In Proceedings of ICANN'99, Edinburgh, Scotland, pp. 251-256, 1999

Modeling the self-organization of directional selectivity in the primary visual cortex

Igor Farkaš
Institute of Measurement Science
Slovak Academy of Sciences
842 19 Bratislava, Slovak Republic

Risto Miikkulainen
Department of Computer Sciences
The University of Texas
Austin, TX 78712, USA

Abstract

A model is proposed to demonstrate how neurons in the primary visual cortex could self-organize to represent the direction of motion. The model is based on a temporal extension of the Self-Organizing Map where neurons act as leaky integrators. The map is trained with moving Gaussian inputs, and it develops a retinotopic map with orientation columns that divide into areas of opposite direction selectivity, as found in the visual cortex.

1 Introduction

Since the pioneering research of Hubel and Wiesel [1], neurons in several areas of the primary visual cortex have been known to be selective to both orientation and direction of movement of the input. Although evidence for columnar organization was originally found only for orientation selectivity, later microelectrode studies suggested that direction selectivity is also arranged in a systematic fashion. The details of this organization were only recently mapped out. Optical imaging methods [2, 3] revealed that there is a mosaic-like map of direction preference, which varies smoothly across the map. Between smooth areas there are line-like areas of discontinuity where the direction abruptly changes. The functional maps for orientation and direction preference are closely related: typically, an iso-orientation patch can be divided into regions that exhibit preference to opposite directions, orthogonal to the orientation.

A number of hand-coded models of directional selectivity have been built (such as [4], or see numerous references in [5]), but we are aware of no work that would demonstrate how such selectivity could arise through self-organization like the other response properties of the neurons. In this paper, we present such a model based on an extension of the self-organizing neural network [6] to input sequences.

2 The model

The standard Self-Organizing Map (SOM) [6, 7] forms a mapping from a high-dimensional input space to a discrete grid of units. If successful, the mapping preserves the local similarities in the data. This model can be used to demonstrate how self-organization of the visual cortex could take place [8].

In our adaptation of SOM to visual cortex modeling, a square retina of $R \times R$ receptors projects onto the cortex modeled by SOM, containing $N \times N$ neurons. Every cortex neuron has a receptive field (RF), which is the set of receptors in the retina from which the neuron receives input. Each neuron is assigned a circular region of receptors of diameter s centered on its projection as its RF. Typically, $s \approx \frac{1}{2}R$, and the RFs of neighboring neurons overlap significantly. The RFs of neurons near the boundary are not fully circular but cut from one side (Fig. 1).

The model preserves the main features of the original SOM, that is, the processes of determining the winner, shrinking the neighborhood, and decreasing the learning rate, but uses a different, temporally enhanced model of the neuron. It is very similar to the RSOM model [9], which has been used for time series prediction.

In the model, every neuron (i, j) has a set

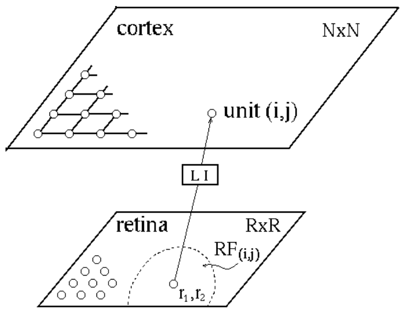

Figure 1: The architecture of the visual cortex model. Every cortex neuron receives inputs from its receptive field in the retina and acts as a leaky integrator.

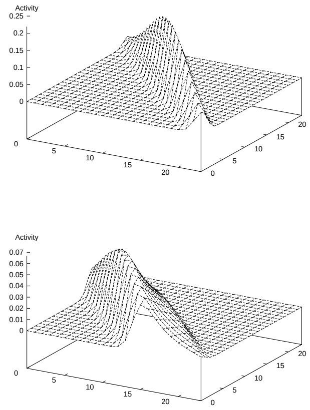

Figure 2: Input activity corresponding to a diagonal direction of motion to the left ($\omega = \pi/4$). *Top:* Initial pattern on the retina (24×24 receptors) at the beginning of the sequence. *Bottom:* Accumulated (leaky-integrated) input that the cortical neurons receive at the end of the sequence (eq. 2).

of parameters, afferent weights μ_{ij,r_1r_2}, corresponding to retinal receptors (r_1, r_2) within its RF. It performs leaky integration of input signals $\xi_{r_1r_2}$. The state s_{ij} of neuron (i,j) at discrete time t is computed as

$$s_{ij}(t) = \gamma \sum_{r_1,r_2} \xi_{r_1r_2}\mu_{ij,r_1r_2} + (1-\gamma)s_{ij}(t-1) \ , \tag{1}$$

where $\gamma \in (0,1)$ is the memory parameter whose value defines the trade-off between the depth and resolution. (For $\gamma = 1$ the state equation reduces to the static, memoryless case.) In view of filtering theory, eq. 1 describes an exponentially weighted IIR filter with the impulse response $h(t) = \gamma(1-\gamma)^t$ (see, e.g. [10]).

The output η_{ij} of neuron (i,j) is defined by standard sigmoid function σ whose nonlinear effect is important for enhancing neuron's the selectivity:

$$\eta_{ij}(t) = \sigma(.) = \frac{1}{1+\exp(-k(s_{ij}(t) - \theta_{ij}(t)))} \ .$$

The k and θ_{ij} are parameters that must be set experimentally. Specifically, if k is high enough, the neuron has the ability to amplify its response to strong stimuli and attenuate it for weaker ones. This effect can be seen as an approximation of lateral connections between units in the map, which in other models serve to sharpen the output responses. The thresholds θ_{ij} must also be properly set to achieve the right amount of activation in the map. The θ_{ij} is updated during training: every neuron remembers its maximum state level $s_{ij}^{max}(t) = \max_{\tau \le t}\{s_{ij}(\tau)\}$, and θ_{ij} is updated at every time step with $\theta_{ij}(t) = \frac{1}{2}s_{ij}^{max}(t)$.

The input to the map consists of sequences of moving normalized Gaussian bars whose direction of motion is always perpendicular to their orientation (Fig. 2). The activity $\xi_{r_1r_2}$ of receptor (r_1, r_2) is given by

$$\xi_{r_1r_2} = \exp\Big(-\frac{(r_{1x_i}\cos\omega - r_{2y_i}\sin\omega)^2}{a^2} - \frac{(r_{1x_i}\sin\omega + r_{2y_i}\cos\omega)^2}{b^2}\Big) \ ,$$

where $r_{1x_i} = r_1 - x_i$, $r_{2y_i} = r_2 - y_i$ are shifted point coordinates, a^2 and b^2 specify the major and minor variances of the Gaussian, $\omega : 0 \le \omega \le \pi$ specifies its orientation, and $(x_i, y_i) : 0 < (x_i, y_i) < R$ specifies its center.

Sequences have a fixed length that covers a part of the retina. They start at randomly chosen positions (x_i, y_i) in the retina and their directions are also randomly chosen. Altogether, 16 possible directions are included corresponding to 8 orientations. The bars move in a constant speed equal to one receptor per time step.

In temporal models of self-organization, the goal is to make neurons sensitive to particular set of sequences, hence sequences running across certain part in the retina and with certain direction of motion. To achieve direction selectivity, it is necessary to look for a *representative winner* [7], which is the neuron (c_1, c_2) whose accumulated response becomes the strongest after the presentation of the complete sequence. Hence,

$$\eta_{c_1 c_2}(T) = \max_{i,j}\{\eta_{ij}(T)\} \ ,$$

where T is the sequence length. Once the winner is found, the neurons in its neighborhood should have their weights rotated towards all the inputs in the sequence. However, this is not possible because these inputs are no longer available. One way to solve this problem is to set up a short-term memory for the buffering of training samples [7]. It is unclear how such a buffer could be implemented in a biological model. However, a more plausible solution is to integrate the incoming samples and produce accumulated input

$$\xi^{ac}_{r1,r_2}(T) = \gamma \sum_{t=1}^{T} (1-\gamma)\ \xi^{T-t}_{r_1 r_2}(t) \ . \quad (2)$$

The accumulated input (Fig. 2b) is then used for weight update. We use the standard Hebbian-type Oja's rule [11]

$$\Delta\mu_{ij,r_1 r_2}(t) = \alpha(t)\ \eta_{ij}\ (\xi^{ac}_{r_1 r_2} - \mu_{ij,r_1 r_2}(t)\ \eta_{ij}) \ ,$$

which implicitly normalizes the weight vectors. This rule is applied to all neurons within the representative winner's neighborhood. (Of course, standard SOM rule would also work.) As usual, both the learning rate $\alpha(t)$ and the neighborhood radius decrease over time.

Even though learning is based on accumulated inputs (eq. 2), neurons generate responses throughout the sequences, not just at the end. This way the model exhibits true temporal recognition behavior, instead of just mapping the input sequence into a spatial representation, as previously reported in literature [12].

3 Experiments

We simulated the temporal SOM model with the following parameters: $R = 24$, $s = 0.6R$, $N = 72$, $a^2 = 1.5$, $b^2 = 160$, $T = 7$, $\gamma = 0.2$, $k = 15$. Learning rate α linearly decreased from 5 to 1 during the first half of self-organization, when also neighborhood radius shrinked linearly from 24 to 1. During the second half, the learning rate decreased to 0, and the neighborhood radius remained unchanged, i.e. equal to 1. Presentation of 6000 sequences was sufficient for self-organization: prolonging the training time did not improve the quality of the final map.

The most important of the parameter settings was that of the memory parameter γ. Range $[0.15, 0.25]$ turned out to be suitable, allowing a unit to integrate input samples within its RF without losing direction information. Too high a value of γ would reduce the map to represent purely orientation, because the "tail" in the weight profiles would be lost. On the other hand, too small a γ would make the final weight profiles more radially symmetric, which would reduce the mapping to retinotopy only.

In demonstration of the final map, methods similar to the evaluation of cortical maps were used [3]. Altogether, there were only 16 sequences (one for each direction), each consisting of a Gaussian that extended across the whole retina, and moved across the whole retina.

During a presentation of a sequence, the maximum output of every neuron was recorded. After presenting all 16 sequences, a directional response profile for every neuron (16-dimensional vector) was obtained. Neuron's *direction preference* was then found as the direction for which neuron's response was the highest. Neuron's *direction selectivity* was calculated as the ratio of neuron's response to its preferred direction and the sum of all its responses. Neuron's *orientation preference* was calculated by first summing the neuron's responses for the opposite directions, and finding the largest of these 8 sums. *Orientation selectivity* was evaluated analogically, as the ratio of the neuron's response to preferred orientation and the sum of all its orientation responses.

4 Results

The final direction and orientation map is shown in Fig. 3. Almost all units are orientation selective, and most of these are also direction selective (with varying degree of selectivity). Typically, a unit that is direction sensitive also has an orientation preference perpendicular to its preferred direction of motion. Orientations vary smoothly across the map,

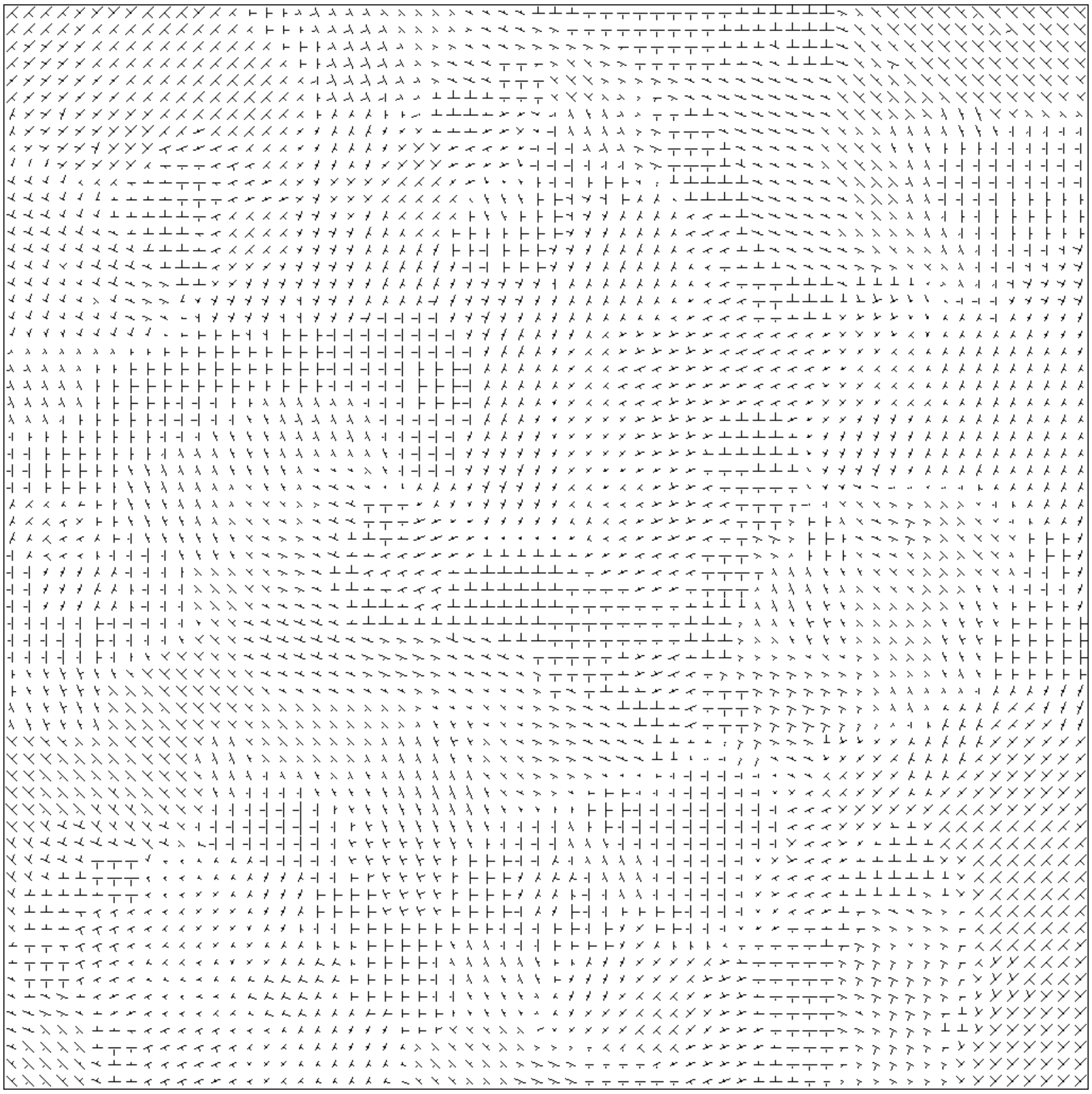

Figure 3: The self-organized direction and orientation map. Each neuron in the inner 64×64 region of the cortex (out of total 72×72) is marked with a line that identifies the neuron's orientation preference. In a similar fashion, (usually) perpendicular to it and touching its center is the shorter line that identifies neuron's directional preference. The length of a line (either orientational or directional) is proportional to neuron's selectivity. Most of the neurons are orientation selective except a few at pinwheel centers. Most of the orientation-selective neurons are also direction selective with varying degree of selectivity. In addition, at most parts of the map, an iso-orientation patch contains subregions that correspond to neurons most responsive to opposite directions, perpendicular to that orientation. All these features have been observed in biological direction maps.

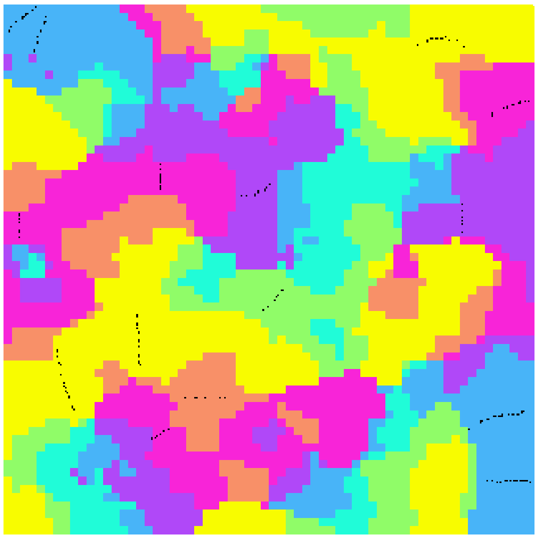

Figure 4: The larger-scale features of the orientation and direction map. The orientation preference is shown in gray scale. The map contains typical features of visuo-cortical maps such as singularities (pinwheel centers), fractures and linear zones. Almost every iso-orientation patch consists of subpatches corresponding to neurons selective to opposite directions. Major direction discontinuities are marked by black dotted lines. They often originate from pinwheels, as is found to be the case in the direction maps in the visual cortex.

and most iso-orientation patches can be subdivided into subpatches with opposite direction preferences. The orientation map has the usual structure found in the visual cortex, including pinwheel centers, fractures and linear zones (Fig. 4).

Neurons in the model can be roughly categorized into three groups, whose representative weight profiles are shown in Fig. 5. Most of the units become both orientation and direction selective, as shown by their asymmetric weight profiles (left). Some neurons are only orientation selective, with symmetric profiles (center). There are also a few nonselective neurons (right) near singularities, as observed also in biological orientation maps.

In terms of neuron's response profiles, the difference between direction selective and nonelective neurons is not that big (Fig. 6). An only-orientation-selective neuron has two peak responses of roughly equal strength; for an orientation and direction-selective neuron, one of the peaks is slightly higher. Such profiles are not surprising. The responses are de-

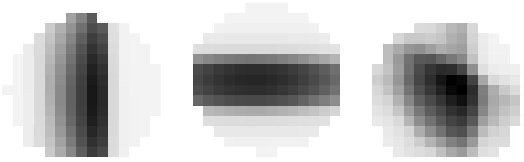

Figure 5: Typical final weight profiles of the neurons. From left to right: both direction and orientation-selective neuron (23 rows from the top and 8 columns from the left), an orientation-selective neuron (12,10), and a non-selective neuron (20,10). A direction-selective neuron typically has a longer tail from the direction to which it is most responsive, whereas a neuron selective to only orientation has a symmetric weight profile.

termined by the weight profile, which matches both directions to some degree. The difference can be made adjusted by tuning the nonlinearity parameter θ of the neuron.

Similarly to biological maps, the neurons in the model also respond to non-oriented moving stimuli, such as Gaussian spots, provided that they move in the preferred direction. The model also makes the prediction that the neuron should respond to a sequence of spots moving in a direction perpendicular to its preferred one, provided that the sequence overlaps considerably with the unit's weight profile. However, there are more such sequences moving in its preferred direction that match the neuron's weight profile well. This suggests that finding this phenomenon experimentally requires a careful study of individual responses, instead of averaging.

5 Conclusion

The model in its current form demonstrates several major characteristics observed in biological direction maps. First, most of the orientation-selective neurons are also direction selective. Second, a neuron's preference to a direction of motion is perpendicular to its preferred orientation. Third, most of the iso-orientation patches contain discontinuities that subdivide them into subpatches with opposite direction selectivity. Forth, these discontinuities have the shape of curved lines within an iso-orientation patch. The model also makes the prediction that the neuron would have a high response to a dot moving perpendicularly to its preferred direction if it

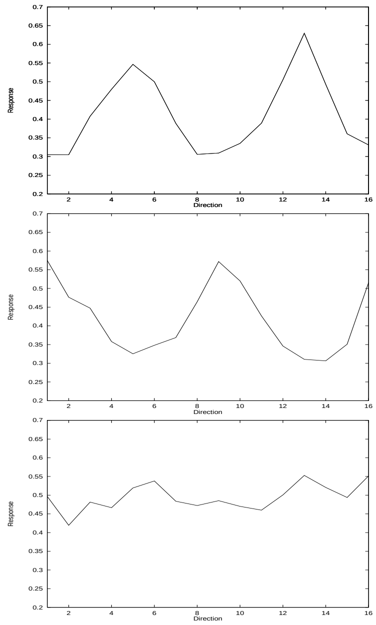

Figure 6: Response profiles of the neurons in Fig. 5. Direction "1" is "up", and index grows counterclockwise. The direction-selective neuron (top) has a clear preference, although it also responds considerably to the opposite direction. The orientation-selective one (middle) responds equally strongly to both, and the non-selective (bottom) has no preference at all.

happens to hit the peak of the weight profile accurately.

On the other hand, there are features observed in biological direction maps that are difficult to reproduce in the model. First, direction discontinuity lines do not run the whole length between pinwheel centers, nor do they tend to run across the center of iso-orientation domains. Second, the neuron's response to a direction of motion opposite to its preferred one is rather high compared to the response measured in biological direction maps. To solve these problems, it may be necessary to increase the resolution by enlarging the retina and the cortex or change the model of the neuron.

6 Acknowledgements

This research was supported by a Fulbright grant 98-22-02 awarded to the first author, in part by GAV under grant 2/5088/98 and in part by NSF under grant 115-9811478.

References

[1] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. of Neurophysiology*, 28:229–289, 1965.

[2] M. Weliky, W. H. Bosking, and D. Fitzpatrick. A systematic map of direction preference in primary visual cortex. *Nature*, 379:725–728, 1996.

[3] A. Shmuel and A. Grinwald. Functional organization for direction of motion and its relationship to orientation maps in cat area 18. *J. of Neuroscience*, 16(21):6945–6964, 1996.

[4] R. Maex and G. A. Orban. Model circuit of spiking neurons generating directional selectivity in simple cells. *J. of Neurophysiology*, 75(4):1515–1545, 1996.

[5] M. E. Sereno. *Neural Computation of Pattern Motion*. The MIT Press, 1993.

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

[7] T. Kohonen. *Self-Organizing Maps*. Springer Verlag, 1997. Second edition.

[8] K. Obermayer, H. J. Ritter, and K. J. Schulten. Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps. *Phys. Rev. A*, 45:7568–7589, 1992.

[9] T. Koskela, M. Varsta, J. Heikkonen, and K. Kaski. Time series prediction using recurrent SOM with local linear models. *Int. Journal of Knowledge-Based Intelligent Eng. Systems*, 2(1):60–68, 1998.

[10] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: Principles, Algorithms, and Applications*. MacMillan Publishing Company, 1992.

[11] E. Oja. A simplified neuron model as a principal component analyzer. *J. of Mathematical Biology*, 15:267–273, 1982.

[12] J. Kangas. Time-dependent self-organizing maps for speech recognition. In *Proc. of ICANN, Espoo, Finland*, pages 1591–1594, 1991.