provide insight into the transformation mechanism itself. The receptive fields and the responses in the hidden layers were unintelligible and proved impossible to analyze. The purpose of hidden layer units remains unclear. The responses of these units did not reveal an obvious relationship between the auditory stimuli and the eye position signal. Although the PlaNet simulator is furnished with a powerful graphical tool in order to investigate error surfaces and hidden layer responses, no progress was made in understanding how and why the transformation is realized in a back propagation network. The only important observation was that hidden layer units reflect the eye position signal. This observation will become very important in chapter 6.

5.3 A biologically more plausible learning rule

Mazzoni, Andersen, and Jordan (1991) recently applied another learning rule to teach the coordinate transformation of the posterior parietal cortex to a three layer network. They trained the network with a variant of the associative reward-penalty learning rule (Barto, 1985; Barto and Jordan, 1987). In a reinforcement learning paradigm, this rule accomplishes the same coordinate transformation as in earlier experiments (Zipser and Andersen, 1988a). Essentially, the same multilayer network structure was implemented as in previous back propagation simulations, except for one important extension. The associative reward-penalty learning rule does not require a precise and local error signal. Instead, a global error signal is generated by a critic and fed back as a reinforcement signal through explicit projections. Because of the distinct feedback this learning is less "artificial" than back propagation learning. The feedback system, with additional synaptic connections, was incorporated in the network structure.

The existence of veto neurons which inhibit entire areas of neurons has been demonstrated in several neurophysiological studies (Crick, 1986). All cortico-cortical and thalamo-cortical connections are especially well-known to have reciprocal feedback pathways (Zipser and Andersen, 1988a). Thus, reinforcement learning is perfectly suited for the posterior parietal cortex. Neurophysiological studies have not demonstrated that an equivalent error signal mechanism could originate in the SC. For example, no direct reciprocal feedback pathways have been reported to exist from deeper layers of the SC to the auditory cortex.

Identical representations encoded the input in the simulations with the associative reward-penalty learning rule (Mazzoni et al., 1991) and with back propagation learning (Zipser and Andersen, 1988a). Stochastic learning automata (Barto, 1989) motivate the hidden and output units in the associative reward-penalty learning rule. The sum of the weighted inputs determines the probability of the unit firing a spike, then a random process settles the activation value for a unit as either 1 or 0. Thus, the hidden and output layers consist of binary stochastic elements. The original output representations (Zipser and Andersen, 1988a) needed modifications in order to incorporate a binary activation. Either a "binary-monotonic" format or a "binary-gaussian" format (Mazzoni et al., 1991) represented the teaching input in the network.

As with back propagation, the associative reward-penalty learning rule managed to establish the correct input-output association to any desired accuracy. Again as with back propagation, however, the simulations did not provide insight into the mechanisms involved. Unfortunately, the associative reward-penalty learning rule also works very slowly as it evaluates a large amount of training examples statistically. A variant of the rule (Barto and Jordan, 1987) improves the speed of learning, but still does not converge with the pace of back propagation (Hertz et al., 1991). Back propagation itself is well-known to be slow because an optimum weight in one layer depends on the configuration of all weights in the network (Hertz et al., 1991). It is important to again note that it takes a very short period of time to develop the auditory receptive fields in the SC (Withington-Wray et al., 1990c). A learning rule like the associative reward-penalty learning rule could not learn the coordinate transformation of the SC in the extremely short "crucial" period discussed in detail in chapter 2.

Chapter 6

Tonic Background Activity

Within 300 msec the head-centered auditory input to the superior colliculus (SC) must be transformed into an oculocentric coordinate system in order to initiate a saccadic eye movement (Lueck et al., 1990). The fact that this interval is so short implies that the transformation process uses information about eye position. In this chapter the biological background of the eye position signal and its influence in back propagation simulations is presented. The signal is then further analyzed using other learning paradigms. Some results about the interactions with other stimuli which could influence the direction of further research are then presented.

6.1 Biological basis of tonic input

In area 7a, and in the lateral intraparietal area (LIP) both cortical fields of the inferior parietal lobule significant effects of eye position were reported in the majority of cells (Andersen et al., 1990). The posterior parietal cortex is a visual area (Andersen and Gnadt, 1989) as lesions did not produce deficits in saccades to auditory or somatosensory stimuli. The parietal cortex and superior colliculus perform different transformations, but both processes might exploit an eye position signal which is equivalent in nature. A modeler is forced to speculate about the influence of eye position because he lacks neurophysiological data about the pathway of auditory input to the SC. Eye position signals must modulate this pathway in some way in order to accomplish the shift of auditory receptive fields in the SC (Jay and Sparks, 1984).

In the parietal cortex, the majority of cells (Area 7a : 53%) respond to a combination of visual and eye position input (Andersen et al., 1985; Andersen and Gnadt, 1989) but some cells exhibit activity exclusively related to eye position (Area 7a: 16%, LIP : 11%). In a recent study of area 7a and the LIP, Andersen and colleagues found 14 out of 409 cells solely responsive to eye position whose "activity varied linearly with eye position" (Andersen et al., 1990). These neurons fire monotonically according to horizontal and vertical orbital angle of the eye. This kind of response presumably reflects a tonic eye position input to the entire area including those neurons with additional visual input.

In the parietal cortex a visual stimulus and an eye position signal are combined, resulting in very characteristic response properties for a cell. Andersen refers to the mode of interaction as "a planar spatial gain field" (Andersen and Zipser, 1988). Unlike a receptive field, a gain field is not intuitively clear. Suppose a cell responds to both eye position and a visual stimulus on the retina. The relation of the visual receptive field to the world (the head-centered coordinate system) is dependent upon eye position. The reference points change every time the retina fixates a different location. Thus with the varying eye position, the visual stimulus must be "re-located" in order to yield the same response in the cell.

A cell may respond to the eye position signal without a visual stimulus. In this case a tonic background activity can be tested in the cell which is due exclusively to an eye position signal. The influence of a visual stimulus alone cannot be determined in an experiment because it is impossible to suppress the eye position signal. But the difference between the response to the visual stimulus combined with eye position and the tonic background activity for eye position includes equivalent information. This difference represents a gain of visual input, which can be computed from experimental data. For recording data at a certain eye position, the tonic background activity must simply be subtracted from the response evoked by a visual stimulus.

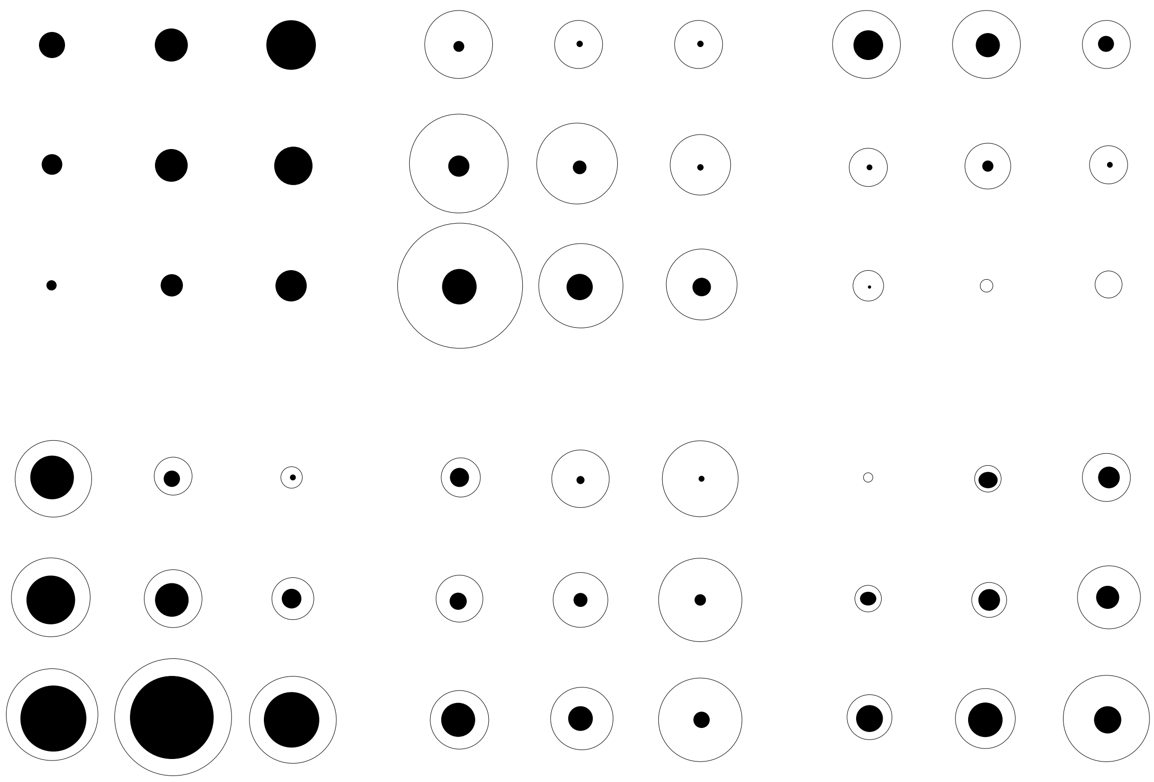

Figure 6.1: Spatial gain fields of six neurons. The diameter of the darkened inner circle represents the visually evoked gain.

In their experiments, Andersen and colleagues (1985) recorded the background activity and visual gain on nine different eye position equally distributed on a 3 x 3 grid with 20 degree spacings. The grid was centered at straight ahead. Figure 6.1 shows a graphical representation of a spatial gain field. For the majority of the cells in area 7a and LIP these gain fields are roughly planar. For 55% of the neurons the visually evoked gain of activity can be fitted to a plane appropriately tilted horizontally and vertically. For 80% of the neurons the total response activity can be modeled by a plane. The crucial point is that the gain fields do not require more complex approximations. The maximum activity at one edge of the field decreases linearly in all directions of the field.

6.2 Back propagation and tonic input

The hidden units of the back propagation model (Zipser and Andersen, 1988a; Andersen and Zipser, 1988) developed spatial gain fields identical to those of neurons in the parietal cortex. The actual appeal of the back propagation simulation arises from this similarity. All of the total-response gain fields in the model were planar independent of the output format. When a monotonic output format was trained, the visual gain fields could more often be approximated with a plane (67%), compared to the result with a gaussian output format (13%). This similarity of spatial gain fields in the hidden layer and in the parietal cortex is quite impressive because the planar shape of the gain fields characterize the interaction of eye position signal and visual stimulus distinctively.

The influence of eye position at early stages of visual processing has recently been reported in several neurophysiological experiments (Galetti and Battaglini, 1989; Lal and Freedlander, 1989; Trotter et al., 1991; Weyland and Malpeli, 1990). Pouget et al. (1992) modeled area V1 and area V3a of the visual cortex as a hierarchy of hidden layers with a back propagation algorithm using shared weights (LeCun et al., 1990). These layers learned to transform a gaussian retinal stimulus into a head centered coordinate system where the stimulus was represented by a monotonic output format. The influence of eye position was simulated with a linear signal to all hidden layers. The same planar shaped gain fields developed for the hidden units as in the back propagation simulations of the parietal cortex.

A three layer network trained with a variant of the associative reward-penalty learning rule (Barto and Jordan, 1987) also generated largely planar gain fields. The correspondence to the neurophysiological experiment was reported to be "approximately equivalent" (Mazzoni et al., 1991) to that of the back propagation model (Zipser and Andersen, 1988a). However, the binary output of the hidden units in an associative reward-penalty network complicates the comparison. Thus, the probability of a unit's firing was used to determine the gain fields in the simulations.

6.3 Feature mapping and tonic input

Feature mapping is an unsupervised learning paradigm for neural networks. The term was introduced by Kohonen (1982) but the concept is closely related to competitive learning (Grossberg, 1976; Rumelhardt and Zipser, 1985; Fukushima, 1975). The idea of competitive learning is extended by feature mapping networks which form a topology preserving map in the output layer. Feature mapping is a very powerful mechanism for clustering or categorizing input data. A short introduction to this learning paradigm is given in Hertz et al. (1991).

Nearly all learning algorithms related to competitive learning have an important, inherent detail in common. During or after the learning step, they normalize the weights of each neuron. From the biological point of view this is a major flaw in the learning paradigm. In the brain only local information is available at a synapse. A normalization requires a synapse to know the strength of other synapses, which is "global" information.

Normalization is crucial to the algorithm. If a winning neuron in a cluster of competitive neurons can increase the total impact of its weights, it will begin to react to all inputs, to win the competition every time, and to suppress the activity of all other neurons. Thus the neurons in the cluster cannot develop any selectivity and will not become feature detectors. Grossberg (1988) proposes a normalization of the input patterns through lateral inhibition. With normalized input patterns, the learning step can apply a local Hebbian rule and still produce normalized weights in each output neuron. However, the application of lateral inhibition is a very sensitive and time-consuming process which makes this kind of normalization mechanism inappropriate in a wide range of situations.

Note that information about eye position may naturally form a normalized pattern. A linear activation function of certain slope and intercept maps eye position into a neuron's response. For every neuron activated by one function, an antagonistic neuron might exist with an antagonistic linear activation function of opposite slope and inverse intercept. Added together, both neuron activations always result in the same total activity. When several pairs of antagonistic neurons are grouped together, the sum of activities also remains constant for different eye positions. Therefore, the activity pattern for such a group of neurons continues to be a normalized input pattern even when eye position changes (Fig. 6.2). In the brain the eye is moved by antagonistic muscle pairs. This suggests a similar organization of antagonistic information in the biological system. If one muscle discharges vigorously the antagonistic muscle will not discharge at all. But the discharge activity of both muscles together always remains constant.

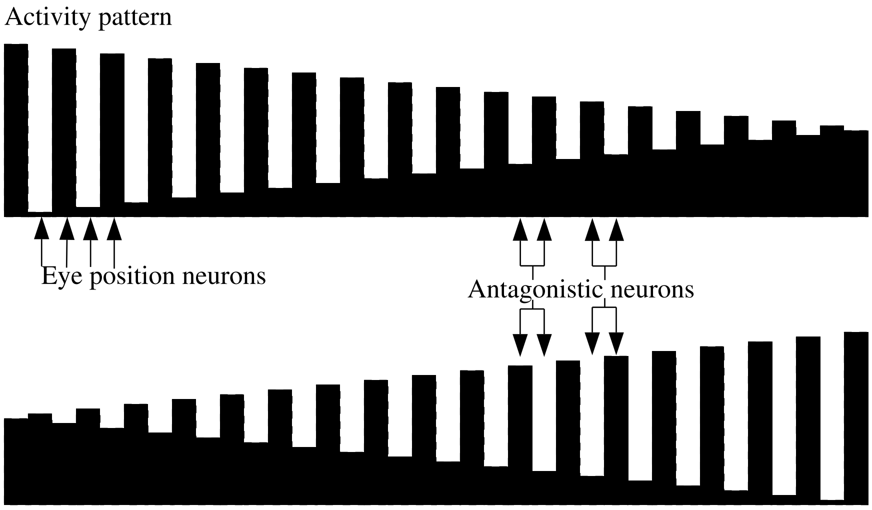

Figure 6.2: A normalized pattern for eye position.

6.3.1 A two layer network for the classification of eye positions

The eye position input appears to form an ideal pattern for a learning paradigm such as feature mapping. Feature detectors for different eye positions should develop easily and result in a network which is indeed capable of discriminating different eye positions. For the present study, a two layer network was constructed to develop feature detectors for horizontal eye positions. The network was fully connected between the input and output layer. Internal connections in the output layer provided the framework in which the neurons interacted in soft competition. A one-dimensional topography in the output layer specified a left and right neighbor for every neuron. A Mexican hat distribution (as it is known in neurophysiology) defined the pattern of connectivity within the output layer (Fig. 6.3). Every neuron excites its neighbor but at the same time inhibits neurons at a greater distance. The inhibition decreases its intensity proportional to the distance (Fig.6.4).

A soft competition was preferred to a winner-take-all scenario for two reasons: First of all, winner-take-all networks have so far not demonstrated a very reliable mechanism.

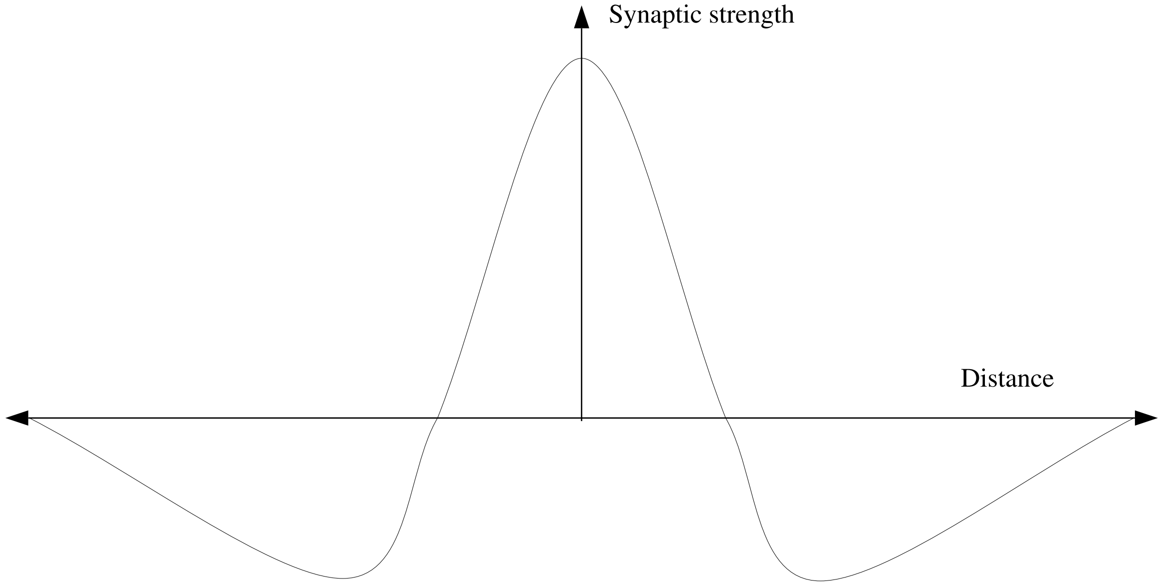

Figure 6.3: The pattern of connectivity for a Mexican hat distribution. The strength of the synaptic weights are related to the distance to postsynaptic neurons.

For certain input configurations, the network activity either starts to oscillate or to cease interaction at an intermediate stage where several neurons are still active. Secondly, the network was intended to learn a topographical map of horizontal eye position in the output layer where neighboring neurons should develop to react to adjacent eye positions. This is the main difference between feature mapping and competitive learning. Competitive learning with its strict winner-take-all specification will not develop a topography in its output layer. Soft competition simply enhances the response properties for neighboring neurons. Thus, the linkage of soft competition with a learning process groups responses together in a topography. However, soft competition is not the mechanism used in Kohonen feature maps (Kohonen, 1982), where a neuron with maximal response is determined and trained on the input pattern. The weights of neighboring neurons are then modified regardless of

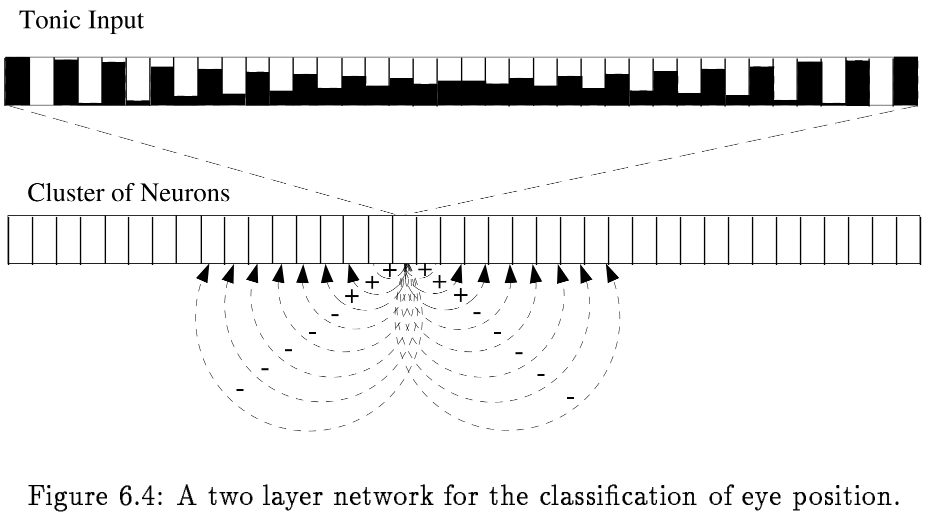

Figure 6.4: A two layer network for the classification of eye position.

their activations.

A single propagation step can implement a soft competition. It simulates a first order approximation of the differential equations involved in the process of lateral inhibition. The activity of all input neurons is temporarily cleared and the output layer activity is propagated through the Mexican hat distribution, resulting in a new activation pattern in the output layer which was computed exclusively from its former activity. An area of highly active neurons in the one-dimensional topography of the output layer will increase its own activity. A weaker area will be inhibited.

The training includes a simple learning cycle:

1. An eye position is chosen at random, determining the input pattern.
2. This input pattern is propagated to the neurons in the output layer, establishing a pattern of activity.
3. Soft competition enhances the contrast in the response pattern.
4. A learning step further increases selectivity among the output layer neurons.

In a Hebbian learning rule synaptic weight modification is proportional to postsynaptic activity. A highly active neuron learns an input pattern better than a neuron with a weak activity. Thus, the process of learning forces each neuron to react more strongly to a certain input pattern, increasing a neuron's selectivity for one of the eye positions.

The network developed selectivity exclusively to the two eye positions at the medial or lateral extreme of the possible orbital angles. Although 40 different eye position patterns were trained, none of the feature detectors reacted to any of the 38 intermediate positions optimally. In the testing phase each input pattern was presented, starting at the extreme medial eye position, moving towards the extreme lateral position (Fig.6.5). A large amount of feature detecting neurons responded vigorously to the extreme medial orbital angle of the eye. These neurons decreased their response activity when eye positions which were less medial represented the input patterns. However, no other feature detecting neurons started to fire as the orbital angle of the eye approached the center position. At center position the overall activity in the output layer was lowest, because feature detectors for the extreme medial eye position were only weakly activated, feature detectors for center positions did not exist, and those feature detectors for the extreme lateral eye position had not yet begun firing. In some of the networks, a zero response activity could be tested for the eye position "straight ahead". As the eye position became more lateral, the feature detectors for the extreme lateral eye position started to exhibit activity. This activity reached its maximum value in accordance with that input pattern which encoded the extreme lateral orbital angle of the eye. This result is very problematic for the task of discriminating different eye positions. It is impossible to classify different eye positions with these feature detectors.

6.3.2 A problem of normalization

The main cause of this "binary" behavior of the feature detecting neurons first seemed to be a problem in the normalization mechanism. Note that eye position patterns are normalized in L1, the Maximum or Manhattan distance norm. The sum of activity in the input pattern remains constant for all eye positions. Normalization in L2, the Euclidean norm, imposes a different condition on the input patterns. Every input pattern with n elements can be viewed as a vector in *n-dimensional* Euclidean space. Vectors with the same length are normalized in Euclidean space. The weighted input to a neuron is (mathematically) the result of the

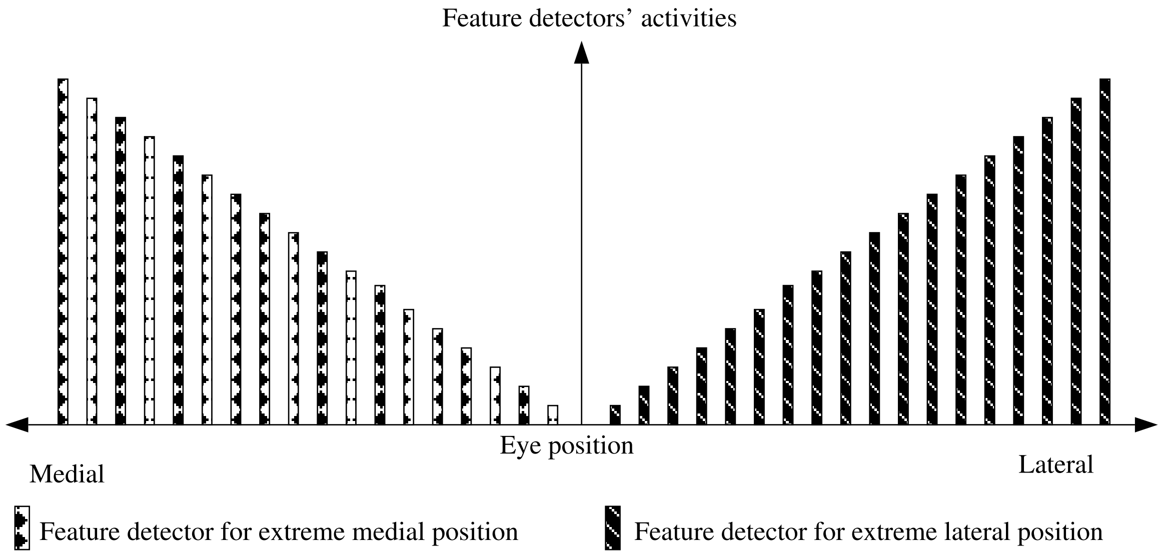

Figure 6.5: The activity of the two classes of feature detectors for all eye positions.

multiplication of the input vector with the neuron's weight vector. With a normalization in L1, extreme eye positions are represented by a longer vector. In L1 the weighted input will always be greater at extreme eye positions than at intermediate positions. A geometrical display of this argument is given in figure 6.6.

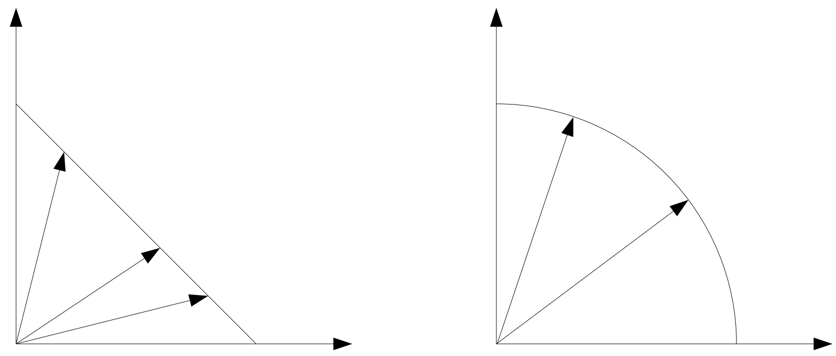

Figure 6.6: Normalization in L1 and in L2 for a pattern with two elements

In subsequent back propagation models of the parietal cortex a squared function was used to map eye position into neuron activities. The researchers explain this squared function as "a rationale to approximate the cumulative response of a group of eye position cells" (Andersen and Zipser, 1988). A squared mapping function offers an interesting perspective because it may produce input patterns which are normalized in L2.

In the following simulations the same training cycle was iterated on an equivalent network. Although the set of 40 input patterns for eye position was normalized in L2, the Euclidean norm, the results are identical to those with a normalization in L1. Again, all output layer neurons exclusively develop a selectivity for the two extreme eye positions. None of the feature detector neurons responds optimally to intermediate positions. A local Hebbian learning rule was used as it is described in chapter 2. This strictly local learning was substituted by a rule which normalizes the weights in L2 after each learning (von der Malsburg, 1973; Kohonen, 1982; Rumelhardt and Zipser, 1985). Even then, the network did not create feature detectors for intermediate eye positions. The output neurons responded exactly as in previous simulations. The more sophisticated normalization schemes did not improve the results, compared to a normalization of input patterns in L1. Thus, the problem is not one of normalization.

6.3.3 A problem of tonic input

The representation of the input patterns is the next suspect for causing the problematic behavior of the feature detecting neurons. In order to confirm this hypothesis, another learning paradigm was tested with tonic input. How would supervised Hebbian learning manage to associate tonic input with a categorizing output? A network equivalent to those used in preceding simulations was trained to answer this question. Input patterns were normalized in L1, but the network performed a different training cycle:

1. An eye position is chosen at random, determining the input pattern as well as an output pattern.
2. A learning step associates input and output patterns with Hebbian learning.

An output pattern represents a localized activation on a map of categorizing neurons. Exactly one neuron encodes the position of the eye on a one-dimensional topographic map. Neighboring neurons represent neighboring eye positions.

In the testing phase the neurons in the output layer responded again with a "binary" pattern of activity. Depending on their response to the tonic input, two different classes of neurons can be distinguished. Any orbital angle of the eye which was oriented medially generated a pattern of activity in the output layer differing from that of an eye position with a lateral orbital angle. Let the input neurons represent an eye position in between the center and the extreme medial eye position. Only that output neuron which was trained to react to exactly this particular input pattern should be active, but the neurons trained to respond to eye positions further medial are even stronger activated. The strongest response is shown by those output neurons coding the extreme medial eye position. The activity in the output layer decreases linearly for neurons trained on extreme positions to neurons trained on center positions (Fig. 6.7). These output patterns did not change significantly as long as an eye position medial to the center represented the input pattern. None of the neurons which encode positions lateral to the center were activated. For a lateral eye position input the same pattern of activity is generated symmetrically. Whenever the input neurons encode an orbital angle lateral to the center position, the same linear slope of activation evolves in the other part of the one-dimensional topographic output map.

In a control simulation the same network was trained with input patterns normalized in L2. The learning rule normalized the synaptic weights after each learning (Kohonen, 1982; von der Malsburg, 1973; Rumelhardt and Zipser, 1985). For the first time, feature detectors for eye positions developed with an enormous amount of training cycles. In the testing phase the same output neuron responded optimally to an input pattern which was trained as a teaching output for that particular eye position. Extreme or intermediate eye positions were represented equally well in the output layer. But these results require a strong teaching input to clamp the activity in the output layer. Neither the feature detectors in an unsupervised learning paradigm nor the hidden units of a back propagation model have

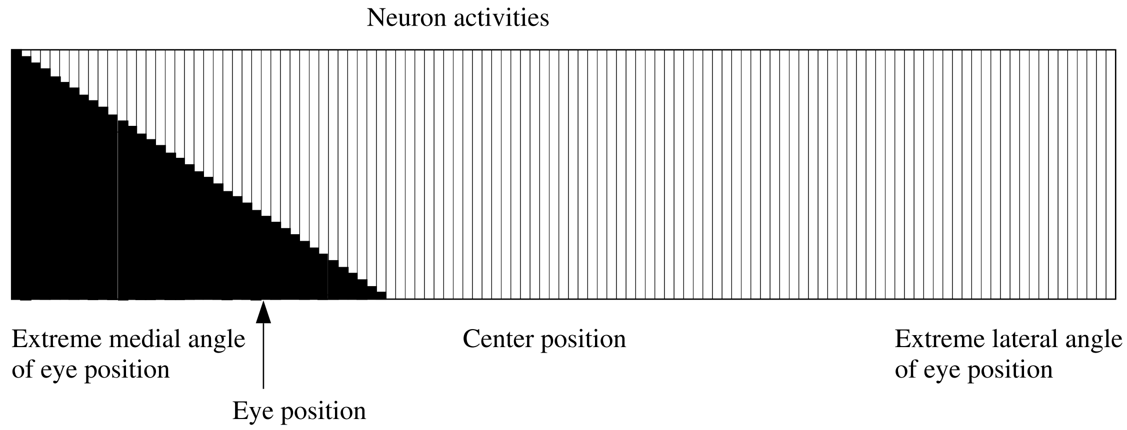

Figure 6.7: A pattern of activity in the output layer

such a strong teaching input. Thus, in both cases the response pattern is shaped by the tonic input. The difference between a normalization in L1 or in L2 merely increases the tendency for such a "binary" response.

A principal component analysis was performed for those input patterns normalized in L1. It revealed a covariance matrix with only two positive eigenvalues. In a 40 x 40 matrix all other eigenvalues are zero. The eye position neurons form highly correlated input patterns which are difficult for a network to learn, and classify. In a successive simulation the output layer neurons were trained without lateral inhibition and without a teaching input. Their weight vectors converge to the eigenvector with the largest eigenvalue as could be concluded from theoretical research on unsupervised learning (Sanger, 1989). This eigenvector represents an intermediate eye position.

6.4 Conclusion

The simulation results are unexpected. From a theoretical point of view, feature detectors should have developed for all eye positions including the center position. The antagonistic activation of neurons produces input patterns which would normally guarantee selective behavior in the output layer. However, feature detectors cannot establish a stable response when balanced between antagonistic synaptic weights. Even if a feature detector encodes an intermediate position at some point of the learning, it will lose this balance between the antagonistic inputs with successive learning trials. It will then shift its weights towards an extreme eye position. The same process happens for the hidden layer units in a back propagation model. Thus, the results of the simulations have one clear interpretation. The neurons in the output layer of a feature mapping network react very similarly to the hidden layer units in the back propagation simulations of the SC in the present work.

The hidden layer of back propagation networks forms an individual internal representation for each network. Although two back propagation networks which are distinguishable only by their initial weights might learn the same input-output associations, they will never develop identical responses in the hidden layers. In the back propagation simulations of the SC, two classes of units always developed in the hidden layer which are equivalent to the two classes in the feature mapping model. One class responds vigorously to the extreme medial eye position inputs. These hidden units decreased activity with more intermediate eye positions. At center position the activity in the hidden layer is minimal. Then, the hidden layer units in the other class start to respond. The more the orbital angle of the eye represents a lateral position, the stronger activity generated in the second class of units becomes. The same behavior was found for the feature mapping networks.

Suppose a feature detector for an intermediate eye position was given. How would the spatial gain field of that detector be shaped? It cannot be planar. The feature detector must respond optimally to an intermediate eye position. Therefore its spatial gain field must have a peak of activity at an intermediate eye position. Such a gain field cannot be approximated by a plane. Instead it has an intermediate location in the field where activity is at its maximum (Fig. 6.8).

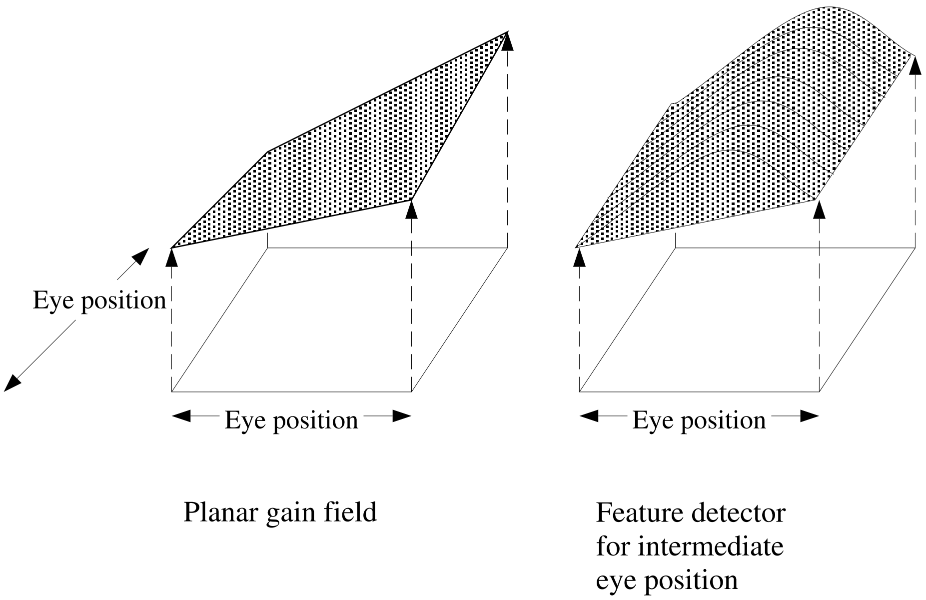

Figure 6.8: A planar spatial gain field compared to the spatial gain field of a feature detector for an intermediate eye position.

A feature detector with a planar spatial gain field must respond similar to the output layer neurons in the feature mapping simulations. Thus the results with unsupervised learning can be correlated to neurophysiological findings (Andersen et al., 1985; Andersen et al., 1990). The influence of the tonic input pushes the spatial gain fields to be planar when the units are not severely restricted in self-organizing their response. Feature mapping networks and back propagation networks have one characteristic in common: they form their own internal representation of the input pattern. If a learning rule is not limited in terms of the formation of an internal representation of the input, it will develop planar spatial gain fields with a tonic input activity. The agreement between model and experiment is a general one, not limited to back propagation simulations.

Chapter 7

A network model for coordinate transformation

To solve the sensor integration in the superior colliculus (SC), a neural network is proposed which learns the coordinate transformation of a head-centered stimulus into an oculocentric coordinate system. The network combines an auditory stimulus and an eye position signal into an internal representation. In different versions of the network these internal representations either self-organize in the learning process, or they are determined by the network structure. A motor error signal from within the eye-centered coordinates is then associated with each internal representation. Thus, the network can establish the appropriate input-output associations without the use of back propagation learning.

A network architecture based on a concept related to the self-organizing version of the network was introduced by Hecht-Nielson (1987,1988) and others (Huang and Lippmann, 1988). There, the architecture is called counterpropagating. The present network is influenced by neurophysiological data, including the results from deprivation experiments (Knudsen et al., 1984a; Withington-Wray et al., 1990d; Withington-Wray et al., 1990a) and the distributed representation of neural activity in the SC (McIlwain, 1991). Thus, the mechanisms implemented here are not as general as those found in counterpropagating networks. Nevertheless, one result from research on counterpropagating network is encouraging: "They are typically 10-100 times faster to train than conventional back propagation networks, with results that are often comparable" (Hertz et al.,1991 p.248). Considering the very short crucial period for the development of auditory receptive fields in the SC (Withington-Wray et al., 1990c), a fast learning procedure is more plausible than back propagation or reinforcement learning.

Most of the simulations were limited to the transformation of stimuli in the horizontal dimension. The vertical component of space was omitted for several reasons: firstly, the neurophysiological basis for vertical localization of acoustic stimuli has not yet been established. Interaural time difference and interaural intensity difference both determine the location of a target on a horizontal axis (Konishi, 1986). A display of the neural circuitry is given in figure 7.1. It is assumed that signals for horizontal and vertical orientation are processed on parallel pathways. The auditory cues used to locate a target vertically are still the subject of major research. The bat computes the vertical location of its prey from the reflections produced in the pinnae by an incoming sound wave (P. Salient, 1992 personal communications). In owls, one ear is shifted on the head with respect to the other ear (Knudsen and Konishi, 1979). The shift generates interaural time differences and interaural intensity differences which can be utilized as vertical cues. How vertical localization information is represented, is still under investigation. Another biological motivation for concentrating only on the horizontal component of the transformation is a lack of data. The shift of auditory receptive fields in the SC (Jay and Sparks, 1984) was proven only for the horizontal axis. The experiment was not designed to reveal anything about the vertical axis.

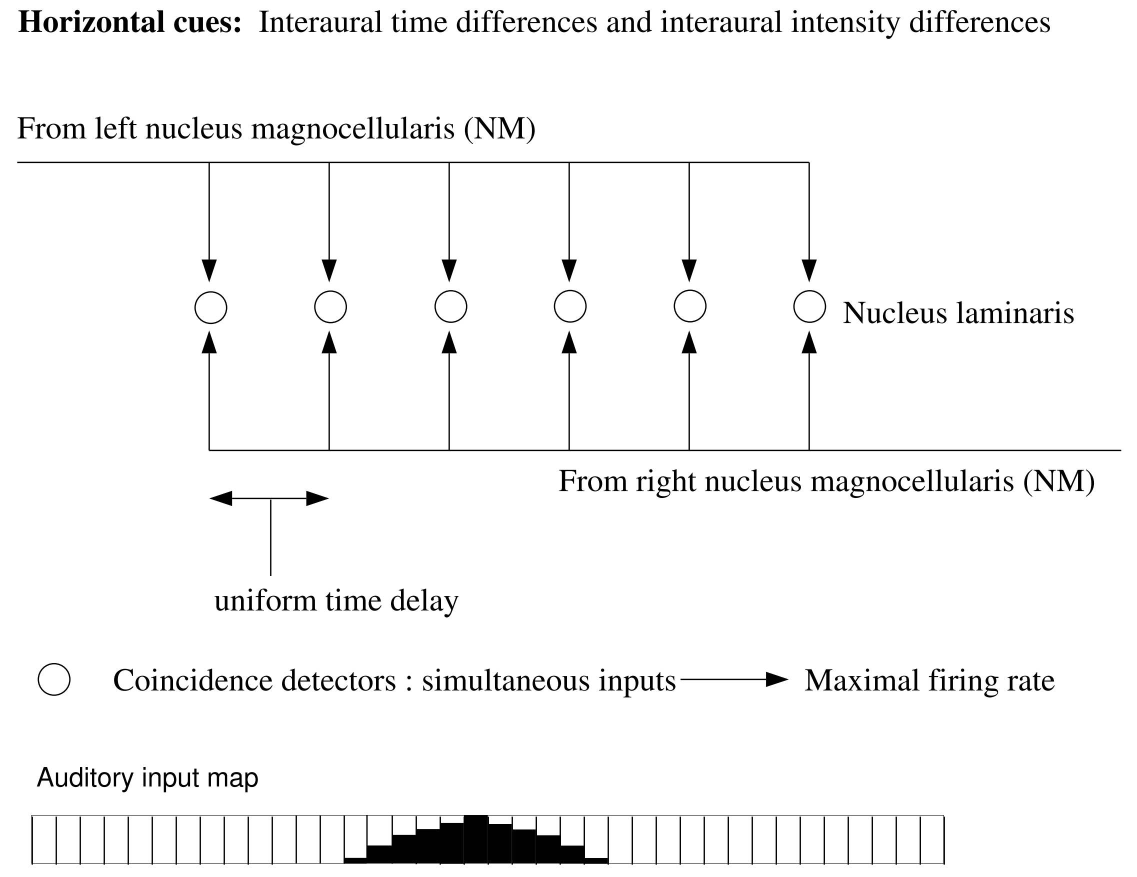

Figure 7.1: The neural network to compute the horizontal location of an auditory target

The third reason to restrict the transformation to the horizontal axis is a matter of simplicity. In previous simulations it proved to be much easier to analyze a network which simulated one dimension. The mechanisms involved are more intelligible, as other inputs (such as neurons which encode the vertical component of eye position or the vertical dimension of the auditory map) merely obscure the transformation process. However, the network proposed is inherently modular and constructed with a vertical extension in mind. Possibilities for incorporating the vertical dimension in the network are discussed throughout this chapter.

7.1 Tonic influence in the layer of internal representations

The network combines two different input sources when constructing an internal representation. One of the sources consists of neurons coding the eye position in orbit. Linear activation functions transform the horizontal orbital angle of the eye into neuron activities.

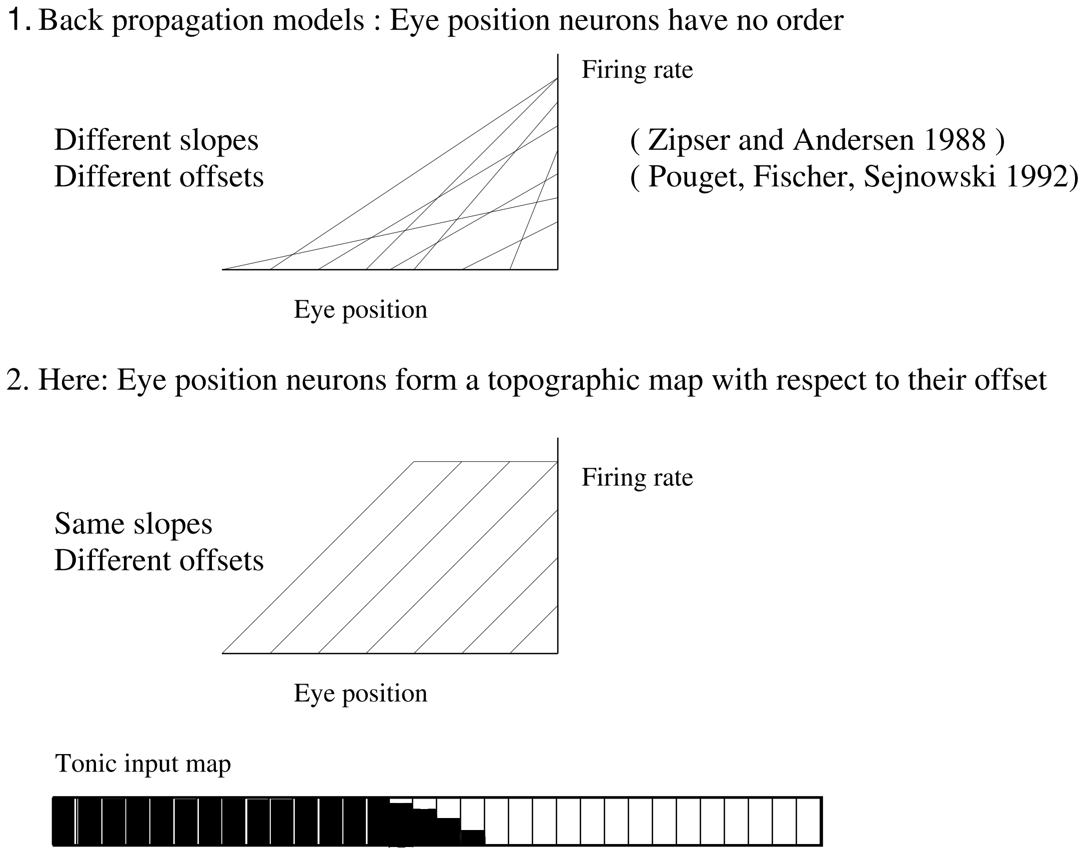

Figure 7.2: Eye position signal in back propagation simulations and in the present model

The same positive slope is applied in each activation function, but the intercept or offset varies with different eye position neurons (Fig. 7.2). With respect to the offset, these neurons form a one-dimensional topographic map. The offset decreases steadily from neurons at the one end of the map to the other. The two extreme neurons have the maximum and minimum offset, respectively. This organization of eye position neurons on a one-dimensional map results in a very specific activity pattern. Every eye position corresponds to an edge of neuron activation as shown in figure 7.3. The location of this edge on the map encodes the horizontal orbital angle of the eye.

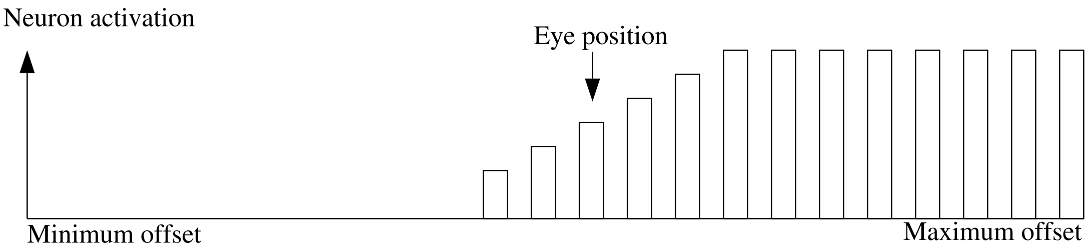

Figure 7.3: The edge of neuron activity on a one-dimensional map

The neurons in the layer of internal representations form distinct clusters. In a cluster all neurons have an identical distribution of synaptic weights which connect to the tonic input map. Therefore, all neurons in one cluster respond optimally to the same eye position. The ratio of synaptic weights to eye position neurons are defined by an on-center-off-surround distribution. This pattern of connectivity was repeatedly found in the brain. It is well known from data on ganglion-cells or Y-cells in vision. The center of the distribution on the tonic input map changes for every cluster, for which the distribution is characterized by near excitation and far inhibition. Neurons in the same cluster are connected excitatorily to their center. All other eye position neurons in the map inhibit the neurons of that same cluster. The distribution of weights is sketched in figure 7.4. Such an organization of synaptic weights operates similarly to a Laplace operator used for edge detection in computer vision. In the network, the distribution extracts the edge of activation from the map of eye position neurons. Neurons in the internal layers will be sensitive to this edge only if excitation and inhibition are chosen appropriately. Only those neurons with excitatory synapses close to the edge of tonic input receive excitation.

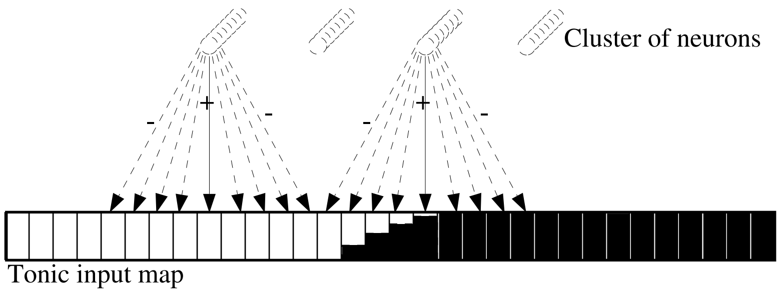

Figure 7.4: An on-center-off-surround distribution

The excitatory connected center of the on-center-off-surround distribution varies from cluster to cluster in the layer of internal representations. Thus, the eye position input selectively suppresses activity in some clusters, but excites others. If the excitatory center in a cluster coincides with the edge of activation on the tonic input map, the cluster's activities will be enhanced. Otherwise the cluster's activation will be inhibited. Cluster neurons are selective to eye position. The central idea is to create internal representations which are as far apart as possible for each input configuration of tonic and auditory input. The association with a motor error signal in the output layer will be learned only if the internal representations are easily distinguishable.

The response of cluster neurons in the layer of internal representations does not produce "planar spatial gain fields" (Andersen et al., 1990) for the combination of tonic and auditory input. Rather for every neuron, the on-center-off-surround distribution generates a peak of activity for a particular eye position. Eye positions more lateral or more medial result in a weaker response. For the internal representations this is crucial. A neuron with a planar spatial gain field cannot respond optimally to an intermediate eye position. Once it is maximal active, a neuron will not decrease its activity again with a greater orbital angle of the eye. And such a response behavior causes a great amount of overlap in the layer of internal representations. It frustrates a correct association of the responses in the layer of internal representations to the output layer itself. A very sophisticated scheme of synaptic weights connecting to tonic input neurons might produce selective responses to different eye positions and still yield the gain field properties of neurons in the parietal cortex (Andersen and Gnadt, 1989). The two goals, however, creating selectivity to distinct eye positions on the one hand and linear response properties on the other seem mutually exclusive.

7.2 Auditory influence in internal representations

Neurons in the layer of internal representations also respond to auditory input. In the network proposed, a one-dimensional map of auditory neurons simulates horizontal auditory space, and an auditory stimulus is represented as an gaussian activity. The location of activity on the map encodes the horizontal stimulus eccentricity with respect to head orientation. In brains, such an activation pattern results from a sophisticated mechanism which detects interaural time differences and interaural intensity differences.

During the learning phase the network is constrained to process a single auditory target. Target selection circuitry must filter auditory input before it reaches the transformation stages. After this developmental period, different auditory stimuli will simultaneously evoke different internal representations, which will inhibit each other's response in the output layer.

Two different network versions exist for the transformation of a head-centered auditory stimulus into a oculocentric coordinate system. These versions differ mainly in the way the auditory input influences the development of internal representations.

7.2.1 A self-organized response to auditory input

In the layer of internal representations several clusters of competitive neurons exist. Each cluster is selective to a specific eye position. The auditory input can be used to control the self-organization of responses within a cluster, which consists of a group of neurons with a one-dimensional topography. The topography defines a "neighborhood" relationship within the cluster. It implements a Mexican hat distribution as the pattern of connectivity between the neurons in the cluster. Neighboring neurons excite each other. But from a certain distance on they exchange inhibitions. The inhibition between cluster neurons decreases proportional to the distance between the neurons. With that pattern of connectivity, the neurons in a cluster interact in soft competition (Fig.7.5).
In the training of a network the following simple cycle was repeated:

1. Both eye position and auditory stimulus location are chosen at random.
2. The input patterns are propagated to the neurons in the layer of internal representations and establish a pattern of activity.
3. Soft competition enhances the contrast of the response pattern.
4. A learning step further increases selectivity among the cluster neurons.

The initially random synaptic weights from the auditory input map to the cluster neurons are modified with local Hebbian learning. Each cluster neuron develops selectivity for another auditory input stimulus. The neuron detects the auditory stimulus as an input feature.

In the cluster architecture the neurons develop an internal representation of auditory space. Different clusters self-organize into different configurations in order to encode auditory space. However, they all form topographic maps of auditory input. The self-organized topographic maps provide a greater flexibility for the transformation process. Environmental damage to the cochlea or to the functional organization of the primary auditory pathways can be compensated through self-organization. Thus, the transformation mechanism adapts better to abnormal, environmental conditions.

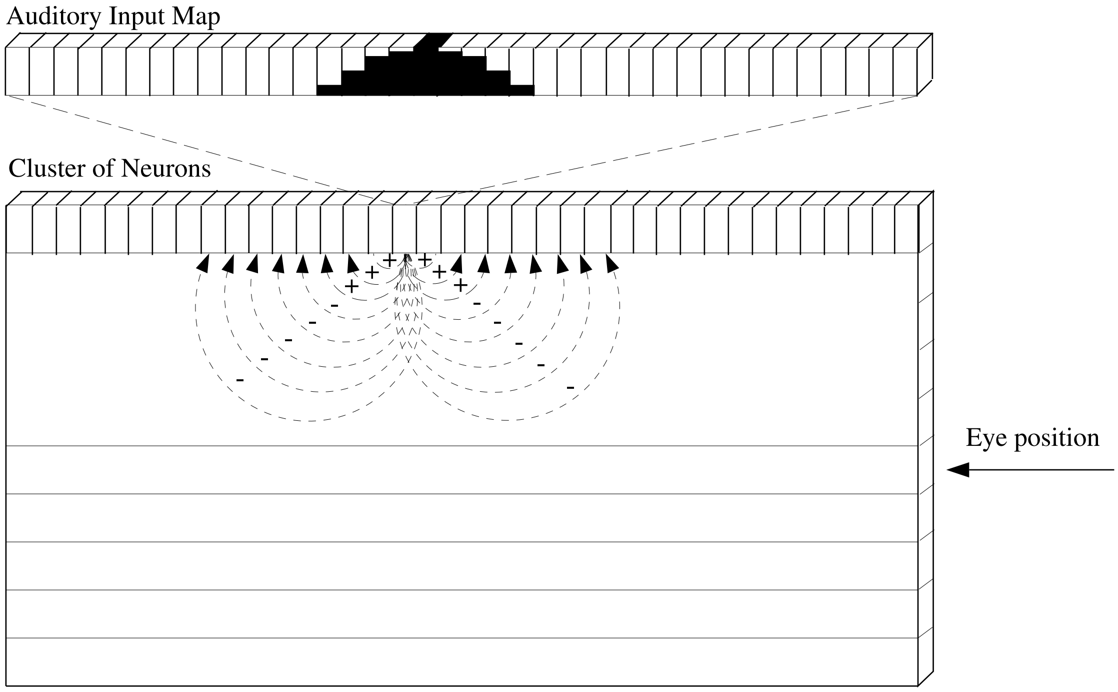

Figure 7.5: A cluster of neurons

7.2.2 A constructed response to auditory input

The response of a cluster neuron to an auditory stimulus can also be determined in advance. In this case, the projections from the auditory input map to the layer of internal representations are not initialized with a random distribution. Instead, the synaptic weights are designed to produce a one-dimensional map of auditory space within each cluster of neurons. The internal layer represents a two-dimensional map where the location of an auditory stimulus is encoded on the horizontal axis. The orbital angle of the eye determines the vertical location of activation in the internal layer. The combination of an eye position signal and an auditory stimulus results in a two-dimensional gaussian activity in the layer of internal representations as in figure 7.6. This gaussian activity moves horizontally with changes in the location of the auditory stimulus; changes in eye position move it vertically.

Note that the network architecture is not restricted to processing the horizontal component of eye position and auditory stimulus. It can also combine the vertical components of the eye position signal and auditory inputs. Thus, several possible extensions exist to overcome the current limitations of the horizontal component of the transformation. A second layer of internal representations may be added for the vertical dimension as in figure 7.7. In this case the association with activity in the output layer must include unique combinations for the horizontal and vertical component. The existing layer may also be extended into a three dimensional cube. The third dimension encodes the vertical dimension of the auditory map. With a fourth axis for the vertical component of the tonic input map, the network can be further expanded to a four dimensional object. A neural network does not have any problems representing multidimensional objects. However, in the brain, multidimensional topographic maps have not yet been established experimentally.

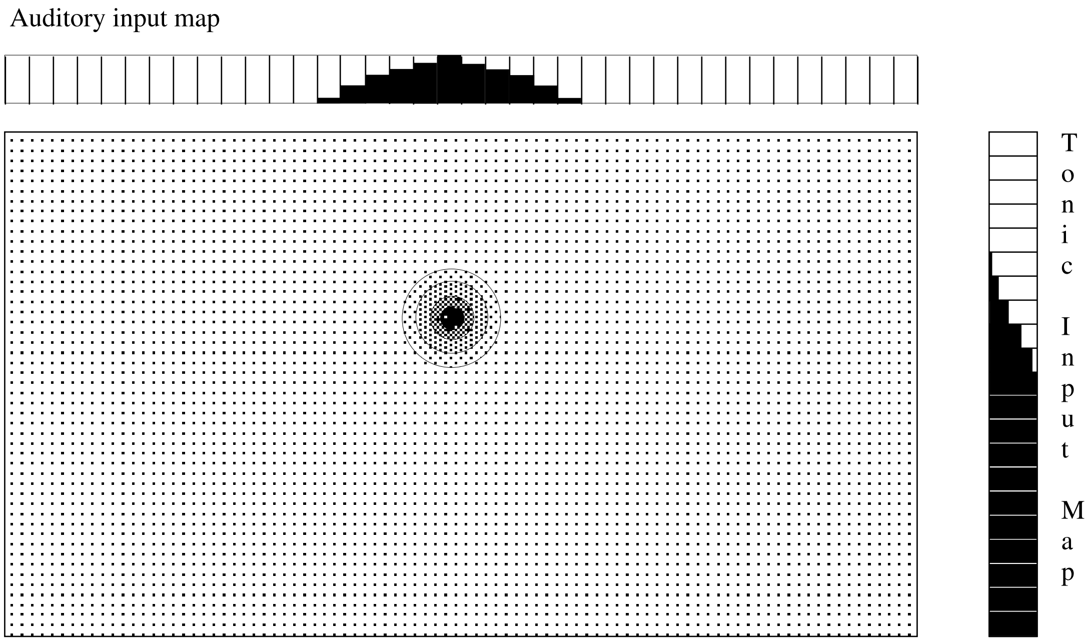

Figure 7.6: A constructed internal representation.

7.3 The output layer

In the deeper layers of the SC auditory and visual stimuli are integrated into a common motor error coordinate system in order to initiate a saccadic eye movement. This coordinate system is modeled by the output layer of the network. Without the vertical component of collicular space, the output layer consists of a one-dimensional map. In a natural environment an auditory stimulus will always originate from a proximal cause which is concurrently perceived as a visual stimulus. The visual stimulus creates an activity in the SC which coincides with the internal representation for the auditory stimulus. The deeper layers of the SC receive an oculocentric signal from the retina via superficial and intermediate layers. This signal instructs the development of the auditory map in the deeper laminae (Knudsen et al., 1984a; Withington-Wray et al., 1990d). In the network the visual stimulus is simulated as a teaching input for the association between the layer of internal representations and the output layer. A gaussian pattern of neuron activities represents this teaching input in the output layer.

The synaptic weights in the connections between neurons in the layer of internal representations and output layer neurons are modified during the training cycles. Through this association, the head-centered auditory stimulus is transformed into an oculocentric coordinate system in the output layer. A Hebbian learning rule establishes the appropriate synaptic weights for the association. The learning corresponds to a developmental process reported in neurophysiological experiments. In the SC of very young Guinea pigs, no topographic order exists for responses to auditory stimuli. The transition to a full topographic order occurs remarkably rapid between 27 and 30 days after birth (Withington-Wray et al., 1990c).

Self-organized and constructed network versions differ with respect to the initial synaptic weights between the layer of internal representations and the output layer. In the networks with a self-organized internal representation every neuron in the output layer must be

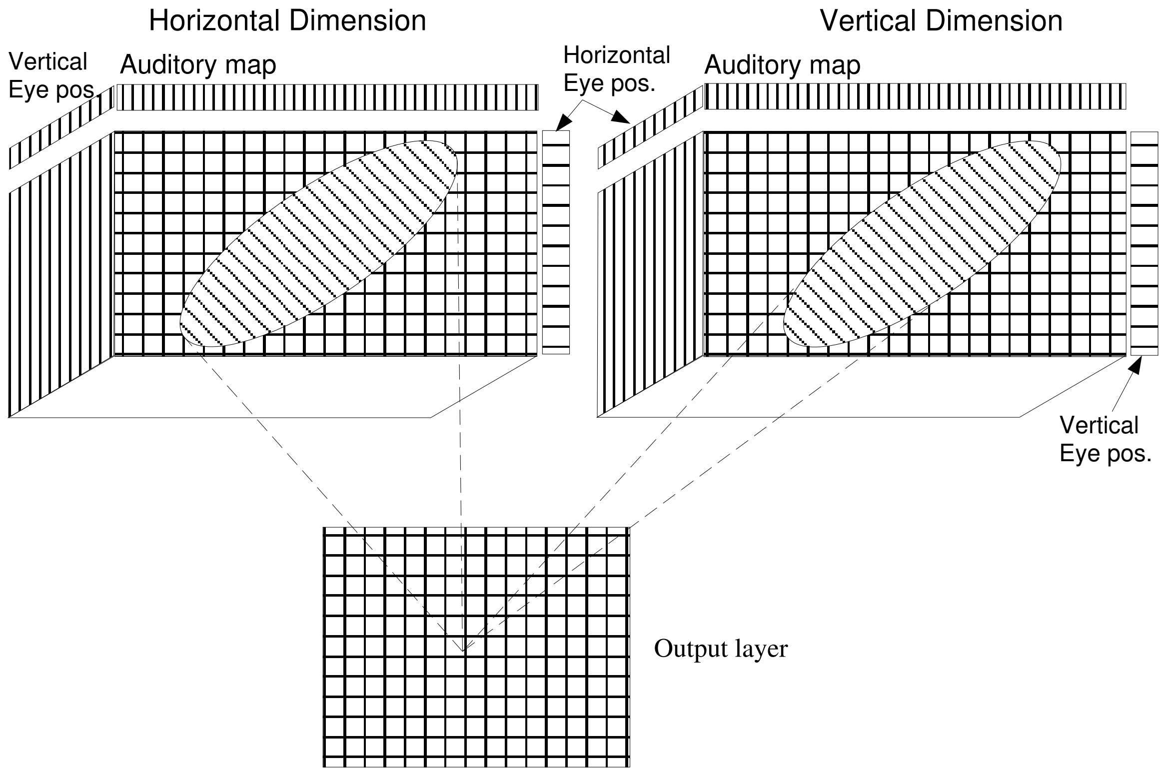

Figure 7.7: A possible network

connected to every neuron in the layer of internal representations. Every connection may be important for the transformation. With internal representations which are constructed into the network, the receptive fields of output neurons may be shaped more precisely from the beginning. This introduces a major computational advantage over the self-organized network. For example, an auditory stimulus at the lateral end of horizontal auditory space will never coincide with a teaching input at the medial end of collicular space. Therefore this synaptic weight is not required for the transformation. The omission of superfluous weights results in smaller networks which process more quickly. However, the resulting networks are also limited in their ability to model inhibitory interactions between auditory and visual stimuli, because they lack the synaptic weights to realize the inhibition. Such interactions were reported in several neurophysiological (Meredith and Stein, 1986a; King and Palmer, 1985) and psychophysical experiments (Lueck et al., 1990).

7.3.1 Learning in the output layer

Hebbian learning increases the strength of a synapse when pre- and postsynaptic neurons are simultaneously active (Hebb, 1949). For uncorrelated pre- and postsynaptic activities the synaptic weight is decreased through a decay term. With this formulation of Hebbian learning, the correct association between internal representation and teaching input in the output layer could not be established. Although the networks trained with simple Hebbian learning first reached a local minimum of 30% error rate, this rate later rose to 50%. There it remained as an equilibrium value.

In subsequent simulations another learning rule was applied, generally known as the

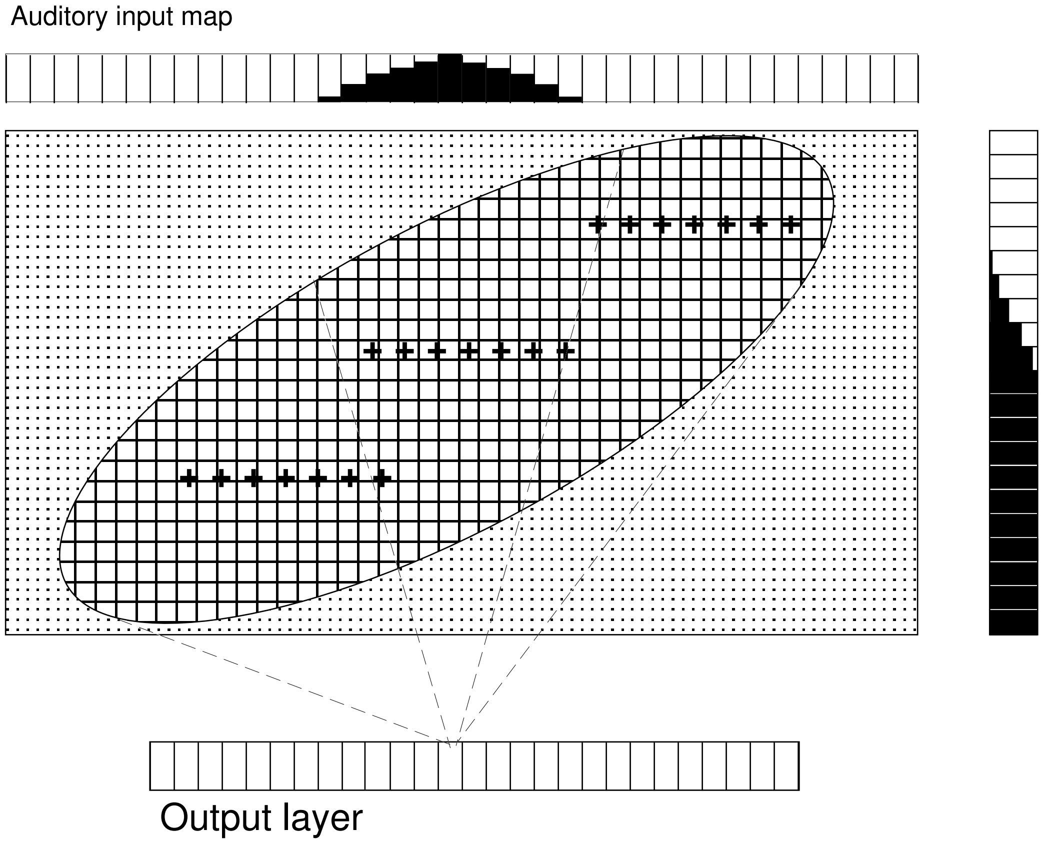

Figure 7.8: The initial synaptic weights of a output layer neuron and neurons in the layer of internal representations (Very large receptive fields).

"Widrow-Hoff" error correction procedure (Widrow and Hoff, 1960). The difference between the actual activity in the output layer and a teaching input results in an error signal. The Widrow-Hoff learning algorithm modifies all synapses with presynaptic activity, distinguishing two cases: in the first, a neuron's activity is higher than determined by the teaching input. The error has a positive sign. The learning rule decreases all excitatory synaptic weights because they are responsible for the error signal. It increases inhibitory weights because a stronger inhibition from the input pattern reduces the error. In the second case, the activity of the teaching input is higher than the neuron's activity. The error has a negative sign. All excitatory weights are reinforced. Inhibitory weights are decreased because they suppress the neuron's activity, resulting in an error. The increase or decrease of synaptic weights might be proportional to postsynaptic activity, to presynaptic activity, or to the error signal. In the simulations it was sufficient to learn proportional to postsynaptic activity, presynaptic activity and the error sign. The formula for weight modifications was presented in chapter 2. However, the association can be trained considerably faster when the synaptic weights are also modified proportional to the amount of error, which is the difference between teaching input and actual neuron activity.

Error correction learning is a biologically plausible model of the processes in the SC. In very young animals, the auditory receptive fields in deeper layer cells of the SC are "typically large, occupying most of the contralateral hemifield"(Withington-Wray et al., 1990d). Thus, an auditory stimulus excites a significant area of deeper layer neurons (Fig

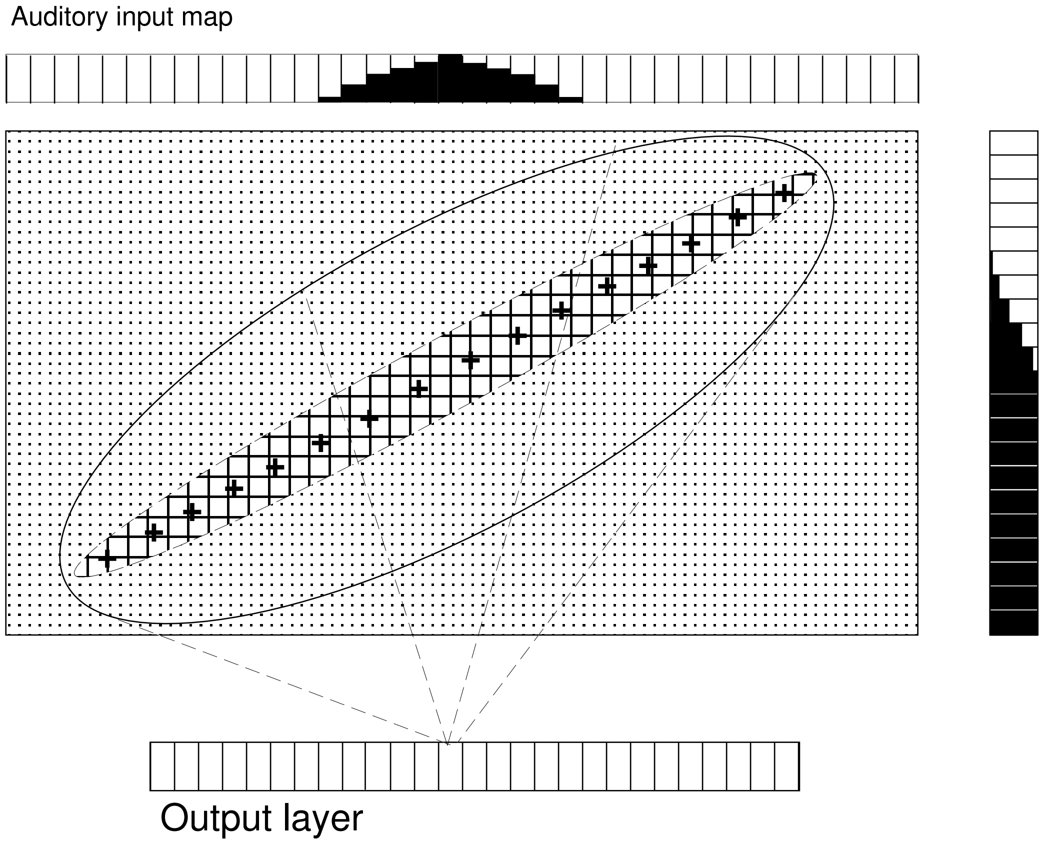

Figure 7.9: The learned synaptic weights: Only those combinations of auditory stimulus and eye position signal are reinforced which prove to be important for the output layer neuron.

7.8). Neurophysiological studies report that the visual signal controls the development of the auditory map (Knudsen et al., 1984b; Withington-Wray et al., 1990b). A coincident visual stimulus enhances some responses to the auditory stimulus, but depresses others. These interactions of visual and auditory stimuli in the SC were established in several studies (Meredith and Stein, 1986a; Meredith and Stein, 1986b; King and Palmer, 1985). Therefore, the visual input shapes the response to an auditory stimulus in the SC. Deeper layer neurons in the SC must modify their auditory receptive fields because a topographic map of auditory space develops there as a result of visual and auditory experience (Knudsen et al., 1984a; Withington-Wray et al., 1990a).

In a learning paradigm, those neurons with enhanced activity increase the weight of a synapse with auditory input. Synapses without presynaptic activity decay. However, this rule of synaptic modification changes for a neuron whose response is suppressed by the visual instruction signal. If a neuron's activity is inhibited below its average response, the synaptic weights will decrease when coincident with presynaptic activity. This scheme of long term potentiation (LTP) and long term depression (LTD) was introduced by Bienenstock et al.(1982). Recent experimental results with the visual cortex of cats and in the rat hippocampus (Bear et al., 1987; Clothiaux et al., 1991) support their theory of synaptic modification. Such a modification rule could explain the very short, crucial period during which a topographic auditory map develops in the SC (Fig. 7.9).

In the brain two different classes of neurons exist: excitatory and inhibitory. An excitatory neuron can never inhibit another neuron directly. An inhibitory synapse never changes its weight. Only the excitatory synapses learn and associate activities. Most neural network simulations go against these biological facts, arguing the following: an inhibitory neuron exists for every excitatory neuron. The inhibitory neuron inhibits the same postsynaptic neurons that the excitatory neuron excites. The inhibitory neuron always responds with the same activity as its excitatory neuron; the inhibition has a constant synaptic weight. If the postsynaptic neurons learn the synaptic weights to the excitatory neurons it will have a similar effect as would modifying the inhibitory synaptic weight. A decrease of excitatory synaptic strength influences the response of a postsynaptic neuron in the same way as an increase of inhibition, and an increase in excitation has the same effect of a decrease of inhibition. For the synaptic weights between the internal representation and the output layer, the same scheme is applicable. The use of this scheme is very important in order to account for depressing interactions of auditory and visual stimuli in the deeper layers of the SC (Meredith and Stein, 1986a; Meredith and Stein, 1986b; King and Palmer, 1985).

7.4 Results

The network architecture transforms a head-centered stimulus into an oculocentric coordinate system. The coordinate transformation evolves through the experience of coincident auditory and visual stimuli. Although the learning scheme is not as general as back propagation, it is well suited for the distributed representations in the SC.

The association in the network produces an activity in the output layer which is similar to the distributed activation of neurons in the SC (McIlwain, 1991). Several tests failed to learn a localized neuron activity in the output layer. The network always developed a distributed response. The great overlap of internal representations prevents a correct association to a localized grandmother cell.

Several networks were trained with self-organizing internal representations. In these networks 1661 neurons and 185,560 connections were simulated. In the auditory input map 40 neurons encoded auditory stimuli as gaussian activities. Seventeen different eye positions were represented as tonic neuron activities. The output map consisted of seventeen different neurons and encoded an oculocentric motor error signal as a gaussian activity. In the testing phase the response in the output layer was compared to the teaching input. In 92% of all tests the peak of the activity generated in the output layer corresponds to the peak of the gaussian activity in the teaching input (Fig. 7.10).

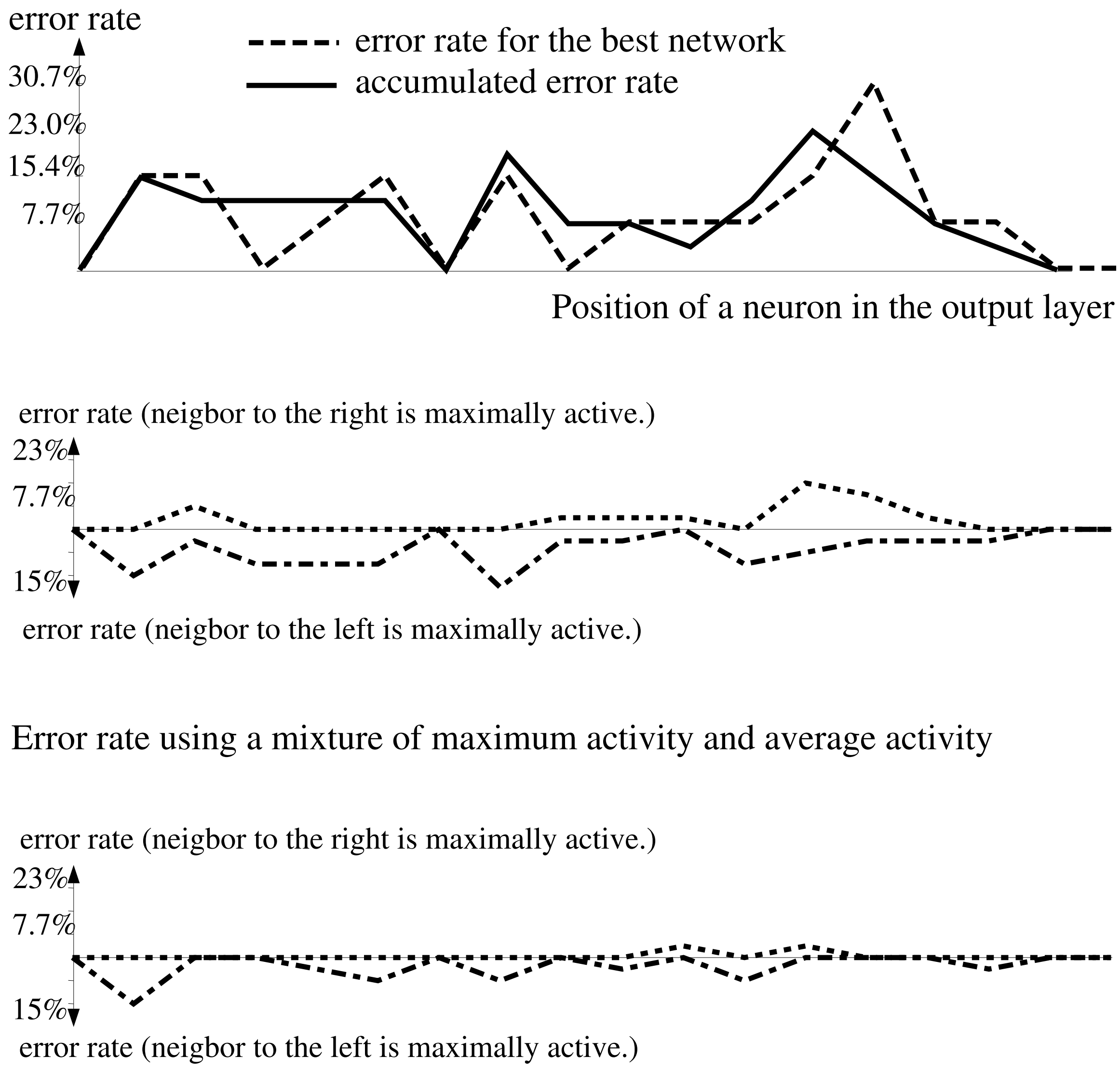

Figure 7.10: The results for the networks with self-organizing internal representations.

The generated neuron activity in the output layer was always shaped similarly to a gaussian activity. Starting at one end of this shape it is possible to combine the neuron activities additively until the middle of total activity is reached. The location of that neuron shall be called average of activity. In more than half of the error cases the average of activity in the output layer encoded the location where the peak should have been. Up to 33% of all errors can be called edge effects. They occurred for an outlying input in possible eye positions. The networks will simulate a perfect transformation if a combination of average and maximum activation in the output layer locates the target position.

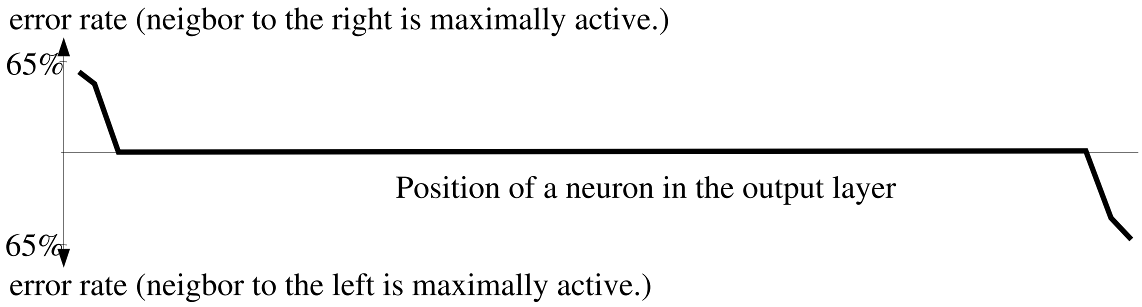

Figure 7.11: The results for the network with a constructed internal representation.

A greater number of possible input combinations can be simulated when the internal representations are constructed into the internal layers. These networks consisted of 2625 neurons and 129,255 connections. Horizontal auditory space consisted of 70 different stimuli. They were simulated by a gaussian activity in a map of 70 auditory neurons. The tonic input map encoded 35 different eye position. In the output layer 35 neurons represented the horizontal motor error coordinate system. The network generated the correct peak of activity in the output layer for 93.25% of the tests. All of the errors were due to edge effects in the output layer. Of the 35 output neurons, 31 neurons never created an error. Only the two leftmost and the two rightmost neurons on the one-dimensional output map are responsible for an incorrect peak of activity in the output layer. The two extreme neurons have the highest error rate: in 60% of the tests they did not respond correctly. Without these four neurons the network is perfect for all test trials. It transforms the auditory input into an oculocentric motor error signal.

Chapter 8

Discussion

This study addresses a very specific problem for a distinct brain structure. A neural mechanism transforms an auditory stimulus from his head-centered perception into an oculocentric coordinate system in the superior colliculus (SC). However, the model of this transformation process presents a first approach to solving a much wider, repeatedly occurring problem: How does the brain manage to map different sensory inputs into one representation of the environment? It must align different perceptual coordinate systems in order to control muscles with a single command signal. An answer to this question contributes to solving a universal puzzle in the field of cognitive science. Cognitive systems combine different inputs from different sources of information. How do they use contradictory information and still yield an optimal response? Why does response time improve when more information is present? The systems must process and merge more information. Thus reaction time should increase.

In the SC the very specific problem is complicated by a developmental process. It is not sufficient to merely construct a neural mechanism to accomplish the coordinate transformation. A constructed neural circuitry would be useless if sensory inputs changed minimally, just as a computer program would crash if a single bit was flipped. However, the brain is capable of adapting to significant damage to sensory organs. Numerous neurophysiological experiments report the developmental emergence of orientation selectivity and binocular interactions in the visual cortex (Singer et al., 1977; Hubel and Wiesel, 1970; Blakemore and Cooper, 1970). In these experiments edge detecting neurons compensate monocular occlusion and adapt to a controlled environment that has exclusively horizontal edges. In the optical tectum, the alignment of auditory and visual maps compensates ear plugs, which produce abnormal binaural cues (Knudsen et al., 1984a). In the SC auditory receptive fields are shaped through visual experience (Withington-Wray et al., 1990b). The coordinate transformation modeled above develops through experience of the environment. Therefore, the task is to create a neural circuitry which can be trained to realize this transformation just as a cognitive system would, by learning from examples and adapting to the environment.

The neural network presented here learns the transformation of a head-centered auditory stimulus into a oculocentric coordinate system. Information about eye position influences this transformation. Every combination of eye position signal and auditory stimulus is mapped onto an unique pattern of activity in a layer of internal representations. This activity is then associated to an output signal in retinotopic coordinates.

Summary and Conlusion

- In the neural network each combination of auditory stimulus and eye position signal is mapped into a unique representation in an internal layer.

- The horizontal and vertical components of auditory space can be treated separatly to minimize the complexity of the mapping.

- Eye position signal is filtered by an on-center-off-surround receptive field in order to transform the frequency code into a place code.

- The association between the internal layer and the output layer can be learned, and it realizes the transformation from one coordinate system into the other.

- The output signal for the transformation process must be distributed.

- It was impossible to associate auditory input, visual input, and eye position signal in a two layer network.

- The tonic input signal for eye position is difficult to process in a neural network. If not transferred from its original frequency code, it will alter internal representations severely.

Figure 8.1: An overview of the most important results of this thesis.

Two different network versions exist. In the first one the internal representation self-organizes throughout the process of learning. The layer of internal representations is divided into different clusters of competitive neurons where each eye position activates another cluster optimally. A soft competition enhances the response properties to auditory stimuli within the clusters. Then the synaptic weights to auditory input neurons are modified. With local Hebbian learning, the cluster neurons develop a selectivity for different locations in auditory space. The resulting internal representations resemble a topographic map of auditory space.

In the second network version, the projection from auditory input to the layer of internal representations is fixed. It is designed to create a unique internal representation for each input combination. These networks are not as adaptive as the networks with a self-organized internal layer. However, on a computer they can be implemented with less synaptic connections. Thus, the information is processed more efficiently.

A distributed activation of neurons in the deeper layer of the SC initiates a saccadic eye movement. It is more the "center of gravity" or center of activation than a localized neuron which constitutes the motor error signal in the SC. This activity, the "output signal" of the SC, excites the premotor areas of the brain stem. Interestingly after learning the output layer always responds with a distributed neural code in accordance with experimental findings (McIlwain, 1976). The association between the internal layer and the output layer in the network requires an output signal with a distributed neuron activation. Due to the great overlap of internal representations, it proved impossible to associate them with a localized output pattern such as a grandmother cell activity. However, a gaussian pattern of neuron activity in the output layer reflected the overlap perfectly. In the simulations the network developed the association between internal representations and output signal in repeated learning trials. In each learning trial an input combination activated an internal representation. Then, the neurons in the output layer were trained to respond to this internal representation. It is this association which realizes the coordinate transformation.

Additionally, the network can account for the neurophysiological results reported by Stein (Meredith et al., 1987; Stein et al., 1989) and psychophysical findings by Lueck (Lueck et al., 1990). Those researchers found that if the auditory and visual input contradict each other, their responses in the SC will be inhibited. If the stimuli from both modalities originate from an identical cause they will enhance the same response in the SC.

A two layer network could not establish the correct association between input stimuli and output signal. Input combinations must map into the internal representations. Back propagation networks also produce internal representations, but in these networks the responses in the internal or hidden layer are obscure. Back propagation compresses the information to an extent where it is impossible to understand the system. Hidden layer neurons function without explanation. For the transformation task the internal representations must provide a coding scheme where each input combination is translated into a unique pattern. Similar patterns in the internal layer must activate similar patterns in the output layer. With these characteristics an internal representation sets up the pre-conditions for a two layer association.

In the simulations one class of input neurons varies its activity linearly with eye position in orbit. For each neuron in this class the activity is computed by a different linear function of eye position. The neurons model the responses of those cells in the parietal cortex which react exclusively to eye position signals. The majority of cells in the parietal cortex combine the tonic background activity for eye position with a visual stimulus. Such a representation of information is not restricted to the parietal cortex. A tonic discharge activity may represent the eye position signal in other brain structures. In addition, it may even encode the feedback signal from the muscles for use by the brain in general.

A set of patterns was constructed where each pattern encoded one eye position. Each pattern consisted of the activation of neurons with different slopes and offsets but it was normalized in the Maximum or Manhattan distance norm short L1. The sum of neural activity was equal for all patterns. For every activation resulting from one combination of slope and offset, the result of the "inverse" slope and offset can be included. No matter which eye position the neurons encode, this pair of activation always has the same amount of excitation.

Unsupervised learning and soft competition among the classifying neurons produced a network in which different input patterns could not be distinguished from one another. Self-organization was unable to develop feature detectors for intermediate eye positions. In another set of patterns, each pattern was normalized in L2, the Euclidean norm. Even then, no selectivity for each eye position developed in the competing neurons. Neither simple Hebbian learning nor a learning rule which normalized the weights of a neuron after each learning trial (Kohonen, 1982) produces feature detectors for each eye position. Instead, only two classes developed among the classifying neurons: either the neuron responded to the one extreme eye position or to the other, but it never specialized to detect a medium eye position.

A principal component analysis of the input data showed only two positive eigenvalues of the correlation matrix. This indicates that the eye position patterns are difficult to classify. Thus, in a neural network approach, the tonic neuron activity creates a difficult input pattern. The frequency of firing contains actual information, which is much harder to process than the usual place coding scheme. It is especially difficult to combine both, frequency and place coding schemes. If a learning paradigm develops internal combinations of these two coding schemes unrestrainedly, the combinations will reflect the linear slope of the tonic input. In an unsupervised learning paradigm such as feature mapping, the linear slope of activity has severe consequences: a network is unable to develop feature detectors for intermediate tonic activities. In a supervised learning paradigm such as back propagation, the tonic input produces planar spatial gain fields in the hidden layers. In both examples an internal representation develops for the combination of inputs and is shaped similarly by the tonic input.

What is encouraging about these results is that they are compatible with the "planar gain fields" which Andersen and Zipser (1988) reported in the parietal cortex and in the back propagation simulations. A neuron selective to an intermediate eye position would not have a "planar gain field"; it would have a gain field with an intermediate maximum response. Therefore, this combination is not specific to neural network simulations. It was also reported for neurons in the parietal cortex (Andersen and Gnadt, 1989).

8.1 Outlook for further research

The proposed network for coordinate transformation requires a great number of neurons, which could cause problems given larger systems to simulate. In the layer of internal representations, the number of required neurons grows with possible combinations of eye position and auditory stimulus. If more than two stimuli were merged the amount of input combinations would result in a combinatory explosion of the number of internal layer neurons. A hierarchy of merging internal layers would have to be organized in a tree structure. However, such an organization lacks biological foundation. The approach has not yet been investigated.

Two alternative extensions of the general concept presented here decrease the number of neurons in the layer of internal representation. The output layer may produce an error signal for feedback. Additional synaptic weights must then propagate the error signal to the internal layer. This error signal influences the self-organization process in the layer of internal representations, and indicates when an internal representation must be re-organized. The signal maximizes the overlap between internal representations associated with the same output signal, and enhances differences between uncorrelated internal representations. Thus, the error signal minimizes the required number of neurons for internal representations.

This extension produces a network related to those trained with the associative reward-penalty learning rule (Barto, 1985; Barto and Jordan, 1987). However, the learning proposed here merely addresses a subset of the problems which reinforcement learning can solve. It groups internal representations associated with the same output signal together, and separates them from internal representations with a different output signal. Also, the distributed activity in the output layer represents a very special class of all possible output signals. An advantage of the approach presented is that the internal representation still bears an intelligible and understandable meaning.

The use of dendritic trees in the network would also reduce the number of neurons in the layer of internal representations. Interactions on the dendrites of a neuron process information. The impact of a particular synapse is determined by the location on the dendrite to which a synapse connects. In the last few years researchers have concentrated on these "sub-processes" with encouraging results. Dendritic tree have even been applied successfully to the SC (Groh and Sparks, 1992). In the present network the tonic input might shunt inhibition or excitation. A synapse processing the eye position signal must connect on a dendritic tree closer to the soma than synapses for auditory inputs. This way eye position signals will enhance or depress auditory inputs nonlinearly. A branch on the dendrite of an output neuron might replace a competitive cluster in the internal layer.

A very interesting result from experimental data supports this concept of "sub-processes" on the dendrite. Meredith and Stein (1986b) report nonlinear multisensory interactions in the SC. They speculate that these interactions are "multiplicative in nature". In deeper layer cells an auditory stimulus can enhance responses to visual stimuli by 1200%. In general the impact of simultaneous multisensory stimuli was significantly greater than the sum of responses for each stimulus. Such a multiplicative influence of stimuli strongly indicates an information processing must be taking place on the dendrite.

The neurophysiological basis of coordinate transformation in the SC has not yet be found. There are many unanswered questions, such as the unknown origin of auditory input to the SC. Auditory cortex and inferior colliculus both project to the SC and both respond to auditory stimuli. But the signal from auditory cortex (Meredith and Clemo, 1989) is diffuse. Meredith and Clemo hypothesize that it merely aids in focusing activity in the SC. Groh and Sparks (1992) argue against the inferior colliculus as a major source of auditory input. Response latencies to auditory stimuli in the inferior colliculus and in the SC differ significantly and make this pathway implausible. Another question is whether the distinction between horizontal cues and vertical cues for auditory stimuli still exists by the time the transformation takes place. Also how does the eye position influence the auditory input to the SC? The experiments of Jay and Sparks (1987a,1987b) show that auditory receptive fields shift with eye position. But they do not reveal how the impact of eye position manifests itself in the auditory pathway.

Bibliography

Andersen, R. A., Bracewell, R., Barash, S., Gnadt, J. W., and Fogassi, L. (1990). Eye position effects on visual, memory, and saccade-related activity in areas lip and 7a of macaque. *The Journal of Neuroscience*, 10:1176–1196.

Andersen, R. A., Essick, G., and Siegel, R. (1985). Encoding of spatial location by posterior parietal neurons. *Science (Washington, DC)*, 230:456–458.

Andersen, R. A. and Gnadt, J. W. (1989). Posterior parietal cortex. In Wurtz, R. and Goldberg, M., editors, *The neurobiology of saccadic eye movements.* Elsevier Science Publisher B.V.

Andersen, R. A. and Zipser, D. (1988). The role of the posterior parietal cortex in coordinate transformations for visual-motor integration. *Canadian Journal of Physiol. Pharmacol.*, 66:488–501.

Anderson, C. and van Essen, D. (1987). Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proccedings of the National Academic Society USA*, 84:6297–6301.

Anderson, J. A., Silverstein, J. W., Ritz, S. A., and Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: some applications of a neural model. *Psychological Review*, 84:413–451.

Barto, A. (1985). Learning by statistical cooperation of self-interested neuron-like computing elements. *Human Neurobiology*, 4:229 – 256.

Barto, A. (1989). From chemotaxis to cooperativity: Abstract exercises in neuronal learning strategies. In Durbin, R., Miall, C., and Mitchinson, G., editors, *The computing neuron.* Addison-Wesley Publishing Company.

Barto, A. and Jordan, M. I. (1987). Gradient following without back-propagation in layered networks. In Caudill, M. and Butler, C., editors, *IEEE First International Conference on Neural Networks volume II (San Diego 1987)*, pages 629 – 636. IEEE.

Bear, M. F., Cooper, L. N., and Ebner, F. F. (1987). A physiological basis for a theory of synapse modification. *Science (Washington, DC)*, 237:42–48.

Blakemore, C. and Cooper, G. (1970). Development of the brain depends on the visual environment. *Nature (Lond.)*, 228:477–478.

Clothiaux, E. E., Bear, M. F., and Cooper, L. N. (1991). Synaptic plasticity in visual cortex: Comparison of theory with experiment. *Journal of Neurophysiology*, 66:1785–1803.

Crick, F. (1986). Certain aspects of the anatomy and physiology of the cerebral cortex. In McClelland, J., Rumelhardt, D., and the PDP Research Group, editors, *Parallel Distributed Processing: Exploration in the Microstructure of Cognition 2: Psychological and Biological Models*, page 371. (MIT Press/Bradford Books).

Draeger, U. C. and Hubel, D. H. (1975). Responses to visual stimulation and relationship between visual, auditory, and somatosensory inputs in mouse superior colliculus. *Journal of Neurophysiology*, 55:690–713.

Fukushima, K. (1975). Cognitron: A self-organized multilayered neural network. *Biological Cybernetics*, 20:121–136.

Galetti, C. and Battaglini, P. (1989). Gaze-dependent visual neurons in the area v3a of monkey prestriate cortex. *Journal of Neuroscience*, 9:1112–1125.

Goodman, S. J. and Andersen, R. A. (1989). Microstimulation of a neural-network model for visually guided saccades. *Journal of Cognitive Neuroscience*, 1(4):317–326.

Gray, C., Koenig, P., Engel, A., and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338:334–336.

Groh, J. M. and Sparks, D. L. (1992). Two models for transforming auditory signals from head-centered to eye centered coordinates. In Press Biological Cybernetics: Running title: Models for coordinate transformations.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I.parallel development and coding of neural feature detectors. *Biological Cybernetics*, 23:121–134.

Grossberg, S. (1988). *The adaptive brain I,II.* Elsevier Science Publisher B.V.

Harris, L., Blakemore, B., and Donaghy, M. (1980). Integration of visual and auditory space in the mammalian superior colliculus. *Nature (Lond.)*, 288:56–59.

Hartline, P., King, A., Kurylo, D., Northmore, D., and Vimal, R. (1989). Effects of voluntary changes in eye positions on auditory localisation and auditory spatial representation in cat superior collculus. *Invest. Ophthalmol. Visual Sci. (Suppl.)*, 30:181.

Hebb, D. O. (1949). *The organization of behavior.* Wiley, New York.

Hecht-Nielson, R. (1987). Counterpropagation networks. *Applied Optics*, 26:4979–4984.

Hecht-Nielson, R. (1988). Applications of counterpropagation networks. *Neural Networks*, 1:131–139.

Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the theory of neural computation.* Addison-Wesley Publishing Componay, Redwood City, California.

Hirsch, J. A., Chan, J. C., and Yin, T. C. (1985). Responses of neurons in the cat's superior colliculus to acoustic stimuli. i. monaural and binaural response properties. *Journal of Neurophysiology*, 53(3):726–745.

Huang, W. and Lippmann, R. (1988). Neural net and traditional classifiers. In Anderson, D., editor, *Neural information processing systems (Denver 1987)*, pages 387–396. American Institute of Physics.

Hubel, D. and Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (Lond.)*, 160:106–154.

Hubel, D. and Wiesel, T. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology (Lond.)*, 206:419 – 436.

Jay, M. F. and Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, 309:345–347.

Jay, M. F. and Sparks, D. L. (1987a). Sensorimotor integration in the primate superior colliculus. a) motor convergence. *Journal of Neurophysiology*, 57(1):22–34.

Jay, M. F. and Sparks, D. L. (1987b). Sensorimotor integration in the primate superior colliculus. b) coordinates of auditory signals. *Journal of Neurophysiology*, 57(1):35–55.

King, A. J., Hutchings, M. E., Moore, D. R., and Blakemore, C. (1988). Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature*, 332:73–76.

King, A. J. and Moore, D. R. (1991). Plasticity of auditory maps in the brain. *Trends in Neuroscience*, 14(1):31–37.

King, A. J. and Palmer, A. R. (1983). Cells responsive to free-field auditory stimuli in guinea-pig superior colliculus: Distribution and response properties. *Journal of Physiology*, 342:361–381.

King, A. J. and Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, 60:492–500.

Knudsen, E. I. and Brainard, M. S. (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. *Science*, 253(5):85–87.

Knudsen, E. I., Esterly, S. D., and Knudsen, P. F. (1984a). Monaural occlusion alters sound localization during a sensitive period in the barn owl. *Journal of Neuroscience*, 4(4):1001–1011.

Knudsen, E. I., Knudsen, P., and Esterly, S. (1984b). A critical period for the recovery of sound localization accuracy following monaural occlusion in the barn owl. *Journal of Neuroscience*, 4:1012–1020.

Knudsen, E. I. and Konishi, M. (1979). Mechanisms of sound localization in the barn owl. *Journal of Comparative Physiology*, 133:13–21.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69.

Konishi, M. (1986). Centrally synthesized maps of sensory space. *Trends in Neuroscience*, 4:163–168.

Lal, R. and Freedlander, M. (1989). Gating of retinal transmission by afferent eye position and movement signals. *Science*, 243:93–96.

LeCun, Y., B.Boser, Denker, J., Henderson, D., Howard, R., and Jackel, L. (1990). Back-propagation applied to to handwritten zip code recognition. *Neural Computation*, 1:540–566.

Lee, C., Rohrer, W. H., and Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*, 332:357–360.

Lueck, C., Crawford, T., and Kennard, C. (1990). Auditory-visual interaction in the generation of saccades in man. *Experimental Brain Research*, 82:149–157.

Marr, D. and Poggio, T. (1976). Cooperative computation of stereo vision. *Science*, 194:283–287.

Mazzoni, P., Andersen, R. A., and Jordan, M. I. (1991). A more biologically plausible learning rule for neural networks. *Proccedings of the National Academic Society USA*, 88:4433–4437.

McIlwain, J. T. (1976). Large receptive fields and spatial transformations in the visual system. *International Review of Physiology, Neurophysiology II*, 10:223–248.

McIlwain, J. T. (1991). Distributed spatial coding in the superior colliculus: a review. *Visual Neuroscience*, 6:3–13.

Meredith, M. and Clemo, H. (1989). Auditory cortical projections from the anterior ectosylvian sulcus (field aes) to the superior colliculus in the cat: An anatomical and electrophysiological study. *Journal of Comparative Neurology*, 288:687–707.

Meredith, M. A., Nemitz, J. E., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. i. temporal factors. *Journal of Neuroscience*, 7(10):3215–3229.

Meredith, M. A. and Stein, B. E. (1986a). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, 365:857–873.

Meredith, M. A. and Stein, B. E. (1986b). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 56(3):640.

Meredith, M. A. and Stein, B. E. (1990). The visuotopic component of the multisensory map in the deep laminae of the cat superior colliculus. *Journal of Neuroscience*, 10(11):3727–3742.

Ottes, F., Gisbergen, J. V., and Eggermont, J. (1986). Visuomotor fields of the superior colliculus: a quantitative model. *Vision Research*, 26:857–874.

Poeppel, E. (1973). Comment on visual system's view of acoustic space. *Nature*, 243:231.

Poeppel, E. and Logothetis, N. (1986). Neuronal oscillations in the human brain. *Naturwissenschaften*, 73:267–268.

Pouget, A., Fischer, S. A., and Sejnowski, T. J. (1992). Hierarchical transformations of space in the visual system. In Moody, J. E., Hanson, S. J., and Lippmann, R. P., editors, *Advances in Neural Information Processing Systems 4*. Morgan-Kauffmann Publishers, San Mateo, CA.

Rumelhardt, D. and McClelland, J. (1986). On learning the past tenses of english verbs. In McClelland, J., Rumelhardt, D., and the PDP Research Group, editors, *Parallel Distributed Processing: Exploration in the Microstructure of Cognition 2: Psychological and Biological Models.* (MIT Press/Bradford Books).

Rumelhardt, D., McClelland, J., and the PDP Research Group (1986). *Parallel Distributed Processing: Exploration in the Microstructure of Cognition 1: Foundations.* (MIT Press/Bradford Books).

Rumelhardt, D. and Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science*, 9:75–112.

Sanger, T. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2:459–473.

Singer, W., Rauschecker, J., and Werth, R. (1977). The effect of monocular exposure to temporal contrasts on ocular dominance in kittens. *Brain Research*, 134:568–572.

Sparks, D. L. and Hartwich-Young, R. (1989). The deeper layers of the superior colliculus. In Wurtz, R. and Goldberg, M., editors, *The neurobiology of saccadic eye movements.* Elsevier Science Publisher B.V.

Stein, B. E., Meredith, M. A., Huneycutt, W., and McDade, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, 1:12–24.

Trehub, A. (1987). Visual-cognitive neural networks. In Arbib, M. A. and Hanson, A., editors, *Vision, brain, and cooperative computation.* MIT Press, Cambridge, MA.

Trotter, Y., Celebrini, S., Thorpe, S., and Imbert, M. (1991). Modulation of stereoscopic processing in primate visual cortex v1 by the distance of fixation. *Society of Neuroscience Abstracts, New-Orleans.*

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14:85–100.

Weyland, T. and Malpeli, J. (1990). Responses of neurons in primary visual cortex are influenced by eye position. *Society of Neuroscience Abstracts, St. Louis*, page 419.7.

Widrow, B. and Hoff, M. E. (1960). Adaptive switching circuits. 1960 IRE WESCON Convention Record, New York: IRE, pp. 96-104.

Wise, L. and Irvine, D. (1983). Auditory response properties of neurons in the deep layers of cat superior colliculus. *Journal of Neurophysiology*, 49(3):674–685.

Withington-Wray, D. J., Binns, K. E., Dhanjal, S. S., Brickley, S. G., and Keating, M. J. (1990a). The maturation of the superior collicular map of auditory space in the guinea pig is disrupted by developmental auditory deprivation. *European Journal of Neuroscience*, 2:693–703.

Withington-Wray, D. J., Binns, K. E., and Keating, M. J. (1990b). The developmental emergence of a map of auditory space in the superior colliculus of the guinea pig. *Developmental Brain Research*, 51:225–236.

Withington-Wray, D. J., Binns, K. E., and Keating, M. J. (1990c). A four-day period of bimodality auditory and visual experience is sufficient to permit normal emergence of the map of auditory space in the guinea pig superior colliculus. *Neuroscience Letters*, 116:280–286.

Withington-Wray, D. J., Binns, K. E., and Keating, M. J. (1990d). The maturation of the superior colliculus map of auditory space in the guinea pig is disrupted by developmental visual deprivation. *European Journal of Neuroscience*, 2:682–692.

Withington-Wray, D. J., Binns, K. E., and Keating, M. J. (1991). The normal emergence of the superior collicular map of auditory space in the guinea pig following developmental removal of visual cortex. *Neuroscience Letters*, 125:209–211.

Zipser, D. and Andersen, R. A. (1988a). A back-propagation network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331:679–684.

Zipser, D. and Andersen, R. A. (1988b). The role of the teacher in learning-based models of parietal area 7a. *Brain Research Bulletin*, 2:505–512.