

Reprinted from

COGNITIVE BRAIN RESEARCH

Cognitive Brain Research 2 (1995) 155-163

Research report

Learning to ignore: psychophysics and computational modeling of fast learning of direction in noisy motion stimuli

Lucia M. Vaina^{a,*}, V. Sundareswaran^a, John G. Harris^b

^a *Intelligent Systems Laboratory, College of Engineering, 44 Cummington Street, Boston University, Boston, MA 02215, USA*

^b *University of Florida, Gainesville, FL 32611, USA*



ELSEVIER

Research report

Learning to ignore: psychophysics and computational modeling of fast learning of direction in noisy motion stimuli

Lucia M. Vaina ^{a,*}, V. Sundareswaran^a, John G. Harris^b

^a Intelligent Systems Laboratory, College of Engineering, 44 Cummington Street, Boston University, Boston, MA 02215, USA

^b University of Florida, Gainesville, FL 32611, USA

Accepted 16 August 1994

Abstract

The effects of practice on the discrimination of direction of motion in briefly presented noisy dynamic random dot patterns are investigated in several forced-choice psychophysical tasks. We found that the percentage of correct responses on any specific task increases linearly with repetition of trials within roughly 200 trials from about chance to a performance of 90% or better. The level of performance remained constant or improved over several days, and in most instances it did not transfer when stimulus parameters changed. We used a modified Radial Basis Function (RBF) representation to model the psychophysical tasks. The performance of the model is functionally similar to the psychophysical results. We propose a Hebbian learning algorithm which deactivates the inputs from neurons responding to motion noise in the stimulus. Our computational model suggests that to solve this task in biological systems, neurons (perhaps in MT) improve their performance by 'learning to ignore' noise in the image.

Keywords: Global motion; Direction discrimination; Middle temporal area; Perceptual learning-psychophysics; Neural network; Hebbian learning model; Radial basis function

1. Introduction

To be a good candidate for modeling, a phenomenon should be robust and well-characterized psychophysically, and have a likely candidate for its neural substructure [26]. These conditions are met by the problem addressed in this study: perception of direction in spatially complex moving patterns. In psychophysics, such patterns are usually represented as dynamic random dot patterns in which a specifiable proportion of the dots provides a correlated motion signal that is spatially dispersed amidst masking motion noise. Since an efficient extraction of the motion signal from such stimuli involves considerable spatial integration, neurons in the macaque middle temporal area (MT or V5) – a region in the extrastriate visual cortex concerned with the perception of motion – are particu-

larly sensitive to such stimuli (e.g. [10]). Psychophysical studies in humans have characterized the limits of spatial and temporal integration in such stimuli [24], and the nature of the underlying motion computations [21].

Since the psychophysical and neural substrate of these global stimuli are fairly well-understood, we were interested to see whether the perception of direction in such motion stimuli can improve with practice. Studies of other early perceptual tasks have shown that improvements of performance obtained in the first experimental session are preserved in a subsequent session and retained over weeks. This is considered as perceptual learning [5]. An important limitation characterizes perceptual learning: the beneficial effects of practice are lost if certain stimulus parameters such as orientation, or location in the visual field are changed. Based on the time scale necessary for the improvement to occur, two major paradigms have been defined in perceptual learning: slow, progressive learning (several

* Corresponding author. Fax: (1) (617) 353-6766.

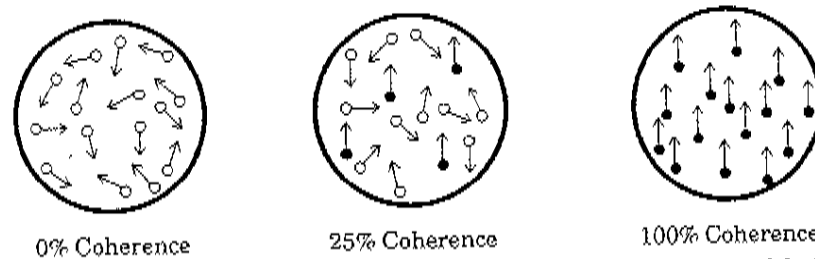


Fig. 1. Schematic drawing of the stochastic stimulus used in the experiments. The stimulus is an adaptation of the Newsome and Pare's stimulus [10] for measuring motion coherence thresholds. The center panel depicts 25% correlation which corresponds to the signal level used in the psychophysical learning tasks.

thousand trials are required to reach stable performance) and fast learning (improvement occurs and stabilizes in the first 100–200 trials). We designed specific psychophysical tasks to study whether fast learning occurs in a certain class of global motion stimuli and to explore the limits of this learning.

The idea of fast learning and the nature of its limits is attractive from a computational point of view because it encourages the exploration of practice-dependent plasticity found in the early visual system [4]. A recent line of research in biologically motivated learning models put forth by Poggio [11] proposes perceptual learning as evidence that "the brain may be able to synthesize – possibly in the cortex – appropriate task-specific modules that receive input from retinotopic cells and learn to solve the task after a short training phase in which they are exposed to examples of the task". Consistent with this view, we developed psychophysically motivated neural network models for fast learning of direction in global motion. Similar to Poggio's theoretical prediction illustrated in comparison with psychophysics in a task of learning vernier acuity [12], our results on learning direction discrimination suggest that the fast learning mechanisms lead to the formation of task-specific modules.

In the modeling, our emphasis has been on developing unsupervised learning algorithms. We use a representation with responses of biologically realistic tuned motion detectors rather than optic flow vectors. Responses of sufficiently large number of such detectors with overlapping receptive fields can be used instead of optic flow in many applications such as moving object detection, determining time to collision, and obstacle avoidance. Coupled with the kind of fast learning we demonstrate here, it is possible to develop specific templates which can be quickly trained to perform such tasks.

In Section 2, we describe the psychophysical learning experiments. In Section 3, we introduce an approach to modeling the psychophysical results, and we show the performance of the model in simulations of the psychophysical tests taken by the human observers. The results and the approach are discussed in Section 4.

2. Psychophysics

In a previous study, Ball and Sekuler [1] showed that the discriminability of the direction of motion of two random dot patterns improved with training. For this task, more than 2000 trials appear to be required for reaching a stable performance. Such 'slow' learning time scale has been reported for the learning of other perceptual tasks, such as vernier acuity [8], stereoacuity [2,14] and discrimination of line orientation [23]. In contrast to this 'slow learning', Fiorentini and Berardi [3] showed that for learning the discrimination of complex gratings with two harmonics of different spatial phase, 100–200 trials suffice. Similarly Poggio et al. [12] showed that a small number of trials suffice for significantly improving performance on a vernier task.

Results on the psychophysical experiments summarized here demonstrate fast learning of direction in a global motion task, and explore some of its critical limitations. A more extensive discussion of these experiments and the limits of learning will appear in [20].

2.1. Visual stimuli

We employed a stochastic motion display in which random dots were displayed in a circular aperture 10° in diameter. The dots were 0.03° in diameter and the dot density was 2 dots/deg²; the dots were of high contrast, producing an average luminance of 0.51 cd/m² against a background luminance of 0.24 cd/m² (measured with a Minolta Luminance meter LS-100). From one frame to the next, all the dots were replotted with a fixed spatial offset of 0.1 deg. Twenty five percent of these dots carried the motion signal¹ that is, from the first frame to the second, they were all replotted in a fixed, predetermined direction: left or right (Expt. 1), or up or down (Expt. 2). The remaining 75% of the dots in the display jumped in random directions producing dynamic masking noise.

Each trial lasted 90 ms during which two frames

¹ This motion strength was well above the threshold (18%) determined from 12 naive subjects.

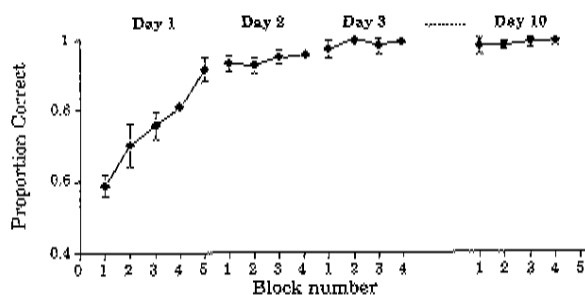


Fig. 2. Proportion of correct discrimination of left vs. right motion during three sessions carried out in three successive days and a fourth session, nine days after the first session. Each point represents the proportion of correct responses relative to a block of 40 trials. The data were pooled together over four subjects. Error bars show one \pm standard deviation of the mean.

were presented (with inter-frame interval equal to zero, and an intertrial interval of 1 s). The observers viewed the display binocularly except when the experimental question addressed interocular transfer in which subjects viewed the display monocularly (with the other eye covered). A session consisted of 4-6 blocks, 40 trials each. Feedback was not provided during the experimental sessions. Prior to the experimental sessions, all the observers took half a block of examples, during which feedback was provided on incorrect answers. Observers were required to maintain fixation on a small white mark placed 2° to the left or right of the imaginary circumference of the stimulus.

2.2. Experiment 1: learning

To investigate the effects of practice and their retention, the discrimination of leftward vs. rightward direction of motion in the display was tested on each experimental session for 3 days and repeated 10 days later. The results are presented in Fig. 2. For most observers, a fast learning effect which saturated within 200 trials or less could be shown. Here we present data from four observers who participated in all the conditions summarized in this study. Learning occurred consistently only during the first blocks, and only when the subjects started from above chance level² ($> 50\%$ correct on the first block). The shape of the learning curve in Fig. 2 proves the hypothesis that learning of motion direction in noise occurs and that the improvement of performance is linear across blocks of trials. The learning achieved on Day 1 was retained by all subjects on Day 2, and it actually continued to improve on Day 3. On Day 10, without any intervening training, the subjects' performance was essentially 100% correct.

² Two subjects started from below-chance and their performance hovered around chance level; they each took 600 trials.

2.3. Experiment 2: direction specificity

Since we wanted to know whether the learning was specific for direction, after learning discrimination of leftward versus rightward motion direction, the subjects were presented with stimuli in which the direction of signal dots was up or down. Fig. 3 shows the comparative results of training on the two conditions. In spite of achieving a high proportion of correct decisions on the horizontal direction of motion discrimination, when the direction of motion was changed to vertical, the subjects' discrimination ability dropped to near chance level. With training, they showed again a gradual improvement and the learning curve is similar to that obtained for discrimination of horizontal directions. This result, consistent with the results reported in other perceptual learning studies, reconfirms that learning is specific to stimulus parameters (direction of motion in this case).

Summary of other experiments testing the limits of the learning. In additional psychophysical learning tasks we showed that (a) the learning is specific to the area of the visual field in which training was performed, (b) there was no transfer when eccentricity was changed from 2° to 5° , or from one visual hemifield to the other, and (c) complete interocular transfer was found (two subjects participated in this experimental condition).

3. Model: learning to ignore

3.1. Related models

Neural network-based learning can be subdivided into supervised or unsupervised. In supervised learning, a set of pattern-outcome pairs is required to train

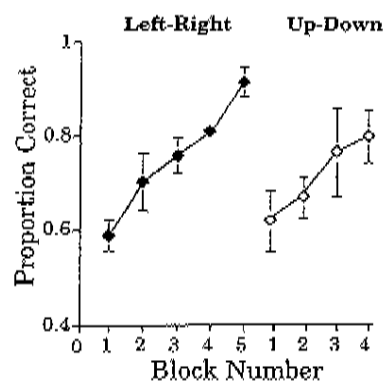


Fig. 3. Learning curves of the four subjects demonstrating the direction specificity of the learning. The left curve shows subjects' performance on the training session of Day 1. The diamond symbols represent in a block of trials (40 trials per block) the average proportion correct responses of four subjects. The graph on the left shows data from left-right direction discrimination; the graph on the right shows results from up-down direction discrimination.

a model. In unsupervised learning, usually no knowledge of the desired outcome is assumed. Backpropagation algorithm is an example of supervised learning. However, this algorithm has been criticized as being 'unbiological' in the sense that it requires (error) information to be propagated backwards along the neuronal connections. Mazzone et al. [7] proposed a more biological alternative based on reinforcement learning. Poggio and Girosi proposed a learning scheme based on the principle that learning from samples is equivalent to interpolation between the samples, using Radial Basis Functions (RBF) [13]; Poggio et al. [12] modeled learning visual hyperacuity using HyperBFs, an extended version of RBFs. All these mechanisms are of the supervised type.

Here we use an RBF approach, with two modifications: the centers of the RBFs are not the samples, but are fixed preferred directions of motion, and the learning mechanism is unsupervised. In our model we use MT-like neurons whose responses are combined by a weighted averaging scheme to determine global motion direction; learning is accomplished by modifying the weights using Hebbian rules. A similar approach applied to learning hyperacuity using basis functions with preferred spatial orientations has been proposed by Weiss et al. [25]. Directionally-tuned neurons, similar to the ones in our model, and a Hebb rule were used by Sereno and Sengco [17] to train neurons to be

sensitive to rotation and dilation like those in the middle superior temporal (MST) region.

3.2. The model

We introduce a computational model that determines direction of global motion given noisy motion vectors. To improve performance over trials, the model uses unsupervised learning and synthesizes a template specific to the stimulus attributes. We use model neurons with properties similar to those in MT. The responses of several neurons (each with a preferred direction of motion) are combined by the model to determine the global motion direction. Only the direction of motion is used in this model.

We consider two types of neurons with different computational properties, to represent the spatial integration performed by large receptive fields in MT: one multiplies localized responses (π -neuron), and another sums over localized responses (σ -neuron). The response of the π -neuron tuned to a direction θ_i is given by

$$x_i = e^{-(1/2\sigma_k^2 \sum_{l=1}^n (\theta_l - \theta_i)^2)}, \quad (1)$$

where σ_k is the standard deviation of the tuning, and information from n moving random dots is integrated. The response of the σ -neuron is expressed as

$$x_i = \sum_{i=1}^n e^{-(1/2\sigma_k^2 (\theta_i - \theta_i)^2)}. \quad (2)$$

Both types of neurons obtain high response to motion in the preferred direction (θ_i), and diminishing responses to directions away from the preferred direction. The π and σ neurons represent two possible ways of integrating motion information over a region.

3.2.1. The architecture

The input pattern consists of the responses x_i of n MT-like neurons, and the associated weights are w_i . The model determines a global motion direction by a weighted combination of the responses of the neurons, as shown schematically in Fig. 4. We treat all vectors to be of constant magnitude. The global motion vector is computed as a weighted combination of the individual neuron's preferred vectors; the weighting is a product of adjustable weights w_i and the neuron responses x_i . If the neuron preferred vector is $v_i = [\cos(\theta_i), \sin(\theta_i)]$, the output vector is calculated as

$$v_0 = \sum_i w_i x_i v_i,$$

or, equivalently, the global motion direction is calculated as

$$\theta_0 = \tan^{-1} \left(\frac{\sum w_i x_i \sin \theta_i}{\sum w_i x_i \cos \theta_i} \right).$$

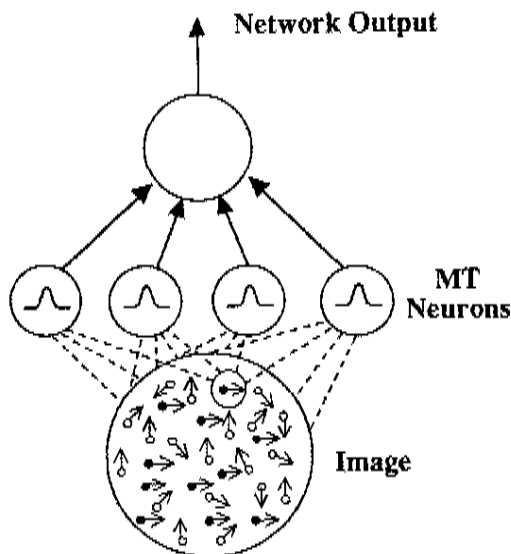


Fig. 4. Network schematic. Simulated MT neurons perform a population coding of the global motion direction. The network output is a weighted average of all the MT neurons. The filled arrows in the circular disk represent signal dots, while the unfilled arrows denote noise.

Clearly, the output direction will be strongly biased towards the preferred direction of a neuron with a high response and a high connection strength (weight). Since symmetric tuning curves are used, and the output is calculated as a linear combination of the basis functions, the model qualifies as an RBF model.

In the model, learning occurs by modification of the weights w_i . Two learning schemes are used: exposure-based, in which weights are adjusted proportional to the current weight, or self-reinforced, in which weights are altered by an amount proportional to the current weight of the neuron and the correlation of the neuron's tuned direction to the output direction calculated by the model. In both cases, we follow Hebb's prescription of increasing the strength of the connection between two neurons one of which consistently activates the other. The model learns to ignore the noisy neurons that do not contribute significantly to the motion direction judgements.

3.2.2. Exposure-based learning

In this scheme, the weight corresponding to a neuron is incremented by an amount proportional to the current weight. Only neurons whose response values are above a certain threshold are allowed to increase their weights; this learning rule favors the neurons that fire consistently. This is in agreement with the Hebb rule because a consistently active input neuron must be contributing to the output, and thereby its connection to the output neuron is strengthened.

The exposure-based learning rule is as follows:

$$w_i \leftarrow w_i + \eta w_i, \text{ if } x_i > r_i, \quad (3)$$

where r_i is a threshold, and η is a small fraction that controls the learning rate. The same learning rule has been used in learning hyperacuity by Weiss et al. [25].

The remarkable aspects of this learning rule are that it is very simple, and local; i.e. no external or internal feedback is necessary. In a forthcoming study [22], we will discuss this implicit learning and present experimental results from psychophysics with human observers, and model simulations. Similarly, Zohary et al. [27] reported that they found neurons in the macaque MT that improved their sensitivity even though the monkey was merely fixating (not doing any task) while being exposed to motion stimuli similar to the ones we have used in our experiments.

3.2.3. Self-reinforced learning

In the self-reinforced learning mechanism, the weight corresponding to a neuron is increased by an amount proportional to the product of the current weight and a decreasing function (exponential) of the angular difference between the calculated global mo-

tion direction and the preferred direction of the neuron:

$$w_i \leftarrow w_i + \eta w_i e^{-(\theta_i - \theta_o)^2 / 2\sigma_i^2}, \text{ if } x_i > r_i. \quad (4)$$

In this case, the model uses its own estimate of the global direction in an internal reinforcement to determine the learning. In fact, for the purpose of learning, the model only requires measures of the discrepancy between the output global motion direction and the individual neuron preferred directions; such a measure may be available even without the knowledge of the precise output global motion direction.

This rule also follows Hebb's rule because the neurons whose weights are increased the most are those neurons which contribute the most to the computed global direction. Neurons whose response is not strong enough to sway the output direction closer to their preferred direction are reduced in importance by a reduction of their weights.

3.3. Experiments

3.3.1. Implementation

The model has been simulated on a workstation. Since we were primarily interested in the underlying learning mechanism, we have not yet implemented a full motion computation algorithm from a raw image sequence. The input (motion) vectors are represented by their angles relative to the positive horizontal axis (i.e. the magnitude is ignored). The responses of the directionally-tuned neurons can be directly computed from the angles (see Eqs. 1 and 2); alternatively, cosine tuning functions were used, and similar results were obtained). For each trial, the coherent motion direction (one of two opposite directions) is randomly decided. In the experiments, σ_h was set to π/n_d , where n_d is the number of neuron preferred directions; the directions θ_i are chosen by uniformly dividing 2π in to n_d parts.

To discriminate between leftward and rightward coherent motion, a simple decision mechanism relies on whether ν_o is in the left half-plane or in the right half-plane. However, this kind of sharp decision-making clearly does not mimic the behavior of human observers performing a 2AFC task. For a more realistic, human-like performance, we use a range of angles around the rightward direction ($-\theta_i < \theta_o < \theta_i$) and around the leftward direction ($\pi - \theta_i < \theta_o < \pi + \theta_i$). If the angle θ_o falls within one of these ranges, the corresponding direction is chosen as the global direction; otherwise, the global direction is decided randomly.

3.3.2. Model justification

Two experiments were done to justify the architecture. The weights were all equal and learning was

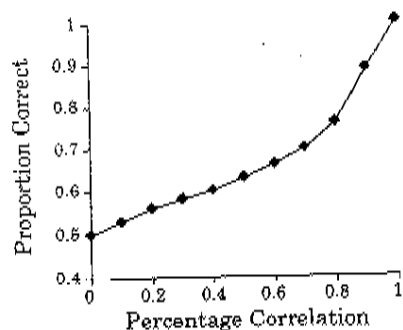


Fig. 5. The graph shows that the network's performance improves linearly with the increase in the proportion of correlated (signal) dots.

disabled (i.e. the weights were not changed). Fig. 5 shows the untrained network response for various levels of correlated signal from 0% correlation (no signal) to 100% correlation (no noise). Each point in the graph represents the performance of 2500 trials of the simulated task using 8 MT neurons with a receptive field sigma of 22.5. The shape of the response is similar to that of real MT-neurons (as described in [10] for example).

Fig. 6 shows improvement in performance as the number of dots is increased, for an untrained network with equal weights. With sufficiently large number of dots, a simple averaging scheme (equal weights) can be expected to perform very well, and this is reflected by this curve.

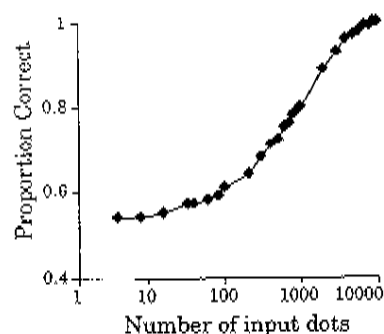


Fig. 6. The graph shows that the performance of the network improves dramatically with the increase in the number of input dots.

3.3.3. Learning

The Hebbian model was implemented with the two different learning rules (exposure-based, and self-reinforced) to modify the weights. Using the learning mechanisms in Eqs. 3, and 4, we ran simulations to study the model's response. Typical learning curves are shown in Fig. 7. The results of a motion vector averaging mechanism (labeled 'avg') and a winner-take-all mechanism ('max') are also shown. The direction judgement of the averaging method is based on the direction of the global average of the motion vectors. The winner-take-all chooses the direction of the neuron with the largest response. These mechanisms are not equipped to learn: the curves should be seen as only providing the statistics of the dataset. Salzman

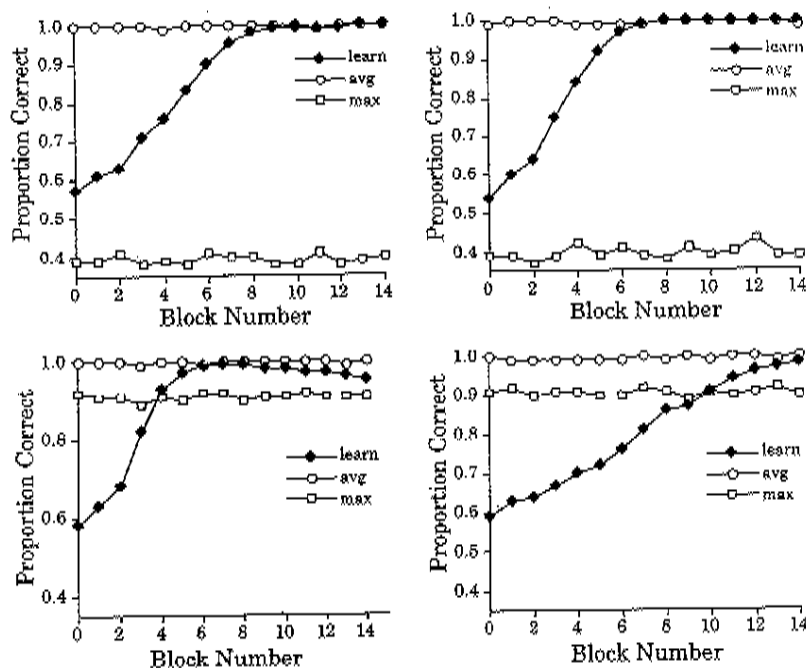


Fig. 7. Network learning curves: on the left, for the exposure-based learning, and on the right, for the self-reinforced learning. The performance of the averaging and the winner-take-all mechanisms are also shown. The top row is for the π -neurons, and the bottom row, for the σ -neurons. All results are obtained by averaging over ten simulations.

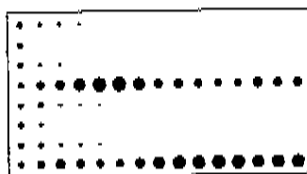


Fig. 8. Pictorial representation of the changing weights; the first column corresponds to the situation before learning (equal weights), the second column for before the second block and so on. As desired, all weights except two go to zero. The exposure-based learning rule was used to train the network with input of π neuron responses.

and Newsome [16] reported evidence that a winner-take-all mechanism (rather than a weighted averaging) mediates judgements by trained monkeys performing a global motion task.

In all the experiments, eight preferred directions were used ($n_d = 8$, with the four cardinal directions and the four oblique directions), each trial containing 40 random dots moving with a correlation of 25% (the same correlation as in the tests with human subjects); each block consisted of 50 trials and the learning rate η was 0.1, except in one case (bottom right in Fig. 7 in which the learning rate had to be kept very low (0.001) due to the average high response of the σ -neurons).

Fig. 8 shows pictorially the change in the weights in a typical simulation in which the network was trained using the exposure-based learning rule. Each weight is represented by a circle whose diameter is proportional to the weight (the weights have been normalized). Each column represents the weights at an instance before a block of trials; the time axis is rightward. As can be seen, most of the weights go rapidly to zero, except the two that carry the signal information to the output.

3.3.4. Transfer

To address computationally an important limit of the learning (stimulus parameter specificity), we carried out simulations in which the model, after improving in performance in the left-right motion task, is

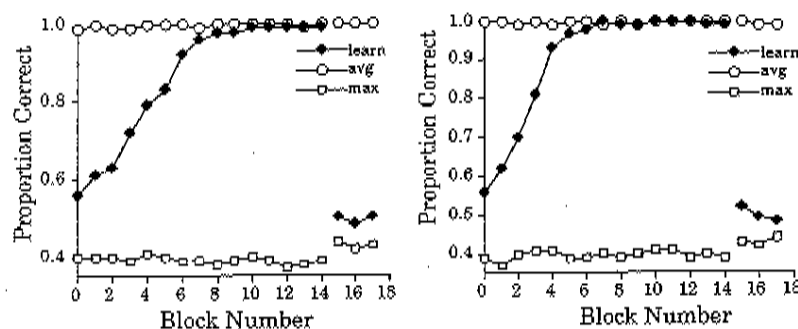


Fig. 9. Network performance showing no transfer between horizontal and vertical motion tasks. The training for the first 15 blocks used a left/right signal, and for the last three blocks, an up/down signal. The performance drops to chance levels, indicating that the learning was specific to stimulus attributes. On the left is the curve for exposure-based learning, and on the right, for self-reinforced learning. This simulation was done using π neurons.

presented with the up-down motion. Consistent with the psychophysical results, there was no transfer of the learning to this new direction (see Fig. 9). In the model, this is to be expected because when the template for one task is 'synthesized', the weights for noisy neurons are reduced to zero; when the task is changed, however, some of these neurons are no longer 'noisy', but instead contain useful (signal) information. For transfer of learning to occur, the corresponding weights must increase; this is not possible when using our learning rules which disfavor a weight from attaining a significant value once the weight has reached a very small value.

4. Discussion

The model proposed here connects psychophysical findings with their putative neurobiological substrate: the goal is to relate the learning of perception of motion direction in stochastic dynamic random dot displays to the computationally relevant characteristics of response patterns of MT neurons in the primate visual cortex. A series of studies of neuronal correlates of the perceptual decision of direction of motion in stochastic motion signals like the stimuli used here, together with studies of performance on these stimuli of monkeys with MT lesions (for a review, Newsome et al. [9]) provide strong support for the hypothesis that perceptual judgements of motion direction are by and large based on the directional signals carried by MT neurons. MT is clearly the first stage in cortical motion analysis where global motion direction is perceived [9].

Our psychophysical studies in human observers clearly demonstrate that perceptual learning of global motion direction occurs. To investigate the possible nature of the essential aspects of this specific type of learning (fast and stimulus parameter-specific), we proposed a model and carried out computer simulations of the same psychophysical experiments performed by the

human observers. The model is Hebbian, and it learns as fast as the human observers do. The model learns to ignore noise in the display and to respond selectively to the direction of signal.

Our current model assumes that MT-like neurons might mediate learning of direction in a global motion task. Recent studies [4,6] demonstrated a surprisingly dynamic nature of receptive fields characteristics and cortical topography in the primary visual cortex, and is very likely that this flexibility is even stronger in the higher visual areas, such as MT. We suggest that MT neurons can undergo functional changes through experience and that these changes will be displaced towards the stimuli paired with the reinforced visual response. We considered two different MT models of spatially integrating local information: summation and product. Summation yields a type of integration that is more tolerant towards noisy information, compared to the product. We proposed two Hebbian learning rules, and demonstrated that the model performs a fast learning of global motion direction.

Another possibility that cannot be ruled out is that the initial network for this psychophysical task automatically wires up a two-neuron style algorithm (the two neurons being tuned to the opposite directions used in the discrimination task). In this case, the fast learning may be due to shrinking the receptive field of the MT neurons with practice. Modification of receptive fields with training has been shown in various situations; for example, neurons responding to repeated tactile stimuli will decrease the size of their receptive fields [15].

Recent experimental studies of Zohary et al. [27] first suggested that MT neurons mediate the motion direction learning, but on further examination, they found that the improvement appears to be transient and they conclude that long-term learning observed in the psychophysical tasks may be mediated by mechanisms outside MT, possibly situated downstream from it [28]. This result is intriguing and will need exploration. The task they used is somewhat different from ours. We allow only two frames of motion, thus truly limiting the problem to a spatial integration, while their stimulus is on for two seconds during which there is a significant amount of temporal integration. They were interested in whether long-lasting improvement in detecting motion direction as found in psychophysical tasks is manifested in the neuronal sensitivity. It is possible that learning is not manifested in a single neuron, but instead in an assembly of neurons. Such an assembly could consist of an integration of the neuronal responses (summation or product), followed by either a weighted averaging similar to the one we proposed, or by a clustering mechanism [19] qualitatively similar to the winner-take-all mechanism favored by Salzman and Newsome [16].

5. Conclusion

Plasticity is a central characteristic of an adaptive system. Engineering an adaptive visual system will undoubtedly be facilitated by an understanding of what visual tasks in the cortex can be learnt, as opposed to being embodied by a specific neural circuitry. In this study, we showed psychophysically that direction in global motion can be learnt, and we proposed a model which learns by ignoring noise in the image and spatially integrates the directional motion signal. Such a model has useful potential applications for computer vision problems, where spatial integration is a strategy used to improve accuracy in the estimation of motion in real images.

We believe that the possibility of using MT-like neurons is perhaps not limited to the learning proposed here. Since optic flow is hard to compute [18], an alternative is to use the responses of MT-like neurons. A sufficiently dense collection of such neurons could provide information similar to optic flow to perform navigational tasks, and to detect moving objects in the scene. In general, a collection of templates that can be quickly trained to learn visual attributes can be harnessed to perform real-world vision tasks.

While significant progress has been made towards understanding the mechanisms and the underlying circuitry of learning, as yet there is no biologically satisfactory computational theory of perceptual learning of direction in noisy motion signals. The physiological data are so far inconclusive. It is not yet known what is the precise problem that the cortex solves when it learns and what specific computations might be involved in solving it. We do not advocate that one or the other of the models sketched here is the biologically correct choice. Nevertheless, the computational models and the psychophysics described here suggest new psychophysical and computational experiments which we are currently conducting in our laboratory [20,22]. We hope that by closely connecting models and psychophysics while keeping in mind neuronal compatibility, we will make progress in understanding how the cortex can so rapidly learn a large variety of perceptual tasks.

Acknowledgements

This research was conducted at the Intelligent Systems Laboratory of Boston University, College of Engineering. L.M.V. and V.S. were supported in part by grants from the Office of Naval Research (N00014-93-1-0381) and the National Institute of Health (EY RO1-07861) to L.M.V. V.S. thanks the Boston University College of Engineering for the support from a Dean's

postdoctoral fellowship. We thank Reynold Comptois for programming assistance.

References

- [1] Ball, K. and Sekuler, R., Direction-specific improvement in motion discrimination, *Vision Res.*, 27 (1987) 953-965.
- [2] Fendick, M. and Westheimer, G., Effects of practice and the separation of test targets on foveal and perifoveal hyperacuity, *Vision Res.*, 23 (1983) 145-150.
- [3] Fiorentini, A. and Berardi, N., Perceptual learning specific for orientation and spatial frequency, *Nature*, 287 (1980) 43-44.
- [4] Fregnac, Y., Shulz, D., Thorpe, S. and Bienstock, E., A cellular analogue of visual cortical plasticity, *Nature*, 333 (1988) 367-370.
- [5] Gibson, E.J., Improvements in perceptual judgement as a function of controlled practice of training, *Psychol. Bull.*, 50 (1953) 402-431.
- [6] Gilbert, C.D. and Wiesel, T.M., Receptive field dynamics in adult primary visual cortex, *Nature*, 356 (1992) 150-152.
- [7] Mazzoni, P., Anderson, R.A. and Jordan, M.I., A more biologically plausible learning rule for neural networks, *Proc. Natl. Acad. Sci. USA*, 88 (1991) 4433-4437.
- [8] McKee, S.P. and Westheimer, G., Improvement in vernier acuity with practice, *Percept. Psychophys.*, 24 (1978) 258-262.
- [9] Newsome, W.T., Britten, K.H. and Movshon, J.A., Neuronal correlates of a perceptual decision, *Nature*, 341 (1989) 52-54.
- [10] Newsome, W.T. and Pare, E.B., A selective impairment of motion perception following lesions of the middle temporal visual area, *J. Neurosci.*, 8 (1988) 2201-2211.
- [11] Poggio, T., A theory of how the brain might work. In *Cold Spring Harbor Symposia on Quantitative Biology*, Cold Spring Harbor Laboratory Press, 1990, pp. 899-910.
- [12] Poggio, T., Fable, M. and Edelman, S., Fast perceptual learning in visual hyperacuity, *Science*, 256 (1992) 1018-1021.
- [13] Poggio, T. and Girosi, F., A theory of networks for approximation and learning. AI Memo 1140, M.I.T., 1989.
- [14] Ramachandran, V.S. and Braddick, O., Orientation-specific learning in stereopsis, *Perception*, 2 (1973) 371-376.
- [15] Recanzone, G.H., Schreiner, C.E. and Merzenich, M.M., Plasticity in the frequency representation in the primary auditory cortex following discrimination training in adult owl monkeys, *J. Neurosci.*, 13 (1993) 87-103.
- [16] Salzman, C.D. and Newsome, W.T., Neural mechanisms for forming a perceptual decision, *Science*, 264 (1994) 231-237.
- [17] Sereno, M.I. and Sereno, M.E., Learning to see rotation and dilation with a hebb rule. In R.P. Lippmann, J.E. Moody and D.S. Touretzky (Eds.), *Advances in Neural Information Processing Systems 3*, Morgan-Kaufman, 1991, pp. 320-326.
- [18] Sundaeswaran, V., *Global methods for image motion analysis*, PhD thesis, Courant Institute, New York University, New York, NY, 1992.
- [19] Sundaeswaran, V. and Vaina, L.M., Learning direction in global motion: two classes of psychophysically-motivated models, Touretzky (Ed.), *Adv. Neur. Inf. Proc.*, Vol. 6, 1994, in press.
- [20] Vaina, L.M., Perceptual learning of direction in global motion: psychophysics, in preparation.
- [21] Vaina, L.M., Grzywacz, N.M. and LeMay, M., Structure from motion with impaired local-speed and global motion-field computations, *Neural Computation*, 2 (1990) 420-435.
- [22] Vaina, L.M. and Sundaeswaran, V., Implicit perceptual learning of direction in stochastic motion displays, in preparation.
- [23] Vogels, R. and Orban, G.A., The effect of practice on the oblique effect in line orientation judgements, *Vision Res.*, 25 (1985) 1679-1687.
- [24] Watamaniuk, S., Sekuler, R. and Williams, D., Direction perception in complex dynamic displays: the integration of direction information, *Vision Res.*, 24 (1984) 55-62.
- [25] Weiss, Y., Edelman, S. and Fahle, M., Models of perceptual learning in vernier hyperacuity, *Neural Computation*, 5 (1993) 695-718.
- [26] Westheimer, G., Neural modeling of a visual perceptual task - a case study, *Seminars in The Neurosciences*, 4 (1992) 75-82.
- [27] Zohary, E., Celebrini, S., Britten, K.H. and Newsome, W.T., Neuronal plasticity that underlies improvement in perceptual performance, *Science*, 263 (1994) 1289-1292.
- [28] Zohary, E. and Newsome, W.T., Perceptual learning in a direction discrimination task is not based upon enhanced neuronal sensitivity in the STS, *Invest. Ophthalmol. Visual Sci.*, Suppl. (1994) 1663.