Modeling Perceptual Learning with Multiple Interacting Elements: A Neural Network Model Describing Early Visual Perceptual Learning

RENANA PERES AND SHAUL HOCHSTEIN
Center for Neural Computation, and Neurobiology Department, Hebrew University, Jerusalem 91904, Israel

Received April 20, 1994; Revised August 15, 1994; Accepted (in revised form) August 25, 1994.

Action Editor: Christof Koch

Abstract. We introduce a neural network model of an early visual cortical area, in order to understand better results of psychophysical experiments concerning perceptual learning during odd element (pop-out) detection tasks (Ahissar and Hochstein, 1993, 1994a).

The model describes a network, composed of orientation selective units, arranged in a hypercolumn structure, with receptive field properties modeled from real monkey neurons. Odd element detection is a final pattern of activity with one (or a few) salient units active. The learning algorithm used was the Associative reward-penalty (Ar-p) algorithm of reinforcement learning (Barto and Anandan, 1985), following physiological data indicating the role of supervision in cortical plasticity.

Simulations show that network performance improves dramatically as the weights of inter-unit connections reach a balance between lateral iso-orientation inhibition, and facilitation from neighboring neurons with different preferred orientations. The network is able to learn even from chance performance, and in the presence of a large amount of noise in the response function. As additional tests of the model, we conducted experiments with human subjects in order to examine learning strategy and test model predictions.

1 Introduction

1.1 Modeling Learning and Visual Perception

When we say learning, we usually mean gaining knowledge, or acquiring new skills and general strategies by practice or experience. Learning includes any change in performance due to experience. Learning occurs at all levels of perception and information processing in the brain (Ahissar and Hochstein, 1994b). It is observed at all stages of development, and continues to be important in adults.

Learning processes have been intensively studied from the psychophysical, physiological, and neural network points of view. Using psychophysical techniques, many experiments have been conducted investigating learning, and in particular early visual perceptual learning, the subject of this study. Over the past 20 years, experiments in early perceptual learning were conducted using a variety of simple tasks. These included hyperacuity tasks such as vernier acuity (Poggio, Fahlke, and Edelman, 1992), discrimination tasks such as discrimination of motion directions (Ball and Sekuler, 1982, 1987), texture segregation (Karni and Sagi, 1991), or global orientation (Ahissar and Hochstein, 1993); and detection tasks such as detection of complex wave gratings (Fiorentini and Berardi, 1980), or of object features (Ahissar and Hochstein, 1993). Performance in all these tasks was subject to substantial improvement. A main result of these studies was that learning is specific to the trained stimuli, namely that learning effects do not transfer across stimuli. Since specificity implies a minimal involvement of top-down mechanisms, which is typical to early areas, these findings indicate that learning in these cases occurs at early processing stages.

Internal representation and learning were investigated also in physiological studies. A huge amount of research has been performed concerning the internal representation of sensory stimuli and their encoding mechanisms in the brain. Each visual cortical area contains its own variety of basic representations, from V1 which codes orientation, color, motion direction, spatial and temporal frequency, binocularity, disparity, etc. (Hubel and Wiesel, 1962, 1968), up to MT, IT and other higher areas which code complex features such as
global motion, and form of complex objects and wholes (see Van Essen, 1985, for review). These higher areas receive input from other modalities and from brain states, which do not relate directly to the retinal image. Physiological evidence for long term plasticity related to visual stimuli in adult animals is found even in the most primary areas such as V1. The change is in the specific features of the cells, such as orientation preference (Creutzfeldt and Heggland, 1975), ocular dominance (Singer et al., 1982), and receptive field size (Chino et al., 1992; Pettet and Gilbert, 1992).

In studies of neural networks and artificial intelligence, learning is a central subject. Numerous models describing various kinds of brain and machine learning have been suggested (Amit, 1989; Sejnowski and Rosenberg, 1987; Hertz, Krogh and Palmer, 1991; Gabriel and Moore, 1990). Learning in neural networks can be defined as a dynamical process, in which the organization of the network (the structure of its connectivity) is modified on the basis of the interaction of the network with external stimuli (Amit, 1989).

Studies in perception and learning in psychophysics, physiology, and neural networks, have much in common with respect to the issues and problems they deal with, and represent different approaches to a common goal. Thus, when building neural networks which mimic brain functions, one must consider constraints on the nature of the representation and the learning imposed by physiological and psychophysical data. This study emphasizes the importance of using an interdisciplinary approach, combining psychophysiological data, and neural network modeling for gaining insights into various brain functions.

1.2 Experimental Paradigm

The study of object identification and perceptual learning is often carried out using visual search tasks, in which a subject has to look for a single target element which differs greatly from a relatively homogeneous background of distractors (non-target items). The target "pops-out" effortlessly from the stimulus, and subjects' reaction time is independent of the number of distractors. This suggests that search is performed at a pre-attentive stage of perception, in which the processing is parallel, and there is little involvement of attentional mechanisms (Treisman and Gelade, 1980; Bergen and Julesz, 1983a, b). Targets that do not "pop-out" are detected in a later, attentive stage, in which the processing is serial, and selective attention shifts between sites in the stimulus (Julesz 1984a, b; Treisman, 1986; Treisman et al., 1990).

Exploration of the underlying representation of the simple, pre-attentively performed tasks may be facilitated by their being "low level," i.e. close to the known physical input stage, and a level of processing for which there is accumulated physiological data and understanding. Thus a behavior-to-neural network correlation may be attempted.

The groundwork for the model presented here is a set of experiments in visual psychophysics in the "pop-out" task (Ahissar and Hochstein 1993, 1994a). The modeled task was to find an odd element (target) in a matrix of tilted line segments (distractors). The stimuli used were arrays of oriented lines, which contained an odd element in one of the sites on half of the trials (Fig. 1a). Following the stimulus a mask was presented to the subjects to limit processing time. The stimulus-to-mask time interval (Stimulus Onset Asynchrony—SOA) was 16–183 milliseconds. The temporal sequence of each trial is illustrated in Fig. 1b. Following each presentation of the stimulation sequence (fixation cross, stimulus, mask), the subject had to press a response key, to indicate whether he or she saw an odd element in the stimulus. Correct responses of the subjects were confirmed by a computer tone. Subjects were presented with these stimuli for 1400 trials in each session, and they participated in 5–10 sessions. For each session, the percent correct as a function of SOA was calculated. (A detailed description of the experimental paradigm is found in Ahissar and Hochstein, 1993).

The basic observation obtained from the experiments is that performance level increases with SOA. When the mask is presented a long time after the stimulus, performance level is high, but for short SOAs it drops to chance. An important result of the experiments is that with practice, all subjects improve in performing the task, as measured by the reduction of the stimulus-mask SOA to produce a fixed performance accuracy; that is, subjects can reach a high performance level even for short SOAs.

In this study, we introduce a neural network model of an early visual cortical area, which exhibits behavior similar to that of human subjects as described above: 1) the network is able to detect an odd element in the array, 2) its performance level depends on the SOA, showing psychometric curves similar to those of human performance, 3) practice leads to improvement in performing the task. The pop-out task is performed by the network in the spirit of the ideas of Koch
and Ullman (1985) concerning pre-attentive processing stages, based on early visual processing, known from physiological recordings.

2 Model Description

2.1 Network Architecture

As mentioned above, the stimuli in Ahissar and Hochstein’s experiments consist of tilted line segments (Fig. 1a). Usually they used arrays of $5 \times 6$ elements. The model network is composed of orientation selective units, analogous to orientation selective neurons of early visual areas (van Essen, 1985), which form a retinotopic representation of the array. Each site in the stimulus, in which a single line segment is located, is processed by four units, each with preferred orientation of one of four possible orientations of 0°, 45°, 90°, 135°, -1/-, respectively. Thus, the four units in each site represent a hypercolumn structure as described in Hubel and Wiesel (1962, 1968).

The network architecture is summarized in Fig. 2.

When a stimulus is presented to the network, each unit is activated according to its tuning curve (response function), which depends on the angle between its preferred orientation and the orientation of the line segment in its receptive field (Fig. 3). For implementing a physiologically related response function we used data from extracellular recordings on alert Rhesus Macaque monkeys, carried out in our laboratory (Zohary, 1991).

The data, summarized in Fig. 3, show the spike count of a cell in the primary visual cortex (V1) when different orientations were presented for time periods of 100–500 ms. The fit function, which is also presented in the figure, was chosen to be periodic, with a cycle of 180°, to have a maximum in the preferred orientation of the unit, and to be symmetric around its peak. The spontaneous activity is described by a constant term (independent of orientation), which depends only on the measurement time.

The response function, satisfying the above requirements, which we chose as the fit to the data presented in Fig. 3 is:

$$I(o, o', \tau) = 0.02 \cdot \tau + 0.1 \cdot \tau$$
$$\times \left[0.5(\cos(2(o - o')) + 1)\right]^3$$
$$= 0.02 \cdot \tau + 0.1 \cdot \tau \cdot f^3$$

where $I$ is the response (spike count), $o'$ is the preferred orientation of the unit ($o' = -15^\circ$ in the cell of Fig. 3), $o$ is the orientation of the line segment in the unit’s receptive field, and $\tau$ is the duration of exposure to the stimulus and the measurement time in milliseconds. The term $0.02 \cdot \tau$ reflects the spontaneous firing of the neuron, present even without stimulation.

The function (1) represents the average value of the response. In fact, the response varies from trial to trial due to noise. Its standard deviation is represented in Fig. 3 by the error bars. The relative noise (standard deviation/mean response) decreases as the response increases. The relations (as modeled from Fig. 3 data) are:

$$D = 2 \cdot I^{0.5}$$

where $D$ is the mean standard deviation (noise), and $I$ is the mean activation. Hence, when a stimulus is presented to the system, the response of every unit in the initial state (time step 0) is:

$$S(T = 0) = I + \text{noise}(\tau)$$

where,

$T$ = the processing time step of the network units,
stimulus

response
\( (T=0) \)

updating according to the activation function (4).

possible final configurations

odd element detected

no odd element detected

**Fig. 4.** Stimulus processing in the network. When the stimulus is presented \((T=0)\), the units respond according to the response function (3). Then, \((1 \leq T \leq 10)\), all the units are updated simultaneously according to (4). At \(T=0\) the activation values are analog, between 0 and 1 (plus noise); on the next time steps they become digital. The 'answer' of the network to the stimulus is determined by the final state: if all the units are 'off', the system did not detect an odd element, and if one (or more) units are 'on', we say that an odd element was detected by the network.

out detection of Koch and Ullman (1985), who proposed that the representation of the odd element remains salient after processing. Following this idea we say that if, at the final time step, all the units in the network are 'off' \((S_{ij0} = 0\) for every \(i, j,\) and \(o\)), the network did not detect an odd element in the stimulus. If, on the other hand, one (or more) units remain 'on' \((S_{ij0} = 1\) for at least one unit \(i, j,\) and \(o\)), we say that an odd element was found by the network. Since we cannot tell whether the subjects saw the odd element at the site where it was presented, and since subjects sometimes report perceiving more than one odd element, we consider any salience in the final activity configuration as pop-out detection.

Figure 4 summarizes the information processing in the network, during a whole trial (similar to the psychophysical experiments, we term such sequence, from the exposure to the stimulus to the decision making point, as one trial).
2.3 Generating Psychometric Curves

For finite SOAs, the 'noise' term in (3) is non-zero. Therefore, a unit that should have received high activation according to the response function, may receive low activation (because of the noise), and, for certain weight configurations, be silenced by inhibition from other units; similarly, units that should have received low activation according to the tuning curve (1) may receive high activation, and remain salient. As a result, we receive the four answer possibilities, that is, Hits, Misses, False Alarms, and Correct Rejections. Since relative noise increases with the decrease of the SOA, we expect high performance for long SOAs, that degrades to chance performance for low SOAs, where the signal to noise ratio is small. Thus, assuming that the weight matrix was properly chosen, either by hardwiring, or as a consequence of the learning process, we achieve psychometric curves with 50% correct responses for short SOAs, and 100% correct performance at long SOAs, similar to those of human subjects.

Figure 5 illustrates the psychometric curve generated by the model network, compared to the curve of a human subject. We see that the shape of the psychometric curve of the model simulates well that of real performance. The two curves differ in their SOA values, with network SOAs being longer. This is explained by the fact that the model assumes one neuron per orientation column. In reality, there are more, so noise would be lower by $\sqrt{N}$, where $N$ is the number of neurons in an orientation column whose combined activity contributes to the effective noise level. It is seen that according to the model, $N$ has an order of magnitude of a few dozen neurons (see Newsome et al., 1989, for supportive physiological evidence from MT).

3 Learning in the Model

Our next step was to implement a learning algorithm into the model. We sought an algorithm which provides maximum correspondence with psychophysical experiments as well as with physiological data.

In the psychophysical experiments, correct responses of the subjects were confirmed by a computer tone (Ahissar and Hochstein, 1993). Accordingly, the learning algorithm used was an algorithm of supervised learning, that is, top-down elements were implemented in the network. In addition, there is firm anatomical and physiological evidence showing the involvement of top-down mechanisms in cortical plasticity, and indicating that cortical plasticity is supervised, in the sense that the retinal input is not sufficient to induce plasticity, and there is a need of an additional, extraretinal input to induce long term changes in the properties of the cells (Kasamatsu and Pettigrew, 1976; Buisseret et al., 1978; Freeman and Bonds, 1979; Singer and Rauschecker, 1982). This input can be in the form of directly applied current (Fregnac et al., 1988; Shulz and Fregnac, 1992; Fregnac and Shulz, 1994), in the form of neuromodulators (Kasamatsu et al., 1979; Greuel et al., 1988) or eye movements and proprioceptive inputs (Freeman and Bonds, 1979; Buisseret and Singer, 1983). Note that human subjects in the experiment are
able to improve even without an external feedback signal. Instead, they develop an internal feeling as to the correctness of their answer on the basis of trials with long SOAs. In this study we deal only with the case of an external teacher.

As mentioned above, the only feedback that subjects receive while performing the task, is a computer tone, confirming a correct response. No information is given about the exact location of the odd element in the presented array, or about the orientations presented in the stimuli (target and distractor orientations are not changed during the experiment, but most subjects are not aware of this). Therefore, we look for an algorithm, in which the system is not provided with the desired output configuration of every unit, but is just informed whether the answer was right or wrong.

In order to enable supervised learning in the model, we add a summing element, which is able to recognize any salience in the final configuration of the network, and to compare it with the external feedback. This unit is called the ‘teacher.’ It is possible to model the teacher as a binary threshold unit, but since its physiological character is not clear, we leave it as an ‘oracle’ which reads the configuration of the network, and sends the proper reinforcement (reward or penalty) signal.

3.1 The Learning Algorithm

The learning algorithm that we chose is an algorithm of reinforcement learning (Widrow et al., 1973; Barto and Anandan, 1985; Barto and Jordan, 1987; Ackley and Littman, 1990; Hertz, Krogh and Palmer, 1991). In this type of supervised learning, the teacher (environment) does not provide the system with the exact desired output configuration. Instead, it gives a reinforcement signal \( r \), telling whether the system was right or wrong.

The weight modification is done according to the Associative reward-penalty (Ar-p) algorithm. This concept is originally due to Barto and Anandan (1985), but here, the presentation of the algorithm from Hertz, Krogh and Palmer (1991) is discussed.

According to the Ar-p algorithm used in the model, the weight modification is given by:

\[
\Delta w_{ij}^{\mu} = \begin{cases} 
\eta^+ [S_{ij}^{\mu} - \langle S_{ij}^{\mu} \rangle] S_{ij}^{\mu} f_{r^+}(T = 0) \\
\eta^- [1 - S_{ij}^{\mu} - \langle S_{ij}^{\mu} \rangle] S_{ij}^{\mu} f_{r^-}(T = 0) 
\end{cases}
\]

where we denote by \( S_{ij}^{\mu} \) the final activation (at the end of a trial) of the unit in site \( i, j \) with preferred orientation \( \mu \) whose activity at \( T = 0 \) was \( S_{ij}^{\mu}(T = 0) \), in response to a particular stimulus pattern \( \mu \).

Determining \( \langle S_{ij}^{\mu} \rangle \) can be done by calculating, or, as in our system, by averaging over a number of trials in a fixed weight configuration. That is, in every learning step, the system calculated \( S_{ij}^{\mu} \) by performing several runs without weight modification.

The weight modification depends on the pre- and post-synaptic activities. The term in the square brackets is the error term. The network compares the average postsynaptic activity \( \langle S_{ij}^{\mu} \rangle \) to the correct answer. Since the exact desired output configuration is not available to the network, it assumes that in case of reward, all the units are correct, (and uses \( S_{ij}^{\mu} \) as the correct answer), and that in the case of penalty, all the units are responsible for the wrong answer (and uses \( 1 - S_{ij}^{\mu} \) as the correct answer).

\( \eta^+ \) and \( \eta^- \) are constants, known as the learning rates. In our system, their values are set to be \( \eta^+ = 1.0 \) and \( \eta^- = 0.01 \), so the major weight changes are done when the network makes the right decision. Note, that if the system reaches its optimal weight configuration, so that \( S_{ij}^{\mu} \approx \langle S_{ij}^{\mu} \rangle \), the weight changes become very small, and a steady state is reached. The major changes occur when the system is usually wrong, but due to its random dynamics, reaches a correct answer; then, both \( [S_{ij}^{\mu} - \langle S_{ij}^{\mu} \rangle] \) and \( \eta \) are large.

3.2 Learning Simulations

A learning simulation in our system consists of stages of weight modification according to the Ar-p algorithm, and of stages in which the performance level is checked, and a psychometric curve, similar to those of Fig. 5, is created. In human subjects, the learning and performance checks are, of course, done together, but in the model we use the advantage of separating learning from testing, and achieve better understanding of the effect of the weight change on performance. As done with the subjects of Ahissar and Hochstein (1993), the network is trained on stimuli with fixed orientations of target and distractors. In the network, the odd element is always in the same location. Thus, during a simulation we teach the network to distinguish between two patterns: one with, and the other without odd element (\( \mu = 1, 2 \)). In each learning step, the stimulus and the SOA for the trial are determined randomly, so learning is done for both types of stimuli, for all the relevant SOAs (100, 200 ... 1500 ms).
During the learning simulation (which includes between 50 and 1500 weight modification stages, or learning steps), the performance level of the network is checked at fixed intervals (every 100 steps, for example). As mentioned above, during these testing modes, no weight modification occurs.

In general, each unit has its own weight matrix, connecting the unit with other units in the eight nearest and next nearest neighboring hypercolumns. Thus, the learning process is an optimization problem of many (4320) free parameters. Since we are interested in following the significant changes during learning, and in the development of global connectivity patterns during the learning process, we use a technique known as weight sharing, in which all the units share the same set of weights (Fukushima, 1980; Rumelhart, Hinton and Williams, 1986; LeCun et al., 1989a, b, 1990; Nowlan and Hinton, 1992).

In the current version of the model, the weight matrix \( w_{ij} \) is identical for every site \( i, j \) and orientation \( o \). It is invariant to location and orientation and depends only on the angle and the relative position between units. The weight matrix \( w \) is shown schematically in Fig. 6.

By using weight sharing we reduce significantly the number of weights in the network, in order to simplify and shorten computation time, and increase the ability to distinguish the meaningful weight change during learning, and to characterize the parameter space. However, none of the conclusions and results depends on the exact form of weight sharing, and whenever it becomes qualitatively important we remove the degeneration in the appropriate module. The weight change \( \Delta w^{ij}_{f, o} \) is taken to be the mean of the change \( \Delta w^{ij}_{f, o} \) that was computed for each unit.

The initial weight configuration in each simulation is determined by introducing a random bias to a primary connectivity. This primary connectivity allows for the presence of lateral iso-orientation inhibition. Assuming lateral iso-orientation inhibition is frequently done when modeling the pop-out task (e.g. Koch and Ullman, 1985), and also has an electrophysiological basis (Blakemore and Tobin, 1972; Maffei and Fiorentini, 1976; Fries et al., 1977; Nelson and Frost, 1978; Van
Fig. 7. Learning in human subjects (a-b) is modeled by the network (c-d). (a, c): The learning curves; the SOA for 82% correct performance (threshold) vs. learning step or session number, according to Quick’s psychometric function. The threshold in (c) is initially 964 ms, 580 ms after 20 learning steps, and 225 ms after 100 steps, indicating that the major improvement (50% of the way to asymptote) occurred between step 0 and step 20 of learning. (b) Psychometric curves in the first, second, and last session in a learning experiment of a human subject. Target orientation is 15°, distractor orientation is 45°. (d) Psychometric curves from a learning simulation in the model before learning, after 20 learning steps and after 100 learning steps. The target orientation is 45° and distractor orientation is 90°. In both human and model behavior, training induces a leftward shift and a steepening of the psychometric curves, substantially decreasing the threshold SOA (a, c).

Essen et al., 1989; Kneirim and Van Essen, 1992). Random numbers with a Gaussian distribution are added to (or subtracted from) every weight of the primary connectivity. The standard deviation of the distribution is proportional to the value of the weight in the primary connectivity.

3.3 Results from Learning Simulations

Simulation results are illustrated in Figs. 7, 8, and 9. Fig. 7 describes an example of learning in the model (7c, 7d), compared with learning of a human subject in a similar task (7a, 7b). It is seen, that the network is able to improve its performance. As with human subjects, the improvement is relatively fast, and a stable state of good performance is reached.

Figures 7c, d describe learning from an initial weight configuration which generates psychometric curves with high percent correct performance at long SOAs. Such curves are characteristic of human behavior. However, the network learning ability is much stronger, and it is able to improve also from initial weight configurations which generate curves of chance (50% correct) performance for all SOAs, using the Ar-p algorithm (5). Human psychometric curves are never flat, since even “naive” subjects show performance which is well above chance for very long SOAs. Experiments have not been conducted to test if humans would learn from trials where their performance is chance.

A typical example for network learning from a flat curve is illustrated in Figs. 8, 9. Figure 8 describes the
Fig. 8. Learning from chance performance, during 1000 learning steps. Target orientation is 45° and distractor orientation is 135° (difference of 90° between target and distractors). The psychometric curves (top), and the learning curve (bottom), show that the improvement from chance performance is a sudden transition.

improvement of performance. Note that the initial performance level was a flat curve at 50% correct, and that in the first 100 learning steps performance improved dramatically to reach 100% correct performance for long SOAs (Fig. 8, top). The network remains in this state of good performance.

The same result is seen in the learning curve (Fig. 8 bottom), showing the SOA of 82% correct performance (threshold) vs. session according to Quick psychometric function (Quick, 1974). The transfer from flat psychometric curve to a psychometric curve with 100% correct performance for long SOAs takes place during a small number of steps (see also Fig. 7, where the major improvement takes place in the first 20 steps of the learning). The sudden transition from flat to non-flat psychometric curve is typical also of the rest of the cases examined, and is an important characteristic of the learning in the system.

When examining the weight modification during learning, we see that while performance improves, the weights also converge to an asymptotic value. Figure 9 illustrates the weight modification during learning in the simulation of Fig. 8. It is seen, that the self excitation ($w^{(s)}_{j,j}$) decreases by about 13% (from 15.1 to 13.2). The lateral inhibition between units with the same preferred orientation ($w^{(v)}_{j,j'}$) greatly increases to 260%, from −11.7 to −31.0, corresponding to strong lateral inhibition being a condition for correct answers for stimuli without odd element. We also see an increase in the connections for $\sigma - \sigma' = \pm 90^\circ$ (Fig. 9d), from −8.1 to −0.3. This increase is due to the units with preferred orientation equal to that of the distractors which interact with the unit with preferred orientation of the target at the target site (the angle between target and distractors is 90° in the simulation of Figs. 8, 9). The weights for $\sigma - \sigma' = \pm 45^\circ$ do not change much, because no elements with orientation difference of 45° are present in the stimuli, and the response values $S_{j,\sigma}(T = 0)$ are small. Thus, the system reaches a balance between excitation, desired
for detecting the odd element, and inhibition, desired for correct response in the cases when there is no odd element in the stimulus. We have performed a multitude of simulations, and found that there are numerous weight configurations of high performance, each with a different balance between excitatory and inhibitory connections. An interesting set of configurations contains those with large facilitory connections from units...
with the preferred orientation of the distractors to units with the preferred orientation of the target.

4 Testing Model Predictions

According to the version of the Ar-p algorithm used in the model (Eq. 5), the weight modification is done according to the error term, in which the average postsynaptic response of every unit is compared to the correct output of the unit, as evaluated by the network. Since the network does not know the correct answer, it assumes that in case of correct answer, all the units were right, and the average output is compared to $S_{ij}$, and in case of mistake, the network assumes that all the units were wrong, and the correct answer is taken to be $1 - S_{ij}$ for all the units. However, it may be that the mistake was due to only a few units, (and not to all of them), which remained ‘on’ ($S_{ij} = 1$), in the case where there was no odd element in the stimulus, so the configuration $1 - S_{ij}$ for every $i, j, o$ can be sometimes remote from the real correct final configuration of the network. In such cases, the weight modification will not be in the optimal direction for improving performance. It follows, that a network with initial connectivity which is excitatory, and whose mistakes are mainly False Alarms, will show slower learning than a network whose mistakes are mainly Misses, and whose connectivity is more inhibitory.

First, we checked that the model really displays this disparate behavior when biased towards False Alarms or Misses. Starting from a network with steady state performance, we corrupted the weights by making

Fig. 10. Model learning dependence on type of errors. Starting with steady state model weights, all weights were changed to make them more excitatory or more inhibitory, in order to increase the mistakes of the False Alarms or the Misses types, respectively. The graphs show two pairs of examples of learning simulations. Initial and final performance were similar for each pair, but learning is faster for the case of Misses than when the mistakes are False Alarms.
them all either more excitatory or more inhibitory, so that the mistakes were all False Alarms or Misses, respectively. Two examples of results are shown in Fig. 10. Initial performance was similar for either change, but learning was faster in the cases of more inhibitory weights (squares), than for more excitatory weights (False Alarms). Both converged eventually to the same level of performance.

We then checked, whether human subjects also use such strategy when learning the task. If, indeed, a similar learning strategy is used by the subjects in the experiment, subjects with tendency to make more False Alarms will show slower/poorer learning than subjects whose mistakes are mainly Misses. Therefore, comparison between groups of subjects with different distributions of Misses and False Alarms will supply us with useful information about the learning strategies used by human subjects in this experiment.

4.1 Methods

10 subjects participated in the experiment, and they were presented with stimuli consisting of a $7 \times 7$ array of line segments, with target orientation at $15^\circ$ and distractor orientation of $45^\circ$. Each subject participated in 5–6 sessions on subsequent days, until reaching the asymptote, according to the standard experimental paradigm in the experiments of Ahissar and Hochstein (1993), described in the introduction. The subjects were divided into two groups: one group (liberals) with the tendency to make False Alarms, and the other (conservatives) with the tendency to make Misses. We generated this bias in the decision making strategy of the subjects through the instructions they received. An instruction such as: 'Press response key "1" only if you are sure that there was an odd element, otherwise press response key "0"; leads to Misses and Correct Rejections, while the instruction 'Press response key "1" only if you are sure that the array of lines is uniform in orientation, otherwise press response key "0", leads to many Hits and False Alarms. (Note that the percent correct will not necessarily be different between these two groups).

4.2 Results

In each session, the subjects were rank-ordered according to their performance in the session, 1 for the best threshold, 10 for the worst. Then, the mean rank for each group was computed in each session. The mean rank vs. session is shown in Fig. 11. It is clearly seen that both groups started with similar average ranking, (the initial threshold values, measured in the first third of the first session of both groups is the same (mean of 124.5 ms for conservatives, 125.3 ms for liberals)), and that the conservatives moved to gain the better ranking. A statistical paired $T$ test for the second and third sessions confirmed that the conservative group is significantly better ($p < 0.01$) since its members have better ranking than those of the liberal group, though the initial ranking for the first session is not statistically different. Of course, eventually, the liberals catch up and final asymptotic performance is again the same (final mean threshold values are 37.8 ms for conservatives, 35.4 ms for liberals).

The results imply, that the learning strategy of the subjects resembles the learning strategy of the network when using the Ar-p algorithm, in the sense, that False Alarms are less informative and contribute less to learning than other kinds of answers. Thus, the learning algorithm that we chose to describe the learning in the network corresponds to the strategy which is used by the subjects.

5 Discussion

We presented a biologically plausible neural network model of an early cortical visual area for modeling perceptual learning in odd-element detection tasks. The network, consisting of orientation selective detection tasks. The network, consisting of orientation selective units analogous to orientation selective neurons of early areas...
in the visual cortex, shows psychometric curves similar to those of human behavior, and simulates well real performance. The network improves in performing the task using the Ar-p algorithm of reinforcement learning. Thus, we succeed in implementing a learning algorithm of supervised learning, as implied by physiological results concerning cortical plasticity, but without the need of a teacher which knows the desired output configuration.

The network is able to learn even from chance performance and in the presence of a large amount of noise in the response function. The main changes in the weight matrix during learning, in the cases that we have studied, were lateral inhibition among units with preferred orientation of the distractors, and facilitation of the unit with preferred orientation of the target at the target site, by neighboring hypercolumn units with preferred orientation of the distractors. The transition between 'poor performance' and 'high performance' weight space regions, takes place very suddenly, in a way reminiscent of a first order phase transition.

As additional tests of the model, we conducted experiments with human subjects in order to examine the learning strategy and test model predictions. We found, that an algorithm such as the version of the Ar-p algorithm presented in (5), for which most of the useful information is in trials resulting in Hits, Misses, and Correct Rejections (and not in False Alarms), suits the general strategy used by human subjects during the learning of the task.

The present work is far from exhausting the potential of the interplay between theory and experiment in the study of performance of this task: on the experimental physiological side, guided by the theoretical results, one can look for the existence of this iso-orientation lateral inhibition, and cross-orientation facilitation. On the theoretical side, the dynamics of the learning process, as a directed walk in the weight space can be further studied with specific concentration on the drastic changes taking place at the boundary of the regions in the weight space of different performance level. In particular one can check dynamical phase transition aspects and other universal properties.

Acknowledgments

We thank Professor Hanoch Gutfreund for his assistance and valuable ideas throughout the period of this study.

We thank Merav Ahissar for collaboration in this theoretical study of her experimental data, and in assistance in performing the additional psychophysical experiments detailed here.

We thank Ehud Zohary for permission to use his physiological data as a basis for model parameters.

We thank Prof. Sorin Solomon and Orna Rosenthal for helpful discussions and comments.

This study was supported by grants from the US-Israel Binational Science Foundation (BSF), the basic research fund of the Israel Academy of Sciences and Humanities, and the Hebrew University Center for Neural Computation.

Notes

1. The exact value of the three coefficients in (1), which were determined by least squares fit to the data of Fig. 3 are (0.02, 0.09, 2.75), instead of (0.02, 0.1, 3) of equation (1). However, trial simulations showed that network performance was not sensitive to the exact value of the coefficients. We also fit the above function (1) to a single neuron example, and there is no reason to assume that these coefficients are identical for all neurons. Therefore, for simplification, we rounded out the parameters and chose the average response function as indicated in (1).

2. Since \( \tau \) includes the retinal after-image duration, \( \tau = \text{SOA} \).

3. The exact fit relations are \( D = 1.94 \cdot \tau^{0.44} \), where again, from the same considerations as in footnote 1, we rounded out the parameter values.

4. Note that we added a row and a column of units to the borders of the 5 \( \times \) 6 array. This was done in order to avoid artificial edge effects from the edge units which receive activation from less than 8 neighbors. The activation of those units \( i = 0, 6, j = 0, 7 \) have no influence in determining the final decision of the system about the presence of odd element, nor to the weight modification during learning.

5. For a given trial, there are four possible answers of the network: Hit, Miss, False Alarm, and Correct Rejection. Hit and Correct Rejection are the correct answers to the cases in which an odd element is, or is not presented in the stimulus, respectively. Miss relates to the case in which the subject did not detect the odd element, though it was presented, and False Alarm is the case where subject reports the presence of odd element, but no odd element was present in the stimulus.

6. The reason that we chose a rank test (Lehmann, 1975), is that a main characteristic of the data is the large variability in the initial performance level between subjects, from 39 to 160 ms. The initial level affects dramatically the learning ability: subjects who began with low threshold (good performance), reach the asymptote already in the first session, while the amount of improvement required from subjects with poor initial performance is much larger. As a result, the definition of learning becomes ambiguous, and various definitions may lead to a bias towards a specific group when interpreting the results. In addition, considering the numerical threshold values is greatly influenced by extreme cases. Therefore, a statistical measure which is not based on the thresholds' numerical values is used.
References


